



Universidade de Brasília - UnB

Instituto de Ciências Biológicas - IB

Programa de Pós Graduação em Biologia Molecular

Laboratório de Biologia Teórica e Computacional - LBTC

# Análise termodinâmica da partição de pequenos ligantes em proteínas de membrana e efeitos da concentração

Leonardo Cirqueira Pimentel

Orientador: Prof. Dr. Werner Treptow

Tese de doutorado apresentada ao Programa de Pós-Graduação em Ciências Biológicas –  
Biologia Molecular do Instituto de Ciências Biológicas da Universidade de Brasília.

Brasília

2023

## Abstract

Many biological processes are regulated by interactions with small ligands. As occupancy states and sites number grow, it becomes harder for nowadays available methods to characterize them. This problem can be solved by considering a new approach that molecule interaction is a partition to a phase that corresponds to the protein interface. In this work, the two-site approach is used for obtaining a partition coefficient from flooding simulations. From the partition coefficient, many properties in different concentrations can be reconstructed, e.g. titration curves and tridimensional ligand densities. Furthermore, the partition process is separated into steps from a thermodynamical cycle and each step and concentration influence are analyzed carefully. The results obtained were satisfactory close when compared to independent systems. The development of the tools in this work has a very promising application potential, being very useful to a myriad of small molecules interactions with proteins.

## Resumo

A interação de pequenas moléculas com proteínas regula diversos processos biológicos. Interações complexas possuem muitos sítios e estados de ocupância possíveis, e os métodos atuais costumam ser insuficientes para realizar uma boa caracterização. Esse problema é resolvido utilizando uma nova abordagem, a qual descreve a interação como a partição para uma fase correspondente à interface proteica. Nesse trabalho, a abordagem da partição em dois sítios é utilizada para obter o coeficiente de partição em simulações de *flooding* e a partir dele obter outras propriedades em diversas concentrações, como curvas de titulação e densidades tridimensionais. Além disso, o processo de partição será separado em etapas através de um ciclo termodinâmico e a influência de cada etapa juntamente com o papel da concentração serão analisados com detalhes. Os resultados obtiveram valores satisfatórios quando comparados com outros sistemas independentes. O desenvolvimento das ferramentas deste trabalho possui potencial de aplicação bastante promissor, sendo útil para diversas interações de pequenas moléculas com proteínas.

## **Agradecimentos**

Agradeço primeiramente à minha família: meus pais Ana e Laécio, por tudo que fizeram por mim, e aos meus irmãos Lucas e Luana. Cada parte dessa conquista também pertence a vocês!

Também sou muito grato à Mari, minha companheira de jornada, principalmente pela compreensão, incentivo e parceria. Estar ao seu lado tornou tudo muito mais alegre, leve e me fez querer dar o meu melhor, até em situações que eu nem sabia que era possível. Muito obrigado meu amor!

Ao meu orientador Werner, por toda a vivência proporcionada nos dias de laboratório, tanto nos momentos de descontração quanto pelo imenso aprendizado que me proporcionou nesse doutorado. Obrigado pela confiança!

Aos amigos do LBTC: Alessandra, Camila, Caio, Fiorote, João, Letícia, Miguel, Mônica, Natália e Vinícius, pelos diversos momentos. Tive o privilégio muito grande de conviver com vocês e aproveitar os momentos o que esse time incrível pôde oferecer. Vocês fizeram toda diferença. Devo agradecer especialmente à Letícia, pela paciência e ajuda em todos esses anos de LBTC.

A todos os amigos da época de escola e da graduação que mantenho contato até hoje. Estar rodeado de boas pessoas faz muita diferença.

Agradeço também aos professores da banca de avaliação deste trabalho pela disponibilidade e contribuições.

Ao Programa de Pós-Graduação em Biologia Molecular, CAPES, FAPDF e CNPq pelo apoio financeiro, que é indispensável pela realização de qualquer atividade.

Por último, quero agradecer à Universidade de Brasília, por ser esse ambiente incrível que transformou minha vida.

# Sumário

<b>I</b>	<b>Introdução</b>	<b>16</b>
I.1	Interação de pequenas moléculas com proteínas . . . . .	16
I.2	Alosteria e descrições macroscópicas . . . . .	17
I.3	Interações de baixa afinidade . . . . .	19
I.4	Modelos de estudo . . . . .	22
<b>II</b>	<b>Objetivos</b>	<b>24</b>
<b>III</b>	<b>Teoria e métodos</b>	<b>26</b>
III.1	Simulações de dinâmica molecular . . . . .	26
III.1.1	Campos de força . . . . .	27
III.1.2	Integração das equações de movimento . . . . .	29
III.1.3	Tratamento das interações de curta e longa distância . . . . .	30
III.1.4	Condições periódicas de contorno . . . . .	31
III.2	Cálculos de energia livre por FEP . . . . .	33
III.3	Descrição termodinâmica das interações . . . . .	34
III.3.1	Coeficiente de partição . . . . .	34
III.3.2	Energia livre da interação . . . . .	39
III.3.3	Restrição espacial da proteína . . . . .	41
III.3.4	Passagem dos ligantes para o vácuo . . . . .	42
III.3.5	Mudança de volumes no vácuo . . . . .	44
III.3.6	Passagem do ligante para o sítio . . . . .	45
III.3.7	Liberação de restrições na proteína . . . . .	46

III.3.8	Reconstrução da energia livre da partição . . . . .	47
III.3.9	Energia livre da partição em condições diluídas . . . . .	48
III.3.10	Projeção de propriedades para outras concentrações . . . . .	49
III.4	Protocolos . . . . .	49
III.4.1	Simulações em geral . . . . .	49
III.4.2	Simulações de FEP . . . . .	50
III.4.3	Simulações com membrana . . . . .	50
III.4.3.1	Criação dos sistemas equilibrados . . . . .	50
III.4.3.2	Simulações de FEP na membrana . . . . .	51
III.4.4	Simulações de <i>flooding</i> Kv1.2-sevoflurano . . . . .	52
III.4.4.1	Simulações de FEP na proteína . . . . .	53
III.4.5	Simulações de <i>flooding</i> da albumina e tricloroetanol . . . . .	54
III.4.6	Análises . . . . .	54
III.4.6.1	Obtenção dos volumes . . . . .	54
III.4.6.2	Obtenção das distribuições de probabilidades e número médio . . . . .	54
III.4.6.3	Obtenção das densidades tridimensionais . . . . .	55
III.4.6.4	Análise da estabilidade por RMSD . . . . .	55
<b>IV</b>	<b>Resultados e discussão</b>	<b>56</b>
IV.1	Sistemas com membrana . . . . .	56
IV.1.1	Equilibração . . . . .	56
IV.1.2	Cálculos de FEP e energia de dessolvatação . . . . .	57
IV.2	Simulações de <i>flooding</i> . . . . .	60
IV.2.1	Equilibração dos sistemas . . . . .	60

IV.2.2	Propriedades das simulações . . . . .	62
IV.3	Propriedades termodinâmicas da interação . . . . .	65
IV.3.1	Coefficiente de partição . . . . .	65
IV.3.2	Perturbação na estrutura interna da proteína . . . . .	66
IV.3.3	Energia de ligação sítio-vácuo . . . . .	68
IV.3.4	Reconstrução do coeficiente de partição . . . . .	70
IV.4	Energia total de ligação . . . . .	72
IV.5	Projeções para outras concentrações . . . . .	74
IV.5.1	Projeções de número médio . . . . .	74
IV.5.2	Projeções de densidades tridimensionais . . . . .	75
IV.5.3	Interação da albumina e tricloroetanol . . . . .	78
<b>V</b>	<b>Conclusões e perspectivas</b>	<b>81</b>
<b>VI</b>	<b>Anexos</b>	<b>90</b>

## Lista de Figuras

- 1 Pequenas moléculas e seu efeito regulatório. Cada ponto representa uma molécula, sua posição no eixo vertical a quantidade de interações ativadoras e no eixo horizontal interações inibitórias. É possível ver que o ATP está envolvido em mais processos, participando tanto como ativador ou inibidor. Adaptado de *The molecular switch*, Philips, 2020[1]. . . . . 16
- 2 Curvas de titulação. Nas curvas de titulação é possível ver qual a porcentagem de receptores ocupados ( $p_{bound}$ ) em relação à concentração. Note que  $K_d$  é o valor de concentração onde metade dos receptores estão ocupados. Quanto menor  $K_d$ , maior a afinidade da proteína pelo substrato. Adaptado de *The molecular switch*, Philips, 2020[1]. . . . . 18
- 3 Procedimento de *docking*. O *docking* é uma ferramenta computacional que varre a superfície da proteína (branco) para encontrar os potenciais sítios de ligação de uma molécula (vermelho). . . . . 20
- 4 Distribuição espacial da molécula em diversas concentrações. Em azul, estão as regiões de maior densidade sugeridas pelo procedimento *docking-FEP* ao longo da superfície proteica (branco) para diversas concentrações. Adaptado de Stock *et al.*, 2017[17]. . . . . 21
- 5 Estado inicial e final da simulação de *flooding*. Uma simulação de *flooding* é uma simulação de dinâmica molecular com presença explícita do ligante no sistema. As interações ocorrem sem potencial artificial, logo as propriedades observadas correspondem às propriedades do equilíbrio. Na esquerda está o sistema inicial de *flooding*, com as moléculas do ligante presentes na água. Após 300 ns de interação, é possível notar que os ligantes estão na maior parte presentes na região transmembrânica. Vermelho: ligante; branco: proteína; laranja: membrana; azul: água. . . . . 22

6	Condutância relativa ( $G/G_{max}$ ) do canal Kv1.2 em função da voltagem na ausência (preto) e na presença (vermelho) de sevoflurano a 1mM. $G_{max}$ é a condutância máxima antes da exposição ao sevoflurano. Adaptado de Liang <i>et al.</i> [11]. . . . .	23
7	Estrutura molecular das pequenas moléculas utilizadas no estudo. Esquerda: Sevoflurano. Direita: Tricloroetanol. . . . .	23
8	Esquematização das condições periódicas de contorno. No centro está o sistema original, rodeado em todos os eixos pelas suas réplicas. A convenção da imagem mínima pode ser notada quando um átomo atravessa a fronteira do sistema e esse movimento é acompanhado pela aparição de outro átomo do lado oposto do sistema, que veio da réplica da extremidade oposta. Retirado de <a href="http://isaacs.sourceforge.net/phys/pbc.html">http://isaacs.sourceforge.net/phys/pbc.html</a> . Acesso em 9/12/2019. . . . .	32
9	Fases da partição do sevoflurano no sistema de <i>flooding</i> . A fase correspondente ao sítio de ligação (roxo) corresponde à região de até 5Å de distância da proteína (branco), enquanto a fase do reservatório corresponde a todo o restante. No caso de sistemas com membrana, o reservatório é composto pela região membranosa (laranja) e a água (azul) que estão a mais de 5Å de distância da proteína. . . . .	35
10	Processo de partição. Nas simulações de <i>flooding</i> , o sistema é composto inicialmente pela proteína vazia (em rosa) e os ligantes (pontos azuis) no reservatório. Com o passar do tempo, alguns ligantes particionam-se para o sítio da interface proteica (etapa K), delimitada na imagem pelo círculo tracejado. O processo de partição pode ser descrito através de um ciclo hipotético com aplicação de potenciais artificiais de cinco etapas: I - restrição da estrutura proteica; II - dessolvatação dos ligantes; III - restrições translacionais; IV - inserção do ligante no sítio; V - liberação das restrições. No caso da partição de n ligantes, o ciclo hipotético é repetido n vezes. . . . .	41



11	Sistemas com membrana e sevoflurano. Seção transversal dos sistemas de <i>flooding</i> com membrana após a equilibração para diversas concentrações. A composição do sistema é de uma bicamada lipídica de POPC (ciano), moléculas de água (azul) e sevoflurano (vermelho). . . . .	56
12	Propriedades da membrana na presença de diversas concentrações de sevoflurano. A - Quantidade relativa de moléculas de ligante que permanecem na região da membrana ao longo do tempo. B - Densidade espacial do sevoflurano ao longo do eixo $z$ . Para ambos os gráficos, cada cor representa um sistema com concentração específica: Verde - Sevoflurano 1mM; Azul - Sevoflurano 25mM; Ciano - Sevoflurano 50mM; Cinza - Sevoflurano 75mM; Rosa - Sevoflurano 100mM; Preto - Sevoflurano 150mM. A região da membrana é delimitada pelos átomos de fósforo da cadeia polar do POPC. . . . .	57
13	Energias de dessolvatação do sevoflurano a partir da membrana de POPC obtidas através dos cálculos de FEP (pontos vermelhos) e valores em função da concentração efetiva no reservatório $\frac{N-n}{V-\nu}$ , obtidos através da regressão do modelo $W_{memb}^* = \bar{\mu} - h^* \times \frac{N-n}{V-\nu} \times c$ (linha preta), onde $\bar{\mu} = 4,3$ kcal/mol, $h^* = 1,9$ kcal/mol e $C^0$ é o inverso da concentração padrão ( $c = (1 \text{ mol/L})^{-1}$ ). . . . .	59
14	Sistema final de uma simulação de <i>flooding</i> do canal de potássio Kv1.2 (branco) com 150mM do anestésico geral sevoflurano (vermelho). O canal iônico está embebido em uma bicamada lipídica de POPC (laranja). A água está representada em azul e o íons foram ocultados para melhor visualização. . . . .	60
15	Valores de RMSD ao longo do tempo de simulação dos sistemas de <i>flooding</i> . Verde - Sevoflurano 1mM; Azul - Sevoflurano 25mM; Ciano - Sevoflurano 50mM; Cinza - Sevoflurano 75mM; Rosa - Sevoflurano 100mM; Preto - Sevoflurano 150mM. . . . .	61

16	Quantidade relativa de moléculas de ligante que permanecem na região da solução aquosa, membrana ou proteína ao longo do tempo. A região da membrana é delimitada pelos átomos de fósforo da cadeia polar do POPC, enquanto as moléculas ligadas à proteína são moléculas que possuem algum átomo a uma distância de até 5Å da proteína. No caso de essas condições serem mutualmente satisfeitas, a molécula é considerada ligada à proteína. Vermelho: ligados à proteína; Verde: presentes na membrana; Azul: presentes na água. . . . .	61
17	Probabilidades dos estados de ocupância do canal Kv1.2 com o anestésicos sevoflurano em diferentes concentrações. Verde - Sevoflurano 1mM; Azul - Sevoflurano 25mM; Ciano - Sevoflurano 50mM; Cinza - Sevoflurano 75mM; Rosa - Sevoflurano 100mM; Preto - Sevoflurano 150mM. . .	63
18	Estrutura média da proteína (em branco) e densidades tridimensionais (em rosa) obtidas a partir dos últimos 100 ns das simulações de <i>flooding</i> . Estão representados os valores de densidades maiores que 0.001 moléculas/Å <sup>3</sup> . .	64
19	Perfis de RMSD de uma molécula de sevoflurano nos diversos ambientes da proteína. De superior ao inferior: Região onde a molécula está presente (wt - água, mb - membrana, pt - proteína); RMSD translacional, rotacional e configuracional. É possível ver que não existe nenhuma correlação entre o valor de RMSD e a região onde o anestésico está presente. . . . .	65
20	Perturbação estrutural na proteína. A - Densidades de probabilidade de ocorrência do RMSD ( $\rho(RMSD)$ ) em diversas concentrações. B - Trabalho de perturbação da estrutura da proteína calculado a partir das simulações de <i>flooding</i> (pontos vermelhos) e regressão (linha preta) segundo modelo quadrático $\Delta W_P = h_p C^2$ , onde $h_p = 200$ kcal/mol e $C$ é a razão da concentração pela concentração padrão . . . . .	67

21	Estados de agregação do sevoflurano em diferentes concentrações. A - Probabilidade de ocorrência do estado de agregação k ao longo do todo tempo da simulação de <i>flooding</i> . B - Energia média calculada via FEP das moléculas de sevoflurano no estado de agregação k. Erro obtido por propagação. . . . .	69
22	Energias de dessolvatação sítio-vácuo de acordo com a concentração. As energias foram calculadas por cálculos de FEP em diversas regiões da proteína e diversos estados de agregação do ligante. Pontos vermelhos: energia calculadas diretamente por FEP. Linha preta: regressão linear utilizando a função $W^{**} = \bar{w} - h^{**} \frac{\langle n \rangle}{\nu} c$ (equação 37), onde $\bar{w} = 5,0$ kcal/mol, $h^{**} = 0,9$ kcal/mol e $c = (1 \text{ mol/L})^{-1}$ . . . . .	70
23	Coeficientes de partição. Preto: Coeficientes de partição calculados a partir das simulações de <i>flooding</i> ( $\varphi_{\text{fld}} = \left(\frac{N - \langle n \rangle}{V - \nu}\right)^{-1} \frac{\langle n \rangle}{\nu}$ ). Vermelho: obtidos a partir dos cálculos de FEP ( $\varphi_{\text{FEP}} = e^{-\beta(W^* - W^{**})}$ ). A linha tracejada horizontal representa a média dos coeficientes obtidos pelo <i>flooding</i> . . . . .	71
24	Contribuições energéticas em função da concentração do sistema. Decomposição de $\Delta W$ nos termos de perturbação da proteína ( $\Delta W_P$ , em laranja), entropia ( $\Delta W_S$ , em vermelho) e energias de dessolvatação ( $\Delta W_L$ , em azul). As linhas sólidas representam as regressões quadráticas de melhor ajuste aos pontos. . . . .	72
25	Variações de energia para a ligação em função da concentração. Vermelho: Variação de energia livre proveniente dos parâmetros termodinâmicos entropia, entalpia e perturbação na proteína. Preto: Variação de energia livre calculada a partir do coeficiente de partição. . . . .	73
26	Número médio de ligantes no sítio. Números obtidos diretamente pelas simulações de <i>flooding</i> (vermelho) e projeção realizada pela equação $\langle n \rangle = \bar{\rho} V \left(1 + \frac{V - \nu}{\nu \varphi}\right)^{-1}$ . Todos os parâmetros (volumes e coeficiente de partição) foram originados da simulação de 150mM. . . . .	75

27	Densidades tridimensionais. Esquerda para direita: densidades tridimensionais obtidos a partir das simulações de <i>flooding</i> , densidades oriundas das projeções de $\langle n \rangle$ e do mapa unitário de 150mM ( $\bar{\rho}'(\mathbf{R})$ ) e sobreposição das nuvens. . . . .	76
28	Densidades tridimensionais obtidos pela simulação de <i>flooding</i> de 5mM. A - Índices de sobreposição das densidades tridimensionais obtidas em diferentes janelas de tempo em relação à projeção para 5mM, sem simetria (vermelho) e após a simetria (cinza). B - Densidades tridimensionais observadas diretamente das simulações de <i>flooding</i> em diversas janelas temporais e a projeção realizada. . . . .	77
29	Partição do tricloroetanol no sistema de albumina e água. A - Número de ligantes e RMSD em comparação à estrutura inicial da albumina ao longo do tempo. B - Estruturas de equilíbrio para as simulações de <i>flooding</i> com 50mM e 100mM. . . . .	78
30	Propriedades calculadas da interação do tricloroetanol com a albumina. A - Pontos vermelhos: Número médio obtido diretamente das simulações de <i>flooding</i> . Linha preta: Projeções de número médio a partir do coeficiente de partição obtido a 100mM ( $\ln \phi = 3,21$ ). Nota-se que a regressão possui uma acurácia muito boa quanto ao número médio em 50mM. B - Pontos vermelhos: Energia livre de ligação nas concentrações de 50 e 100 mM. A energia foi calculada utilizando a estimativa do coeficiente de partição com as contribuições entrópicas ( $\Delta W = \Delta W_S - \langle n \rangle \beta \ln \phi$ ). Linha preta: regressão quadrática das energias em função das concentrações ( $\Delta W = -583,45\bar{\rho}/C^0$ ). . . . .	79
31	Projeções da distribuição espacial do tricloroetanol na albumina sérica. As projeções foram obtidas a partir do reescalonamento do mapa unitário de densidade tridimensional da simulação de 100mM. . . . .	80

- 32 Probabilidades de abertura do Kv1.2 em diversas condições. Curva preta - probabilidades em função da voltagem para o Kv1.2 sem anestésico. Curva azul - probabilidades obtidas experimentalmente com a presença de 1mM de sevoflurano. Curvas vermelha, verde e amarela - Curvas projetadas teoricamente para diferentes concentrações. Na caixa inferior estão representados os valores de  $V_{1/2}$  das curvas pxV. Adaptado de Stock *et al.* [16] . . . . . 82

## Lista de Tabelas

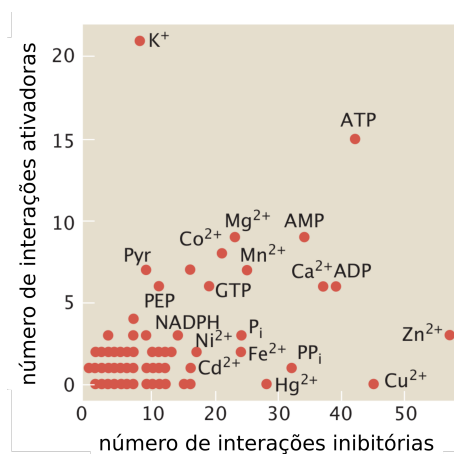
1	Concentração desejada, fração molar ligante-POPC, número de ligantes adicionados e tempo de simulação nos sistemas com sevoflurano. . . . .	51
2	Concentração da desejada, fração molar ligante-POPC, e número total de ligantes dos sistemas de <i>flooding</i> com sevoflurano. . . . .	53
3	Média ( $\bar{z}$ ) e desvio padrão ( $\sigma_z$ ) da posição em $z$ dos ligantes presentes na região superior da membrana, após o equilíbrio. . . . .	57
4	Variações de energia para o desacoplamento do sevoflurano na membrana através de cálculos de FEP ( $W_{memb}^*$ ) e erro estatístico associado, estimadas pelo método SOS. A variação de energia calculada corresponde à energia necessária para a transferência de uma molécula de sevoflurano partindo do ambiente para o vácuo. . . . .	58
5	Energia média de dessolvatação resultante para os dois componentes do reservatório ( $W^*$ ) e erro estatístico associado, em kcal/mol. A energia resultante é oriunda do escalonamento das energias de dessolvatação e do volume de cada componente. . . . .	59
6	Volumes (em $\text{\AA}^3$ ) do sistema calculados a partir do início das simulações de <i>flooding</i> . Legenda: $V_{total}$ = volume total; $V_{prot}$ = volume da proteína; $\nu$ = volume do sítio proteico; $V_{memb}$ = volume da membrana e $V_{wat}$ = volume da água. . . . .	62
7	Número médio ( $\langle n \rangle$ ) de ligantes em cada concentração, obtidos diretamente nas simulações de <i>flooding</i> , a partir dos 100 ns de simulação. . .	63
8	Coefficientes de partição ( $\varphi$ ) de cada concentração, obtidos diretamente das simulações de <i>flooding</i> , a partir dos 100 ns de simulação. . . . .	66
9	Perturbações estruturais realizadas pela proteína ( $\Delta W_P$ ), em kcal/mol, para diversas concentrações. . . . .	66

10	Energias de dessolvatação partindo da interface proteica para cada concentração. Valores em kcal/mol. . . . .	70
11	Diferença das energias de dessolvatação ( $W^* - W^{**}$ , em kcal/mol) e coeficientes de partição ( $\phi_{FEP}$ ) de cada concentração, obtidos pela diferença nas energias de dessolvatação ( $\phi_{FEP} = e^{-\beta(W^* - W^{**})}$ ). . . . .	71
12	Propriedades no equilíbrio das simulações de <i>flooding</i> da albumina bovina sérica com o sedativo tricloroetanol. . . . .	79

# I Introdução

## I.1 Interação de pequenas moléculas com proteínas

A interação de proteínas com pequenas moléculas é um campo de estudo de extrema importância para a compreensão da vida: um receptor pode estar ativado ou inativado; um canal aberto ou fechado; a proteína pode estar fosforilada ou não. Cada estado corresponderá a uma sinalização diferente, e por consequência, respostas diferentes por parte do organismo. Na maioria das vezes, o fator determinante para qual estado a proteína assume é a ligação com pequenas moléculas. É possível encontrar inúmeros exemplos onde a associação com essas moléculas de menor escala são importantes: regulação gênica, motilidade celular, controle metabólico, sinalização celular, entre outros. O controle da estrutura da proteína inclusive pode ser visto como um elemento lógico que determina a ocorrência ou não de algum processo, permitindo a adaptação da vida às mais diversas condições[1]. Muitas das pequenas moléculas já são conhecidas pelo controle de diversos processos celulares, como ATP, cofatores e íons, e a Figura 1 mostra a presença dessas moléculas com o tipo de interação realizada.



**Figura 1:** Pequenas moléculas e seu efeito regulatório. Cada ponto representa uma molécula, sua posição no eixo vertical a quantidade de interações ativadoras e no eixo horizontal interações inibitórias. É possível ver que o ATP está envolvido em mais processos, participando tanto como ativador ou inibidor. Adaptado de *The molecular switch*, Philips, 2020[1].

A necessidade de caracterizar quais moléculas participam da regulação desses pro-



cessos é tamanha que estudos vêm se preocupando em mapear as moléculas e seus efeitos nas proteínas em estruturas complexas chamadas de interatomos[2]. Um ligante pode agir tanto para estabilizar um estado quanto para desestabilizá-lo. A interação de pequenas moléculas com proteínas não se limita somente às enzimas: estima-se que cerca de 40% dos pequenos fármacos atualmente utilizados possuem interação com proteínas de membrana, como canais, bombas e receptores celulares[3]. Além disso, esses ligantes também são cruciais para outros processos, como quimiotaxia[4].

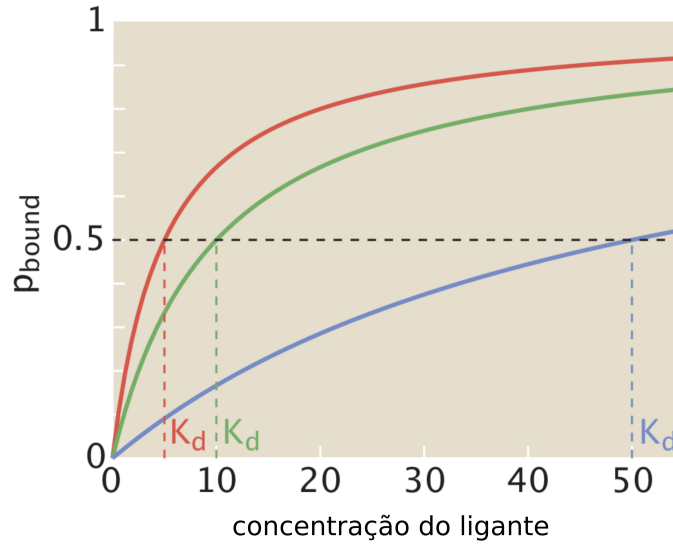
A principal maneira de representar a ligação de uma pequena molécula em uma proteína é por uma curva de titulação, mostrada na Figura 2. Na curva de titulação, é mostrada a fração do receptor que interage com o ligante em função da concentração disponível de ligantes. Diversas informações podem ser obtidas a partir dessa curva, como a comparação entre condições, obtenção da concentração necessária para saturar metade dos receptores e até o coeficiente de cooperatividade[5]. Um dos primeiros sistemas de interação utilizados para ilustrar a importância da interação de pequenas moléculas com proteínas foi a interação da mioglobina com o oxigênio[6]. Anos mais tarde, foi possível descrever mais minuciosamente essa interação e obter valiosas informações a partir de curvas de titulação, como a influência do pH e temperatura na interação, além de até realizar uma regressão da curva de titulação[7].

## I.2 Alosteria e descrições macroscópicas

A medida que outras caracterizações ocorreram, observou-se que a modulação por pequenos ligantes pode ser muito complexa. Por exemplo, foi visto que algumas moléculas possuíam efeitos na proteína, porém sem exibir o comportamento de uma inibição competitiva, o que sugeriu outra maneira de interação. Assim surgiu a ideia da alosteria, uma ação à distância da molécula, onde sua interação induziria uma mudança conformacional na proteína, incluindo no sítio de ligação[1].

A descrição analítica da interação de pequenos ligantes com proteínas é semelhante ao de um equilíbrio químico:





**Figura 2:** Curvas de titulação. Nas curvas de titulação é possível ver qual a porcentagem de receptores ocupados ( $p_{bound}$ ) em relação à concentração. Note que  $K_d$  é o valor de concentração onde metade dos receptores estão ocupados. Quanto menor  $K_d$ , maior a afinidade da proteína pelo substrato. Adaptado de *The molecular switch*, Philips, 2020[1].

onde R é o receptor proteico, L é o ligante, RL é o complexo ligante-receptor e K é a constante de equilíbrio. A obtenção de K é muito importante para descrever parâmetros termodinâmicos da interação, como a variação da energia livre padrão ( $K = e^{-\frac{\Delta G^0}{RT}}$ ). Experimentalmente, a constante de ligação pode ser obtida de diversas maneiras, por calorimetria, principalmente pelo método ITC[8] ou pelas curvas de saturação[9]. Em trabalhos envolvendo abordagens computacionais, é preferível encontrar K partindo da abordagem microscópica:

$$K = \frac{[RL]}{[R][L]} \quad (1)$$

$$K = \frac{n_{RL}}{n_L} \frac{1}{\bar{\rho}}$$

onde  $n_{RL}$ ,  $n_L$  e  $\bar{\rho}$  são os números de receptores complexados com o ligantes, vazios e a densidade do ligante, respectivamente. Dividindo a equação acima pelo número total de receptores, obtemos:

$$K = \frac{p_{occ}}{p_{free}} \frac{1}{\bar{\rho}} \quad (2)$$

$$\frac{p_{occ}}{p_{free}} = K \bar{\rho}$$

onde  $p_{occ}$  e  $p_{free}$  são as probabilidades de um receptor estar ocupado, e vazio, respectiva-

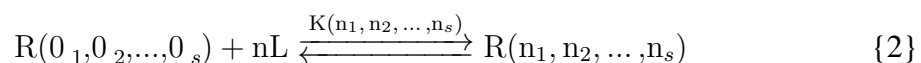
mente. A razão de probabilidades também é abordada na teoria da mecânica estatística como uma função da diferença de energia entre dois estados ( $\Delta\mu$ ). Logo:

$$e^{-\beta\Delta\mu} = \frac{p_{occ}}{p_{free}} = K \bar{\rho} \quad (3)$$

onde  $\beta = 1/RT$ . Essa equivalência evidencia o quanto o cálculo de K pode ser flexível. Cada tipo de experimento possuirá uma maneira mais eficiente de encontrar a constante de ligação. Por exemplo, a abordagem estatística é bastante utilizada em trabalhos envolvendo simulações de Monte Carlo e de dinâmica molecular, pois é mais fácil realizar o mapeamento dos estados ocupado e desocupado[10]. O cálculo da constante através da diferença de energia também é amplamente usado em técnicas computacionais que calculam a variação de energia livre, como o *Docking*[11], *Linear Interaction Energy* (LIE)[12, 13] e o *Free Energy Perturbation* (FEP)[14, 15, 16].

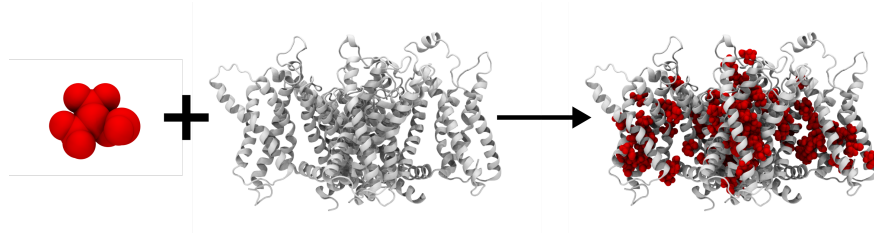
### I.3 Interações de baixa afinidade

As interações de baixa afinidade têm uma caracterização bastante desafiadora. Devido à pouca especificidade ao longo da superfície da proteína, os pequenos ligantes podem se associar em diversas regiões, o que dificulta a busca por sítios através de ferramentas computacionais, pois o espaço amostral é muito grande e homogêneo. Para solucionar essa questão, foi desenvolvido um protocolo que combina técnicas computacionais com o intuito de facilitar a busca dos sítios e calcular a respectiva constante de equilíbrio para as ditas interações de baixa afinidade (ou reversíveis)[17]. O primeiro passo foi encontrar os candidatos a sítios de ligação utilizando a técnica de *docking* molecular. Caracterizados os  $s$  sítios, o próximo objetivo consistiu no cálculo da constante de ligação agregada da seguinte reação:



onde  $R(0_1, 0_2, \dots, 0_s)$  é o receptor vazio nos  $s$  sítios,  $nL$  são os  $n$  ligantes livres e  $R(n_1, n_2, \dots, n_s)$  é o receptor com  $n_1$  moléculas no sítio 1,  $n_2$  moléculas no sítio 2, e assim por diante e  $K(n_1, n_2, \dots, n_s)$  é a constante de equilíbrio agregada.

Essa constante agregada corresponde a um estado de múltiplas ocupâncias em diversos sítios. Caso os sítios possuíssem uma correlação, seria necessário obter todas as com-



**Figura 3:** Procedimento de *docking*. O *docking* é uma ferramenta computacional que varre a superfície da proteína (branco) para encontrar os potenciais sítios de ligação de uma molécula (vermelho).

binacões possíveis dos estados, o que seria impraticável. Para simplificar, considerou-se que os sítios são independentes, podendo assim realizar a fatoração:

$$K(n_1, n_2, \dots, n_s) = K(n_1) \times K(n_2) \times \dots \times K(n_s) \quad (4)$$

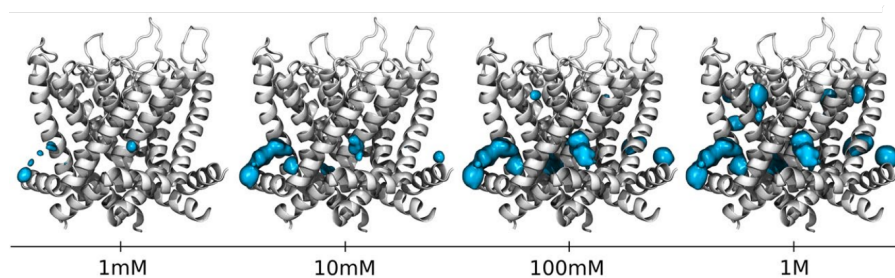
onde  $K(n_s)$  é a constante de ligação de  $n$  ligantes no sítio  $s$  com todos os outros vazios. As constantes para os sítios individuais foram encontradas a partir de dois cálculos de FEP. No primeiro cálculo é obtido a energia de retirar a molécula da água ( $\mu$ ), enquanto no segundo obtém-se a energia de retirada de um ligante do sítio sugerido ( $W_i^*$ ). Entretanto, para os cálculos de FEP amostrarem a transformação desejada é necessária uma aplicação de um potencial artificial que mantém o ligante dentro do sítio com constante de força  $k_i$ . A constante de ligação depende desses três parâmetros, possuindo a forma:

$$K(i) = \frac{2\pi}{\beta k_i} e^{-\beta(W_i^* - \mu)} \quad (5)$$

Como um sítio pode comportar até  $n_s$  moléculas, a constante de ligação até a ocupância máxima é dada por  $K(n_s) = \prod K(i|i-1)$ , onde  $K(i|i-1)$  é a constante relacionada à ligação do  $n$ -ésimo ligante com o sítio pré ocupado por  $n-1$  ligantes.

Com as constantes agregadas, foi possível realizar projeções de propriedades para qualquer concentração desejada, incluindo as densidades tridimensionais do ligante, mostrados na Figura 4.

O procedimento *docking-FEP* mostra-se bem-sucedido quando existe um pequeno número de estados possíveis. A medida que a interação degenerada possui novos estados de ocupância, a consideração de que os sítios são independentes vai tornando-se improvável. Assim, seria necessária a amostragem de todos os estados possíveis, que crescem numa ordem geométrica, o que torna o cálculo das constantes de ligação via FEP im-



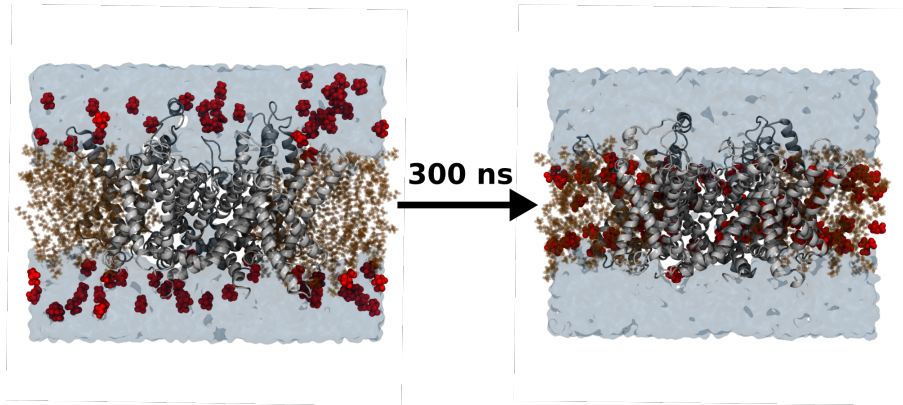
**Figura 4:** Distribuição espacial da molécula em diversas concentrações. Em azul, estão as regiões de maior densidade sugeridas pelo procedimento *docking-FEP* ao longo da superfície proteica (branco) para diversas concentrações. Adaptado de Stock *et al.*, 2017[17].

praticável. Para contornar essas limitações, deve-se analisar o fenômeno sob uma nova abordagem. A interação entre as moléculas e a proteína pode ser vista como um fenômeno de partição entre duas fases: a superfície proteica e o reservatório.

Nesse trabalho, a caracterização como um processo de partição será aproveitada para desenvolver um conjunto de ferramentas teóricas nas quais seria possível calcular uma propriedade análoga a uma constante de ligação agregada a partir de altas concentrações e realizar os estudos das propriedades em concentrações fisiológicas. Além disso, também foi estudada a influência da concentração nessas propriedades a partir da decomposição dos fatores que influenciam a ligação de uma molécula em um sítio: afinidade pelo reservatório, perturbação estrutural da proteína e afinidade pela proteína. Ao final, será possível obter de maneira muito simples diversas propriedades do regime diluído - número de ligantes complexados à proteína, densidades tridimensionais, variação de energia livre - ao entender o comportamento de cada termo em função da concentração, sem necessidade de simulações adicionais.

As propriedades em altas concentrações serão obtidas a partir de simulações de *flooding*, que são simulações de dinâmica molecular da proteína e presença do ligante numa concentração fixa. No início do *flooding*, as moléculas estão dispersas no ambiente e movimentam-se de acordo com suas interações, particionando-se na membrana e certas regiões da superfície proteica, como mostrado na Figura 5. No *flooding* não existe a adição de potenciais artificiais, logo a permanência dos ligantes em alguma região proteica indica uma região de maior afinidade, tornando as propriedades mais fidedignas. A principal vantagem desse tipo de simulação consiste em obter distribuições exatamente de

acordo com as interações no equilíbrio. Porém, a qualidade dos resultados está atrelada ao número de eventos de amostragem, tornando assim o *flooding* um procedimento muito custoso computacionalmente, exigindo simulações em escalas de mais de 100ns, podendo chegar a escala de  $\mu s$ .

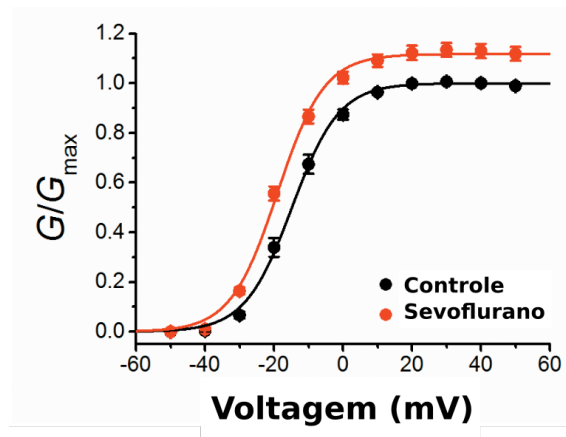


**Figura 5:** Estado inicial e final da simulação de *flooding*. Uma simulação de *flooding* é uma simulação de dinâmica molecular com presença explícita do ligante no sistema. As interações ocorrem sem potencial artificial, logo as propriedades observadas correspondem às propriedades do equilíbrio. Na esquerda está o sistema inicial de *flooding*, com as moléculas do ligante presentes na água. Após 300 ns de interação, é possível notar que os ligantes estão na maior parte presentes na região transmembrânica. Vermelho: ligante; branco: proteína; laranja: membrana; azul: água.

## I.4 Modelos de estudo

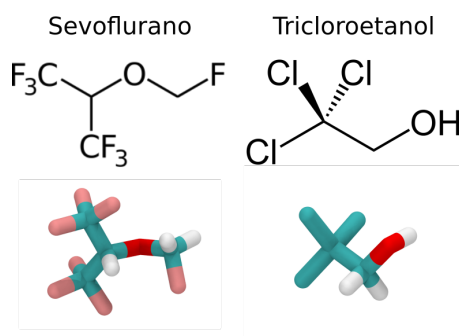
No presente trabalho, será utilizado como modelo de interação entre pequenos ligantes e proteínas o anestésico geral sevoflurano e o canal iônico dependente de voltagem Kv1.2. A escolha dessa interação se dá pelo fato da interação com sevoflurano cumprir os requisitos da interação de pequenos ligantes em sítios degenerados: o sevoflurano está presente em diversas regiões da proteína[18, 19, 16], as interações ocorrem em concentrações baixas[20] e existe a multiplicidade dentro de cada sítio[17]. A ativação do sevoflurano e o canal Kv1.2 é mostrada na Figura 6.

A consistência da teoria e dados será realizada comparando as projeções diretamente com os dados de simulações independentes de *flooding* do sevoflurano com o Kv1.2 aberto em diversas concentrações. Além disso, o modelo desenvolvido também será utilizado para estudar outra interação: da albumina bovina sérica com o sedativo tricloroetanol,



**Figura 6:** Condutância relativa ( $G/G_{max}$ ) do canal Kv1.2 em função da voltagem na ausência (preto) e na presença (vermelho) de sevoflurano a 1mM.  $G_{max}$  é a condutância máxima antes da exposição ao sevoflurano. Adaptado de Liang *et al.* [11].

uma conhecida interação de pequenos ligantes com uma proteína globular[21]. O trabalho já está publicado em um periódico[22] e está presente no Anexo I.



**Figura 7:** Estrutura molecular das pequenas moléculas utilizadas no estudo. Esquerda: Sevoflurano. Direita: Tricloroetanol.

## II Objetivos

### Objetivo geral

Este trabalho visa realizar uma descrição termodinâmica da interação degenerada de pequenos ligantes com proteínas sob a ótica de um processo de partição entre duas fases: a interface proteica e o reservatório. Juntamente, será analisado o papel da concentração nesse tipo de processo. Realizada a prova de conceito, esse estudo será útil para outros estudos envolvendo a interação de pequenas moléculas com proteínas.

### Objetivos específicos

**Objetivo 1: Desenvolvimento do arcabouço teórico para análises do coeficiente de partição.**

O coeficiente de partição é uma maneira de descrever analiticamente o processo de ligação de uma molécula à fase do sítio proteico. A partir de princípios da mecânica estatística, está sendo desenvolvido um arcabouço teórico para obter essas constantes em simulações do tipo *flooding*.

**Objetivo 2: Desenvolvimento de ciclo termodinâmico para comparação com o coeficiente de partição.**

A partição entre dois estados também será descrita como um ciclo termodinâmico hipotético, onde os princípios da mecânica estatística e desenvolvimento das integrais de configuração trarão a uma alternativa para o cálculo do coeficiente de partição que utilizará parâmetros energéticos obtidos por outros métodos, como FEP.

**Objetivo 3: Caracterização da influência da concentração no processo de partição**

Muitas teorias desenvolvidas assumem que o ligante esteja em concentrações diluídas ou que não interajam entre si. Porém, as simulações de *flooding* realizadas nesse trabalho possuem alta concentração, tornando necessário descrever o comportamento do coeficiente de partição e outras propriedades da interação em concentrações diferentes, a fim de compreender como as interações entre moléculas semelhantes podem alterar a



partição entre dois estados.

#### **Objetivo 4: Prova de conceito com sistema independente**

Uma vez finalizada a caracterização da interação sevoflurano-Kv1.2, a teoria desenvolvida será aplicada aos sistemas com tricloroetanol e albumina, com o objetivo de observar a reprodutibilidade das ferramentas desenvolvidas e das características do coeficiente de partição.

### **III Teoria e métodos**

A regulação de proteínas por pequenas moléculas no geral pode apresentar diversos graus de complexidade: os sítios podem estar distribuídos em diferentes regiões do receptor, pode haver a possibilidade de mais de uma molécula ocupar cada sítio, além de geralmente as interações terem baixa afinidade. Essas características fazem com que seja difícil obter uma descrição a nível atômico das interações entre receptor e ligante. Assim, os métodos computacionais mostram-se como ótimas ferramentas para obter detalhes sobre o comportamento das interações moleculares, que regem diversos processos biológicos.

O uso de técnicas computacionais permite a realização de estudos em condições consideradas extremas em experimentos de bancada, como experimentos em baixas densidades, temperaturas muito baixas, pressão elevada, altas diferenças de potencial elétrico, por exemplo.

As técnicas de dinâmica molecular, cálculos de energia livre por FEP mostram-se bastante adequadas ao estudo, devido à escala de espaço (Å) e de tempo (ns) necessários para um estudo de interações entre moléculas no nível atômico. Nesse estudo, serão utilizadas estruturas cristalográficas do canal Kv1.2, campos de força de anestésicos parametrizados disponíveis na literatura, além de modelos teóricos que consigam descrever as propriedades nessas escalas, mas ainda lidam com a complexidade do problema.

#### **III.1 Simulações de dinâmica molecular**

A dinâmica molecular é uma técnica onde é possível simular a movimentação de partículas em um dado sistema [23], obtendo como resultado uma trajetória atômica. A movimentação dos átomos é regida pelas interações inter e intramoleculares, que possuem seus potenciais calculados, e através da integração das equações de Newton por um certo intervalo de tempo, chamado de passo de integração, é possível calcular a próxima posição de cada átomo [24]. Esse ciclo de cálculos se repete até o fim da simulação. Os potenciais são calculados como função das posições atômicas, tomando como parâmetros

de interação os campos de força. Devido ao pequeno passo de integração e grande quantidade de interações a serem calculadas, essa técnica é realizada em computadores através de *softwares* especializados.

O principal objetivo das simulações de dinâmica molecular é a obtenção de propriedades do sistema que correspondam às que seriam obtidas numa escala macroscópica. Assim, deve-se obter as propriedades médias da simulação. A comparação entre todos os microestados do *ensemble* ofereceria um valor médio bem acurado, porém os sistemas de dinâmica molecular possuem um grande número de átomos, o que levaria a amostragem de todos os microestados algo inviável. Entretanto, as trajetórias de dinâmica molecular possuem a característica da ergodicidade, ou seja, suas propriedades obtidas por uma média temporal, correspondem à média entre os estados possíveis no *ensemble*. Então, para uma trajetória muito longa:

$$\lim_{t \rightarrow \infty} \bar{X}(t) = \langle X \rangle$$

a média no tempo da propriedade  $X$  tende à média de todos os microestados possíveis no *ensemble*.

Diversas características dos sistemas podem ser controladas nas simulações de dinâmica molecular, como volume, temperatura e pressão. Assim, diferentes tipos de *ensemble* podem ser criados: Canônico (Número de moléculas, volume e temperatura constantes – NVT) ; Microcanônico (Número de moléculas, volume e energia constantes – NVE) e Isotérmico-isobárico (Número de moléculas, volume e pressão constantes – NPT).

### III.1.1 Campos de força

O cálculo dos potenciais de interação é a etapa principal da dinâmica molecular. Nessa etapa é calculada a energia potencial total  $U(\vec{r})$ , que é a soma de diversos termos relativos às maneiras de interação entre átomos ou grupos de átomos. Todas as energias potenciais são funções da distância.

Cada grupo de interação (por exemplo: um par de átomos ou três átomos da mesma molécula) possui seus parâmetros específicos, que definem o campo de força. Os campos de força geralmente são obtidos por estudos de parametrização envolvendo cálculos *ab*

*initio* [15, 25]. A qualidade de um campo de força é dada pela capacidade de reproduzir as propriedades macroscópicas. Dentre os diversos campos de força disponíveis, alguns foram desenhados especificamente para biomoléculas e moléculas orgânicas, como CHARMM [26], AMBER [27] e GROMOS [28]. Nesse trabalho será utilizado o campo de força CHARMM, por ser o campo de força utilizado nas trajetórias já realizadas, e ser o tipo de campo de força disponibilizado em moléculas que fazem parte do estudo.

No CHARMM as interações são classificadas em interações não-ligadas, ou interações ligadas. As interações não-ligadas são calculadas entre pares de átomos, não importando se pertencem à mesma molécula, denominados a seguir como  $i$  e  $j$ :

### **Energia eletrostática**

$$U_{elet}(\vec{r}) = \sum \frac{q_i q_j}{\epsilon r_{ij}}$$

Potencial referente à Lei de Coulomb. Dependente da distância  $r_{ij}$  entre os átomos, da constante dielétrica do meio  $\epsilon$  e das cargas  $q$  dos átomos.

### **Energia de van der Waals**

$$U_{vdW}(\vec{r}) = \sum \epsilon_{ij} \left[ \left( \frac{r_0}{r_{ij}} \right)^{12} - 2 \left( \frac{r_0}{r_{ij}} \right)^6 \right]$$

Potencial que considera as interações de van der Waals, na forma do potencial aproximado de Lennard-Jones 12-6, sendo  $\epsilon_{ij}$  o módulo do potencial mínimo, e  $r_0$  a distância de equilíbrio e  $r_{ij}$ . No primeiro termo, quando a distância entre os átomos tende a zero, a energia aumenta, representando uma repulsão entre os átomos. Quando  $r_{ij} = r_0$ , a energia atinge seu menor valor, equivalendo a  $-\epsilon_{ij}$ . Já quando  $r_{ij} \gg r_0$ , a energia tende a zero, representando uma diminuição de interação entre átomos devido à grande distância.

As interações ligadas ocorrem necessariamente entre átomos da mesma molécula. Vale ressaltar que dependendo da geometria da molécula ou átomo pertencente, alguns termos podem ser nulos. São potenciais ligados:

### **Energia de ligação**

$$U_{lig}(\vec{r}) = k_b (r_{ij} - r_0)^2$$

Potencial harmônico relativo à distância entre dois átomos que realizam uma ligação covalente.  $k_b$  é a constante de força da ligação, e  $r_0$  a distância de equilíbrio.

### Energia de ângulo

$$U_{\text{ângulo}}(\vec{r}) = k_{\theta}(\theta_{ij} - \theta_0)^2$$

Potencial harmônico relativo ao ângulo entre três átomos ligados covalentemente.  $k_{\theta}$  é a constante de força da ligação, e  $\theta_0$  o ângulo de equilíbrio.

### Energia de diedro

$$U_{\text{diedro}}(\vec{r}) = k_d [1 + \cos(n\phi - \delta)]$$

Potencial responsável por levar em conta a influência da torção nas interações. Os quatro átomos ligados formam dois planos, com ângulo  $\phi$ . Os termos  $k_d, n$  e  $\delta$  são a constante de força, a multiplicidade e o ângulo de equilíbrio, respectivamente.

### Impróprio

$$U_{\text{imp}}(\vec{r}) = k_i(\omega - \omega_0)^2$$

Termo artificial adotado pelo CHARMM relativo à planaridade de moléculas. É definido como um potencial harmônico relacionado com o ângulo entre dois planos  $\omega$  e um ângulo de equilíbrio  $\omega_0$ , com a constante de força  $k_i$ .

## III.1.2 Integração das equações de movimento

Uma trajetória de dinâmica molecular consiste na movimentação dos átomos ao longo do tempo. Para um átomo  $i$ , a trajetória pode ser descrita como a resolução das seguintes etapas:

$$\vec{r}_i(t_0) \rightarrow \vec{r}_i(t_0 + \delta t) \rightarrow \vec{r}_i(t_0 + 2\delta t) \rightarrow \dots \rightarrow \vec{r}_i(t_0 + n\delta t)$$

Os átomos são movimentados conforme a força, que possui a seguinte relação com a energia potencial:

$$\vec{F} = -\frac{dU(\vec{r})}{dr} \vec{n}$$

onde  $\vec{n}$  é o vetor unitário do vetor distância ( $\vec{n} = \vec{r}/r$ ). A partir das leis de Newton, também é possível calcular a aceleração:

$$\vec{a} = \frac{d^2\vec{r}}{dt^2} = \frac{\vec{F}}{m}$$

Essa sucessão dos passos da trajetória pode ser resolvida por diversos algoritmos, que visam ser eficientes computacionalmente, além de serem condizentes com princípios físicos, como respeitar as leis de movimento e conservação de energia. Os mais frequentes são os de Verlet [29], *leap-frog* [30], e Beeman [31]. O algoritmo de Verlet, que é o utilizado no programa Nanoscale Molecular Dynamics (NAMD - [32]), utiliza parâmetros como a força ( $\vec{F}$ ), massa ( $m$ ), posição atual das moléculas ( $\vec{r}(t)$ ), posição anterior ( $\vec{r}(t - \delta t)$ ), e o intervalo de tempo de aplicação da força ( $\delta t$ ), que são definidos antes da simulação, ou no caso dos que dependem de  $\vec{r}$ , calculados durante a simulação, para calcular qual a próxima posição da trajetória:

$$\vec{r}(t + \delta t) = 2\vec{r}(t) - \vec{r}(t - \delta t) + \frac{\vec{F}(t)}{m}\delta t^2$$

O intervalo de tempo  $\delta t$ , que também é chamado de passo de integração, é um elemento crucial para a qualidade da trajetória. Um valor muito pequeno leva a uma simulação mais precisa, onde a reversibilidade dos processos é mais assegurada. Porém, isso pode elevar bastante o custo computacional da simulação. Já valores de  $\delta t$  grandes podem levar a uma amostragem incorreta em eventos que possuem o tempo característico menor que o passo de integração, como vibrações, tornando a trajetória carregada de artefatos. Assim, o passo de integração desejado é o mais próximo do evento mais rápido que pode ser observado. No caso da dinâmica molecular, o valor utilizado é de 2fs (1fs =  $10^{-15}$ s), que é igual ao período da vibração dos átomos de água.

### III.1.3 Tratamento das interações de curta e longa distância

Calcular as energias potenciais entre todos os átomos é a etapa que possui a maior demanda computacional. A quantidade de cálculos de interações não-ligadas a serem feitos cresce na ordem de  $N^2$ . Assim, é possível notar que pequenos acréscimos nas quantidades de átomos podem levar a um grande aumento nos cálculos a serem realizados. Outro problema é que muitas interações tendem a zero quando os átomos estão muito afastados, levando a cálculos desnecessários, pois possuem pouca contribuição na energia potencial total. Dessa maneira, é adotada uma distância máxima de interação, chamada de *cutoff*, onde as interações calculadas são somente interações que possuam a distância menor que

esse raio de corte. Para distâncias maiores que o *cutoff*, os termos são desconsiderados. Essa medida reduz muito a quantidade de cálculos a serem feitos numa rodada de passo de integração, economizando tempo computacional.

Porém, a adoção do *cutoff* leva a alguns problemas, como a descontinuidade abrupta dos valores de interação na vizinhança do raio de corte. Considerando um valor de *cutoff*  $R_c$ , os potenciais terão seus valores originais em  $R \leq R_c$ , mas quando  $R > R_c$ , o valor subitamente é zero. Essa mudança dos potenciais calculados até o *cutoff* pode ser suavizada adotando uma função de ajuste (*switching function*), que é definida como:

$$U_{trunc} = \begin{cases} U(r) - U(R_c) & , r \leq R_c \\ 0 & , r > R_c \end{cases}$$

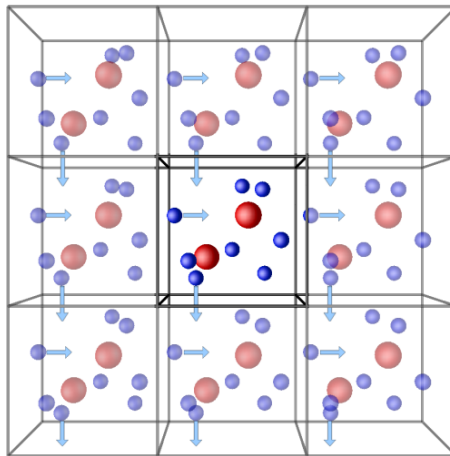
Outro problema é o tratamento das interações eletrostáticas. Enquanto o potencial de van der Waals decai na ordem de  $r^{-6}$ , o potencial eletrostático decai em ordem de  $r^{-1}$ . Assim, distâncias de *cutoff* que podem já ter a componente de van der Waals perto de zero ainda podem levar a um termo eletrostático com um valor significativo. Assim, o truncamento pode levar a um tratamento errôneo das interações não-ligadas. A solução pode ser dada adicionando um termo de longo alcance para descrever as interações eletrostáticas. Dentre várias descrições de longo alcance disponíveis na literatura, será utilizado nesse trabalho o método PME (*Particle Mesh Ewald* - [33]), que leva em conta a influência de átomos distantes para definir um potencial implícito dependente da densidade de carga  $\rho(\mathbf{k})$ . Dessa maneira, o potencial eletrostático é definido como:

$$U_{elet} = U_{Coulomb}(r) + U_{PME}(\rho(\mathbf{k}))$$

### III.1.4 Condições periódicas de contorno

Um sistema de dinâmica molecular é composto por um conjunto de átomos distribuídos em um dado volume. Assim, existem átomos que ficam nos limites desse volume definido, que estariam em contato com o vácuo. Isso faria com que esses átomos difundissem para fora do sistema, entrando em fase gasosa, além de poder causar uma tensão superficial. Na maioria das vezes, esses fenômenos relacionados ao efeito de borda são

totalmente indesejados, sendo considerados como artefatos. Esse problema pode ser solucionado quando adotadas as condições periódicas de contorno (*periodic boundary conditions - PBC*), que consiste em replicar infinitamente o sistema em todos os eixos, onde os limites de cada célula estariam em contato com os limites opostos das réplicas adjacentes. Assim, os átomos que estariam na fronteira do volume interagiriam com os átomos da réplica ao lado, extinguindo qualquer efeito de borda. Qualquer movimentação de átomos na caixa original é replicada em todas as células.



**Figura 8:** Esquemática das condições periódicas de contorno. No centro está o sistema original, rodeado em todos os eixos pelas suas réplicas. A convenção da imagem mínima pode ser notada quando um átomo atravessa a fronteira do sistema e esse movimento é acompanhado pela aparição de outro átomo do lado oposto do sistema, que veio da réplica da extremidade oposta. Retirado de <http://isaacs.sourceforge.net/phys/pbc.html>. Acesso em 9/12/2019.

A infinidade de réplicas faz com que adotar o PBC torne o número de interações a serem calculadas impraticável para a dinâmica molecular. Porém, muitas das interações seriam redundantes, pois ocorreriam entre átomos muito distantes, e até entre réplicas. Assim, no PBC é adotada a convenção da imagem mínima, que estabelece que só ocorram cálculos entre os átomos da réplica mais próxima e além disso, a distância de *cutoff* deve ser menor que metade do comprimento do sistema, para assegurar que as interações estão ocorrendo apenas uma vez com cada átomo.



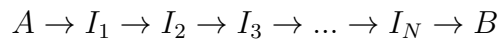
## III.2 Cálculos de energia livre por FEP

No estudo da interação entre proteínas e ligantes, a obtenção da energia livre de ligação é uma das etapas cruciais para determinar a acurácia dos resultados, pois a partir da energia é possível obter diversas outras propriedades do sistema. As variações de energia podem ser calculadas através de ferramentas computacionais, utilizando como base teórica as teorias de perturbação provenientes da mecânica estatística. Suponha que existam dois estados: um estado inicial A, e um estado final B. A variação de energia livre entre o estado A e B é dada por [34]

$$\Delta A = -k_b T \ln \left\langle e^{-\frac{(E_B - E_A)}{k_b T}} \right\rangle$$

onde  $k_b$  é a constante de Boltzmann,  $T$  a temperatura,  $E_B$  e  $E_A$  as energias dos estados B e A, respectivamente.

Apesar da descrição teórica suficiente, amostrar corretamente vários estados para calcular as diferenças de energia pode ser uma tarefa difícil. Para solucionar esta questão, a mudança de A para B é dividida em pequenas mudanças por estados intermediários  $I$ , mostrado abaixo. Essa alteração gradual assegura uma maior convergência e acurácia das diferenças de energia dos estados, além de garantir a reversibilidade do processo.



Dessa maneira, a variação total de energia da transformação de A para B ( $\Delta G_{A \rightarrow B}$ ) é dada pela soma da variação de energia de cada transformação:

$$\Delta G_{A \rightarrow B} = \Delta G_{A \rightarrow I_1} + \Delta G_{I_1 \rightarrow I_2} + \Delta G_{I_2 \rightarrow I_3} + \dots + \Delta G_{I_N \rightarrow B}$$

No *Free Energy Perturbation* (FEP), os estados intermediários são mapeados seguindo um parâmetro de controle  $\lambda$ . Esse parâmetro descreve o quanto as interações intermoleculares dos átomos que sofrem a transformação estão desligadas. Ou seja, uma transformação completa parte de  $\lambda = 0$ , onde todas as interações com o ambiente estão sendo consideradas, até  $\lambda = 1$ , onde as interações com o ambiente não existem. Se desejado, também é possível realizar transformações de  $\lambda = 1$  para  $\lambda = 0$ , no qual  $\Delta G(0 \rightarrow 1) = -\Delta G(1 \rightarrow 0)$ .

No caso da dessolvatação de uma molécula, que será o objetivo principal dos cálculos de FEP nesse estudo, o estado inicial corresponde a uma molécula inserida no ambiente e interagindo normalmente com ele ( $\lambda = 0$ ), que será transformada em estados com menos interações com os átomos e moléculas ao redor. As transformações ocorrem sucessivamente até atingir o estado final sem interações. A variação de energia resultante é a energia de dessolvatação da molécula naquele ambiente. Vale relembrar que a energia de solvatação é o valor oposto ( $\Delta G_{dessolv} = -\Delta G_{solv}$ ).

O processo é repetido várias vezes de maneira independente, para obter um maior número de estados e conseqüentemente um valor mais acurado. As estimativas da variação de energia livre e erros estatísticos são obtidas pelo método estatístico SOS [35].

### III.3 Descrição termodinâmica das interações

#### III.3.1 Coeficiente de partição

A ferramenta utilizada para descrever as interações entre pequenos ligantes e sítios é a simulação de *flooding*. Um sistema completo de uma simulação de *flooding* é composto por um volume  $V_{total}$  ocupado por  $M$  moléculas: um receptor proteico,  $N$  ligantes, além das moléculas que farão parte do ambiente de solvatação, como água, membrana (quando necessário) e íons. Os ligantes podem acessar qualquer local do sistema, com exceção da proteína. Dessa maneira, é definido um volume acessível para os ligantes  $V = V_{total} - V_{prot}$ , no qual  $V_{prot}$  é o volume proteico. A região de até 5 Å de distância da superfície da proteína é definida como o sítio de ligação, que possui volume  $\nu$ . A escolha dessa distância ocorre pelo fato do sevoflurano possuir o maior valor de densidade local[36]. A quantidade média de ligantes que ocupa esse sítio é denominado como  $\langle n \rangle$ . O ligante também possui uma distribuição espacial, descrita pela sua densidade de probabilidade  $\bar{\rho}(\mathbf{R})$ , que relaciona-se com a quantidade de ligantes  $n_X$  dentro de qualquer volume  $X$  do sistema:

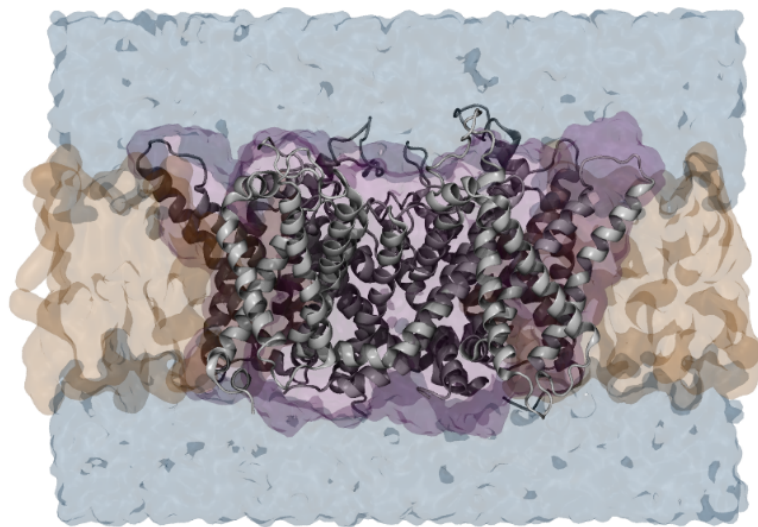
$$n_X = N \int_X d\mathbf{R} \bar{\rho}(\mathbf{R}) \quad (6)$$

Onde  $\mathbf{R} = (R_x, R_y, R_z)$  é o centro geométrico do ligante. Assim, é possível observar que apenas os termos de translação são suficientes para definir a fase em que um ligante está

presente. Essa definição é condizente com a ideia de uma interação degenerada, a qual não possui nenhuma preferência em outros graus de liberdade, como rotacionais.

As definições acima são importantes para considerar o processo de ligação do pequeno ligante aos sítios proteicos como um processo de partição entre duas fases, mostrados na Figura 9. A primeira fase é o ambiente onde o ligante não está em contato com a proteína, denominada como reservatório, de volume  $V - \nu$ . A segunda fase é composta pelo sítio de ligação, com volume  $\nu$ . No início, todos os ligantes estão presentes no reservatório, mas com o passar do tempo e consequente alcance do equilíbrio, eles particionam-se no sistema e podem interagir com a proteína. O sítio é considerado como uma região aproximadamente homogênea, o que faz com que a interação com ligantes seja considerada degenerada. O coeficiente de partição é definido pela razão das densidades em cada fase no equilíbrio:

$$\varphi = \left( \frac{N - \langle n \rangle}{V - \nu} \right)^{-1} \frac{\langle n \rangle}{\nu} \quad (7)$$



**Figura 9:** Fases da partição do sevoflurano no sistema de *flooding*. A fase correspondente ao sítio de ligação (roxo) corresponde à região de até 5Å de distância da proteína (branco), enquanto a fase do reservatório corresponde a todo o restante. No caso de sistemas com membrana, o reservatório é composto pela região membranosa (laranja) e a água (azul) que estão a mais de 5Å de distância da proteína.

O coeficiente de partição é elemento chave para a compreensão de mecanismos de funcionamento a nível atômico de pequenas moléculas com proteínas, tornando sua obtenção um dos principais objetivos desse trabalho, pois diversas propriedades das interações

no equilíbrio podem ser descritas em função desse coeficiente, como variação de energia livre, probabilidades de estados e a influência da concentração. Além disso, toda essa descrição feita acima é compatível com o observado em simulações de *flooding*, onde os ligantes movem-se no sistema de acordo com suas interações, podendo interagir com o sítio proteico caso exista alguma afinidade. Também deve-se considerar que a concentração do ligante não altere a integridade do sistema e nem altere a natureza da partição.

Para validar a descrição do fenômeno como uma partição é necessário comparar com outras abordagens já consolidadas, as quais utilizam as energias livres de solvatação do ligante[14, 17, 37].

A definição do coeficiente de partição pode ser estendida para obter relações entre outras propriedades termodinâmicas, como a variação de energia necessária para a transferência de um ligante do reservatório para a interface proteica. Utilizando uma descrição microscópica do fenômeno, será definido um potencial efetivo  $U'(\mathbf{R})$ , que dependerá do centro geométrico do ligante que está sendo transferido ( $\mathbf{R}_1$ )

$$e^{-\beta U'(\mathbf{R})} = \int d\mathbf{r}_1 \delta[\mathbf{R}_1(\mathbf{r}_1) - \mathbf{R}] \int d\mathbf{r}^{N-1} \int d\mathbf{r}^{M-N-1} e^{-\beta U(\mathbf{r}^M)} \quad (8)$$

onde os graus de liberdade de todos os átomos do sistema são denotados pela notação compacta  $\mathbf{r}^M$  e os graus de liberdade puramente translacionais são denominados como  $\mathbf{R}$ . Cabe ressaltar que neste potencial todos os outros graus de liberdade do sistema já foram integrados de acordo com a energia potencial  $U(\mathbf{r}^M)$ , restando apenas os termos relacionados à translação do ligante que está sendo transferido ( $\mathbf{R}$ ). A densidade de probabilidade é definida como a razão do potencial efetivo  $U'(\mathbf{R})$  pela integral de configuração, que equivale à integração do potencial efetivo em todo volume do sistema ( $V$ ):

$$\bar{\rho}(\mathbf{R}) = \frac{e^{-\beta U'(\mathbf{R})}}{\int_V d\mathbf{R} e^{-\beta U'(\mathbf{R})}} = \frac{e^{-\beta U'(\mathbf{R})}}{Z} \quad (9)$$

Utilizando a definição do coeficiente de partição (equação 7) e da relação do número de ligante em um volume específico ( $\langle n \rangle = N \int d\mathbf{R} \bar{\rho}(\mathbf{R})$  - equação 6), obtém-se:

$$\wp = \left( \frac{N - \langle n \rangle}{V - \nu} \right)^{-1} \frac{\langle n \rangle}{\nu} \quad (10)$$

$$\wp = \frac{V - \nu}{\nu} \times \frac{\int_{\nu} d\mathbf{R} \bar{\rho}(\mathbf{R})}{\int_{V-\nu} d\mathbf{R} \bar{\rho}(\mathbf{R})}$$

Substituindo as densidades de probabilidade pela razão do potencial efetivo pela integral de configuração (equação 9):

$$\wp = \frac{V - \nu}{\nu} \times \frac{\int_{\nu} d\mathbf{R} e^{-\beta U'(\mathbf{R})}}{\int_{V-\nu} d\mathbf{R} e^{-\beta U'(\mathbf{R})}} \quad (11)$$

Adicionalmente, será definido um novo potencial efetivo  $U''(\mathbf{R})$  que dependerá do centro geométrico do ligante e corresponde à integração dos graus de liberdade de acordo com a energia potencial do estado desacoplado ( $U_0(r^M)$ ):

$$e^{-\beta U''(\mathbf{R})} = \int d\mathbf{r}_1 \delta[\mathbf{R}_1(\mathbf{r}_1) - \mathbf{R}] \int d\mathbf{r}^{N-1} \int d\mathbf{r}^{M-N-1} e^{-\beta U_0(\mathbf{r}^M)}$$

Como o estado desacoplado corresponde à ausência de interações, o ligante poderia assumir qualquer posição no vácuo. Para evitar esse problema e manter o ligante no mesmo volume correspondente ao da integração, será necessária a adição de um potencial de restrição espacial  $\mu(\mathbf{R})$ . Com essa descrição, será possível utilizar a definição que relaciona a energia de dessolvatação do ligante ( $W$ ) e os potenciais efetivos [17]:

$$e^{-\beta U'(\mathbf{R})} = e^{-\beta[U''(\mathbf{R}) + \mu(\mathbf{R})]} e^{+\beta W} \quad (12)$$

Substituindo na equação 11, para os termos relacionados ao reservatório (\*) e interface proteica (\*\*):

$$\wp = \frac{V - \nu}{\nu} \times \frac{\int_{\nu} d\mathbf{R} e^{+\beta W^{**}} e^{-\beta[U''(\mathbf{R}) + \mu^{**}(\mathbf{R})]}}{\int_{V-\nu} d\mathbf{R} e^{+\beta W^*} e^{-\beta[U''(\mathbf{R}) + \mu^*(\mathbf{R})]}} \quad (13)$$

O potencial efetivo no estado desacoplado não depende da posição da molécula, pois as interações no vácuo são iguais independentemente da posição dos átomos. Adicionalmente, devido ao fato das interações serem degeneradas, será assumido que a energia de dessolvatação independe da posição do ligante. Logo:

$$\wp = \frac{V - \nu}{\nu} \frac{e^{+\beta W^{**}}}{e^{+\beta W^*}} \frac{e^{-\beta U''(\mathbf{R})}}{e^{-\beta U''(\mathbf{R})}} \frac{\int_{\nu} d\mathbf{R} e^{-\mu^{**}(\mathbf{R})}}{\int_{V-\nu} d\mathbf{R} e^{-\mu^*(\mathbf{R})}} \quad (14)$$

Para todo domínio de integração  $\nu$ ,  $\mu^{**}(\mathbf{R}) = 0$ . O mesmo pode ser afirmado de maneira análoga para  $\mu^*(\mathbf{R})$  em  $V - \nu$ , o que simplifica a resolução das integrais:

$$\begin{aligned} \wp &= \frac{V - \nu}{\nu} \frac{e^{+\beta W^{**}}}{e^{+\beta W^*}} \frac{e^{-\beta U''(\mathbf{R})}}{e^{-\beta U''(\mathbf{R})}} \frac{\int_{\nu} d\mathbf{R} e^{-\mu^{**}(\mathbf{R})}}{\int_{V-\nu} d\mathbf{R} e^{-\mu^*(\mathbf{R})}} \\ \wp &= \frac{V - \nu}{\nu} \times \frac{e^{+\beta W^{**}}}{e^{+\beta W^*}} \frac{e^{-\beta U''(\mathbf{R})}}{e^{-\beta U''(\mathbf{R})}} \frac{\nu}{V - \nu} \end{aligned} \quad (15)$$

$$\wp = e^{-\beta(W^* - W^{**})}$$

Com as condições mostradas acima, foi possível obter a relação entre as energias de dessolvatação e o coeficiente de partição:

$$\wp = e^{-\beta(W^* - W^{**})} \quad (16)$$

onde  $W^{**}$  é a energia correspondente à transferência do ligante do sítio para o vácuo e  $W^*$  é a energia correspondente à transferência do reservatório para o vácuo.

A partição do ligante em duas fases leva em conta diversas interações: interação das moléculas com a proteína, com o ambiente de solvatação e com outras moléculas da mesma natureza em diversos regimes de concentração [37]. Assim, a equação 7 não consegue contemplar a influência individual desses fatores no processo de interação entre proteína e ligante, enquanto a relação do coeficiente de partição com os potenciais (equação 16) consegue delimitar a influência de cada etapa.

A interação que está sendo estudada envolve uma proteína de membrana, o que requer um reservatório composto por água e membrana, levando a duas energias de dessolvatação diferentes. Assim, é necessário obter a relação entre a energia de dessolvatação ( $W^*$ ) resultante e a energia de dessolvatação de cada componente do reservatório ( $W_{wat}$  e  $W_{memb}$ ). Partindo da princípio da conservação da matéria no reservatório, onde  $\langle n_{wat} \rangle$  é o número médio de ligantes presentes na água e  $\langle n_{memb} \rangle$  presentes na membrana:

$$N - \langle n \rangle = \langle n_1 \rangle + \langle n_2 \rangle$$

O número de ligantes em cada região pode ser encontrado a partir da densidade de

probabilidade do ligante (equação 6):

$$N \int_{V-\nu} d\mathbf{R} \bar{\rho}(\mathbf{R}) = N \int_{v_{wat}} d\mathbf{R} \bar{\rho}(\mathbf{R}) + N \int_{v_{memb}} d\mathbf{R} \bar{\rho}(\mathbf{R}) \quad (17)$$

Utilizando a relação  $\bar{\rho}(\mathbf{R}) = e^{-\beta U'(\mathbf{R})}/Z$  (equação 9) e a definição da energia de dessolvatação da equação 12 :

$$\int_{V-\nu} \frac{e^{-\beta[U''(\mathbf{R})+\mu^*(\mathbf{R})]}}{e^{-\beta W^*}} d\mathbf{R} = \int_{v_{wat}} \frac{e^{-\beta[U''(\mathbf{R})+\mu^*(\mathbf{R})]}}{e^{-\beta W_{wat}^*}} d\mathbf{R} + \int_{v_{memb}} \frac{e^{-\beta[U''(\mathbf{R})+\mu^*(\mathbf{R})]}}{e^{-\beta W_{memb}^*}} d\mathbf{R} \quad (18)$$

É considerado que as interações desacopladas não dependem da posição assumida pelo ligante, além de que a energia de dessolvatação também não sofre influência da posição ocupada no reservatório:

$$\frac{e^{-\beta U''(\mathbf{R})}}{e^{-\beta W^*}} \int_{V-\nu} e^{-\beta \mu^*(\mathbf{R})} d\mathbf{R} = \frac{e^{-\beta U''(\mathbf{R})}}{e^{-\beta W_{wat}^*}} \int_{v_{wat}} e^{-\beta \mu^*(\mathbf{R})} d\mathbf{R} + \frac{e^{-\beta U''(\mathbf{R})}}{e^{-\beta W_{memb}^*}} \int_{v_{memb}} e^{-\beta \mu^*(\mathbf{R})} d\mathbf{R} \quad (19)$$

De maneira semelhante à equação 15, o potencial  $\mu^*(\mathbf{R}^n)$  é nulo nos limites de integração. Assim:

$$(V - \nu)e^{+\beta W^*} = v_{wat}e^{+\beta W_{wat}^*} + v_{memb}e^{+\beta W_{memb}^*} \quad (20)$$

Reagrupando os termos:

$$e^{+\beta W^*} = \frac{v_{wat}}{V - \nu} e^{+\beta W_{wat}^*} + \frac{v_{memb}}{V - \nu} e^{+\beta W_{memb}^*} \quad (21)$$

onde  $v$  corresponde ao volume e  $W^*$  às energias de solvatação da água e membrana. Vale lembrar que  $v_{wat} + v_{memb} = V - \nu$ . A demonstração acima poderá ser estendida para quaisquer fases compostas por mais de um ambiente.

### III.3.2 Energia livre da interação

O coeficiente de partição é um parâmetro que depende somente de propriedades do ligante, como sua concentração em cada fase ou sua afinidade. Além disso, as propriedades do coeficiente de partição são obtidas no equilíbrio, após a estabilização do sistema. Agora é desejado obter a variação de energia livre correspondente à transferência

de  $n = \lfloor \langle n \rangle \rfloor$  ligantes, localizados inicialmente no reservatório para o volume do sítio, como mostrado na etapa K da Figura 10.

É possível obter a variação de energia livre da interação ( $\Delta W$ ) utilizando uma descrição microscópica do sistema juntamente com princípios de mecânica estatística. Vale ressaltar que as características da interação devem ser iguais às da seção III.3.1. Será definido um potencial efetivo  $U'$  onde diversos graus de liberdade serão integrados de acordo com a energia potencial do sistema  $U(\mathbf{r}^M)$ :

$$e^{-\beta U'(\mathbf{r}^P, \mathbf{R}^N)} = \int d\mathbf{r}_1 \delta[\mathbf{R}'(\mathbf{r}_1) - \mathbf{R}_1] \dots \int d\mathbf{r}_N \delta[\mathbf{R}'(\mathbf{r}^N) - \mathbf{R}^N] \int d\mathbf{r}^{M-N-P} e^{-\beta U(\mathbf{r}^M)} \quad (22)$$

onde  $\mathbf{r}$  corresponde aos graus de liberdade configuracionais e translacionais, enquanto  $\mathbf{R}$  corresponde somente aos translacionais. O potencial efetivo ainda não oferece uma descrição completa do sistema, pois ainda são necessárias informações sobre o estado translacional dos ligantes ( $\mathbf{R}^N$ ) e configuracional da proteína ( $\mathbf{r}^P$ ) para a obtenção da integral de configuração completa.

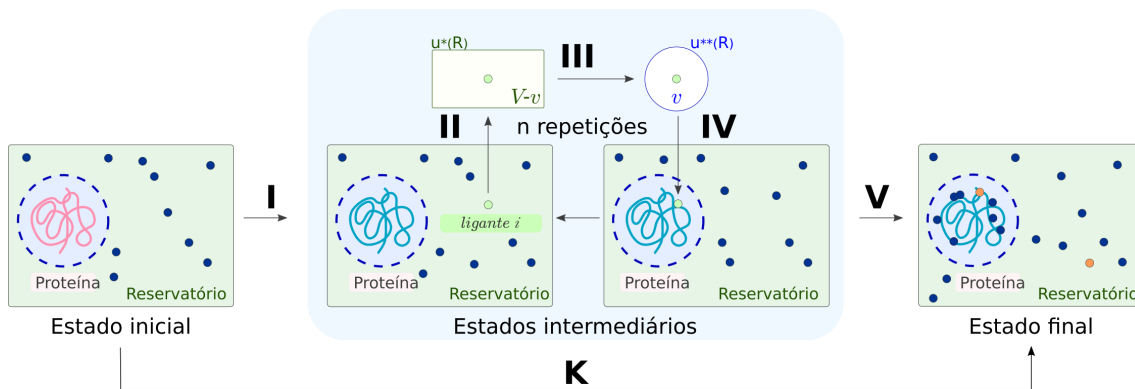
A variação de energia livre  $\Delta W$  pode ser expressa em função das integrais de configuração entre o estado inicial e final:

$$e^{-\beta \Delta W} = \frac{N!}{(N-n)!n!} \frac{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta U'(\mathbf{r}^P, \mathbf{R}^N)}}{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta U'(\mathbf{r}^P, \mathbf{R}^N)}} \quad (23)$$

onde  $N!/(N-n)!n!$  corrige a indistinguibilidade do ligante, o termo  $\beta = 1/k_B T$  depende da constante de Boltzmann. No denominador está presente a integral de configuração completa para o estado inicial, onde todos os  $N$  ligantes estão no reservatório e o sítio está vazio, enquanto no numerador a integral de configuração corresponde ao sítio ocupado por  $n$  ligantes.

O cálculo direto das integrais de configuração apresentadas na forma da equação 23 seria muito difícil e custoso em termos de tempo computacional, o que não é interessante. Porém, a Figura 10 mostra que é possível reinterpretar o processo de partição como um ciclo termodinâmico composto por várias etapas: I - restrição da estrutura proteica; II - dessolvatação dos ligantes; III - restrições translacionais; IV - inserção do ligante no sítio; V - liberação das restrições. Essas etapas ocorrem com o auxílio de potenciais externos e possuem integrais de configuração mais simples, facilitando os cálculos.





**Figura 10:** Processo de partição. Nas simulações de *flooding*, o sistema é composto inicialmente pela proteína vazia (em rosa) e os ligantes (pontos azuis) no reservatório. Com o passar do tempo, alguns ligantes particionam-se para o sítio da interface proteica (etapa K), delimitada na imagem pelo círculo tracejado. O processo de partição pode ser descrito através de um ciclo hipotético com aplicação de potenciais artificiais de cinco etapas: I - restrição da estrutura proteica; II - dessolvatação dos ligantes; III - restrições translacionais; IV - inserção do ligante no sítio; V - liberação das restrições. No caso da partição de  $n$  ligantes, o ciclo hipotético é repetido  $n$  vezes.

Assim, as seções a seguir terão como objetivo realizar a descrição teórica separadamente de cada passo para facilitar o processo de compreensão da equação 23 e sua posterior simplificação. As etapas do ciclo conseguem ser fielmente reproduzidas pelas simulações de dinâmica molecular, onde inclusive existe a possibilidade de inserção de condições especiais para corresponderem aos sistemas moleculares das etapas do ciclo.

### III.3.3 Restrição espacial da proteína

A proteína pode possuir uma estrutura interna diferente de quando está em contato com os ligantes. Assim, o primeiro passo do ciclo termodinâmico consiste na transformação de uma estrutura interna correspondente ao do estado vazio para uma estrutura interna correspondente ao estado final, ilustrado como a etapa I da Figura 10. Essa transformação é regida por um potencial artificial  $u(\mathbf{r}^P)$ , da seguinte forma:

$$u(\mathbf{r}^P) = \begin{cases} 0 & , \text{ se } \text{RMSD} \leq \text{RMSD}^* \\ \frac{1}{2}k_{\text{RMSD}}(\text{RMSD} - \text{RMSD}^*)^2 & , \text{ se } \text{RMSD} \geq \text{RMSD}^* \end{cases} \quad (24)$$

onde  $\mathbf{r}^P$  são as coordenadas dos átomos da proteína,  $\text{RMSD}^*$  é o RMSD de referência adotado (2,5 Å) e  $k_{\text{RMSD}}$  é a constante de força utilizada no potencial (100 kcal/mol Å<sup>2</sup>).

O potencial  $u(\mathbf{r}^P)$  mantém a estrutura da proteína dentro de um valor de referência de RMSD em comparação à sua estrutura inicial,  $\text{RMSD}^* = 2,5\text{\AA}$ . Caso a estrutura atinja um RMSD maior, ela sofrerá um acréscimo energético dependente da constante de força e desvio do RMSD referência. Esse potencial artificial é bastante utilizado em protocolos de simulação onde há a transformação de estruturas para uma estrutura alvo [38]. Esse termo  $u(\mathbf{r}^P)$  pode ser calculado diretamente das simulações de dinâmica molecular, a partir do *ensemble* de estruturas. No caso da transformação dessa etapa, o cálculo ocorreu utilizando estruturas de uma simulação de dinâmica molecular de uma proteína vazia.

A integral de configuração é:

$$\frac{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta[U'(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P)]}}{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta U'(\mathbf{r}^P, \mathbf{R}^N)}} = \langle e^{-\beta u(\mathbf{r}^P)} \rangle_P \quad (25)$$

a qual está relacionada com o trabalho externo necessário para a passagem da estrutura interna da proteína vazia para a estrutura interna correspondente ao estado ligado, porém sem a presença de ligantes.

### III.3.4 Passagem dos ligantes para o vácuo

A passagem dos ligantes partindo do reservatório para a fase gasosa é o segundo passo do ciclo termodinâmico idealizado, correspondente à etapa II da Figura 10. A etapa inicial desse passo consiste na proteína presa pelo potencial  $u(\mathbf{r}^P)$  e os ligantes livres no sistema, tendo disponível todo seu volume acessível  $V - \nu$ . No estado final, os  $n$  ligantes situam-se no vácuo sob efeito de um potencial artificial  $\mu^*(\mathbf{R}^n)$  que restringe sua translação apenas ao volume da fase acessível. Esse potencial possui a seguinte forma geral:

$$\mu^*(\mathbf{R}^n) = \begin{cases} 0 & , \text{ se } \mathbf{R}_i \subset V - \nu \\ \gg 0 & , \text{ se } \mathbf{R}_i \not\subset V - \nu \end{cases} \quad (26)$$

no qual a molécula é mantida no volume acessível por sofrer punições energéticas caso acesse alguma região fora desse volume. Dessa maneira, a integral de configuração para essa etapa é escrita como:

$$\frac{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P) + \mu^*(\mathbf{R}^n)]}}{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta[U'(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P)]}} \quad (27)$$

Além da adição do potencial de restrição de volume  $\mu^*(\mathbf{R}_i)$ , o potencial efetivo do estado final foi alterado para  $U''(\mathbf{r}^P, \mathbf{R}^N)$ , que é um potencial efetivo semelhante a  $U'(\mathbf{r}^P, \mathbf{R}^N)$  (definido na equação 22), porém integrado para a energia potencial das interações correspondentes ao vácuo ( $U_0(\mathbf{r}^P)$ ):

$$e^{-\beta U''(\mathbf{r}^P, \mathbf{R}^N)} = \int d\mathbf{r}_1 \delta[\mathbf{R}'(\mathbf{r}_1) - \mathbf{R}_1] \dots \int d\mathbf{r}_N \delta[\mathbf{R}'(\mathbf{r}_N) - \mathbf{R}_N] \int d\mathbf{r}^{M-N-P} e^{-\beta U_0(\mathbf{r}^M)}$$

O numerador da equação 27 pode passar por uma simplificação: como a integração das moléculas está definida no volume  $V - \nu$ , o potencial  $\mu^*(\mathbf{R}^n)$  é igual a zero:

$$\begin{aligned} \int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P) + \mu^*(\mathbf{R}^n)]} \\ \approx \\ \int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P)]} \end{aligned} \quad (28)$$

Logo:

$$\frac{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P)]}}{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta[U'(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P)]}} = e^{-\beta W_n^*} \quad (29)$$

As funções de partição tanto do numerador quanto do denominador possuem os mesmos limites de integração, diferindo apenas na energia potencial da molécula. Como o processo descrito é de uma dessolvatação,  $W_n^*$  é definida como a energia de dessolvatação.

Devido à composição heterogênea do reservatório, a energia de dessolvatação resultante dependerá das energias e dos volumes de cada componente do reservatório, como mostrado na equação 21:

$$e^{+\beta W^*} = \frac{V_{wat}}{V - \nu} e^{+\beta W_{wat}^*} + \frac{V_{memb}}{V - \nu} e^{+\beta W_{memb}^*}$$

Cabe observar que a variação de energia  $W_n^*$  é dependente da quantidade de moléculas de ligante que estão sendo retiradas do reservatório e a simulação de *flooding* é realizada em altas concentrações, com um alto número de ligantes, tornando a remoção ou inserção de moléculas um evento que altera a concentração, e por consequência pode alterar a energia de remoção. Ou seja, as interações do tipo ligante-ligante existem e são explicitamente consideradas, podendo levar a uma alteração na energia potencial do sistema e em outras propriedades[39], inclusive na energia necessária em remover moléculas do ambiente. O valor dessa alteração de energia é específico para cada ligante,

variando principalmente de acordo com os grupos químicos. No equilíbrio, poucos ligantes permanecem na água, assim pode-se considerar que não há influência da concentração para a energia de dessolvatação na água, que já foi calculada em estudos anteriores ( $W_{wat}^* = 0,10 \pm 0,09$ )[17].

No caso da partição membrana-vácuo, será utilizada uma adaptação linear de um modelo existente na literatura que descreve a energia de dessolvatação de um ligante em função de sua concentração [37]. Esse modelo leva em conta a concentração do ligante no reservatório e o coeficiente  $h^*$ , que dita se a presença de mais ligantes facilita ( $h^* < 0$ ) ou dificulta ( $h^* > 0$ ) a remoção.

$$W_{memb}^* = \bar{\mu} - h_{memb}^* \frac{N - n}{V - \nu} c \quad (30)$$

onde  $c$  é o inverso da concentração padrão ( $c = (1 \text{ mol/L})^{-1}$ ) e  $\bar{\mu}$  é a energia de dessolvatação no sistema diluído ( $N \rightarrow 0$ ). Para a remoção de  $n$  ligantes, é considerado que  $W_n^* = n \times W^*$ .

Para obter os valores de  $h_{memb}^*$  é necessário calcular diversos valores da variação de energia livre de solvatação em diversas de concentrações. Daí em diante, realiza-se uma regressão desses parâmetros, seguindo a equação 30. Nesse trabalho, a regressão será realizada através do módulo *optimize* do pacote *Scipy* [40]. O protocolo utilizado para os cálculos de energia livre será explicado na seção III.4.3.2.

### III.3.5 Mudança de volumes no vácuo

A terceira parte do ciclo - etapa III da Figura 10 - corresponde à passagem de  $n$  ligantes do volume acessível  $V - \nu$  ao volume  $\nu$  do sítio, ainda permanecendo no vácuo. Nesse novo estado o potencial anterior  $\mu^*(\mathbf{R}^n)$  é extinguido e é aplicado aos  $n$  ligantes outro potencial artificial  $\mu^{**}(\mathbf{R}^n)$  para mantê-los no volume correspondente ao do sítio. A forma geral desse potencial é semelhante ao potencial anterior (representado pela equação 26), onde a molécula sofre uma punição energética quando se afasta do volume  $\nu$ . Dessa maneira, a integral de configuração dessa etapa é dada por:

$$\frac{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P) + \mu^{**}(\mathbf{R}^n)]}}{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P) + \mu^*(\mathbf{R}^n)]}} \quad (31)$$

Ao desenvolver a integral, é possível aproveitar-se do fato que  $U''(\mathbf{r}^P, \mathbf{R}^N)$  independe da posição assumida pelo ligante, afinal não existem interações entre os ligantes e outras moléculas no vácuo. A partir daí, os termos relativos à energia potencial total restante - tanto no numerador quanto no denominador - dependem somente da posição das  $n$  moléculas, levando a uma reordenação dos integrandos sem causar prejuízos à resolução da integral. O potencial  $u(\mathbf{r}^P)$  e os graus de liberdade dos átomos da proteína são idênticos entre os estados. Assim, a razão das integrais podem ser reescrita como:

$$\frac{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N)+u(\mathbf{r}^P)+\mu^{**}(\mathbf{R}^n)]}}{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^0 \int_{V-\nu} d\mathbf{R}^{N-0} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N)+u(\mathbf{r}^P)+\mu^*(\mathbf{R}^n)]}} = \frac{\int_{\nu} d\mathbf{R}^n e^{-\beta\mu^{**}(\mathbf{R}^n)}}{\int_{V-\nu} d\mathbf{R}^n e^{-\beta\mu^*(\mathbf{R}^n)}} \quad (32)$$

Como  $\mu^*(\mathbf{R}^n) = 0$  se os ligantes estão em  $V - \nu$  e  $\mu^{**}(\mathbf{R}^n) = 0$  se os ligantes estão em  $\nu$ :

$$\frac{\int_{\nu} d\mathbf{R}^n}{\int_{V-\nu} d\mathbf{R}^n} = \left( \frac{\nu}{V - \nu} \right)^n \quad (33)$$

Dessa maneira, é possível ver que a razão das integrais de configuração da etapa de mudança de volumes no vácuo tem como resultado a razão do volume do sítio pelo volume do reservatório. Cabe ressaltar que devido à natureza da interação com pequenos ligantes, os potenciais artificiais são puramente translacionais, já que não há nenhuma restrição quanto à orientação da molécula (ver seção IV.2.2).

### III.3.6 Passagem do ligante para o sítio

A penúltima parte do ciclo, correspondente à etapa IV da Figura 10, é passagem dos  $n$  ligantes do estado de volume restrito no vácuo para o sítio, com as interações seguindo a energia potencial  $U(\mathbf{r}^M)$ . A integral de configuração dessa etapa é:

$$\frac{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta[U'(\mathbf{r}^P, \mathbf{R}^N)+u(\mathbf{r}^P)]}}{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N)+u(\mathbf{r}^P)+\mu^{**}(\mathbf{R}^n)]}} \quad (34)$$

De maneira semelhante à equação 28, como o limite de integração de  $d\mathbf{R}^n$  é  $\nu$ , o denominador pode ser reescrito da seguinte maneira:

$$\begin{aligned} & \int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N)+u(\mathbf{r}^P)+\mu^{**}(\mathbf{R}^n)]} \\ & \approx \\ & \int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N)+u(\mathbf{r}^P)]} \end{aligned} \quad (35)$$

Reescrevendo a integral de configuração:

$$\frac{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta[U'(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P)]}}{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta[U''(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P)]}} = e^{+\beta W_n^{**}} \quad (36)$$

As integrais de configuração possuem os mesmos limites de integração, porém diferem na energia potencial. Denominando a energia de dessolvatação de  $n$  ligantes no sítio como  $W_n^{**}$ , a diferença de energia entre os dois estados é de  $-W_n^{**}$ . Semelhante ao caso da retirada dos ligantes do reservatório (seção III.3.4), a maior presença de ligantes no sítio pode alterar a energia de dessolvatação, logo a dependência da energia em relação à concentração será dada por:

$$W^{**} = \bar{w} - h^{**} \frac{N-n}{V-\nu} c \quad (37)$$

onde  $c$  é o inverso da concentração padrão ( $c = (1 \text{ mol/L})^{-1}$ ) e  $\bar{w}$  é a energia de dessolvatação no sistema diluído. De maneira semelhante à seção III.3.4, o termo  $h^{**}$  também é encontrado a partir da regressão linear das energias de dessolvatação do ligante obtidas em várias concentrações. Para a remoção de  $n$  ligantes do sítio, é considerado que  $W_n^{**} = n \times W^{**}$ .

Existem diversas maneiras de se calcular essa energia, principalmente pelo uso de ferramentas computacionais, como o *docking* molecular [11], método de *Linear Interaction Energy* [41, 13] e pelo método de FEP [16, 15], que será o utilizado neste estudo (mais detalhes na seção III.4.4.1).

### III.3.7 Liberação de restrições na proteína

A última parte do ciclo (etapa V) consiste na extinção do potencial  $u(\mathbf{r}^P)$ , o que libera a estrutura interna da proteína para o estado final, com ligantes. A integral de configuração é:

$$\frac{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta U'(\mathbf{r}^P, \mathbf{R}^N)}}{\int d\mathbf{r}^P \int_{\nu} d\mathbf{R}^n \int_{V-\nu} d\mathbf{R}^{N-n} e^{-\beta[U'(\mathbf{r}^P, \mathbf{R}^N) + u(\mathbf{r}^P)]}} = \frac{1}{\langle e^{-\beta u(\mathbf{r}^P)} \rangle_{Pn}} \quad (38)$$

a qual está relacionada com a energia liberada quando a estrutura interna da proteína torna-se a estrutura final. Nesse etapa, a média termodinâmica ocorreu segundo um *ensemble* de estruturas de uma simulação de dinâmica molecular com ligantes na concentração  $\bar{\rho}$ .

### III.3.8 Reconstrução da energia livre da partição

Após a releitura do processo de partição em passos hipotéticos mais simples, a variação de energia livre ( $\Delta W$ ) pode ser reescrita em função desses parâmetros, ao juntar as equações 23, 25, 29, 33, 36 e 38:

$$\begin{aligned}
 e^{-\beta\Delta W} &= \frac{N!}{(N-n)!n!} \times \langle e^{-\beta u(\mathbf{r}^P)} \rangle_P \times e^{-\beta W_n^*} \times \left( \frac{\nu}{V-\nu} \right)^n \times e^{+\beta W_n^{**}} \times \frac{1}{\langle e^{-\beta u(\mathbf{r}^P)} \rangle_{Pn}} \\
 &= \\
 e^{-\beta\Delta W} &= \frac{N!}{(N-n)!n!} \left( \frac{\nu}{V-\nu} \right)^n \times \frac{\langle e^{-\beta u(\mathbf{r}^P)} \rangle_P}{\langle e^{-\beta u(\mathbf{r}^P)} \rangle_{Pn}} \times e^{-\beta(W_n^* - W_n^{**})}
 \end{aligned} \tag{39}$$

É possível ver os termos em três grupos, cada qual com sua influência específica na energia livre do processo. No primeiro grupo estão os termos relacionados à entropia:

$$\frac{N!}{(N-n)!n!} \left( \frac{\nu}{V-\nu} \right)^n = e^{-\beta\Delta W_S}.$$

Pode-se notar pelos termos entrópicos que quanto maior o número total de ligantes ou maior o volume do sítio  $\nu$ , mais fácil a ocorrência do processo. No caso do limite termodinâmico ( $N \gg n$ ), a expressão  $\frac{N!}{(N-n)!n!} = \omega$  pode ser aproximada para  $\frac{N^n}{n!}$ . A energia livre correspondente ao grupo entrópico é dada por  $\Delta W_S = -\frac{1}{\beta} \ln \frac{N!}{(N-n)!n!} \left( \frac{\nu}{V-\nu} \right)^n$ .

O segundo grupo é composto pelos termos relativos à mudança de estrutura interna da proteína:

$$\frac{\langle e^{-\beta u(\mathbf{r}^P)} \rangle_P}{\langle e^{-\beta u(\mathbf{r}^P)} \rangle_{Pn}} = e^{-\beta\Delta W_P}$$

onde  $\Delta W_P$  é o trabalho reversível realizado pela perturbação imposta pela presença dos ligantes, que pode influenciar alguma mudança de estruturas internas na proteína.

Por fim, o último grupo possui os termos relacionados à transferência dos  $n$  ligantes entre os ambientes (reservatório, sítio e vácuo):

$$e^{-\beta(W_n^* - W_n^{**})} = e^{-\beta\Delta W_L}$$

$\Delta W_L$  é a diferença das energias de dessolvatação. Quanto mais difícil dessolvatar o ligante do sítio, maior o valor de  $W_n^{**}$ , e por consequência mais espontânea é a ligação.

A remoção gradual de ligantes altera a concentração tanto do reservatório quanto da fase de interface proteica. Logo, o cálculo do  $\Delta W_L$  deve levar o efeito de concentração para a transferência de cada ligante individualmente:

$$\Delta W_L = W_n^* - W_n^{**} = \sum_{i=1}^n W_i^* - W_i^{**} \quad (40)$$

onde  $W_i$  é a energia necessária para a transferência do  $i$ -ésimo ligante após  $i - 1$  ligantes já terem sido transferidos.

Assim, a estabilidade geral pode ser reescrita como:

$$e^{-\beta\Delta W} = \omega \left( \frac{\nu}{V - \nu} \right)^n e^{-\beta[\Delta W_P + W_n^* - W_n^{**}]} \quad (41)$$

### III.3.9 Energia livre da partição em condições diluídas

Na condição de diluição, há ocorrência do limite termodinâmico ( $\omega \approx \frac{N^n}{n!}$ ), a perturbação dos ligantes na estrutura da proteína torna-se nula ( $\Delta W_P = 0$ ) e as energias livres de dessolvatação convergem para os valores diluídos, ou seja  $W_n^* = n\bar{\mu}$  e  $W_n^{**} = n\bar{w}$ . Substituindo na equação 41:

$$e^{-\beta\Delta W} = \frac{N^n}{n!} \left( \frac{\nu}{V - \nu} \right)^n e^{-n\beta[\bar{\mu} - \bar{w}]} \quad (42)$$

Substituindo  $e^{-n\beta[\bar{\mu} - \bar{w}]} = \wp^n$  (equação 7):

$$e^{-\beta\Delta W} = \frac{N^n}{n!} \left( \frac{\nu\wp}{V - \nu} \right)^n = \left( \frac{N}{V - \nu} \right)^n \frac{\nu^n}{n!} \wp^n \quad (43)$$

É possível ver que a relação entre constante de ligação conseguiu ser resgatada quando o sistema está na condição de baixas concentrações, mostrando que o coeficiente de partição pode ser uma propriedade bastante útil ao estudar interações de pequenas moléculas com proteínas nas condições pouco concentradas, oferecendo uma ótima aproximação nessas condições.

A aproximação nas condições de diluição também mostram que o coeficiente de partição pode ser visto como uma espécie de constante de ligação agregada, o que torna possível a caracterização termodinâmica do processo utilizando toda teoria já desenvolvida para os casos de ligação clássicos[5, 17].

$$\left( \frac{N}{V - \nu} \right)^n \frac{\nu^n}{n!} \wp^n = \left( \frac{N}{V - \nu} \right)^n K(n) = \bar{\rho}^n K(n) \quad (44)$$



### III.3.10 Projeção de propriedades para outras concentrações

Um dos principais objetivos na obtenção do coeficiente de partição é a capacidade de descrever facilmente as propriedades termodinâmicas desse processo. Nesse trabalho, juntamente com o conjunto de ferramentas teóricas e computacionais para o cálculo de  $\varphi$ , também são trazidas as ferramentas para reconstrução de propriedades sem a necessidade de realizar novas simulações, principalmente em concentrações mais diluídas, levando a uma grande economia de tempo computacional. Primeiramente, podemos encontrar  $\langle n \rangle$  para qualquer concentração desejada a partir da equação 7:

$$\langle n \rangle \approx N \left( 1 + \frac{V - \nu}{\nu\varphi} \right)^{-1} = \bar{\rho}V \left( 1 + \frac{V - \nu}{\nu\varphi} \right)^{-1} \quad (45)$$

Com  $\langle n \rangle$  em mãos, também é possível calcular como será a distribuição do ligante naquela concentração, utilizando a densidade espacial. Considerando uma densidade unitária  $\bar{\rho}'(\mathbf{R})$ , ou seja,  $\int \bar{\rho}'(\mathbf{R})d\mathbf{R} = 1$ , basta multiplicar a densidade espacial unitária pelo número médio:

$$\bar{\rho}(\mathbf{R}) = \langle n \rangle \times \bar{\rho}'(\mathbf{R}) \quad (46)$$

A densidade unitária é obtida a partir das densidades tridimensionais da simulação de 150mM. Detalhes sobre as ferramentas utilizadas para obtenção dos mapas tridimensionais são dados na seção III.4.6.3.

## III.4 Protocolos

### III.4.1 Simulações em geral

O programa Visual Molecular Dynamics (VMD)[42], seus diversos *plugins* associados, e um pacote em Python acessório [11] foram utilizados para a criação de todos dos sistemas.

Todas as simulações de dinâmica molecular foram realizadas utilizando o NAMD 2.12 [32], com o campo de força CHARMM36 [26] para água, proteínas e lipídios. O modelo de água utilizado foi o modelo TIP3 [43]. Foram utilizadas as condições periódicas de contorno nos eixos  $x$ ,  $y$  e  $z$ . O raio de *cutoff* utilizado foi de 11Å, e os

cálculos eletrostáticos de longa distância foram realizados pelo método PME. O passo de integração adotado foi de 2fs. O controle de temperatura foi realizado pela dinâmica de Langevin. Todas as simulações, nos primeiros 0,08ns, sofrem uma termalização gradual de 273K para 300K.

### III.4.2 Simulações de FEP

Os cálculos de energia livre via Free Energy Perturbation (FEP) foram realizados no NAMD. A simulação realizou o desacoplamento em 47 transformações progressivas:  $\Delta\lambda = 0,05$  quando  $0 \leq \lambda < 0,5$ ,  $\Delta\lambda = 0,025$  quando  $0,5 \leq \lambda < 0,8$ ,  $\Delta\lambda = 0,01$  quando  $0,8 \leq \lambda < 0,95$  e  $\Delta\lambda = 0,005$  até  $\lambda = 1$ . A energia de cada microestado é coletada a cada 100 passos e cada janela de transformação possui 275000 passos (0,55ns), onde os 25000 primeiros passos correspondem à equilibração pós-transformação e a coleta dos dados ocorre ao longo dos outros 250000 passos. Ou seja, numa réplica, são coletados por transformação, 2500 estados, totalizando 129250 estados para a transformação completa. Neste, trabalho, não será feita a transformação reversa, de  $\lambda = 1$  para  $\lambda = 0$ , pois a amostragem partindo de um estado desacoplado para um estado com todas as interações pode sofrer efeito de histerese, um fenômeno que mesmo realizando o processo inverso, não há garantia de retorno para o estado inicial. Dessa maneira, a comparação entre os estados de ida e volta não seria equivalente, levando à um resultado inacurado. Além disso, utilizar réplicas somente para a transformação de ida não causa prejuízos na estimativa de energia livre, pois o método SOS [35] consegue levar em conta transformações de somente um sentido [17, 16, 15].

### III.4.3 Simulações com membrana

#### III.4.3.1 Criação dos sistemas equilibrados

A membrana utilizada nos sistemas é uma bicamada fosfolipídica de 1-palmitoil-2-oleoil-sn-glicero-3-fosfatidilcolina (POPC). Um pequeno segmento de 58 moléculas de POPC (29 em cada camada) já realizado em outros estudos [15] foi replicado 2 vezes nas dimensões  $x$  e  $y$  pelo *plugin* TopoTools [44] do VMD, totalizando uma membrana com

232 moléculas. Na sequência, o segmento de membrana foi solvatado por moléculas de água através do *plugin* Solvate, e ionizado com uma solução de 150mM de *KCl* usando o *plugin* Autoionize.

Com a membrana já feita, resta adicionar as moléculas de sevoflurano, de acordo com a simulação desejada. Os sistemas foram submetidos à dinâmica molecular seguindo os parâmetros da seção III.4.1 nas características NPT: número de moléculas, temperatura (300 K) e pressão (1 atm) constantes, onde a manutenção da pressão é realizada pelo pistão de Langevin. Cada sistema possui seu respectivo tempo de simulação, que é encerrada após algum tempo dos ligantes equilibrarem sua distribuição no sistema. O final das simulações corresponde ao *input* para a criação dos sistemas utilizados em outras técnicas. Foi utilizado um campo de força do modelo CHARMM36 disponível para o sevoflurano [25]. A Tabela 1 lista as características dos sistemas membrana-sevoflurano em relação ao número de moléculas adicionadas e seu tempo total de simulação.

**Tabela 1:** Concentração desejada, fração molar ligante-POPC, número de ligantes adicionados e tempo de simulação nos sistemas com sevoflurano.

Concentração	Fração molar	Ligantes adicionados	Tempo de simulação
1mM	0,0043	1	150ns
25mM	0,06	14	140ns
50mM	0,11	28	100ns
75mM	0,16	44	120ns
100mM	0,21	61	95ns
150mM	0,29	94	276ns

#### III.4.3.2 Simulações de FEP na membrana

Com o objetivo de obter as energias de dessolvatação (vide seção III.3.4), três réplicas da configuração do final da simulação de cada sistema foram utilizadas como estrutura inicial para os cálculos de FEP (vide seção III.2). A molécula que sofre a transformação é escolhida aleatoriamente, com a única exigência que esteja presente na membrana. Para

assegurar que a molécula permaneça na membrana durante todo o cálculo, principalmente quando está muito desacoplada, foi adicionado por meio da seção de *colvars* do NAMD um potencial harmônico do tipo *flat-well* ( $U_{fw}(\mathbf{z}_i)$ ) na molécula  $i$  que sofre a transformação, para mantê-la na membrana. O potencial possui forma semelhante ao da equação 26:

$$U_{fw}(\mathbf{z}_i) = \begin{cases} 0, & \text{se } \mathbf{z}_{down} < \mathbf{z}_i < \mathbf{z}_{up} \\ k_f(\mathbf{z}_i - \mathbf{z}_{down})^2, & \text{se } \mathbf{z}_i \leq \mathbf{z}_{down} \\ k_f(\mathbf{z}_i - \mathbf{z}_{up})^2, & \text{se } \mathbf{z}_i \geq \mathbf{z}_{up} \end{cases} \quad (47)$$

onde  $\mathbf{z}_{down}$  e  $\mathbf{z}_{up}$  representam os limites inferior e superior em  $z$  da região nula do potencial. Eles são calculados de acordo com as posições dos ligantes em  $z$  ao longo da simulação de equilíbrio: é escolhida uma camada lipídica (superior ou inferior), onde a média  $\bar{z}$  dessa distribuição juntamente com o desvio padrão  $\sigma$  delimitam as coordenadas do cálculo do FEP, no qual  $\mathbf{z}_{down} = \bar{z} - \sigma$  e  $\mathbf{z}_{up} = \bar{z} + \sigma$ .

Na intenção de realizar uma equilibração prévia, cada réplica foi submetida a uma dinâmica molecular com a presença dos *colvars*, porém sem a realização dos cálculos de FEP, de 400 passos de minimização e 120000 passos de integração antes do início do FEP.

#### III.4.4 Simulações de *flooding* Kv1.2-sevoflurano

Uma simulação de 300ns do canal iônico de potássio Kv1.2 na conformação aberta, embebido em uma bicamada lipídica de POPC com 150mM de sevoflurano foi obtida de trabalhos anteriores já realizados no laboratório. Essa simulação cedida possui 426 moléculas de POPC e 174 moléculas de sevoflurano, resultando numa fração molar de 0,29. Usando o VMD, a criação dos sistemas correspondentes às outras concentrações desejadas tomou como ponto de partida o sistema inicial de 150mM, com a remoção dos ligantes até alcançar a concentração desejada. A Tabela 2 enumera as características dos sistemas iniciais de *flooding* com o sevoflurano que foram simuladas. Todas as simulações tiveram a duração de 300ns cada e ocorreram em condições semelhantes às citadas na seção III.4.1, na característica NPT: pressão de 1 atm e temperatura de 300 K mantidas constantes. O campo de força para o sevoflurano utilizado foi do tipo CHARMM [25].

**Tabela 2:** Concentração da desejada, fração molar ligante-POPC, e número total de ligantes dos sistemas de *flooding* com sevoflurano.

Concentração	Fração molar	Número total de ligantes
1mM	0,0047	2
25mM	0,06	29
50mM	0,11	58
75mM	0,16	87
100mM	0,21	116
150mM	0,29	174

#### III.4.4.1 Simulações de FEP na proteína

Com o objetivo de obter as energias de dessolvatação sítio-vácuo, três réplicas da configuração do final da simulação de cada sistema foram utilizadas como estrutura inicial para os cálculos de FEP (vide seção III.2). A molécula que sofre a transformação é escolhida aleatoriamente, com a única exigência que esteja presente na superfície da proteína. Para assegurar que a molécula permaneça na vizinhança do sítio durante todo o cálculo, principalmente quando está muito desacoplada, foi adicionado por meio da seção de *colvars* do NAMD um potencial harmônico do tipo esférico na molécula  $i$  que sofre a transformação, para mantê-la no sítio. O potencial possui forma:

$$U_{sphere}(\mathbf{R}) = \begin{cases} 0, & \text{se } |\mathbf{R}_i - \mathbf{R}_c| < R^* \\ k_f(|\mathbf{R}_i - \mathbf{R}_c| - R^*)^2, & \text{se } |\mathbf{R}_i - \mathbf{R}_c| \geq R^* \end{cases} \quad (48)$$

onde  $\mathbf{R}_i$  e  $\mathbf{R}_c$  representam as coordenadas do centro do ligante e do centro do sítio, respectivamente.  $k_f$  é a constante de força e  $R^*$  é o raio do sítio. A constante de força adotada e o tamanho do sítio vieram de trabalhos que já realizaram cálculos de FEP entre sevoflurano e Kv1.2. Assim como nos FEPs da membrana, cada réplica foi submetida a uma dinâmica molecular com a presença dos *colvars*, sem a realização dos cálculos de FEP, de 400 passos de minimização e 120000 passos de integração antes do início do FEP.

### III.4.5 Simulações de *flooding* da albumina e tricloroetanol

Para realizar a prova de conceito da obtenção do coeficiente de partição utilizando um sistema independente, foi escolhido o par albumina-tricloroetanol. Os sistemas consistiam na estrutura da albumina bovina sérica, obtida pelo do banco de dados PDB - ID: 4F5S[45], uma caixa de solvatação de água e 43 ou 85 moléculas de tricloroetanol, correspondendo às concentrações de 50 e 100mM respectivamente. O volume total do sistema corresponde a  $1.361.310 \text{ \AA}^3$ . O campo de força utilizado para o tricloroetanol foi oriundo de um estudo de parametrização no qual fui coautor[15] (Anexo II). As simulações ocorrem por 300 ns com os mesmos parâmetros citadas no protocolo geral para as simulações (seção III.4.1).

### III.4.6 Análises

As análises das simulações de *flooding* foram realizadas utilizando o programa VMD [42], seus *plugins* e o pacote em *Python* acessório [11].

#### III.4.6.1 Obtenção dos volumes

Todos os volumes foram calculados utilizando o sistema inicial da simulação de *flooding* de 150mM. O volume total ( $V_{total}$ ) do sistema foi calculado a partir da célula unitária das condições periódicas de contorno da simulação. O volume da proteína ( $V_{prot}$ ) foi calculado utilizando o MSMS[46], com a sondagem (*probe*) de  $1,4 \text{ \AA}$ . O volume da superfície proteica ( $\nu$ ) foi calculado como o volume de proteína com o *probe* de  $5 \text{ \AA}$  decrescido por  $V_{prot}$ . O volume da membrana ( $V_{memb}$ ) foi calculado com o *probe* de  $1,4 \text{ \AA}$ , enquanto o volume da água foi considerado como o volume restante ( $V_{wat} = V_{total} - V_{prot} - V_{memb} - \nu$ ).

#### III.4.6.2 Obtenção das distribuições de probabilidades e número médio

As distribuições de probabilidades foram obtidas através da leitura direta dos sítios, onde foi contada qual a fração da simulação o sítio estava ocupado por  $n$  ligantes -  $p(n)$ .

O número médio de ligantes no sítio ( $\langle n \rangle$ ) é dado pela média da distribuição:

$$\langle n \rangle = \sum_{i=0}^{n_{max}} p(i) \times i \quad (49)$$

### III.4.6.3 Obtenção das densidades tridimensionais

As densidades tridimensionais ( $\rho(\mathbf{R})$ ) foram obtidas através do *plugin* Volmap utilizando o parâmetro de densidade, com o mapa discretizado em cubos de  $1 \text{ \AA}^3$ . A análise dos *grids* gerados foi realizada pelo pacote GridDataFormats[47]. A comparação entre duas densidades tridimensionais **A** e **B** é feita através da análise de sobreposição (ou *overlap*):

$$o(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A} \cap \mathbf{B}}{\mathbf{A} \cup \mathbf{B}} \quad (50)$$

onde o valor de 100% indica sobreposição absoluta e 0% nenhuma sobreposição.

### III.4.6.4 Análise da estabilidade por RMSD

A estabilidade dos sítios pode ser mensurada através de uma medida de desvio chamada de raiz do desvio quadrático médio, ou RMSD (*root mean square deviation*). O RMSD indica quão distante em média as coordenadas estão de uma posição referência. O RMSD é calculado como:

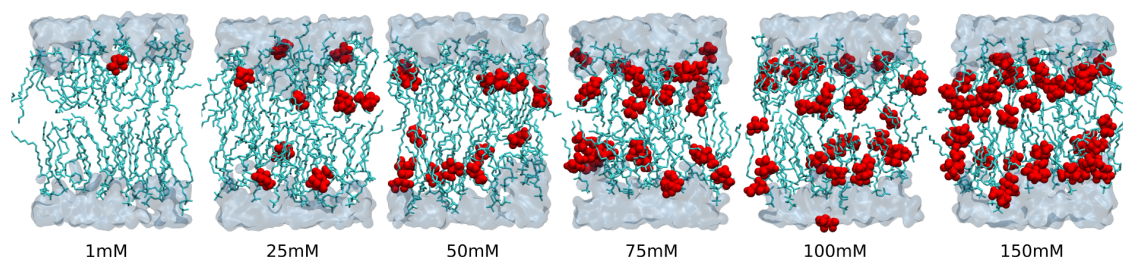
$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N |\mathbf{R}_i - \mathbf{R}_0|^2} \quad (51)$$

onde  $\mathbf{R}_0$  é a coordenada do átomo  $i$  na posição de referência e  $\mathbf{R}_i$  é a coordenada desse mesmo átomo no instante do cálculo. Nesse trabalho, as análises de RMSD são feitas utilizando o VMD, e a posição de referência utilizada sempre é a primeira posição da simulação.

## IV Resultados e discussão

### IV.1 Sistemas com membrana

#### IV.1.1 Equilibração

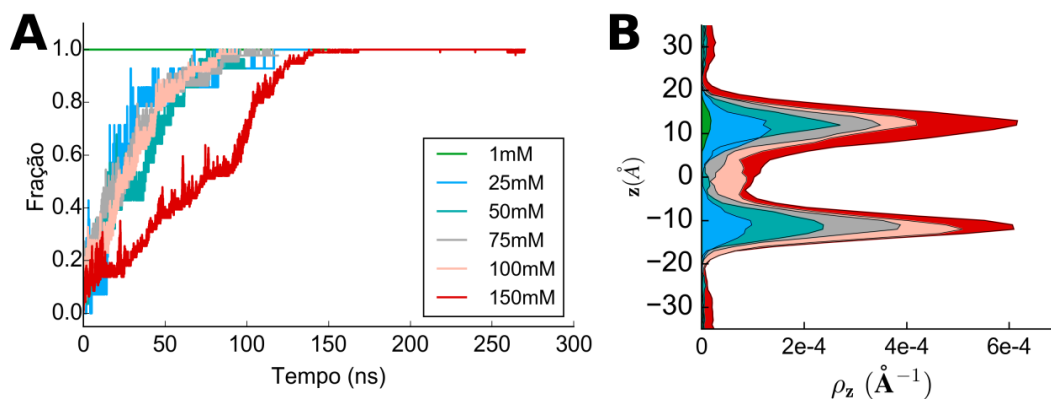


**Figura 11:** Sistemas com membrana e sevoflurano. Seção transversal dos sistemas de *flooding* com membrana após a equilibração para diversas concentrações. A composição do sistema é de uma bicamada lipídica de POPC (ciano), moléculas de água (azul) e sevoflurano (vermelho).

Apesar da membrana ter origem em um segmento previamente equilibrado, a replicação, adição de moléculas de ligante e a solvatação levam à necessidade de avaliar a estabilidade do novo sistema. Não foi observada nenhuma passagem indesejada de íons através da membrana ou nenhum evento de extravasamento de água. As simulações no geral tiveram no início a agregação dos ligantes na solução aquosa no início, seguido pelo particionamento gradual para a membrana, que é completado quando o sistema atinge equilíbrio. A divergência do tempo de simulação dos sistemas (Tabela 1) é oriunda do fato dos ligantes apresentarem mudanças de taxas de particionamento na membrana em concentrações diferentes. Uma vez alcançado o equilíbrio, a trajetória ainda é estendida por um tempo para a coleta dos dados. A Figura 11 ilustra os sistemas no equilíbrio e a Figura 12 mostra a fração de ligantes que particionam na membrana ao longo do tempo. Por essas figuras é possível afirmar que todas as simulações entraram em equilíbrio.

Considerando o tempo que a simulação está em equilíbrio, foi analisada a distribuição no eixo  $z$  dos ligantes em cada simulação, para determinar os limites do potencial *flat-well* que seria aplicado nas simulações de FEP (seção III.4.3.2). Foi escolhida a distribuição da camada superior da membrana, para fins de tornar a região de amostragem mais homogênea entre as réplicas do FEP. A média e desvio padrão das distribuições são





**Figura 12:** Propriedades da membrana na presença de diversas concentrações de sevoflurano. A - Quantidade relativa de moléculas de ligante que permanecem na região da membrana ao longo do tempo. B - Densidade espacial do sevoflurano ao longo do eixo  $z$ . Para ambos os gráficos, cada cor representa um sistema com concentração específica: Verde - Sevoflurano 1mM; Azul - Sevoflurano 25mM; Ciano - Sevoflurano 50mM; Cinza - Sevoflurano 75mM; Rosa - Sevoflurano 100mM; Preto - Sevoflurano 150mM. A região da membrana é delimitada pelos átomos de fósforo da cadeia polar do POPC.

encontradas na Tabela 3.

**Tabela 3:** Média ( $\bar{z}$ ) e desvio padrão ( $\sigma_z$ ) da posição em  $z$  dos ligantes presentes na região superior da membrana, após o equilíbrio.

Simulação	$\bar{z}$	$\sigma_z$
Sevoflurano 1mM	11,83Å	2,71Å
Sevoflurano 25mM	10,36Å	3,3Å
Sevoflurano 50mM	13,23Å	6,68Å
Sevoflurano 75mM	11,63	5,39Å
Sevoflurano 100mM	10,89Å	5,05Å
Sevoflurano 150mM	11,80Å	6,45Å

#### IV.1.2 Cálculos de FEP e energia de dessolvatação

Após a equilibração, foram feitos três cálculos de FEP independentes para calcular a energia de dessolvatação de uma molécula de sevoflurano na membrana. O cálculo de  $W^*$  e o erro estimado ocorreram segundo o protocolo na seção III.4.3.2 e estão listados

na tabela 4.

**Tabela 4:** Variações de energia para o desacoplamento do sevoflurano na membrana através de cálculos de FEP ( $W_{memb}^*$ ) e erro estatístico associado, estimadas pelo método SOS. A variação de energia calculada corresponde à energia necessária para a transferência de uma molécula de sevoflurano partindo do ambiente para o vácuo.

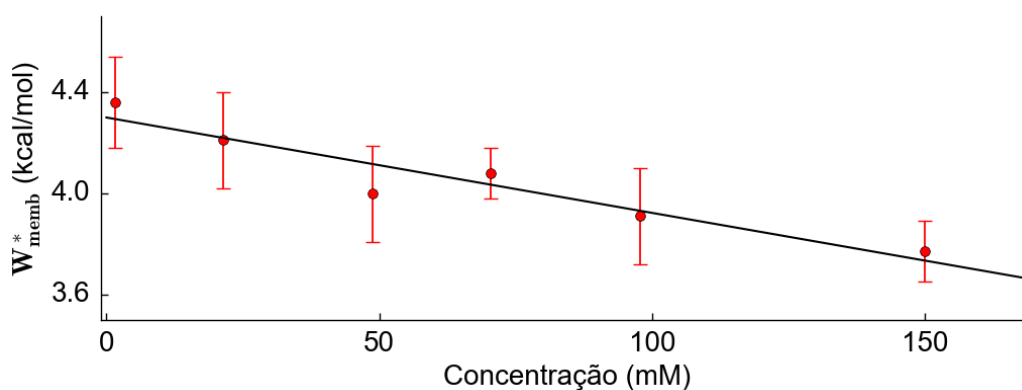
Concentração (mM)	$W_{memb}^*$ (kcal/mol)
1	$4,36 \pm 0,18$
25	$4,21 \pm 0,19$
50	$4,00 \pm 0,19$
75	$4,08 \pm 0,10$
100	$3,91 \pm 0,19$
150	$3,77 \pm 0,12$

Com os valores de  $W_{memb}^*$  para cada concentração, foi possível realizar a regressão de acordo com a equação 30:

$$W_{memb}^* = \bar{\mu} - h_{memb}^* \frac{N - n}{V - \nu} c$$

A regressão retornou os valores de  $\bar{\mu}$  e  $h_{memb}^*$  como 4,3 e 1,9 kcal/mol, respectivamente. Assim, o  $\bar{\mu}$  possui um valor próximo à energia de partição do sevoflurano na membrana obtida através de outra metodologia [15], corroborando a qualidade dos resultados obtidos e do modelo teórico utilizado. Já o fato de  $h_{memb}^* > 0$  pode ser interpretado como o consequência das interações ligante-ligante não serem favoráveis, o que pode ser mostrado pelo valor de  $W^*$  diminuir a medida que a fração molar aumenta, tornando a ligação na solução não-ideal um fenômeno que necessita de menos energia comparado à situação de soluções ideais ou suficientemente diluídas, onde  $N \rightarrow 0$  por definição[39]. A Figura 13 ilustra a distribuição das energias de dessolvatação e os valores obtidos pela regressão.

É importante ressaltar que o fato de apesar da retirada do ligante ser mais fácil à medida que a concentração aumenta, não há garantia de que o processo completo de partição seja facilitado, pois ainda há a necessidade de conferir a influência que a etapa



**Figura 13:** Energias de dessolvatação do sevoflurano a partir da membrana de POPC obtidas através dos cálculos de FEP (pontos vermelhos) e valores em função da concentração efetiva no reservatório  $\frac{N-n}{V-n}$ , obtidos através da regressão do modelo  $W^*_{memb} = \bar{\mu} - h^* \times \frac{N-n}{V-n} \times c$  (linha preta), onde  $\bar{\mu} = 4,3$  kcal/mol,  $h^* = 1,9$  kcal/mol e  $C^0$  é o inverso da concentração padrão ( $c = (1 \text{ mol/L})^{-1}$ ).

de ligação do sítio sofre da concentração.

Totalizada a caracterização energética das duas fases e usando a equação 21, as energias de dessolvatação resultantes para cada concentração estão listadas na Tabela 5.

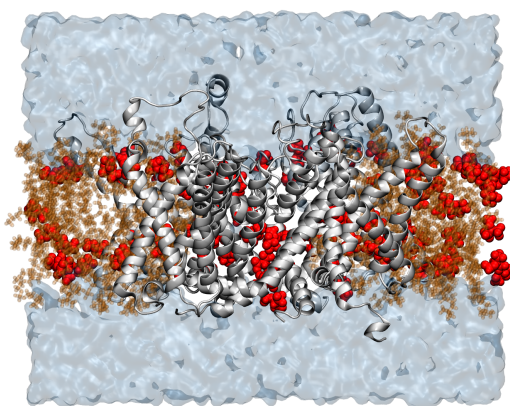
**Tabela 5:** Energia média de dessolvatação resultante para os dois componentes do reservatório ( $W^*$ ) e erro estatístico associado, em kcal/mol. A energia resultante é oriunda do escalonamento das energias de dessolvatação e do volume de cada componente.

Concentração (mM)	$W^*$ (kcal/mol)
1	3,80 ± 0,13
25	3,64 ± 0,13
50	3,43 ± 0,13
75	3,51 ± 0,09
100	3,34 ± 0,13
150	3,2 ± 0,10

## IV.2 Simulações de *flooding*

### IV.2.1 Equilibração dos sistemas

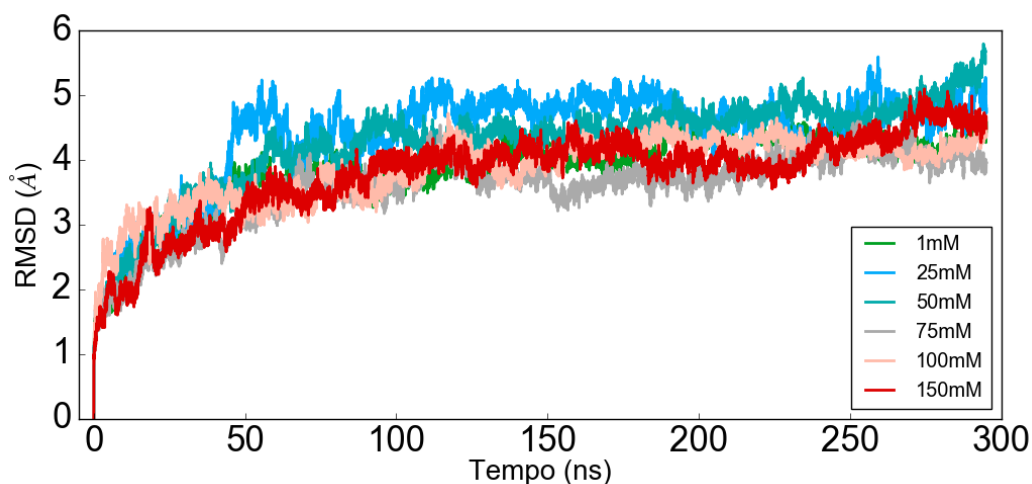
A análise de estabilidade de uma simulação de *flooding* passa por algumas etapas. Primeiramente, é visto se a água não apresenta algum comportamento anormal, como algum extravasamento causado por uma delimitação de volume errônea. Todas as simulações foram conferidas visualmente, e não houve nenhum movimento de moléculas de água ou íons para locais indesejados. A Figura 14 ilustra o final de uma simulação de *flooding*. Nela é possível ver que o sistema possui todos seus componentes íntegros.



**Figura 14:** Sistema final de uma simulação de *flooding* do canal de potássio Kv1.2 (branco) com 150mM do anestésico geral sevoflurano (vermelho). O canal iônico está embebido em uma bicamada lipídica de POPC (laranja). A água está representada em azul e o íons foram ocultados para melhor visualização.

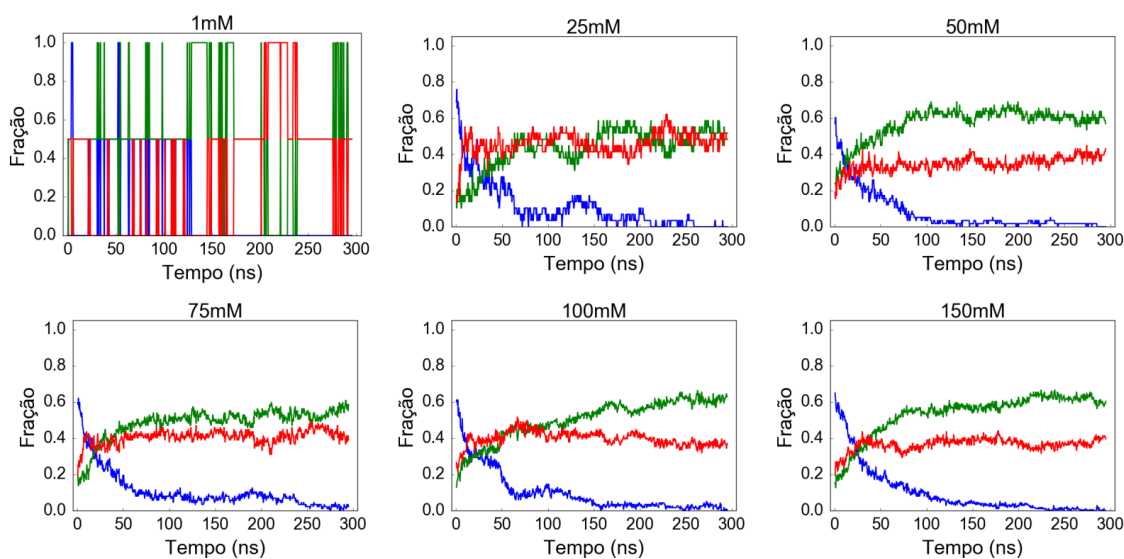
Porém, o critério visual não é suficiente para avaliar a estabilidade do sistema como um todo. Dependendo das condições de equilíbrio, a proteína pode passar por mudanças conformacionais que podem afetar sua estabilidade, levando a fenômenos indesejados. Assim, é desejável observar a análise da estabilidade da proteína utilizando medidas de dispersão, como o RMSD (seção III.4.6.4). Os valores de RMSD para todas as simulações estão listadas na Figura 15. Por ela é possível ver que, a medida que as simulações avançam, as proteínas possuem um RMSD flutuando em torno de um valor médio, indicando o equilíbrio conformacional das proteínas.

Resta avaliar o comportamento dos ligantes para definir o equilíbrio de todo o *flooding*. Será mensurada a fração de ligantes presentes na solução aquosa, particionados na



**Figura 15:** Valores de RMSD ao longo do tempo de simulação dos sistemas de *flooding*. Verde - Sevoflurano 1mM; Azul - Sevoflurano 25mM; Ciano - Sevoflurano 50mM; Cinza - Sevoflurano 75mM; Rosa - Sevoflurano 100mM; Preto - Sevoflurano 150mM.

membrana (como realizado na seção IV.1.1), e a fração que está ligada no canal iônico ao longo de toda simulação (Figura 16).



**Figura 16:** Quantidade relativa de moléculas de ligante que permanecem na região da solução aquosa, membrana ou proteína ao longo do tempo. A região da membrana é delimitada pelos átomos de fósforo da cadeia polar do POPC, enquanto as moléculas ligadas à proteína são moléculas que possuem algum átomo a uma distância de até 5Å da proteína. No caso de essas condições serem mutuamente satisfeitas, a molécula é considerada ligada à proteína. Vermelho: ligados à proteína; Verde: presentes na membrana; Azul: presentes na água.

Pelos perfis de partição ao longo do tempo, é possível confirmar que em todas as simulações, os ligantes entraram em equilíbrio com o sistema. Além disso, é possível ver o alto grau de hidrofobicidade das moléculas, pois elas rapidamente saem da água e particionam-se na membrana e proteína, sem grandes passagens de moléculas de volta para a água.

Tomando em conjunto as análises realizadas acima, é possível afirmar que todos os componentes do sistema, e por consequência todas as simulações de *flooding* realizadas, estão estáveis e no equilíbrio, principalmente a partir dos 100 ns de simulação, intervalo de tempo que iniciará a coleta de dados.

Os volumes calculados a partir do estado inicial da simulação de 150mM estão listados na Tabela 6.

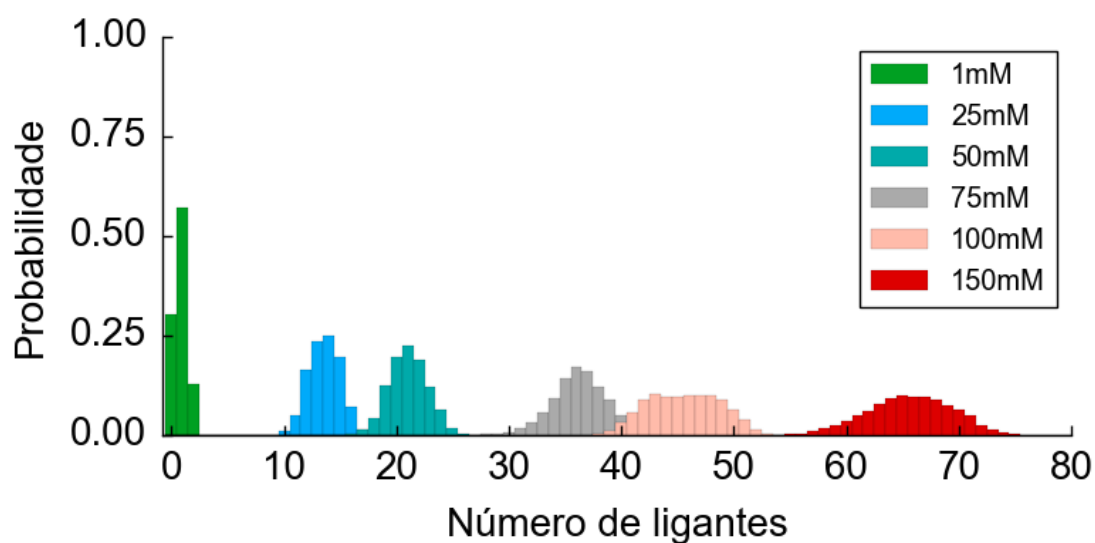
**Tabela 6:** Volumes (em  $\text{\AA}^3$ ) do sistema calculados a partir do início das simulações de *flooding*. Legenda:  $V_{total}$  = volume total;  $V_{prot}$  = volume da proteína;  $\nu$  = volume do sítio proteico;  $V_{memb}$  = volume da membrana e  $V_{wat}$  = volume da água.

$V_{total}$	$V_{prot}$	$\nu$	$V_{memb}$	$V_{wat}$
1.872.081 $\text{\AA}^3$	142.755 $\text{\AA}^3$	159.713 $\text{\AA}^3$	602.027 $\text{\AA}^3$	967.586 $\text{\AA}^3$

## IV.2.2 Propriedades das simulações

Com todos os sistemas considerados estáveis, o próximo passo é obter as propriedades da interação para cada simulação. O canal Kv1.2 possui vários candidatos a sítios de ligação [16, 18, 11], distribuídos praticamente ao longo de toda superfície proteica. Para uma molécula ser considerada ligada no sítio, é necessário que um átomo esteja a menos de 5  $\text{\AA}$  de distância de algum aminoácido. As probabilidades de cada estado de ocupação do sítio ( $p(n)$ ) a partir dos 100 ns de simulação e o número médio ( $\langle n \rangle = \sum p(n) \times n$ ) em cada uma delas estão presentes na Figura 17 e na Tabela 7.

Um grande diferencial das simulações de *flooding* é a possibilidade de obter a distribuição da densidade tridimensional do ligante nos sítios ( $\rho(\mathbf{R})$ ). Essas regiões costumam oferecer importantes informações acerca da interação, principalmente sobre a afinidade



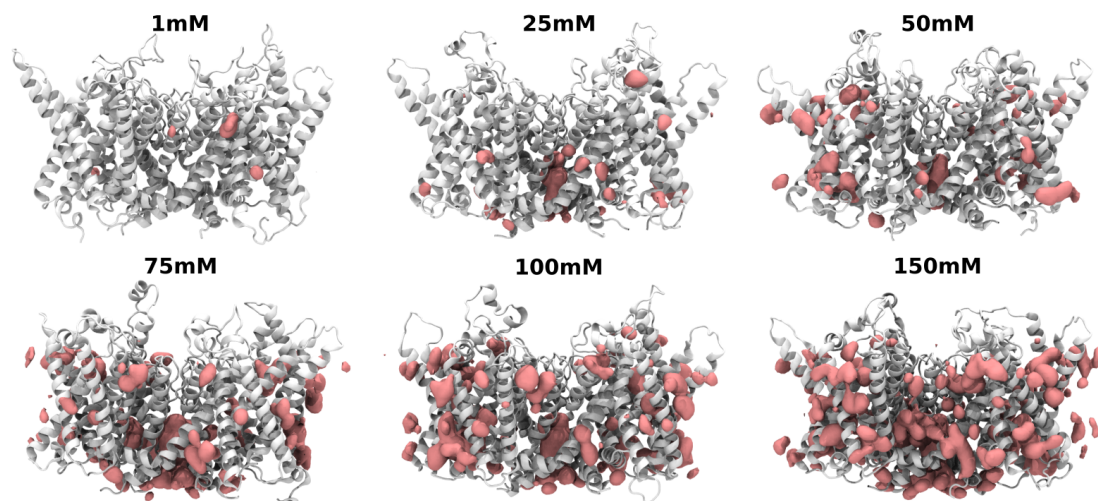
**Figura 17:** Probabilidades dos estados de ocupância do canal Kv1.2 com o anestésicos sevoflurano em diferentes concentrações. Verde - Sevoflurano 1mM; Azul - Sevoflurano 25mM; Ciano - Sevoflurano 50mM; Cinza - Sevoflurano 75mM; Rosa - Sevoflurano 100mM; Preto - Sevoflurano 150mM.

**Tabela 7:** Número médio ( $\langle n \rangle$ ) de ligantes em cada concentração, obtidos diretamente nas simulações de *flooding*, a partir dos 100 ns de simulação.

	1mM	25mM	50mM	75mM	100mM	150mM
$\langle n \rangle$	$0,82 \pm 0,23$	$13,67 \pm 0,99$	$22,87 \pm 1,04$	$36,30 \pm 1,25$	$45,57 \pm 1,16$	$66,10 \pm 1,29$

do sítio, detalhes sobre a dispersão do ligante no sítio e número médio de ligante. Cabe lembrar que, a densidade tridimensional é uma propriedade que também pode ser reconstruída através das projeções provenientes do coeficiente de partição (equação 46, seção III.3.10). Na Figura 18 estão mostradas as densidades tridimensionais de todas as simulações de *flooding*.

Os ligantes associam-se à interface proteica em quase toda integridade da região transmembrânica, onde inclusive diversos pontos possuem densidades comparáveis. Além disso, fica evidente o motivo de que o melhor mapa de densidade unitário  $\bar{\rho}'(\mathbf{R})$  deve ser oriundo da simulação de 150mM. Uma alta concentração é necessária devido à necessidade de convergência espacial, para uma maior qualidade do mapa. Simulações de menor concentração, como 1mM por exemplo, necessitariam de escalas de tempo muito dispendiosas para atingir mapas com dispersão semelhantes ao de altíssimas concentrações.

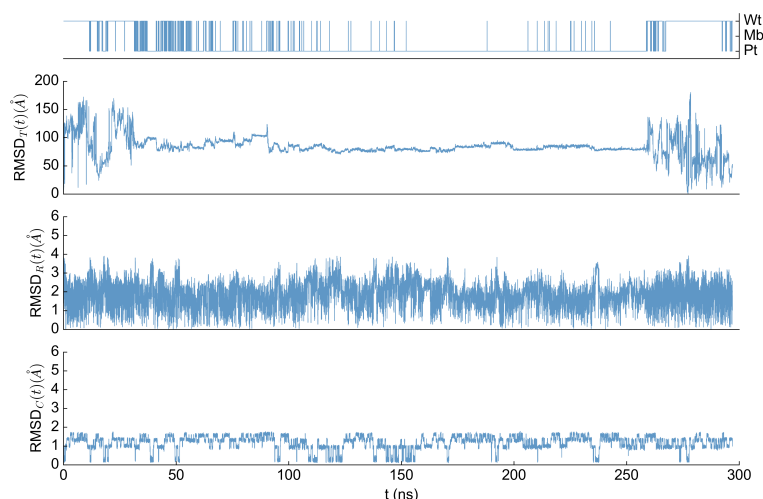


**Figura 18:** Estrutura média da proteína (em branco) e densidades tridimensionais (em rosa) obtidas a partir dos últimos 100 ns das simulações de *flooding*. Estão representados os valores de densidades maiores que 0.001 moléculas/Å<sup>3</sup>.

Em todas as simulações, o poro de condução (PC) possui uma densidade de ligantes, mesmo em baixas concentrações. Além disso, como mostrado na Figura 18, a densidade desse sítio ocupa todo o caminho de condução da proteína, fazendo com que o fluxo iônico seja impedido, algo totalmente oposto com o efeito do anestésico, que é de potenciação do fluxo iônico [11]. Essa incoerência foi alvo de um estudo envolvendo a interação do sevoflurano no poro na presença de voltagem, que comprovou que a ligação de íons potássio no PC impede a ligação do sevoflurano no PC [16]. Assim, pode-se concluir que em sistemas vivos, esse fenômeno não ocorreria devido à presença da corrente de  $K^+$  no canal [48]. Dessa maneira, o sítio PC é considerado como um artefato das simulações sem presença de voltagem.

Em relação à configuração interna do ligante, uma comparação de RMSD com sua estrutura inicial mostra que não existe nenhuma configuração presencial em nenhuma das fases. No RMSD translacional, todos os graus de liberdade são considerados, enquanto no rotacional há a sobreposição dos centros geométricos dos ligantes e no configuracional ocorre a máxima sobreposição das estruturas. A Figura 19 mostra os perfis de RMSD de um ligante que esteve presente na vizinhança da proteína e nos dois componentes do reservatório (água e membrana). Esse comportamento é indispensável para considerar que ocorre uma ligação de pequenos ligantes com proteínas, principalmente nos termos





**Figura 19:** Perfis de RMSD de uma molécula de sevoflurano nos diversos ambientes da proteína. De superior ao inferior: Região onde a molécula está presente (wt - água, mb - membrana, pt - proteína); RMSD translacional, rotacional e configuracional. É possível ver que não existe nenhuma correlação entre o valor de RMSD e a região onde o anestésico está presente.

relacionados à entropia, onde não foi necessário desenvolver um potencial artificial para restringir a molécula em uma pose específica (seção III.3.5).

## IV.3 Propriedades termodinâmicas da interação

### IV.3.1 Coeficiente de partição

A interação agora será analisada sob a ótica de um fenômeno de partição, dado que o ligante possui baixa especificidade ao longo da interface proteica. A vantagem de utilizar este método é de que todos os parâmetros para o cálculo do coeficiente de partição ( $\varphi$ ) foram obtidos diretamente da simulação de maneira bem simples (vide metodologia). O cálculo de  $\varphi$  ocorreu utilizando a equação 7 e os dados das tabelas 2, 6 e 7. A Tabela 8 contém os valores de  $\varphi$  calculados nas concentrações das simulações de *flooding*.

É possível notar que o coeficiente de partição possui pouca variação de acordo com as concentrações, o que sugere fortemente uma independência em relação à concentração. Apenas o coeficiente da simulação de 25mM foi destoante em relação aos outros, fato provavelmente ocorrido por uma amostragem enviesada de  $\langle n \rangle$  devido ao complexo

**Tabela 8:** Coeficientes de partição ( $\varphi$ ) de cada concentração, obtidos diretamente das simulações de *flooding*, a partir dos 100 ns de simulação.

	<b>1mM</b>	<b>25mM</b>	<b>50mM</b>	<b>75mM</b>	<b>100mM</b>	<b>150mM</b>
$\varphi$	6,89 $\pm$ 0,7	8,75 $\pm$ 0,14	6,42 $\pm$ 0,13	7,02 $\pm$ 0,10	6,35 $\pm$ 0,08	6,05 $\pm$ 0,07

pentamérico presente no poro central, como discutido na seção IV.3.3

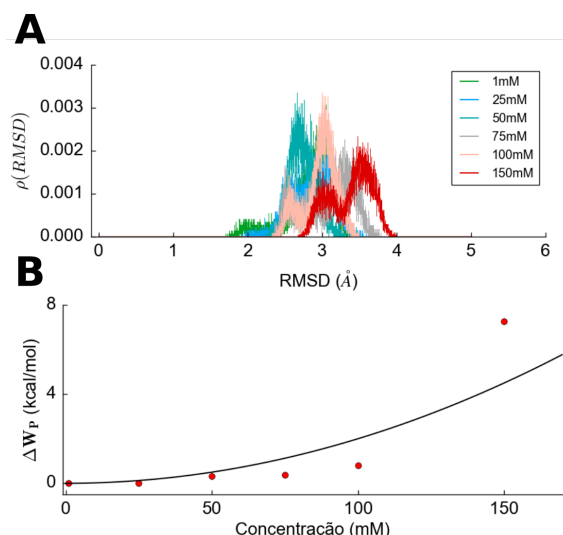
A relativa insensibilidade de  $\varphi$  com a concentração a princípio não é compatível com o fato da energia de dessolvatação ( $W^*$ ) variar com a concentração, como já mostrado na seção IV.1.2. Dessa maneira, é necessária uma maior investigação também de outros parâmetros que influenciam o processo de ligação, como a energia de dessolvatação a partir do sítio ( $W^{**}$ ) e as perturbações estruturais causadas pela presença do ligante para a estrutura da proteína ( $\Delta W_P$ ). Para a consistência do modelo de partição ser mantida, algum desses outros parâmetros deve sofrer perturbação semelhante a de  $W^*$ , resultando no cancelamento das influências causadas pela concentração, e assim obtendo um  $\Delta W$  invariável.

### IV.3.2 Perturbação na estrutura interna da proteína

Como mencionado na seção III.3.3, a presença dos ligantes pode perturbar a estrutura interna da proteína, onde o trabalho realizado ( $\Delta W_P$ ) é calculado a partir das variações de RMSD que a estrutura vazia possui comparado com a estrutura ocupada por ligantes (equação III.3.8). Ou seja, se a flutuação da estrutura em uma concentração for maior que a flutuação da estrutura vazia, é considerado que houve um trabalho mecânico realizado pelos ligantes. A distribuição de probabilidades do RMSD e o valor de  $\Delta W_P$  para cada concentração estão mostrados na Figura 20 e Tabela 9.

**Tabela 9:** Perturbações estruturais realizadas pela proteína ( $\Delta W_P$ ), em kcal/mol, para diversas concentrações.

	<b>1mM</b>	<b>25mM</b>	<b>50mM</b>	<b>75mM</b>	<b>100mM</b>	<b>150mM</b>
$\Delta W_P$	0	0,01	0,31	0,37	0,79	7,25



**Figura 20:** Perturbação estrutural na proteína. A - Densidades de probabilidade de ocorrência do RMSD ( $\rho(RMSD)$ ) em diversas concentrações. B - Trabalho de perturbação da estrutura da proteína calculado a partir das simulações de *flooding* (pontos vermelhos) e regressão (linha preta) segundo modelo quadrático  $\Delta W_P = h_p C^2$ , onde  $h_p = 200$  kcal/mol e  $C$  é a razão da concentração pela concentração padrão .

As distribuições de RMSD estão contidas em intervalos considerados como normais para flutuações do Kv1.2[49], ressaltando a estabilidade das estruturas utilizadas. Pode-se ver que o comportamento esperado para soluções diluídas ( $\Delta W_P \rightarrow 0$ ) também ocorre nas simulações. É notório o deslocamento das curvas para valores maiores a medida que a concentração aumenta, implicando em uma maior perturbação estrutural causada pelas moléculas de sevoflurano. Apesar disso, apenas a simulação de 150mM possui uma contribuição energética relevante quando comparado às outras concentrações. A diferença de aproximadamente dez vezes quando comparado à de 100mM é explicada pela distribuição do RMSD da proteína a 150mM estar completamente além do valor de tolerância de RMSD (2,5  $\text{\AA}$ ). Essa faixa de valores de RMSD distantes da tolerância juntamente com o alto valor da constante  $k$  do potencial  $u(\mathbf{r}^P)$  (equação 24) causaram um trabalho de maior magnitude na maior concentração. Porém, o trabalho realizado na concentração de 150mM ainda está abaixo do que ao trabalho de outros estudos envolvendo transformações de estruturas internas [50, 51], o que sugere que os ligantes de sevoflurano não realizam perturbações estruturais em escala significativa na proteína.

### IV.3.3 Energia de ligação sítio-vácuo

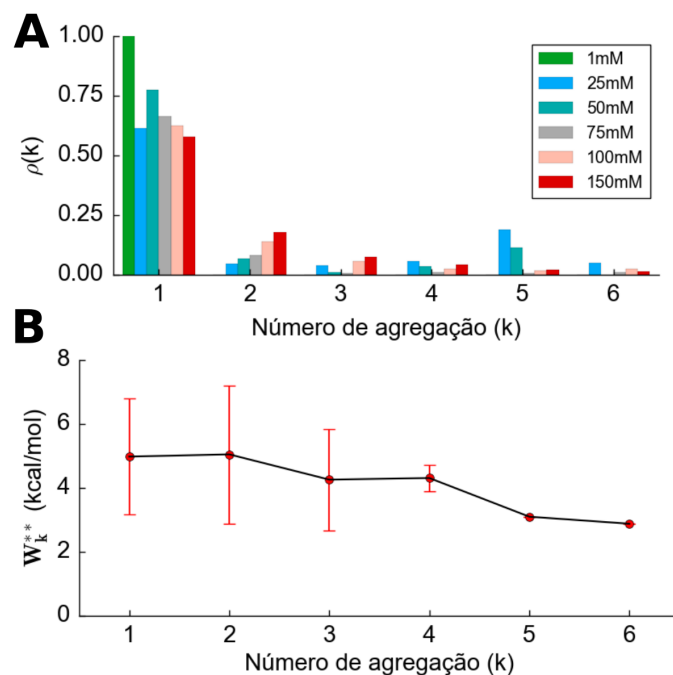
O parâmetro restante para a reconstrução da energia livre de partição é a energia de dessolvatação a partir do sítio ( $W^{**}$ ). A energia de dessolvatação pode ser influenciada pela região da proteína onde ocorre a transformação e também pela quantidade de ligantes que estão próximas do ligante que estão sendo transferidas (de maneira semelhante à dessolvatação membrana-vácuo). Para avaliar o primeiro problema, foram realizados 47 cálculos de FEP independentes distribuídos nos pontos de maior densidade da superfície proteica: poro de condução, Linker, interface S5-S6 e hélice S4.

A maneira mais direta de resolver o segundo problema seria realizar os cálculos de FEP em várias regiões da proteína a diversas concentrações. Porém, isso requer muito tempo computacional, tornando os cálculos impraticáveis. Assim, na busca de um procedimento mais simples, porém fidedigno, as interações em várias concentrações foram reproduzidas através do estado de agregação de cada concentração. O estado de agregação ( $n_k$ ) consiste em quantas moléculas de ligante estão em volta da molécula que será transformada, por exemplo:  $k=1$  - monômeros;  $k=2$  - dímeros, e assim por diante. Caracterizados os estados de agregação, a energia de dessolvatação média em uma concentração fixa foi definida como:

$$W^{**} = \sum \rho(k) \times W_k^{**} \quad (52)$$

onde  $\rho(k)$  é a frequência que o estado de agregação  $k$  ocorreu ao longo da simulação e  $W_k^{**}$  é a energia média que ocorreu no estado de agregação  $k$ . A Figura 21 ilustra a distribuição de  $\rho(k)$  e a energia média para cada estado de agregação.

Como esperado, a concentração de 1mM somente possui monômeros. As outras concentrações possuem diversos estados de multimerização das moléculas de sevoflurano, ficando cada vez mais raras quando  $k$  aumenta. Uma exceção visível é o estado  $k = 5$  nas simulações de 25 e 50mM. Esse estado pentamérico ocorreu na região do poro central, onde o agregado ficou preso nela por muito tempo, enviesando a amostragem. Caso a simulação ocorresse por mais tempo, esses ligantes separariam-se e esse estado seria mais improvável, como aconteceu nas concentrações maiores. Os valores da energia de dessolvatação possuem um comportamento semelhante ao encontrado na membrana, onde



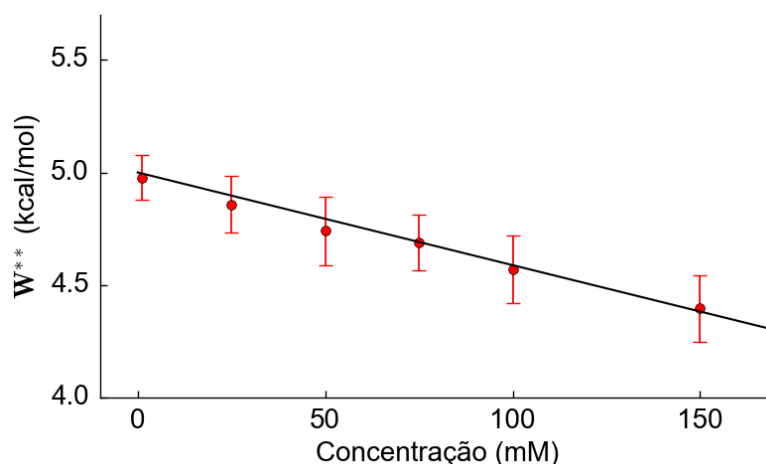
**Figura 21:** Estados de agregação do sevoflurano em diferentes concentrações. A - Probabilidade de ocorrência do estado de agregação  $k$  ao longo do todo tempo da simulação de *flooding*. B - Energia média calculada via FEP das moléculas de sevoflurano no estado de agregação  $k$ . Erro obtido por propagação.

a maior presença de ligantes facilita a dessolvatação. Além disso, os valores encontrados para o desacoplamento de monômeros e dímeros são muito condizentes com os obtidos de outros estudos [17, 16, 11]. O erro estatístico está superestimado, pois foi obtido a partir da propagação de erros, principalmente para  $k=1$  e  $k=2$ , onde foram feitas 20 réplicas cada. Melhores valores quanto à incerteza dos dados seriam obtidos utilizando outros procedimentos, como *Bootstrap*.

As energias resultantes de dessolvatação foram calculadas seguindo a equação 52 e estão listadas na Tabela 10. A regressão linear, mostrada na Figura 22, apontou os coeficientes como  $\bar{w} = 5,0$  kcal/mol,  $h^{**} = 0,9$  kcal/mol, ou seja, o efeito de concentração é semelhante à situação do ligante na membrana, onde a maior presença de ligantes facilita a remoção do ambiente.

**Tabela 10:** Energias de dessolvatação partindo da interface proteica para cada concentração. Valores em kcal/mol.

	1mM	25mM	50mM	75mM	100mM	150mM
$W^{**}$	$4,97 \pm 0,1$	$4,85 \pm 0,13$	$4,74 \pm 0,15$	$4,68 \pm 0,12$	$4,56 \pm 0,15$	$4,40 \pm 0,15$



**Figura 22:** Energias de dessolvatação sítio-vácuo de acordo com a concentração. As energias foram calculadas por cálculos de FEP em diversas regiões da proteína e diversos estados de agregação do ligante. Pontos vermelhos: energia calculadas diretamente por FEP. Linha preta: regressão linear utilizando a função  $W^{**} = \bar{w} - h^{**} \frac{\langle n \rangle}{\nu} c$  (equação 37), onde  $\bar{w} = 5,0$  kcal/mol,  $h^{**} = 0,9$  kcal/mol e  $c = (1 \text{ mol/L})^{-1}$ .

#### IV.3.4 Reconstrução do coeficiente de partição

Como mostrado na equação 16, o coeficiente de partição pode ser obtido a partir da diferença das energias de dessolvatação do reservatório e do sítio:

$$\phi = e^{-\beta(W^* - W^{**})}$$

Assim, com os valores das Tabelas 10 e 5 é possível reconstruir os coeficientes de partição, listados na Tabela 11.

Um dos principais objetivos do cálculo de  $\phi$  utilizando a diferença de energia consiste na possibilidade de avaliar a influência da concentração em cada etapa do ciclo termodinâmico proposto, algo que era uma limitação do cálculo utilizando o número de moléculas e volumes. Através dos cálculos de FEP independentes, foi possível ver que tanto a retirada do reservatório quanto do sítio sofreram efeitos da concentração, mas não havia

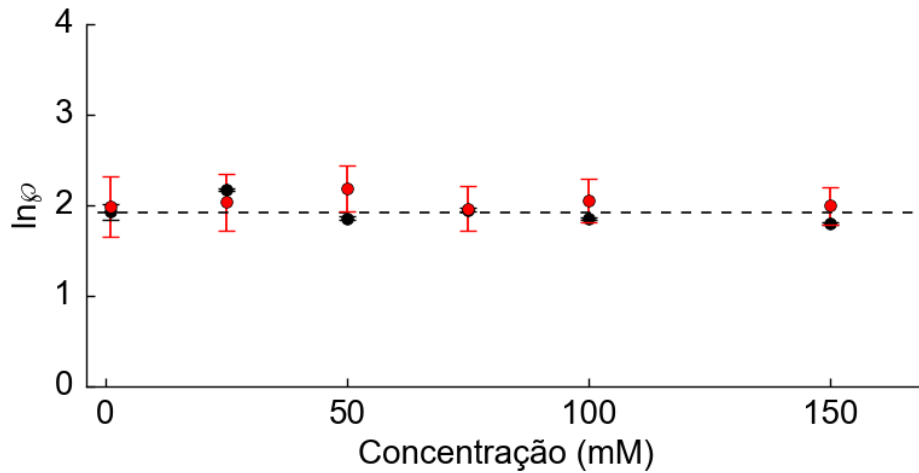
**Tabela 11:** Diferença das energias de dessolvatação ( $W^* - W^{**}$ , em kcal/mol) e coeficientes de partição ( $\varphi_{FEP}$ ) de cada concentração, obtidos pela diferença nas energias de dessolvatação ( $\varphi_{FEP} = e^{-\beta(W^* - W^{**})}$ ).

	1mM	25mM	50mM	75mM	100mM	150mM
$W^* - W^{**}$	1,18	1,21	1,30	1,17	1,22	1,19
$\varphi_{FEP}$	7,27	7,66	8,92	7,14	7,78	7,34

nenhuma garantia de que essas influências seriam semelhantes. Apenas após os cálculos em várias concentrações pôde-se constatar que os efeitos causados pela concentração em ambas as situações se anulam:

$$h_{memb}^* \frac{N - n}{V - \nu} c - h^{**} \frac{\langle n \rangle}{\nu} c \approx 0 \quad (53)$$

Além disso, o valor absoluto dos coeficientes de partição são muito próximos, o que reforça a qualidade e consistência do ferramental teórico deste trabalho, afinal os resultados foram bastante consistentes. A semelhança entre o coeficiente de partição obtido pela simulação de *flooding* ( $\varphi_{fld}$ ) e a obtida pelos cálculos de FEP ( $\varphi_{FEP}$ ) é evidenciada na Figura 23.



**Figura 23:** Coeficientes de partição. Preto: Coeficientes de partição calculados a partir das simulações de *flooding* ( $\varphi_{fld} = \left(\frac{N - \langle n \rangle}{V - \nu}\right)^{-1} \frac{\langle n \rangle}{\nu}$ ). Vermelho: obtidos a partir dos cálculos de FEP ( $\varphi_{FEP} = e^{-\beta(W^* - W^{**})}$ ). A linha tracejada horizontal representa a média dos coeficientes obtidos pelo *flooding*.

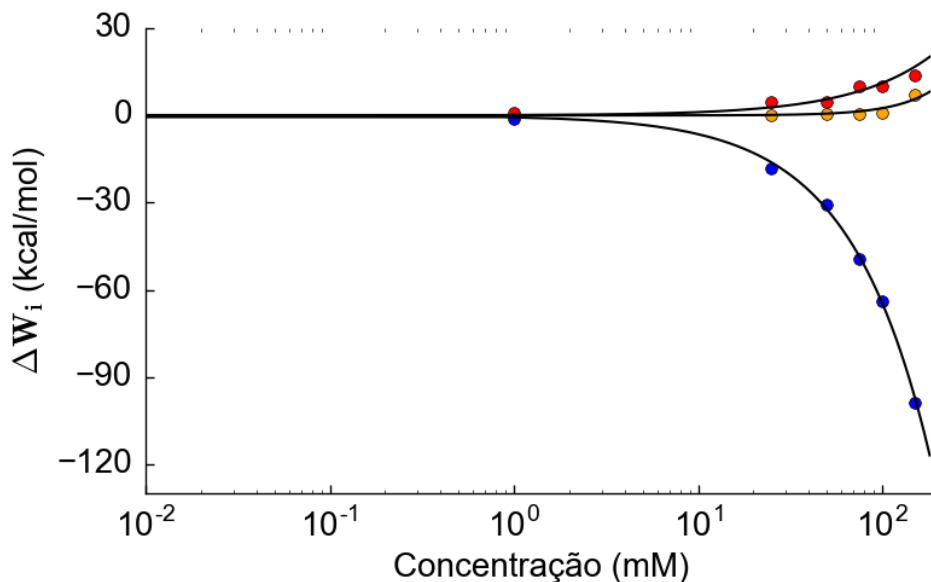
#### IV.4 Energia total de ligação

O coeficiente de partição é um termo que depende das afinidades do ligante com o ambiente e com o sítio, sendo assim o termo correspondente à contribuição entálpica do processo de ligação. Como discutido na seção III.3.8, existem outras contribuições energéticas que devem ser levadas em conta para calcular a energia total de ligação ( $\Delta W$ ):

$$\Delta W = \Delta W_S + \Delta W_P + \Delta W_L \quad (54)$$

onde  $\Delta W_S$ ,  $\Delta W_P$  e  $\Delta W_L$  são as contribuições da entropia, perturbação da estrutura e entalpia, respectivamente.

A obtenção de  $\Delta W$  também é importante para caracterizar a energia total de ligação de  $n$  moléculas, afinal as simulações de *flooding* mostraram que a interação sevoflurano-Kv1.2 envolve diversas moléculas do anestésico. Logo, o interesse em obter  $\Delta W_n$  é pertinente, pois será de grande valia para a obtenção de outras propriedades da interação [16, 17]. É possível ver na Figura 24 o peso de cada tipo de energia na energia resultante e como se comporta em várias concentrações.

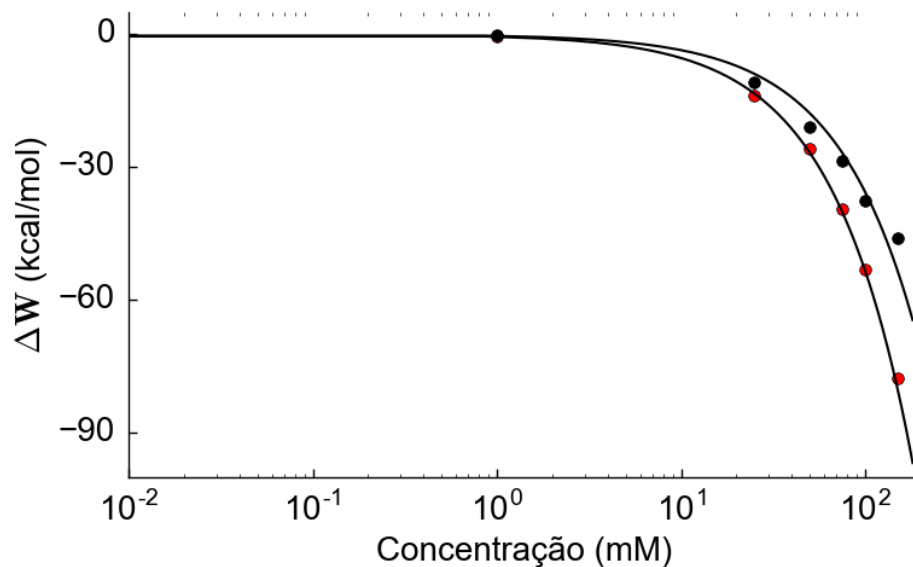


**Figura 24:** Contribuições energéticas em função da concentração do sistema. Decomposição de  $\Delta W$  nos termos de perturbação da proteína ( $\Delta W_P$ , em laranja), entropia ( $\Delta W_S$ , em vermelho) e energias de dessolvatação ( $\Delta W_L$ , em azul). As linhas sólidas representam as regressões quadráticas de melhor ajuste aos pontos.



A maior presença dos ligantes de fato leva a uma maior perturbação da estrutura interna ( $\Delta W_P$ ) na proteína, a qual chega a alcançar 10 kcal/mol em concentrações grandes. A contribuição entrópica ( $\Delta W_S$ ) também é punitiva para o processo, pois os  $n$  ligantes partem de uma fase com maior espaço e acabam ligando-se ao sítio, que é mais restrito espacialmente. Cabe lembrar que a restrição é apenas translacional, pois os demais graus de liberdade continuam inalterados, logo a punição entrópica pode ser considerada relativamente baixa quando comparada às situações onde a orientação do ligante também é restrita [37, 52, 53]. O que torna a ligação do sevoflurano na superfície do Kv1.2 espontânea é a diferença das afinidades do ligante na superfície proteica e no reservatório ( $\Delta W_L$ ). Observa-se que a contribuição de  $\Delta W_L$  em módulo é bastante superior às outras, na ordem de 120 kcal/mol em concentrações altas. Essa decomposição é essencial para mostrar que a energia livre negativa vem justamente da afinidade da superfície proteica do Kv1.2 e sua afinidade com o ligante.

Na Figura 25, é mostrada a comparação de eficácia da estimativa da variação da energia livre utilizando o coeficiente de partição ( $\Delta W = \Delta W_S - \langle n \rangle \beta \ln \phi$ ) com a energia obtida pelo ciclo termodinâmico ( $\Delta W = \Delta W_P + \Delta W_S + \Delta W_L$ ).



**Figura 25:** Variações de energia para a ligação em função da concentração. Vermelho: Variação de energia livre proveniente dos parâmetros termodinâmicos entropia, entalpia e perturbação na proteína. Preto: Variação de energia livre calculada a partir do coeficiente de partição.

Os modelos possuem divergências em altas concentrações, principalmente devido à estimativa utilizando o coeficiente de partição levar uma média de energias, enquanto a estimativa utilizando as contribuições individuais é feita utilizando a energia de dessolvatação para cada ligante transferido ( $\Delta W_n = \sum W_i^* - W_i^{**}$ ), além de que a pequena diferença numérica foi amplificada nas maiores concentrações. Porém, a medida que a concentração é reduzida, a estimativa fornecida pelo coeficiente de partição torna-se bem próxima da calculada pelas contribuições individuais, mostrando que a aproximação torna-se mais coerente com os valores obtidos em concentrações menores, que estão na magnitude da maior parte dos fenômenos de interesse no estudo da interação de pequenas moléculas com proteínas[54]. Ou seja, o uso do coeficiente de partição, mesmo que obtido em concentrações altas, ainda consiste em uma ótima ferramenta para o estudo dessas interações em concentrações baixas.

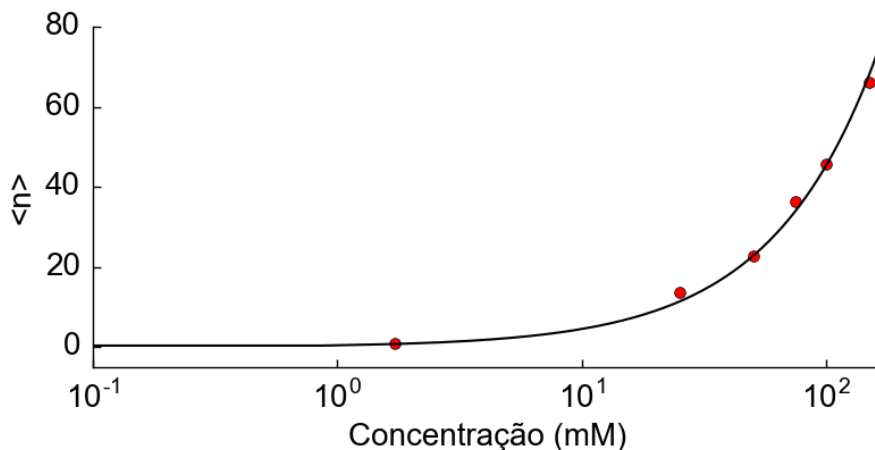
## IV.5 Projeções para outras concentrações

O procedimento de calcular os coeficientes de partição, como mostrado nesse trabalho, pode envolver diversos tipos de simulações, como os cálculos de energia livre por FEP ou apenas o cálculo pela simulação de *flooding*. Independente do método de obtenção, com  $\varphi$  e utilizando o ferramental teórico da seção III.3.10, é possível reconstruir diversas propriedades dependentes do número de ligantes para qualquer concentração desejada, sem a necessidade de realizar simulações adicionais, que possuiriam alto custo computacional. Foi escolhido o coeficiente de partição da simulação de 150mM como referência para a realização dos cálculos ( $\varphi'$ ). Assim, a prova de conceito foi realizada através da comparação das projeções com as propriedades obtidas nas simulações de 1, 25, 50, 75 e 100mM.

### IV.5.1 Projeções de número médio

As projeções de  $\langle n \rangle$ , análogas às curvas de titulação, ocorreram de acordo com a equação 45 e são mostradas na Figura 26. A concordância com os valores foi bastante satisfatória, mostrando que o modelo abrange bem as interações tanto no regime diluído

quanto no concentrado.

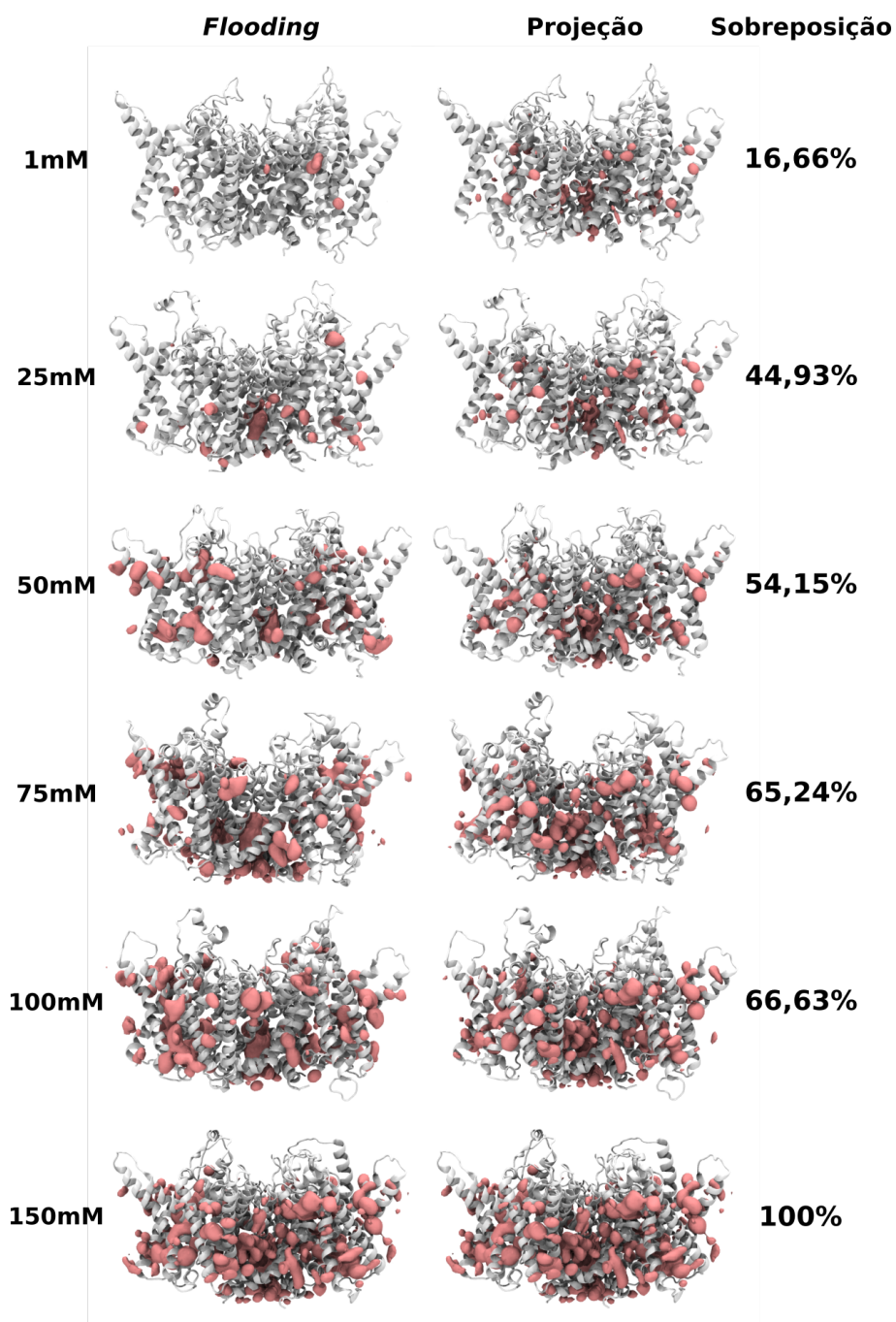


**Figura 26:** Número médio de ligantes no sítio. Números obtidos diretamente pelas simulações de *flooding* (vermelho) e projeção realizada pela equação  $\langle n \rangle = \bar{\rho}V \left(1 + \frac{V-\nu}{\nu\phi}\right)^{-1}$ . Todos os parâmetros (volumes e coeficiente de partição) foram originados da simulação de 150mM.

#### IV.5.2 Projeções de densidades tridimensionais

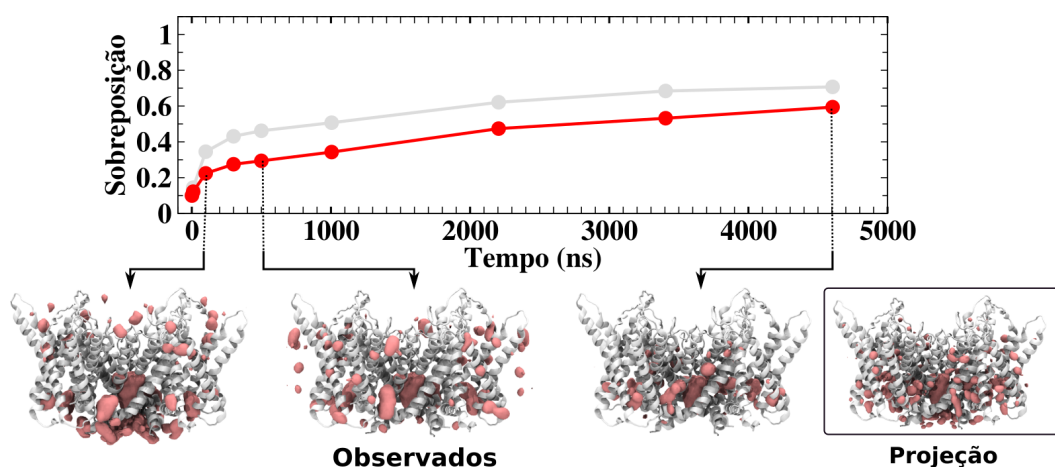
As projeções tridimensionais ocorreram de acordo com a equação 45. O número médio utilizado foi o  $\langle n \rangle$  calculado pela equação 45 e a densidade unitária  $\bar{\rho}'(\mathbf{R})$  foi oriunda da simulação de 150mM. A comparação entre a distribuição das densidades tridimensionais projetadas com as nuvens obtidas no *flooding* é dada através do cálculo da sobreposição (*overlap*) entre as duas nuvens. Na Figura 27 encontram-se as projeções realizadas, as densidades obtidas pela simulação e o valor de sobreposição entre as nuvens. Os valores de sobreposição mostram que as nuvens de densidade projetadas conseguiram reproduzir a distribuição espacial que ocorreu na simulação de *flooding* em um grau satisfatório. Observa-se também, que as interações remanescentes nos regimes de menor concentração estão nas regiões do Linker e hélice S4, fato já mostrado de maneira independente por experimentos de fotomarcção [18].

As projeções mostram-se bastante satisfatórias quando a comparação ocorre entre altas concentrações. A medida que as concentrações tornam-se menores, a qualidade das projeções cai substancialmente. A origem dessa má sobreposição pode ser tanto um mapa unitário de qualidade inferior quanto o fato dos mapas obtidos das simulações de *flooding* não ter atingido o equilíbrio. Para solucionar essa questão, foi realizada a comparação



**Figura 27:** Densidades tridimensionais. Esquerda para direita: densidades tridimensionais obtidos a partir das simulações de *flooding*, densidades oriundas das projeções de  $\langle n \rangle$  e do mapa unitário de 150mM ( $\bar{\rho}'(\mathbf{R})$ ) e sobreposição das nuvens.

das projeções com uma simulação de 5mM de sevoflurano com duração de 5000 ns, em diversas janelas de tempo, como mostrados na Figura 28.

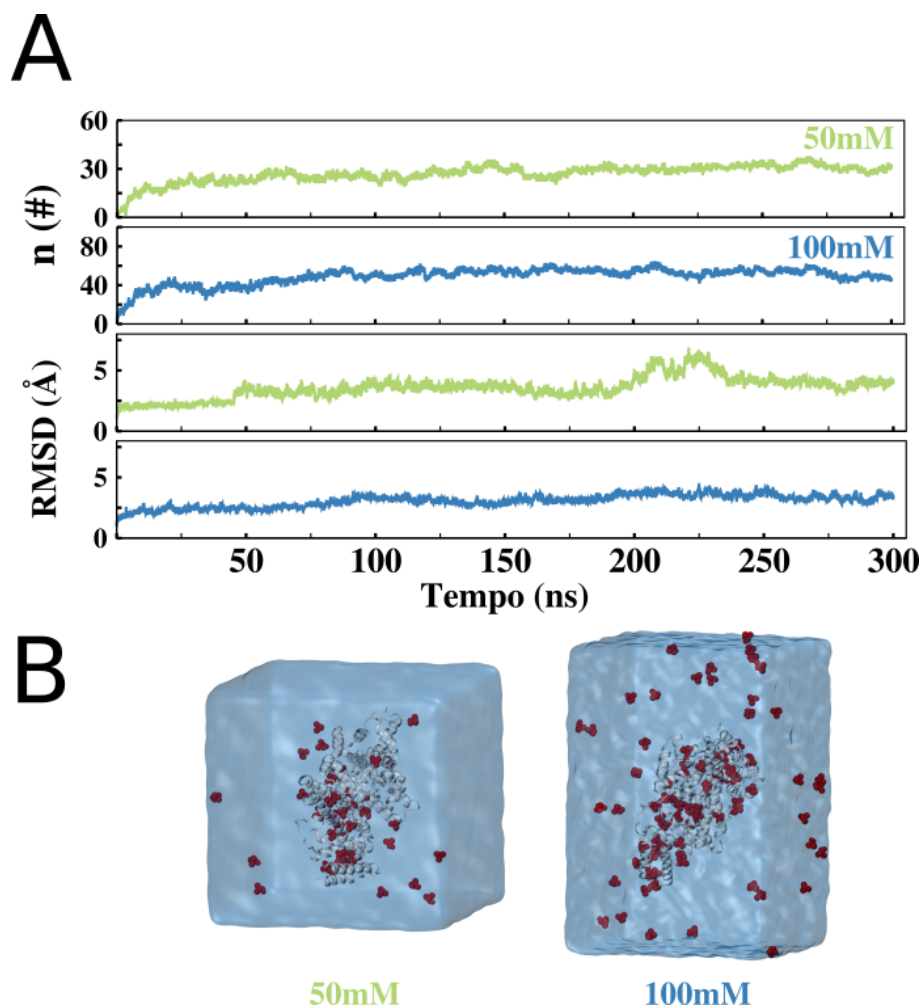


**Figura 28:** Densidades tridimensionais obtidos pela simulação de *flooding* de 5mM. A - Índices de sobreposição das densidades tridimensionais obtidas em diferentes janelas de tempo em relação à projeção para 5mM, sem simetrização (vermelho) e após a simetrização (cinza). B - Densidades tridimensionais observadas diretamente das simulações de *flooding* em diversas janelas temporais e a projeção realizada.

É possível notar que a sobreposição aumenta a medida que o tempo de simulação aumenta, o que indica que as densidades tridimensionais da simulação de 5mM alcançam progressivamente a densidade obtida a partir das projeções. A aproximação ocorre de maneira assintótica, o que sugere que maiores escalas de tempo ainda são necessárias para atingir melhores índices de sobreposição. Assim, é possível afirmar que as projeções obtidas a partir de concentrações maiores possuem boa qualidade, enquanto as simulações de menor concentração devem possuir escalas de tempo superior a dez vezes para obter qualidade semelhante. Como o Kv1.2 é um homotetrâmero, a simetrização torna-se uma alternativa para melhorar a qualidade das densidades tridimensionais. Dessa maneira, o procedimento de projeções tridimensionais a partir de concentrações altas mostra-se bastante sucedido e vantajoso para estudar os aspectos estruturais das interações a baixas concentrações, oferecendo enorme economia de recursos computacionais para alcançar a convergência.

### IV.5.3 Interação da albumina e tricloroetanol

Como prova de que a teoria pode ser utilizada para outros casos de interação de pequenos ligantes, foram calculados os coeficientes de partição para o tricloroetanol e a albumina em diferentes concentrações. A Figura 29 mostra a quantidade de ligantes na vizinhança da proteína e o RMSD da albumina ao longo do tempo.



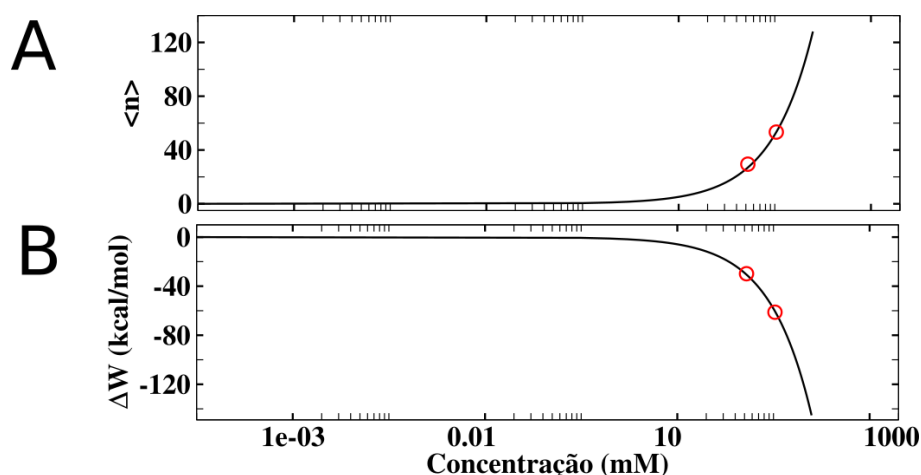
**Figura 29:** Partição do tricloroetanol no sistema de albumina e água. A - Número de ligantes e RMSD em comparação à estrutura inicial da albumina ao longo do tempo. B - Estruturas de equilíbrio para as simulações de *flooding* com 50mM e 100mM.

Os dados obtidos pela simulação de *flooding* estão na Tabela 12. É possível ver que os coeficientes de partição possuem valores muito próximos, o que é condizente com a ideia de que o coeficiente de partição é concentração-independente, fato já mostrado nos estudos usando sevoflurano. A regressão de  $\langle n \rangle$  também obteve valores muito próximos

do observado, como mostrado na Figura 30.

**Tabela 12:** Propriedades no equilíbrio das simulações de *flooding* da albumina bovina sérica com o sedativo tricloroetanol.

Concentração	N	$\langle n \rangle$	V	$\nu$	$\varphi$	$\ln \varphi$
50mM	43	29,47	1.361.310 Å <sup>3</sup>	86.306Å <sup>3</sup>	32,18	3,47
100mM	85	53,32	1.361.310 Å <sup>3</sup>	86.306Å <sup>3</sup>	24,87	3,21

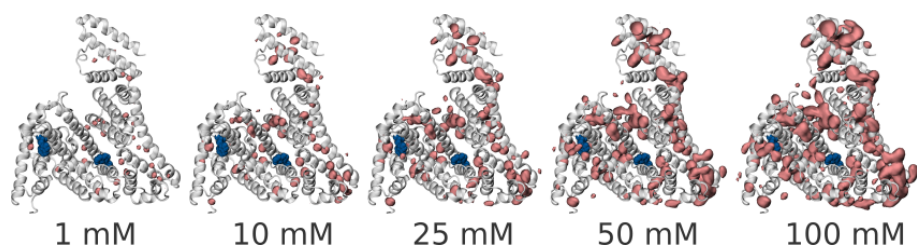


**Figura 30:** Propriedades calculadas da interação do tricloroetanol com a albumina. A - Pontos vermelhos: Número médio obtido diretamente das simulações de *flooding*. Linha preta: Projeções de número médio a partir do coeficiente de partição obtido a 100mM ( $\ln \varphi = 3,21$ ). Nota-se que a regressão possui uma acurácia muito boa quanto ao número médio em 50mM. B - Pontos vermelhos: Energia livre de ligação nas concentrações de 50 e 100 mM. A energia foi calculada utilizando a estimativa do coeficiente de partição com as contribuições entrópicas ( $\Delta W = \Delta W_S - \langle n \rangle \beta \ln \varphi$ ). Linha preta: regressão quadrática das energias em função das concentrações ( $\Delta W = -583,45 \bar{\rho} / C^0$ ).

Com o mapa unitário obtido a partir da simulação de 100mM, também é possível fazer as projeções de  $\bar{\rho}(\mathbf{R})$  para outras concentrações, mostrado na Figura 31.

O comportamento dos coeficientes de partição é semelhante ao descoberto na interação sevoflurano-Kv1.2. Além disso, a albumina é uma proteína globular, ou seja, a teoria continuou mostrando sua consistência mesmo em ambientes onde o reservatório possui outra composição. A facilidade da obtenção de  $\varphi$  e a independência em relação a concentração são fatores muito importantes ao estender a teoria aqui desenvolvida para outros sistemas, o que vem se mostrando bastante sucedido. Os resultados obtidos a partir da

interação da albumina com tricloroetanol mostra que esse procedimento pode ser usado para quaisquer tipos de interação de pequenas moléculas com ligantes.



**Figura 31:** Projeções da distribuição espacial do tricloroetanol na albumina sérica. As projeções foram obtidas a partir do reescalonamento do mapa unitário de densidade tridimensional da simulação de 100mM.



## V Conclusões e perspectivas

A interação de pequenas moléculas com proteínas possui papel chave na regulação de diversos processos biológicos, o que sempre tornou esse campo alvo de bastante interesse, principalmente da farmacologia. Vistas as limitações das técnicas utilizadas atualmente para esse tipo de interação, nesse trabalho o processo foi abordado de uma nova maneira: como uma partição entre duas fases. Com isso, foi desenvolvido um conjunto de ferramentas teóricas que visam descrever a ligação a partir de uma propriedade termodinâmica mais simples, o coeficiente de partição ( $\varphi$ ).

As interações degeneradas são melhor reproduzidas em simulações de *flooding*, logo o ferramental teórico utilizou-se de propriedades que podem ser obtidas diretamente dessas simulações. Além disso, como as simulações de *flooding* consideram explicitamente a interação ligante-ligante, a influência da concentração também foi estudada. Os coeficientes de partição calculados em diversas simulações independentes mostraram que  $\varphi$  não possui variações significativas em relação à concentração, trazendo a ideia de que o coeficiente de partição é concentração-independente.

Apesar da boa descrição, outras contribuições energéticas existiam além da oferecida pelo coeficiente de partição. A proteína poderia sofrer uma perturbação estrutural ou a restrição de volumes poderia afetar a energia livre do processo. Dessa maneira, a ligação de uma molécula foi descrita em um ciclo termodinâmico que mostrou que a perturbação causada na proteína não é significativa em baixas concentrações e que a retirada do ligante do reservatório e a inserção no sítio são influenciados de maneira parecida pela concentração do ligante, anulando as influências.

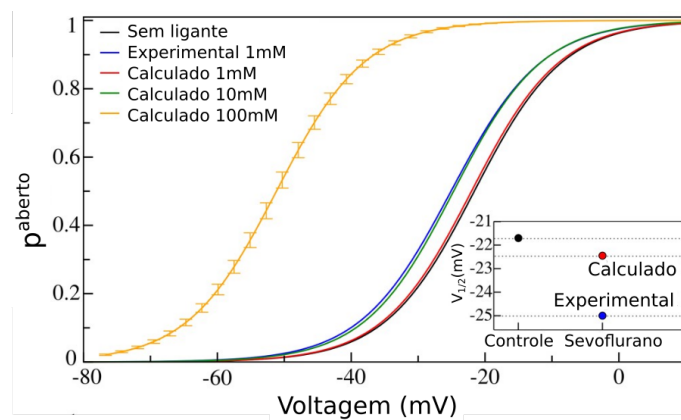
Adaptando as ferramentas teóricas já desenvolvidas para as constantes de ligação, a teoria aqui desenvolvida também contempla a projeção de propriedades para diversas concentrações. Mostrado bem exitoso, esse conjunto de projeções pode ser usado para o estudo de interações em concentrações baixas a partir de simulações com concentração alta, assim permitindo uma grande economia em recursos computacionais.

Caracterizar via coeficiente de partição uma interação é apenas o começo. Diversos processos biológicos envolvem o equilíbrio entre dois estados, e a modulação do li-

gante perturba esse equilíbrio. Um dos primeiros modelos alostéricos que descreve esse processo, o modelo de Monod-Wyman-Changeux leva em conta a afinidade pelo estado ativado e a afinidade pelo estado inativado[55]. Uma generalização dessa teoria envolvendo constantes de ligação agregadas já foi realizada em outro trabalho[22], onde as constantes de ligação para os estados aberto e fechado foram encontradas pelo método *docking-FEP*, utilizando como modelo de interação o sevoflurano e o Kv1.2. Assim, foi possível reconstruir a probabilidade do canal estar aberto em função da voltagem:

$$\rho_{open} = \left[ 1 + \frac{Z_{closed}}{Z_{open}} e^{\beta \Delta Q (V_m - V)} \right]^{-1} \quad (55)$$

onde  $Z$  é a função de partição do estado, que é encontrada a partir das constantes de ligação ( $Z = \sum \bar{\rho}^n K(n)$ ). As curvas de abertura podem ser comparadas diretamente com as curvas obtidas pela eletrofisiologia, como mostrado na Figura 32. Esse trabalho, o qual participei como coautor, pode ser encontrado no Anexo III.



**Figura 32:** Probabilidades de abertura do Kv1.2 em diversas condições. Curva preta - probabilidades em função da voltagem para o Kv1.2 sem anestésico. Curva azul - probabilidades obtidas experimentalmente com a presença de 1mM de sevoflurano. Curvas vermelha, verde e amarela - Curvas projetadas teoricamente para diferentes concentrações. Na caixa inferior estão representados os valores de  $V_{1/2}$  das curvas  $p \times V$ . Adaptado de Stock *et al.* [16]

Como no trabalho atual a caracterização dos coeficientes de partição ocorreu somente no estado aberto, basta calcular os coeficientes de partição para o canal fechado e desenvolver uma função de partição análoga para realizar a reconstrução da curva de probabilidade de abertura.

Foi mostrado que as ferramentas desenvolvidas funcionam tanto para proteínas de

membrana quanto globulares. Isso ressalta a importância do estudo e o seu grande potencial para ser utilizado em uma miríade de interações de pequenas moléculas com ligantes, onde os alvos variam em diversos tipos de proteínas, principalmente receptores, canais iônicos e enzimas[1]. Uma abordagem utilizando coeficientes de partição poderá trazer uma inovação muito importante para o estudo de outros processos relevantes que envolvam a interação de pequenas moléculas com proteínas, principalmente por trazer consigo ideias como interação degenerada e de baixa afinidade, que poderão mudar a visão sobre alguns processos e assim trazer ideias de novos experimentos e explicações para diversos fenômenos.

## Referências

- [1] Rob Phillips. *The Molecular Switch*. en. Set. de 2020. ISBN: 978-0-691-20024-8.
- [2] Ed Reznik et al. “Genome-Scale Architecture of Small Molecule Regulatory Networks and the Fundamental Trade-Off between Regulation and Enzymatic Activity”. en. Em: *Cell Reports* 20.11 (set. de 2017), pp. 2666–2677. ISSN: 2211-1247.
- [3] John P. Overington, Bissan Al-Lazikani e Andrew L. Hopkins. “How many drug targets are there?” en. Em: *Nature Reviews Drug Discovery* 5.12 (dez. de 2006). Number: 12 Publisher: Nature Publishing Group, pp. 993–996. ISSN: 1474-1784.
- [4] William Bialek. *Biophysics: Searching for Principles*. Inglês. Annotated edição. Princeton (N. J.): Princeton University Press, 2012. ISBN: 978-0-691-13891-6.
- [5] Jeffries Wyman e Stanley J. Gill. *Binding and Linkage: Functional Chemistry of Biological Macromolecules*. Inglês. Mill Valley, Calif: Univ Science Books, 1990. ISBN: 978-0-935702-56-9.
- [6] H Theorell. “Kristallinisches myoglobin”. Em: *Biochem. Z.* 268 (1934), pp. 73–82.
- [7] A. Rossi-Fanelli e E. Antonini. “Studies on the oxygen and carbon monoxide equilibria of human myoglobin”. en. Em: *Archives of Biochemistry and Biophysics* 77.2 (out. de 1958), pp. 478–492. ISSN: 0003-9861.
- [8] Ernesto Freire, Obdulio L. Mayorga e Martin Straume. “Isothermal titration calorimetry”. Em: *Analytical Chemistry* 62.18 (set. de 1990). Publisher: American Chemical Society, 950A–959A. ISSN: 0003-2700.
- [9] John Tileston Edsall e Jeffries Wyman. *Biophysical chemistry. Vol. 1, Thermodynamics, electrostatics, and the biological significance of the properties of matter*. Academic Press, 1958.
- [10] Vincent J. Hilser e Ernesto Freire. “Structure-based Calculation of the Equilibrium Folding Pathway of Proteins. Correlation with Hydrogen Exchange Protection Factors”. en. Em: *Journal of Molecular Biology* 262.5 (out. de 1996), pp. 756–772. ISSN: 0022-2836.
- [11] Qiansheng Liang et al. “Positive Allosteric Modulation of Kv Channels by Sevoflurane: Insights into the Structural Basis of Inhaled Anesthetic Action”. eng. Em: *PLoS One* 10.11 (2015), e0143363. ISSN: 1932-6203.

- [12] Ariane Nunes-Alves e Guilherme Menegon Arantes. “Ligand–Receptor Affinities Computed by an Adapted Linear Interaction Model for Continuum Electrostatics and by Protein Conformational Averaging”. en. Em: *Journal of Chemical Information and Modeling* 54.8 (ago. de 2014), pp. 2309–2319. ISSN: 1549-9596, 1549-960X.
- [13] Isabel T. G. Silva et al. “Biophysical studies of cholesterol effects on chromatin”. eng. Em: *Journal of Lipid Research* 58.5 (2017), pp. 934–940. ISSN: 1539-7262.
- [14] B Roux et al. “Thermodynamic stability of water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study.” Em: *Biophysical Journal* 71.2 (ago. de 1996), pp. 670–681. ISSN: 0006-3495.
- [15] Alessandra S. Kiametis et al. “Atomistic Model for Simulations of the Sedative Hypnotic Drug 2,2,2-Trichloroethanol”. en. Em: *ACS Omega* (nov. de 2018).
- [16] Letícia Stock et al. “Binding of the general anesthetic sevoflurane to ion channels”. en. Em: *PLOS Computational Biology* 14.11 (nov. de 2018). Ed. por Rebecca C. Wade, e1006605. ISSN: 1553-7358.
- [17] Letícia Stock, Juliana Hosoume e Werner Treptow. “Concentration-Dependent Binding of Small Ligands to Multiple Saturable Sites in Membrane Proteins”. en. Em: *Scientific Reports* 7.1 (jul. de 2017). Number: 1 Publisher: Nature Publishing Group, p. 5734. ISSN: 2045-2322.
- [18] Kellie A. Woll et al. “Photoaffinity Ligand for the Inhalational Anesthetic Sevoflurane Allows Mechanistic Insight into Potassium Channel Modulation”. Em: *ACS Chemical Biology* 12.5 (mai. de 2017), pp. 1353–1362. ISSN: 1554-8929.
- [19] Annika F. Barber, Qiansheng Liang e Manuel Covarrubias. “Novel Activation of Voltage-gated K<sup>+</sup> Channels by Sevoflurane\*”. en. Em: *Journal of Biological Chemistry* 287.48 (nov. de 2012), pp. 40425–40432. ISSN: 0021-9258.
- [20] Annika F. Barber et al. “Molecular Mapping of General Anesthetic Sites in a Voltage-Gated Ion Channel”. English. Em: *Biophysical Journal* 101.7 (out. de 2011), pp. 1613–1622. ISSN: 0006-3495.
- [21] K. Solt e J.S. Johansson. “Binding of the active metabolite of chloral hydrate, 2,2,2-trichloroethanol, to serum albumin demonstrated using tryptophan fluorescence quenching”. English. Em: *Pharmacology* 64.3 (2002), pp. 152–159. ISSN: 0031-7012.

- [22] Leonardo Cirqueira, Letícia Stock e Werner Treptow. “Concentration-dependent thermodynamic analysis of the partition process of small ligands into proteins”. en. Em: *Computational and Structural Biotechnology Journal* 20 (jan. de 2022), pp. 4885–4891. ISSN: 2001-0370.
- [23] Martin Karplus e Gregory A. Petsko. “Molecular dynamics simulations in biology”. en. Em: *Nature* 347.6294 (out. de 1990), pp. 631–639. ISSN: 1476-4687.
- [24] Eric Paquet e Herna L. Viktor. *Molecular Dynamics, Monte Carlo Simulations, and Langevin Dynamics: A Computational Review*. en. Research article. 2015.
- [25] Jérôme Hénin et al. “An Atomistic Model for Simulations of the General Anesthetic Isoflurane”. en. Em: *The Journal of Physical Chemistry B* 114.1 (jan. de 2010), pp. 604–612. ISSN: 1520-6106, 1520-5207.
- [26] Jing Huang e Alexander D. MacKerell. “CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data”. en. Em: *Journal of Computational Chemistry* 34.25 (set. de 2013), pp. 2135–2145. ISSN: 1096-987X.
- [27] Junmei Wang et al. “Development and testing of a general amber force field”. en. Em: *Journal of Computational Chemistry* 25.9 (jul. de 2004), pp. 1157–1174. ISSN: 1096-987X.
- [28] Lukas D. Schuler, Xavier Daura e Wilfred F. van Gunsteren. “An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase”. en. Em: *Journal of Computational Chemistry* 22.11 (ago. de 2001), pp. 1205–1218. ISSN: 1096-987X.
- [29] Loup Verlet. “Computer” experiments” on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules”. Em: *Physical review* 159.1 (1967), p. 98.
- [30] R. W. Hockney. “Potential Calculation and Some Applications.” English. Em: *Methods Comput. Phys.* 9: 135-211(1970). (jan. de 1970).
- [31] D. Beeman. “Some multistep methods for use in molecular dynamics calculations”. Em: *Journal of Computational Physics* 20.2 (fev. de 1976), pp. 130–139. ISSN: 0021-9991.
- [32] James C. Phillips et al. “Scalable molecular dynamics with NAMD”. eng. Em: *Journal of Computational Chemistry* 26.16 (dez. de 2005), pp. 1781–1802. ISSN: 0192-8651.
- [33] Tom Darden, Darrin York e Lee Pedersen. “Particle mesh Ewald: An  $N \bullet \log(N)$  method for Ewald sums in large systems”. en. Em: *The Journal of Chemical Physics* 98.12 (jun. de 1993), pp. 10089–10092. ISSN: 0021-9606, 1089-7690.

- [34] Christophe Chipot, Andrew Pohorille e Chipot-Pohorille, ed. *Free energy calculations: theory and applications in chemistry and biology ; with 2 tables*. en. Springer series in chemical physics 86. OCLC: 255480142. Berlin: Springer, 2007. ISBN: 978-3-540-38447-2.
- [35] Nandou Lu, David A. Kofke e Thomas B. Woolf. “Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods”. en. Em: *Journal of Computational Chemistry* 25.1 (jan. de 2004), pp. 28–40. ISSN: 0192-8651, 1096-987X.
- [36] Natália Mendes de Oliveira. “Análise do impacto da ligação de anestésicos no equilíbrio entre estados do canal Kv1.2 diretamente de simulações de flooding em altas concentrações”. Português. Em: (dez. de 2021). Accepted: 2022-03-25T22:04:25Z.
- [37] Reza Salari et al. “A Streamlined, General Approach for Computing Ligand Binding Free Energies and Its Application to GPCR-Bound Cholesterol”. en. Em: *Journal of Chemical Theory and Computation* 14.12 (dez. de 2018), pp. 6560–6573. ISSN: 1549-9618, 1549-9626.
- [38] J. Schlitter, M. Engels e P. Krüger. “Targeted molecular dynamics: A new approach for searching pathways of conformational transitions”. en. Em: *Journal of Molecular Graphics* 12.2 (jun. de 1994), pp. 84–89. ISSN: 0263-7855.
- [39] Jean-Louis Burgot. *The Notion of Activity in Chemistry*. en. Ed. por Jean-Louis Burgot. Cham: Springer International Publishing, 2017. ISBN: 978-3-319-46401-5.
- [40] Pauli Virtanen et al. “SciPy 1.0—Fundamental Algorithms for Scientific Computing in Python”. Em: *arXiv:1907.10121 [physics]* (jul. de 2019). arXiv: 1907.10121.
- [41] Johan Åqvist, Carmen Medina e Jan-Erik Samuelsson. “A new method for predicting binding affinity in computer-aided drug design”. en. Em: *Protein Engineering, Design and Selection* 7.3 (mar. de 1994), pp. 385–391. ISSN: 1741-0126.
- [42] William Humphrey, Andrew Dalke e Klaus Schulten. “VMD: Visual molecular dynamics”. Em: *Journal of Molecular Graphics* 14.1 (fev. de 1996), pp. 33–38. ISSN: 0263-7855.
- [43] William L. Jorgensen et al. “Comparison of simple potential functions for simulating liquid water”. en. Em: *The Journal of Chemical Physics* 79.2 (jul. de 1983), pp. 926–935. ISSN: 0021-9606, 1089-7690.
- [44] Axel Kohlmeyer e Josh Vermaas. *topotools Release 1.7*. Abr. de 2017.

- [45] A. Bujacz. “Structures of bovine, equine and leporine serum albumin”. en. Em: *Acta Crystallographica Section D: Biological Crystallography* 68.10 (out. de 2012). Number: 10  
Publisher: International Union of Crystallography, pp. 1278–1289. ISSN: 0907-4449.
- [46] Michel F. Sanner, Arthur J. Olson e Jean-Claude Spehner. “Reduced surface: An efficient way to compute molecular surfaces”. en. Em: *Biopolymers* 38.3 (1996). \_eprint:  
<https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291097-0282%28199603%2938%3A3%3C305%3BIP4%3E3.0.CO%3B2-Y>, pp. 305–320. ISSN: 1097-0282.
- [47] Oliver Beckstein et al. *MDAnalysis/GridDataFormats: release 0.4.1*. Abr. de 2019.
- [48] Leonardo Cirqueira. “Condução de íons pelo canal neuronal Kv1.2 na presença do anestésico geral sevoflurano”. Português. Em: (fev. de 2018).
- [49] Pär Bjelkmar et al. “Conformational Changes and Slow Dynamics through Microsecond Polarized Atomistic Molecular Simulation of an Integral Kv1.2 Ion Channel”. Em: *PLoS Computational Biology* 5.2 (fev. de 2009). ISSN: 1553-734X.
- [50] Philippe Ferrara, Joannis Apostolakis e Amedeo Caflisch. “Computer simulations of protein folding by targeted molecular dynamics”. en. Em: *Proteins: Structure, Function, and Bioinformatics* 39.3 (2000). \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291097-0134%2820000515%2939%3A3%3C252%3A%3AAID-PROT80%3E3.0.CO%3B2-3>, pp. 252–260. ISSN: 1097-0134.
- [51] C. Amaral et al. “Exploring conformational states of the bacterial voltage-gated sodium channel NavAb via molecular dynamics simulations”. en. Em: *Proceedings of the National Academy of Sciences* 109.52 (dez. de 2012), pp. 21336–21341. ISSN: 0027-8424, 1091-6490.
- [52] Jiyao Wang, Yuqing Deng e Benoît Roux. “Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials”. en. Em: *Biophysical Journal* 91.8 (out. de 2006), pp. 2798–2814. ISSN: 0006-3495.
- [53] David L. Mobley, John D. Chodera e Ken A. Dill. “On the use of orientational restraints and symmetry corrections in alchemical free energy calculations”. Em: *The Journal of Chemical Physics* 125.8 (ago. de 2006). Publisher: American Institute of Physics, p. 084902. ISSN: 0021-9606.



- [54] David G. McLaren et al. “High-Throughput Mass Spectrometry for Hit Identification: Current Landscape and Future Perspectives”. en. Em: *SLAS Discovery*. Special Issue: Hit Discovery Methodology 26.2 (fev. de 2021), pp. 168–191. ISSN: 2472-5552.
- [55] Jacque Monod, Jeffries Wyman e Jean-Pierre Changeux. “On the nature of allosteric transitions: A plausible model”. en. Em: *Journal of Molecular Biology* 12.1 (mai. de 1965), pp. 88–118. ISSN: 0022-2836.

## **VI Anexos**

### **Anexo I**

Leonardo Cirqueira, et al. “Concentration-dependent thermodynamic analysis of the partition process of small ligands into proteins”. en. Em: Computational and Structural Biotechnology Journal 20 (jan. de 2022), pp. 4885–4891. ISSN:2001-0370.

### **Anexo II**

Alessandra S. Kiametis et al. “Atomistic Model for Simulations of the Sedative Hypnotic Drug 2,2,2-Trichloroethanol”. en. Em: ACS Omega (nov. de 2018).

### **Anexo III**

Letícia Stock et al. “Binding of the general anesthetic sevoflurane to ion channels”. en. Em: PLOS Computational Biology 14.11 (nov. de 2018). ISSN: 1553-7358.



## Concentration-dependent thermodynamic analysis of the partition process of small ligands into proteins

Leonardo Cirqueira, Letícia Stock, Werner Treptow\*

Laboratório de Biologia Teórica e Computacional (LBTC), Universidade de Brasília, DF CEP 70904-970, Brazil



### ARTICLE INFO

#### Article history:

Received 2 June 2022

Received in revised form 19 August 2022

Accepted 20 August 2022

Available online 1 September 2022

#### Keywords:

Non-specific low-affinity interactions

Partition coefficient

Free energy stability

Flooding MD

Cosolvent MD

### ABSTRACT

In the category of functional low-affinity interactions, small ligands may interact with multiple protein sites in a highly degenerate manner. Better conceived as a partition phenomenon at the molecular interface of proteins, such low-affinity interactions appear to be hidden to our current experimental resolution making their structural and functional characterization difficult in the low concentration regime of physiological processes. Characterization of the partition phenomenon under higher chemical forces could be a relevant strategy to tackle the problem provided the results can be scaled back to the low concentration range. Far from being trivial, such scaling demands a concentration-dependent understanding of self-interactions of the ligands, structural perturbations of the protein, among other molecular effects. Accordingly, we elaborate a novel and detailed concentration-dependent thermodynamic analysis of the partition process of small ligands aiming at characterizing the stability and structure of the dilute phenomenon from high concentrations. In analogy to an “aggregate” binding constant of a small molecule over multiple sites of a protein receptor, the model defines the stability of the process as a macroscopic equilibrium constant for the partition number of ligands that can be used to analyze biochemical and functional data of two-component systems driven by low-affinity interactions. Acquisition of such modeling-based structural information is expected to be highly welcome by revealing more traceable protein-binding spots for non-specific ligands.

© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

As we expand our scientific knowledge and gain insight about previously unknown phenomena, it becomes evident that low-affinity non-specific interactions may play functional roles in the realm of molecular biology. Recent examples include reactions of chemicals, lipids and disordered proteins with a variety of macromolecular targets ranging from proteins [1], chromatin [2,3] to membraneless compartments [4]. With implications for drug repurposing and side-effects [5], enhancement of proteasome based therapies [6] and pathologies [7], low-affinity non-specific interactions have just started to be uncovered across a number of emergent research fields.

In the category of functional low-affinity interactions, small ligands may interact with multiple sites of a protein target in a highly degenerate manner to a degree that it is better conceived as a partition phenomenon at the molecular interface of proteins [8]. Such low-affinity interactions are structurally and functionally challenging to characterize and thus might be hidden in the most relevant physiological, low-concentration regime [9]. Investigation

of the partition phenomenon under higher chemical thermodynamic forces or potentials could be one strategy to tackle the problem provided that the translation of results back to the low-concentration regime is made possible. The latter is a non-trivial task that requires among other molecular effects ligand self-interactions and protein structural perturbations to be well understood in a concentration-dependent manner.

To assess the feasibility of said strategy, we elaborate a novel and detailed concentration-dependent thermodynamic analysis of the partition process of small ligands to proteins, aiming at characterizing the free-energy stability and structure of the dilute phenomenon from high concentration sampling. According to the model, acquisition of dilute structural information derives directly from the spatial distribution of the ligand at high concentrations. The high-concentration distribution helps to define the free-energy stability of the dilute partition process as a macroscopic equilibrium constant that in analogy to an “aggregate” binding constant of a small molecule over multiple microscopic sites of a protein receptor is expected to be useful for analysis of biochemical and functional data of two-component systems governed by low-affinity non-specific interactions. In more technical terms,

\* Corresponding author.

E-mail address: [treptow@unb.br](mailto:treptow@unb.br) (W. Treptow).

the thermodynamic analysis establishes energetic conditions in which concentration effects are ineffective and the dilute partition coefficient of the ligand into the protein site becomes a concentration invariant quantity of the molecular distribution across two well-defined phases of the system. The outcome is a simple formulation that allows the partition number, free-energy stability and structure of the dilute partition process to be predicted from the unitary probability density of the ligand at high concentrations  $\bar{\rho}'(\mathbf{R})$ . Illustration and discussion of the approach is shown in the context of flooding or cosolvent Molecular Dynamics (MD) which is a widespread method that can be efficiently applied to resolve  $\bar{\rho}'(\mathbf{R})$  for a variety of systems under concentration effects and/or mixed-solvent compositions [10]. Here, we focus on the general anesthetic sevoflurane and the neuronal membrane protein Kv1.2, a mammalian ion channel for which structural data is available [11]. The fact that one or more anesthetic molecules do interact with the channel over multiple degenerate sites makes the system an important benchmark for numerical validation of our formulation [12].

## 1. Theory and methods

We consider a macromolecular system comprised of  $M$  chemical species, including a single protein  $P$  embedded in a large solvent volume  $V$  that contains  $N$  indistinguishable ligands  $L$ . The volume  $v$  within a certain cut-off distance of the protein is assumed to be a continuous phase occupied by a certain number of molecules

$$\langle n \rangle = N \int d\mathbf{R} \bar{\rho}(\mathbf{R}) \quad (1)$$

dictated by the spatial unitary density of the ligand across the system  $\bar{\rho}(\mathbf{R})$ .

Next, we rely on that definition to devise a thermodynamic model for the partition process of non-specific low-affinity ligands. Because such ligands occupy the protein site in a highly degenerate manner, the model assumes by construction that volume  $v$  is a nearly-homogeneous phase characterized by position-independent desolvation free-energy of the ligand.

### 1.1. Low-concentration partition coefficient, free-energy stability and structure

From Eq. (1), a partition coefficient may be defined relative to the number of ligand molecules  $N - \langle n \rangle$  in the bulk  $V - v$

$$\wp = \left( \frac{N - \langle n \rangle}{V - v} \right)^{-1} \frac{\langle n \rangle}{v} \quad (2)$$

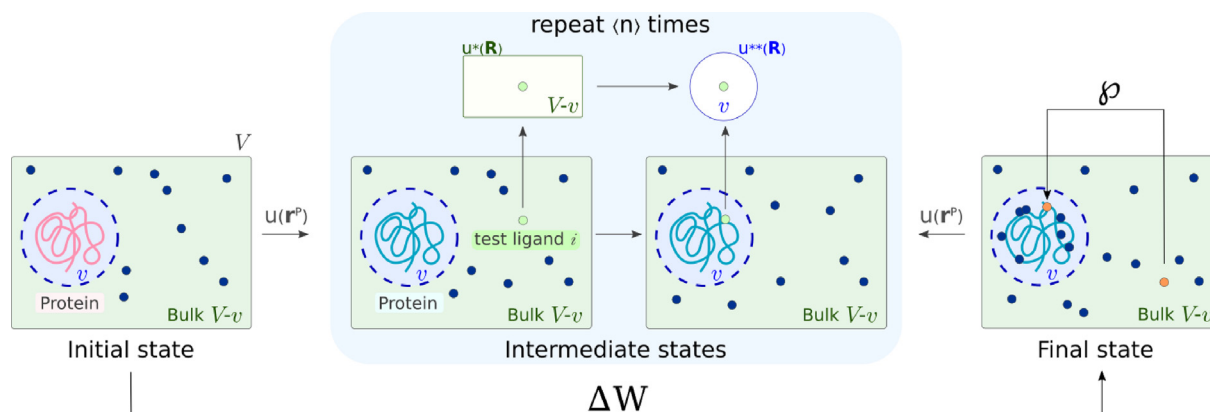
In the present form, Eq. (2) does not clarify any potential dependence that  $\wp$  may have with concentration. To analyze that aspect more carefully, we reinterpret in Fig. 1 the partition coefficient along an idealized thermodynamic cycle in which a single ligand molecule is reversibly decoupled from the bulk (\*) and the protein interface (\*\*)) into the gas phase with the use of auxiliary external potentials [13]. In case of small ligands, the molecule typically does not adopt any special configuration along the partition process and the external potentials  $\{u^*, u^{**}\}$  are purely translational flat wells applied to confine the ligands within three-dimensional volumes. According to that construction, the partition coefficient may then be directly linked to the desolvation free-energy of the ligand in each phase of the system  $\{W^*, W^{**}\}$

$$\wp = e^{-\beta[W^* - W^{**}]} \quad (3)$$

at a fixed temperature  $\beta = (k_B T)^{-1}$ . As detailed in Supporting Information, desolvation free energies are computed along decoupling transformations in FEP simulations and  $W^* - W^{**}$  reports the resulting free-energy difference to transfer the ligand from the bulk into the layer around the protein. Because each free-energy corresponds to reversible external work against non-bonded molecular forces imposed by the local environment including other ligand molecules,  $\{W^*, W^{**}\}$  are concentration-dependent quantities that according to a linear version of the partition model for binary mixtures [14]

$$\begin{cases} W^* = \bar{\mu} - h^* \frac{N - \langle n \rangle}{V - v} c \\ W^{**} = \bar{w} - h^{**} \frac{\langle n \rangle}{v} c \end{cases} \quad (4)$$

might deviate from their dilute reference values  $\{\bar{\mu}, \bar{w}\}$  at high molar concentrations ( $c = 1, 660 \text{ \AA}^3$ ), in consequence of self interactions of the ligand in the bulk and in the protein site  $\{h^*, h^{**}\}$ . Self-



**Fig. 1.** Partition process. In a typical isothermal flooding MD simulation, one simulates a molecular system comprised of a protein embedded in a ligand-rich environment. As the simulation proceeds, ligands dissolve across the system to reversibly partition into the protein under equilibrium conditions, regardless of any assumptions about the underlying phenomenon. The partition process can be described along an idealized thermodynamic cycle in which (i) the protein  $P$  is first restrained with energy  $u$  to its final equilibrium structure; (ii)  $\langle n \rangle$  ligands are decoupled from the bulk and the protein interface into the gas phase under applied restraints  $\{u^*, u^{**}\}$ ; and (iii) restraints are released in the final state. Implicit in step (ii), and demonstrated in the results, is the consideration that in case of small ligands, the molecule typically does not adopt any special configuration in the final state. External potentials are then purely translational flat wells applied to confine the ligands within their equilibrium three-dimensional volumes in the bulk  $V - v$  and in the protein interface  $v$ . As discussed in the text, the overall free-energy stability of the process  $\Delta W$  can be described in terms of the partition coefficient of the ligand  $\wp$ .

interaction energies of the ligand,  $\{h^*, h^{**}\}$ , can be either attractive ( $<0$ ) or repulsive ( $>0$ ). In the approximated limit of comparable self-interaction energies in both phases

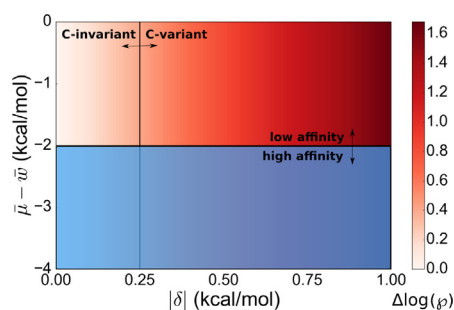
$$\delta \equiv \left[ h^* \frac{N - \langle n \rangle}{V - v} c - h^{**} \frac{\langle n \rangle}{v} c \right] \approx 0, \quad (5)$$

such concentration effects are expected to cancel out over the free-energy difference  $W^* - W^{**}$  and the partition coefficient to be conveniently re-expressed as a concentration invariant quantity of the desolvation free energies of the ligand under dilution i.e.,  $\varphi \approx e^{-\beta(\bar{\mu} - \bar{w})}$  (Fig. 2). Note that invariance of  $\varphi$  may hold true only at non-saturating concentrations in which the protein site is not maximally occupied by the ligand. At saturating conditions, the partition coefficient is expected to become concentration dependent and to decrease monotonically with increasing concentrations of the ligand in the bulk – a conclusion that agrees with the recent work by Carlson and coworkers dealing with binding of a single probe molecule in the microscopic cavity of a protein receptor under saturation conditions [15].

Extension of the thermodynamic cycle shown in Fig. 1 over additional intermediate states constructed by the use of external potentials [13] allows the overall stability of the partition process  $\Delta W$  to be properly evaluated relative to an initial state of the system in which all ligands are in the bulk. More specifically,  $\Delta W$  can be evaluated along a reversible path in which (i) the protein  $P$  is first restrained with energy  $u$  to its final equilibrium structure, (ii)  $\langle n \rangle$  ligands are decoupled from the bulk and the protein interface under restraints  $\{u^*, u^{**}\}$  and (iii) restraints are released in the final state. According to that construction, the overall stability of the partition process with  $\omega$  degenerate states then writes

$$e^{-\beta \Delta W} = \omega \left( \frac{v}{V - v} \right)^{\langle n \rangle} e^{-\beta [\Delta W_P + W_{(n)}^* - W_{(n)}^{**}]} \quad (6)$$

in terms of free-energy variations of the protein and ligands between initial and final states. For better readership of the work, rigorous derivation of Eq. (6) is presented on SI Eq. S5 through S13 in which each step of the thermodynamic cycle in Fig. 1 is properly defined as a reversible external work under restraints (see SI section-II for details). Briefly,  $\Delta W_P$  is the net amount of reversible work coupled to the transformation of the protein internal structure between initial and final states and as such, it is a concentration dependent quantity that results from structural perturbations imposed by multi-ligand interactions. On the other hand,



**Fig. 2.** Energetic conditions for the concentration-dependence of the partition coefficient of a small low-affinity molecule at non-saturating conditions. For typical low-affinity interactions  $\bar{\mu} - \bar{w} \geq -2.0$  kcal/mol, the partition coefficient  $\log(\varphi) \leq 3.5$  is expected to be concentration independent  $\Delta \log(\varphi) \leq 0.5$  in case of minor concentration effects in the range of thermal energy  $|\delta| \leq 0.25$  kcal/mol. Above that threshold, concentration effects are important and the partition coefficient depends on concentration. Color bar shows the variation of the partition coefficient relative to the case in which concentration effects are absent i.e.,  $\Delta \log(\varphi) \equiv \log(\varphi(|\delta|)) - \log(\varphi(0))$ .

$W_{(n)}^* - W_{(n)}^{**}$  is the net reversible work to decouple ligands from the system and as a state function, it can be expressed from Eq. (4)

$$W_{(n)}^* - W_{(n)}^{**} = \sum_{0 \leq i \leq \langle n \rangle} W^* \left( \frac{N - i}{V - v} c \right) - W^{**} \left( \frac{i}{v} c \right) \quad (7)$$

along a step-wise equilibrium process in which a single molecule  $i$  is decoupled from the bulk and re-coupled into the protein interface occupied by  $i - 1$  molecules [16]. In the low-concentration regime, structural perturbations of the protein are expected to vanish  $\Delta W_P \rightarrow 0$  and desolvation free-energies of the ligand to converge  $W_{(n)}^* - W_{(n)}^{**} \rightarrow \langle n \rangle (\bar{\mu} - \bar{w})$  implying that,

$$e^{-\beta \Delta W} = \omega \left( \frac{v \varphi}{V - v} \right)^{\langle n \rangle} = \left( \frac{N}{V - v} \right)^{\langle n \rangle} K \quad (8)$$

is the overall stability of the partition process of the ligands under dilution. In analogy to an aggregate binding constant of a small molecule over multiple microscopic binding sites of a protein receptor [8], Eq. (8) defines a macroscopic equilibrium constant for the average partition number that can be useful to analyze biochemical and functional data of the ligand. By taking into consideration the number of degenerate states  $\omega \approx N^{\langle n \rangle} / \langle n \rangle!$  in the thermodynamic limit  $\langle n \rangle \ll N$ , note that proper regularization of Eq. (8) by concentration  $(N/V - v)^{\langle n \rangle}$  provides us with the familiar binding constant equation i.e.,  $K \equiv \frac{v^{\langle n \rangle}}{\langle n \rangle!} \exp[-\beta \langle n \rangle (\bar{\mu} - \bar{w})]$ . The analogy is particularly important as knowledge of the binding constant ensures the occupancy probability of the protein to be known and therefore, quantification of any ensemble average thermodynamic property of the system with biochemical and functional implications [8].

For completeness, note further that under moderate structural perturbations of the protein and surroundings, the unitary density of the ligand in the final partition state is expected to be a concentration-independent three-dimensional map that satisfies  $\int d\mathbf{R} \bar{\rho}(\mathbf{R}) = 1$ . The spatial distribution of the ligands in the low-concentration regime may be then conveniently approximated from Eq. (1) as a weighted function of the unitary density nearby the protein i.e.,  $\rho(\mathbf{R}) = \langle n \rangle \bar{\rho}(\mathbf{R}) \forall \mathbf{R} \in v$ . For large protein receptors embedded in their native environment, structural perturbations are expected to be moderate within typical rmsd equilibrium values of the reference state free of ligands  $\leq 5.0$  Å.

## 1.2. Reconstruction of partition process from high concentrations

Taken together, the present analysis offers a self-consistent formulation for investigation of the dilute partition phenomenon of small low-affinity ligands from high concentrations. In practice, if the unitary-three-dimensional density of the ligand  $\bar{\rho}'(\mathbf{R})$  and its derived partition coefficient  $\varphi'$  are known from high concentrations then a number of ligands is expected to partition into the protein when the total number  $N$  or their concentration is decreased in the system  $N/V$

$$\langle n \rangle \approx \left( 1 + \frac{V - v}{v \varphi'} \right)^{-1} N \quad (9)$$

with an approximate free-energy stability and spatial density respectively given by

$$\Delta W \approx -\beta^{-1} \ln(\omega) - \beta^{-1} \langle n \rangle \ln \left( \frac{v \varphi'}{V - v} \right) \quad (10)$$

and

$$\rho(\mathbf{R}) \approx \langle n \rangle \bar{\rho}'(\mathbf{R}) \quad (11)$$

**Table 1**  
Equilibrium properties of flooding MD simulations.

Simulation	$N(\#)$	$V(\text{\AA}^3)$	$\frac{N}{V}(\text{M})$	$\langle n \rangle(\#)$	$v(\text{\AA}^3)$	$\frac{\langle n \rangle}{v}(\text{M})$	$N - \langle n \rangle(\#)$	$V - v(\text{\AA}^3)$	$\frac{N - \langle n \rangle}{V - v}(\text{M})$	$\log(\varphi)$
1	2	$1,729.326 \times 10^3$	0.002	0.82	$159.713 \times 10^3$	0.008	1.18	$1,569.613 \times 10^3$	0.001	1.93
2	29	$1,729.326 \times 10^3$	0.028	13.67	$159.713 \times 10^3$	0.142	14.33	$1,569.613 \times 10^3$	0.015	2.17
3	58	$1,729.326 \times 10^3$	0.056	22.87	$159.713 \times 10^3$	0.238	35.13	$1,569.613 \times 10^3$	0.037	1.86
4	87	$1,729.326 \times 10^3$	0.083	36.30	$159.713 \times 10^3$	0.377	50.70	$1,569.613 \times 10^3$	0.054	1.95
5	116	$1,729.326 \times 10^3$	0.111	45.57	$159.713 \times 10^3$	0.474	70.43	$1,569.613 \times 10^3$	0.074	1.85
6	174	$1,729.326 \times 10^3$	0.167	66.10	$159.713 \times 10^3$	0.687	107.90	$1,569.613 \times 10^3$	0.114	1.80

### 1.3. Computational methods

We rely on Eqs. (9), (10) and (11) to investigate the molecular partition of sevoflurane into Kv1.2 in the context of flooding or cosolvent MD simulations at high concentrations. Details of the calculations are provided as [Supporting Information](#). Scripts for analysis and molecular configurations for flooding MD simulations of the sevoflurane/kv1.2 system can be downloaded from github ZENODO repository under DOI number (<https://doi.org/10.5281/zenodo.6964766>).

## 2. Results and discussion

The main goal here is to investigate the partition phenomenon of small ligands into proteins. The work is illustrated and discussed in the context of flooding or cosolvent MD simulations of the general anesthetic sevoflurane and the voltage-gated channel Kv1.2 [11].

### 2.1. Partition coefficient

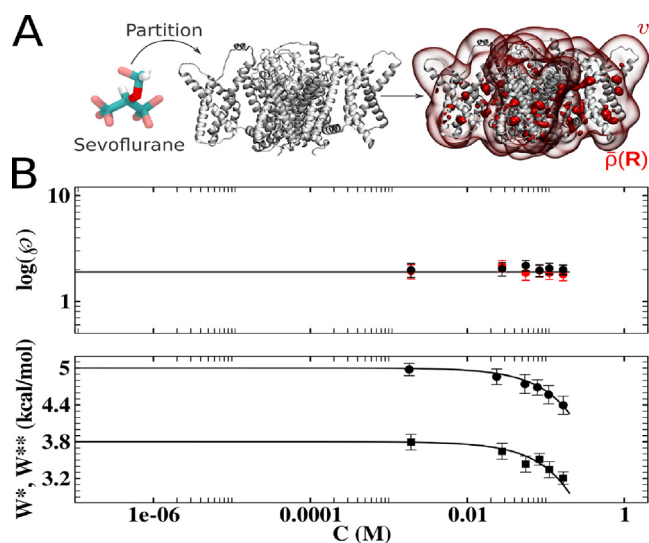
As detailed in Table 1, Fig. 3 and Fig. S1, independent MD simulations were carried out with the channel embedded in a phospholipid bilayer and flooded with sevoflurane at concentrations ranging from 0.02 to 0.167 M. Consistent with the previously reported finding that sevoflurane may impact the ion channel function by interacting with one or more binding sites in the low-affinity mM concentration range [12], the simulated three-dimensional distribution of the ligand nearby the protein was found to be highly degenerate, occurring over multiple spots of comparable unitary probability density. Non-specific interactions of sevoflurane were then analyzed as a partition phenomenon at the molecular interface of the channel.

For each simulation system, the bulk volume  $V - v$  was determined from the interface volume  $v$ ; whereas  $v$  is defined to be within typical ( $\leq 5.0$  Å) non-bonded distances of the protein [17]. Analysis of the simulations according to that volume decomposition reveals that ligand partition into the protein converges at a characteristic timescale of  $t^* \geq 0.1 \mu\text{s}$ , with an average concentration-independent partition coefficient  $\log(\varphi)$  of  $\sim 1.93$ . That invariant property of  $\log(\varphi)$  was independently investigated in terms of bulk (\*) and protein (\*\*) desolvation energies of the ligand at the same local concentrations of the flooding-MD simulations at equilibrium. Because the bulk of membrane proteins is made itself of aqueous (subscript 1) and lipid (subscript 2) regions:

$$e^{-\beta W^*} = \left(\frac{v_1}{V - v}\right)e^{-\beta W_1^*} + \left(\frac{v_2}{V - v}\right)e^{-\beta W_2^*}, \quad (12)$$

desolvation free energies  $\{W^*, W^{**}\}$  were estimated across each of the system's phases accordingly. Structural analysis of the ligand supports that no specific internal conformation or spatial orientation is adopted by the molecule across the system. Energy estimates were then obtained in the context of FEP calculations under applied translational flat wells (see [SI section-I](#) for further details; [Fig. S2](#) and [S3](#); [Table-S1](#) and [Table-S2](#)).

The desolvation free-energy of sevoflurane in the bulk is concentration-dependent and largely dominated by favorable interactions with lipids, in agreement with past studies [20]. FEP calculations further support that  $W^{**}$  is another concentration-dependent quantity characterized by non-specific ligand interactions with distinct regions of the protein interface. Adjustment of the partition model in Eq. (4) to FEP estimates shows that  $\{W^*, W^{**}\}$  respectively converges to their dilute values,  $\bar{\mu} = 3.8$  kcal.mol<sup>-1</sup> and  $\bar{w} = 5.0$  kcal.mol<sup>-1</sup>, in the low-concentration range ( $\leq 0.05$  M) in which repulsive self interactions of the ligand are absent ( $h^* = 3.8$  kcal.mol<sup>-1</sup>;  $h^{**} = 0.90$  kcal.mol<sup>-1</sup>). Because self-



**Fig. 3.** Partition coefficient and desolvation free energies as a function of the total molar concentration  $C$  of the ligand. (A) Molecular structures of sevoflurane, Kv1.2 and isovalue representation of the unitary-three-dimensional distribution  $\bar{\rho}(\mathbf{R})$  of the ligand within the interface volume  $v$  of the protein (red blobs). (B) Partition coefficient  $\log(\varphi)$  and desolvation free energies  $\{W^*, W^{**}\}$ . Estimates of  $\log(\varphi)$  (red circles) were computed from Eq. (2) by taking into consideration equilibrium MD properties reported in Table 1. Average value is  $\sim 1.93$  (continuous line). Estimates of  $\{W^*, W^{**}\}$  and statistical errors were determined using the simple overlap sampling formula based on at least 10 independent FEP simulations *per phase per concentration* [18].  $W^*$  was estimated by decoupling the ligand from bulk waters and lipids as devised in Eq. (12) (squares).  $W^{**}$  was estimated by decoupling the ligand from distinct regions of the channel including the S4 helix, S4S5 linker, S5S6 interface and central cavity (circles). According to the partition model (continuous line), the desolvation energy deviates from its dilute value as a result of unfavorable self interactions of the ligand at high concentration. The partition model was resolved as a linear fit of Eq. (4) to FEP estimates. Adjusted values with best regression coefficient ( $R > 0.99$ ) are: ( $\bar{\mu} = 3.8$  kcal.mol<sup>-1</sup>,  $h^* = 3.8$  kcal.mol<sup>-1</sup>) and ( $\bar{w} = 5.0$  kcal.mol<sup>-1</sup>,  $h^{**} = 0.90$  kcal.mol<sup>-1</sup>). Note that concentration effects cancel out over the free-energy difference  $W^* - W^{**}$  and the derived partition coefficient (black circles) is a concentration-independent quantity that agrees with estimates obtained from flooding-MD. Statistical errors of  $\log(\varphi)$  from flooding MD and FEP simulations were respectively obtained by bootstrap analysis of the partition number ( $n$ ) and free-energy estimates  $\{W^*, W^*\}$ . Analysis involved 100 sets of 50 resampling values with replacement in each concentration. Analysis of the MD trajectories was performed in VMD [19]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

interaction energies are comparable between phases  $\delta \approx 0$  (Table-S3), such concentration effects tend to cancel out over the free-energy difference  $W^* - W^{**}$  and the derived partition coefficient  $\log(\phi)$  nicely agrees with the flooding-MD estimates in a concentration-independent manner. Note that, FEP and flooding-MD estimates are independent themselves thus strengthening the conclusion that  $\log(\phi)$  is invariant across the range of non-saturating concentrations under investigation.

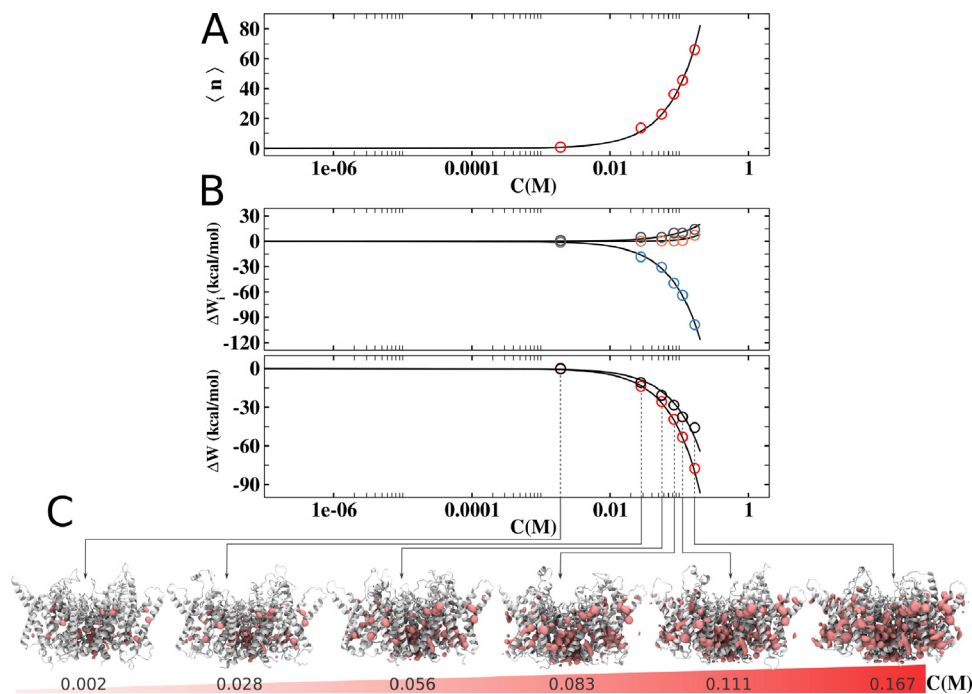
## 2.2. Reconstruction of the partition process from high concentrations

From Fig. 3, the conclusion that  $\log(\phi)$  is concentration invariant allows the stability and structure of the dilute partition process to be known from high concentrations as devised in Eqs. (9), (10) and (11).

As a saturation curve, the number of protein-interacting ligands versus concentration  $\langle n \rangle \times \log(C)$  is a sigmoid function that reaches a plateau number of ligands at high concentrations. In Fig. 4A, the partition number resolved from flooding MD increases with concentration thus indicating that simulations were indeed carried out at non-saturating conditions. Eq. (9) provides the prediction of the number of ligands  $\langle n \rangle$  that according to the high concentration partition coefficient  $\phi'$  is expected to partition into the protein as a function of the total concentration of the system  $N/V$ . Based on the partition coefficient  $\log(\phi')$  of 1.80 known from flooding-MD at 0.167 M, Fig. 4A then shows the solution of Eq. (9) as a function of the number of molecules  $0 \leq N \leq 200$  in the total solvent volume

of the simulation box i.e.,  $V=1,729,326 \times 10^3 \text{ \AA}^3$ . While agreeing with the flooding-MD estimates (regression coefficient  $R = 0.98$ ), the model predicts that the number of partition events must vanish in the  $\mu\text{M}$  range as a consequence of the low affinity of the molecule to the protein – a conclusion that is highly consistent with experimental findings [21].

Given the small low-affinity nature of the ligand, the structural conditions underlying its partition across the system appear to fulfill the molecular premises in Eq. (6) thus allowing the stability of the process to be modeled accordingly, in terms of perturbations of the internal structure of the protein P, desolvation energies of  $\langle n \rangle$  ligands L and entropic contribution S (see SI section-III for further details; Figs. S4, S5, S6 and S7). Despite the finite timescales of MD simulations, the protein structure shows a moderate yet clear tendency to deviate from its initial state with increasing ligand concentrations. Consistent with that structural perturbation induced by multi-ligand interactions, the free-energy change of the protein along the partition process  $\Delta W_p$  was found to be a positive and concentration-dependent quantity that amount up to  $+10.0 \text{ kcal.mol}^{-1}$  at high concentration and vanishes with depletion of the ligand. The total desolvation energy of the ligands along the partition process  $\Delta W_L = W_{(n)}^* - W_{(n)}^{**}$  is in contrast a more significant, negative and concentration-dependent function that totals a minimum of  $-120.0 \text{ kcal.mol}^{-1}$  at high concentration and vanishes with depletion of the ligand. Including entropic contributions  $\Delta W_S$ , the total energetic outcome then points to a quite stable partition process  $\Delta W = \sum \Delta W_i \leq 0$  primarily determined by favorable



**Fig. 4.** Partition number, free-energy stability and structure as a function of the total molar concentration  $C$  of the ligand. (A) Partition number  $\langle n \rangle$ . According to Eq. (9), prediction of the partition number (line) agrees with flooding-MD estimates (circles). Eq. (9) was resolved for the partition coefficient  $\log(\phi') = 1.80$  as a function of the total concentration i.e., number of molecules  $N$  in the total solvent volume of the simulation system  $V$ . (B) Partial  $\Delta W_i$  and total  $\Delta W$  free-energy variations. Estimates of  $\Delta W_i$  stem from structural perturbations of the protein P (orange circles), desolvation energies of the ligand L (blue circles) and entropic S contributions (gray circles). Individual contributions were combined in the total free-energy variation of the partition process  $\Delta W = \sum \Delta W_i$  (red circles) and subsequently compared to estimates obtained from Eq. (10) in which the total stability of the partition process is approximated from entropic contributions and the partition coefficient of the ligand i.e.,  $\Delta W' \approx \Delta W_S - \beta^{-1} \langle n \rangle \ln(\phi')$  (black circles). As respectively detailed in Figs. S4–S7, free-energy estimates  $\Delta W_{i=(P,L,S)}$ ,  $\Delta W$  and  $\Delta W'$  were linearly plotted as a function of the total molar concentration of the system  $C = (N/V)c$  and a linear or quadratic function (line) was adjusted to the data i.e.,  $\Delta W_p = h_p C^2$ ,  $\Delta W_L = h_L C$ ,  $\Delta W_S = h_S C$ ,  $\Delta W = hC$  and  $\Delta W' = h' C$ . Adjusted values with best regression coefficient ( $R > 0.93$ ) are: ( $h_p=200.0 \text{ kcal.mol}^{-1}$ ,  $h_L=-580.0 \text{ kcal.mol}^{-1}$ ,  $h_S=100.0 \text{ kcal.mol}^{-1}$ ,  $h=-480.0 \text{ kcal.mol}^{-1}$  and  $h'=-320.0 \text{ kcal.mol}^{-1}$ ). (C) Three-dimensional distribution  $\rho(\mathbf{R})$  at isovalues of  $1.0 \times 10^{-4} \text{ \AA}^{-3}$ . Concentration-dependent spatial densities of the ligand (red) at the interface of the protein (white) were reconstructed at 0.002, 0.028, 0.056, 0.083, 0.111 and 0.167 M. Reconstruction followed Eq. (11), by respectively reweighing the unitary density of the molecule resolved from flooding MD at high concentration (0.167 M) by the partition number in each concentration. Because every reconstructed density derives from the same unitary density  $\bar{\rho}(\mathbf{R})$  at 0.167 M, differences between any two concentration-dependent distributions result from the respective reweighing partition numbers  $\langle n \rangle$  considered in the procedure (Table 1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

desolvation energies of the ligands that surmount minor destabilizing perturbations of the protein structure across the entire concentration range. Because concentration effects accounting for structural perturbations of the protein and self-interactions of the ligand disappear in the dilute regime,  $\Delta W$  becomes a function of the partition coefficient of the molecule and its partition number, and as such, it can be well approximated at low concentrations by Eq. (10). According to that approximation, sevoflurane partition into Kv1.2 appears to be stabilized by less than 10.0 kcal.mol<sup>-1</sup> across the range of experimental concentrations ( $\leq 0.01$  M) in which general anesthetics affect ion channels [22].

Because perturbations of the protein structure induced by ligand partition are moderate within typical root-mean-square deviations in solution  $\leq 5.0$  Å, the equilibrium spatial distribution of the ligand at the protein interface  $\rho(\mathbf{R})$  was reconstructed from Eq. (11) by reweighing the unitary density of the molecule resolved from MD at high concentrations. Better conveying the partition phenomenon under investigation,  $\rho(\mathbf{R})$  is uniformly spread over multiple concentration dependent sites that are preferentially localized at the protein interface with lipids as a consequence of favorable interactions of the ligand with that moiety of the bulk. Careful inspection of the distribution model supports that multiple interaction spots may equally contribute to drug modulation of the channel in the low-concentration regime ( $\leq 0.01$  M), including the gating-implicated S4S5 linker previously reported in photolabeling experiments [23,24]. Supporting the conclusion of a degenerate mode of action, the ability of general anesthetics to modulate Kv channels was indeed found not to depend on specific interactions of the molecule at established cavities of the protein [25]. Further implications for the action of general anesthetics on ion channels will be discussed elsewhere (manuscript in preparation).

### 3. Concluding remarks

Proteins are targets for a large family of ligands, including small low-affinity molecules featuring a wide spectrum of biological roles. How such ligands modulate protein function must build on understanding their highly degenerate atomic-level interactions under dilution, thus currently challenging the resolution of both theoretical and experimental routines. Looking for new developments in the field, we conceive the interaction mode of small low-affinity ligands as a partition phenomenon governed by non-specific interactions at the molecular surface of proteins. According to that molecular description, thermodynamic analysis of concentration-dependent self-interactions of the ligand and structural perturbations of the protein allows the partition coefficient, stability and structure of the dilute partition process to be reconstructed from the unitary equilibrium probability density of the ligand  $\bar{\rho}(\mathbf{R})$ . The result stems essentially from the fact that under non-disruptive chemical forces on the protein structure and surroundings, the spatial distribution of the ligand is expected to be a concentration-independent three-dimensional map that can be efficiently learned at high concentrations.

Illustration and discussion of the thermodynamic analysis is made here in the context of flooding-MD simulations of the general anesthetic sevoflurane and the important neuronal membrane protein Kv1.2, illuminating its great utility for the structural biology field. Ideally, the dilute partition process of a small low-affinity molecule would be resolved from MD directly if such simulations did not require very long time scales to converge at low concentrations. By allowing the dilute partition process to be reconstructed directly from high-concentration conditions in which sampling is significantly enhanced over shorter time scales, our thermodynamic analysis is then expected to unlock the power of MD in the low concentration regime. Especially true for the more sampling-

sensitive structural properties, the reconstructed spatial distribution of the ligand in the low-concentration regime provides us with a larger set of interaction spots nearby the protein than the same estimates directly resolved from MD in consequence of poor sampling of the partition process over the finite timescale of the simulation. Supporting that conclusion (Fig. S8), the low-concentration mismatch between the volume occupied by the ligand in the observed and predicted spatial probability densities is significantly decreased by enhancing sampling either via symmetrization across the fourfold symmetry of the channel or substantial extension of simulation time (over  $\sim 5$   $\mu$ s at 5 mM of sevoflurane).

Considering the complete space of physical chemical characteristics of ligands, protein and environment, and the combinatorial arrangements thereof, a myriad of scenarios are possible. While it would be difficult to list and evaluate all of them, the present formulation, devised to study low-affinity ligands, has underlying assumptions that are required for its proper applicability in the context of flooding MD simulations. First, we hereby define low-affinity ligand as a class of molecules that interact with proteins in the mM concentration range. Second, as a consequence of low affinity, the ligand's distribution nearby the protein should be degenerate, occurring over multiple spots of comparable unitary spatial probability density  $\bar{\rho}(\mathbf{R})$ . If both conditions are true, then the ligand interaction is likely better described by a partition phenomenon, in which the molecule occupies the protein with a characteristic partition coefficient  $\log(\phi) \leq 3.5$  and predominant loss of translational freedom. Third, if the partition phenomenon is indeed concentration independent,  $\log(\phi)$  calculated in a second, lower, concentration provides a  $\Delta\log(\phi) \leq 0.5$  (cf. Fig. 2), indicating that the probability density of the ligand and the derived partition coefficient are concentration-invariant and can be efficiently learned from high concentrations.

Naturally, invariance of the unitary distribution and the derived partition coefficient only hold true at non-saturating concentrations in which the partition number  $\langle n \rangle$  does not saturate and perturbations of the protein structure are moderate, within typical RMSD equilibrium values of the reference state free of ligands  $\leq 5.0$  Å. As formulated herein, it is important to clarify that the low-affinity description is not suitable for processes involving high affinity binders such as, charged molecules, lipids and peptides. In such high-affinity scenario, the binder interacts preferentially at one specific microscopic cavity of the protein receptor with a position-dependent desolvation energy  $W^{**}(\mathbf{R})$  and loss of translational, rotational and conformational freedom. The implications of the latter are important for ligand interaction and its description must include additional gas-phase intermediates to define a proper binding constant for accurate description of the binding process and the derived equilibrium properties at the dilute low concentration range [13,14]. An aggregate constant may be defined accordingly in case of two or more high-affinity sites [8].

Abiding by the same requirements, independent flooding-MD calculations for the sedative hypnotic drug trichloroethanol [26] and the aqueous-soluble bovine serum albumin [27] further support that the thermodynamic model may be applicable to other two-component interacting systems guided by non-bonded low-affinity interactions (Table-S4, Fig. S9 and S10). Consistent with fluorescence quenching experiments showing that trichloroethanol interacts with albumin in the low-affinity mM range [28], the spatial distribution of the ligand resolved from simulation  $\bar{\rho}(\mathbf{R})$  is degenerate over multiple spots of comparable unitary probability density at the protein layer. Better conveying a partition phenomenon, the partition coefficient of the ligand  $\log(\phi) \approx 3.21$  is shown to be clearly concentration-independent  $\Delta\log(\phi) \approx 0.25$  across two non-saturating conditions 50 and 100 mM in which the partition number  $\langle n \rangle$  does not saturate and structural



deviations of the protein are within equilibrium RMSD values of the reference state free of ligands  $\leq 5.0$  Å. Solution of the thermodynamic model then allows the partition coefficient, stability and structure of the dilute partition process to be reconstructed from the unitary density of the ligand at high concentration. Careful inspection of the distribution model supports that multiple interaction spots may equally contribute for drug interaction in the low-concentration regime, including tryptophan regions previously reported in fluorescence experiments [28].

Because proteins are primary targets for small low-affinity molecules that can be widely investigated by the scientific community in the context of a partition phenomenon at high concentrations, we thus believe the study is of broad interest and likely useful in producing new results in the field. With promising prospects, implementation and convergence of the model in the context of MD simulations appears easier than other more elaborate and time consuming approaches [29], making it especially attractive for a large number of applications. By revealing more traceable atomic-level spots at the molecular surface of proteins, we anticipate that modeling-based structural information might be especially useful to delineate and interpret novel high-throughput mass spectrometry experiments [9] aimed at characterizing functional interactions of non-specific ligands under physiological conditions. Reconstruction of the dilute spatial distribution of two distinct solvents from high concentrations might be of utility to potentiate identification of active and allosteric binding sites in protein receptors in the context of mixed-solvent simulations [30].

#### CRedit authorship contribution statement

**Leonardo Cirqueira:** Investigation, Software, Visualization, Validation, Writing – review & editing. **Letícia Stock:** Investigation, Software, Visualization, Validation, Writing – review & editing. **Werner Treptow:** Conceptualization, Methodology, Software, Visualization, Validation, Writing – original draft, Supervision, Funding acquisition.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

We would like to thank B. Roux for helpful discussions. The work was supported by National Council for Scientific and Technological Development CNPq [grant number 302089/2019-5 and 200114/2020-4 (WT)], Coordenação de Aperfeiçoamento de Pessoal de Nível Superior CAPES [grant number 23038.010052/2013-95 (WT)], and Fundação de Apoio a Pesquisa do Distrito Federal FAPDF [grant number 193.001.202/2016 (WT)]. LS thanks CAPES for postdoctoral fellowship (grant number 88882.463151/2019-01).

#### Author contributions

LC and LS contributed equally to this work.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2022.08.049>.

#### References

- [1] Uversky VN. Intrinsic disorder-based protein interactions and their modulators. *Curr Pharm Des* 2013;19(23):4191–213.
- [2] Silva ITG, Fernandes V, Souza C, Treptow W, Santos GM. Biophysical studies of cholesterol effects on chromatin. *J Lipid Res* 2017;58(5):934–40.
- [3] Fernandes V, Teles K, Ribeiro C, Treptow W, Santos G. Fat nucleosome: Role of lipids on chromatin. *Prog Lipid Res* 2018;1(70):29–34.
- [4] Musacchio A. On the role of phase separation in the biogenesis of membraneless compartments. *EMBO J* 2022:e109952.
- [5] Chartier M, Morency LP, Zylber MI, Najmanovich RJ. Large-scale detection of drug off-targets: hypotheses for drug repurposing and understanding side-effects. *BMC Pharmacol Toxicol* 2017;18(1):18.
- [6] George DE, Tepe JJ. Advances in Proteasome Enhancement by Small Molecules. *Biomolecules* 2021;11(12):1789.
- [7] Uversky VN. Intrinsically disordered proteins and their (disordered) proteomes in neurodegenerative disorders. *Front Aging Neurosci* [Internet]. 2015 [cited 2022 Feb 7]; Available from: <https://www.frontiersin.org/article/10.3389/fnagi.2015.00018>.
- [8] Stock L, Hosoume J, Treptow W. Concentration-dependent binding of small ligands to multiple saturable sites in membrane proteins. *Sci Rep* 2017;7(1):5734.
- [9] McLaren DG, Shah V, Wisniewski T, Ghislain L, Liu C, Zhang H, et al. High-throughput mass spectrometry for hit identification: current landscape and future perspectives. *SLAS Discov Adv Sci Drug Discov* 2021;26(2):168–91.
- [10] Guvench Jr O, Adm. Computational fragment-based binding site identification by ligand competitive saturation. *PLOS Comput Biol* 2009;5(7):e1000435.
- [11] Long SB, Campbell EB, MacKinnon R. Crystal structure of a mammalian voltage-dependent Shaker family K<sup>+</sup> channel. *Science* 2005;309:897–903.
- [12] Stock L, Hosoume J, Cirqueira L, Treptow W. Binding of the general anesthetic sevoflurane to ion channels. *PLOS Comput Biol* 2018;14(11):e1006605.
- [13] Woo HJ, Roux B. Calculation of absolute protein–ligand binding free energy from computer simulations. *Proc Natl Acad Sci U S A* 2005;102(19):6825–30.
- [14] Salari R, Joseph T, Lohia R, Héning J, Brannigan G. A streamlined, general approach for computing ligand binding free energies and its application to GPCR-bound cholesterol. *J Chem Theory Comput* 2018;14(12):6560–73.
- [15] Ghanakota P, DasGupta D, Carlson HA. Free Energies and Entropies of Binding Sites Identified by MixMD cosolvent simulations. *J Chem Inf Model* 2019;59(5):2035–45.
- [16] Roux B, Nina M, Pomès R, Smith JC. Thermodynamic stability of water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study. *Biophys J* 1996;71(2):670–81.
- [17] Sanner MF, Olson AJ, Spehner JC. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers* 1996;38(3).
- [18] Lu N, Kofke DA, Woolf TB. Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods. *J Comput Chem* 2004;25(1):28–40.
- [19] Humphrey W, Dalke A, Schulten K. VMD - visual molecular dynamics. *J Molec Graph* 1996;14(1):33–8.
- [20] Pickholz M, Saiz L, Klein ML. Concentration effects of volatile anesthetics on the properties of model membranes: A coarse-grain approach. *Biophys J* 2005;88(3):1524–34.
- [21] Barber AF, Liang Q, Covarrubias M. Novel activation of voltage-gated K<sup>+</sup> channels by sevoflurane. *J Biol Chem* 2012;4(287):40425–32.
- [22] Liang Q, Anderson WD, Jones ST, Souza CS, Hosoume JM, Treptow W, et al. Positive Allosteric Modulation of Kv Channels by Sevoflurane: Insights into the Structural Basis of Inhaled Anesthetic Action. *PLoS ONE* 2015 Nov 24;10(11):e0143363.
- [23] Bu W, Liang Q, Zhi L, Maciunas L, Loll PJ, Eckenhoff RG, et al. Sites and functional consequence of alkylphenol anesthetic binding to Kv1.2 channels. *Mol Neurobiol*. 2017;1–11.
- [24] Woll KA, Peng W, Liang Q, Zhi L, Jacobs JA, Maciunas L, et al. Photoaffinity ligand for the inhalational anesthetic sevoflurane allows mechanistic insight into potassium channel modulation. *ACS Chem Biol* 2017;12(5):1353–62.
- [25] Bu W, Liang Q, Zhi L, Maciunas L, Loll PJ, Eckenhoff RG, et al. Sites and functional consequence of alkylphenol anesthetic binding to Kv1.2 channels. *Mol Neurobiol* 2018;55(2):1692–702.
- [26] Kiametis AS, Stock L, Cirqueira L, Treptow W. Atomistic model for simulations of the sedative hypnotic drug 2,2,2-trichloroethanol. *ACS Omega* 2018;3(11):15916–23.
- [27] Bujacz A. Structures of bovine, equine and leporine serum albumin. *Acta Crystallogr D Biol Crystallogr* 2012;68(Pt 10):1278–89.
- [28] Solt K, Johansson JS. Binding of the active metabolite of chloral hydrate, 2,2,2-trichloroethanol, to serum albumin demonstrated using tryptophan fluorescence quenching. *Pharmacology* 2002;64(3):152–9.
- [29] Durrant JD, McCammon JA. Molecular dynamics simulations and drug discovery. *BMC Biol* 2011;9(1):71.
- [30] Ghanakota P, Carlson HA. Moving beyond active-site detection: MixMD applied to allosteric systems. *J Phys Chem B* 2016;120(33):8685–95.

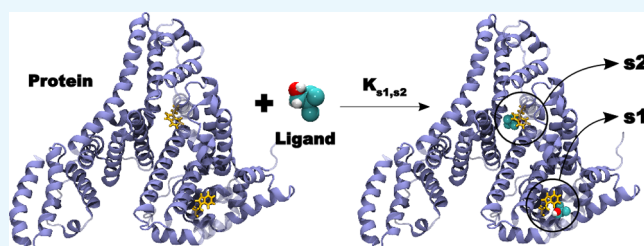
# Atomistic Model for Simulations of the Sedative Hypnotic Drug 2,2,2-Trichloroethanol

Alessandra S. Kiametis, Letícia Stock, Leonardo Cirqueira, and Werner Treptow\*<sup>1b</sup>

Laboratório de Biologia Teórica e Computacional, Departamento de Biologia Celular, Universidade de Brasília, DF, 70910-900 Brasília, Brazil

## Supporting Information

**ABSTRACT:** 2,2,2-Trichloroethanol (TCE) is the active form of the sedative hypnotic drug chloral hydrate, one of the oldest sleep medications in the market. Understanding of TCE's action mechanisms to its many targets, particularly within the ion channel family, could benefit from the state-of-the-art computational molecular studies. In this direction, we employed de novo modeling aided by the force field toolkit to develop CHARMM36-compatible TCE parameters. The classical potential energy function was calibrated targeting molecular conformations, local interactions with water molecules, and liquid bulk properties. Reference data comes from both tabulated thermodynamic properties and ab initio calculations at the MP2 level. TCE solvation free energy calculations in water and oil reproduce a lipophilic, yet nonhydrophobic, behavior. Indeed, the potential mean force profile for TCE partition through the phospholipid bilayer reveals the sedative's preference for the interfacial region. The calculated partition coefficient also matches experimental measures. Further validation of the proposed parameters is supported by the model's ability to recapitulate quenching experiments demonstrating TCE binding to bovine serum albumin.



## INTRODUCTION

Chloral hydrate is a sedative hypnotic drug used for short-term insomnia treatment. First synthesized in 1832,<sup>1</sup> chloral hydrate is the oldest sleep medication in the market. Chloral hydrate has been widely used as sedative in children undergoing clinical procedures. Its prescription in pediatrics is recommended for certain diagnostic procedures such as neurological imaging,<sup>2–5</sup> echocardiography,<sup>6</sup> and auditory brainstem response testing<sup>7</sup> when the patients do not respond to other agents.<sup>8</sup>

After oral administration, the chloral hydrate prodrug is rapidly converted into 2,2,2-trichloroethanol (C<sub>2</sub>H<sub>3</sub>Cl<sub>3</sub>O<sub>2</sub>). This active metabolite acts at the barbiturate recognition site on  $\gamma$ -aminobutyric acid (GABA<sub>A</sub>) receptors, eliciting sedation.<sup>9</sup> Whole-cell patch-clamp recording has shown that 2,2,2-trichloroethanol (TCE) potentiates GABA<sub>A</sub>-activated chloride current in mouse hippocampal neurons,<sup>10</sup> enhancing synaptic transmission. In addition to GABA<sub>A</sub> receptors, electrophysiology essays reported that TCE can interact with several other molecular targets. For example, TCE inhibits excitatory N-methyl-D-aspartate receptor and kainate-activated currents in mouse hippocampal neurons at clinical concentrations, more potently than ethanol.<sup>11</sup> Recently, it has also been shown that TCE modulates the recombinant human two-pore-domain potassium channels TREK-1 and TRAAK in a reversible, concentration-dependent manner.<sup>12</sup> Furthermore, tryptophan fluorescence quenching experiments demonstrated that TCE binds to bovine (BSA) and human (HSA) serum albumin plasma proteins.<sup>13</sup>

Despite the fact that TCE potentiates a range of protein receptors, its action mechanisms at the microscopic level are not clear. In the context of molecular dynamics (MD) simulations, understanding how TCE modulates these proteins requires an accurate atomistic model for both receptor and ligand. Although high-resolution crystallographic structures for some of these targets are available in the Protein Data Bank, an all-atom TCE model is still missing. We therefore present TCE parameters compatible with the CHARMM additive force field for biomolecules, useful for MD simulations.<sup>14</sup> The model is based on target quantities including molecular conformations, bulk phase properties, partition coefficient in water and oil, and TCE's binding affinity to BSA.

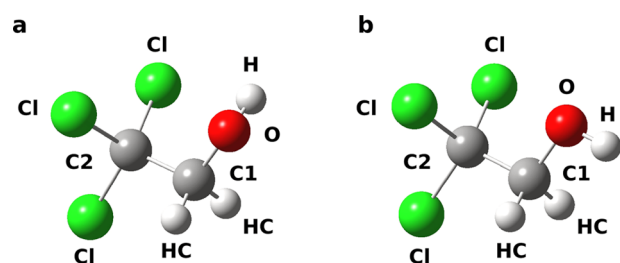
## RESULTS AND DISCUSSION

**TCE Parameters.** TCE presents two main conformers, trans and gauche (Figure 1), with electronic properties calculated at the MP2 level shown in Table 1. The magnitude of the trans conformation dipole moment  $|\vec{\mu}|$  differs significantly from that of the gauche one (3.2 D against 1.6 D, respectively). Despite that, the isotropic portion of the polarizability tensor  $\bar{\alpha}$  (au) shows almost the same magnitude for both geometries (53.0 and 52.1 au, respectively), with all individual components being very similar. The values for nondiagonal components suggest that both TCE conformers

Received: August 13, 2018

Accepted: November 8, 2018

Published: November 26, 2018



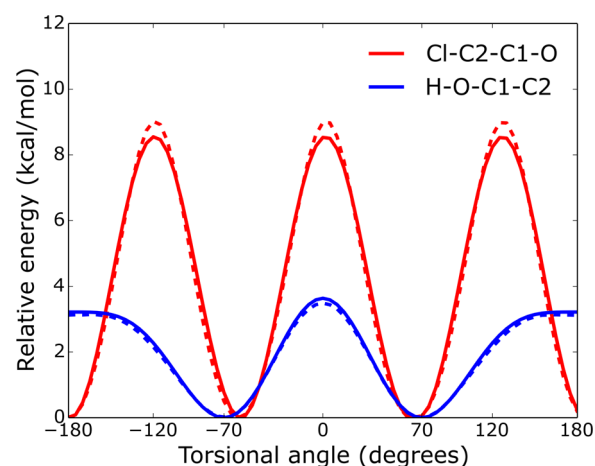
**Figure 1.** Ball-and-stick representation of TCE conformers (A) gauche- and (B) trans-optimized at the MP2 level. Atom names are indicated. Graphics were rendered using GaussView.<sup>15</sup>

display an isotropic polarizable character. Gauche is the most stable at the MP2 level; the energy difference between geometries is 3.22 kcal/mol. All-atom parameters were thus calibrated by taking the gauche conformer as the molecular target.

The dipole moment of TCE at the quantum mechanical (QM) level is overestimated by 43% in the mechanical molecular (MM) model, preserving vector orientation.<sup>14</sup> MM parameters for bond and valence angles reproduce QM geometry within an acceptable margin of error with deviations up to 0.03 Å and 3°, respectively.<sup>14</sup> TCE intramolecular interactions also agree with same data obtained from large-angle X-ray scattering experiments.<sup>16</sup> Torsional angles Cl–C2–C1–O and H–O–C1–C2 are used as conformational descriptors, the latter distinguishing trans and gauche conformers. As shown in Figure 2, their MM potential energy surfaces (PESs) fit QM data within 0.125 kcal/mol, comparable to  $k_B T$ . Parameters and associated errors for partial charges and Lennard-Jones (LJ) and bond parameters are presented as Supporting Information Tables S1–S3. Note that Lennard-Jones parameters used to treat van der Waals interactions of TCE were assigned by transferability of similar chemical types available at the CHARMM force field (cf. Computational Methods).

**Pure Solvent Properties.** As shown in Table 2, calculated density and enthalpy of vaporization of TCE agree with experimental reference values, reproducing physicochemical properties of the bulk phase. Radial pair distribution functions (RDFs) as computed from pure solvent MD simulations of TCE show well-resolved peaks at 1.79, 2.72, 2.92, 3.12, and 3.94 Å, respectively, assigned to bond (–) and nonbonded (⋯) interactions  $C_2\text{--}Cl_1$ ,  $C_1\cdots Cl_1$ ,  $Cl_1\cdots Cl_3$ ,  $O_1\cdots Cl_1$ , and  $O_1\cdots Cl_3$  (Figure 3). A hydrogen bond between  $O_1$  and  $O_1$  is also resolved at 2.92 Å. All peaks coincide with the reference RDF description for pure TCE, indicating that intra- and intermolecular interactions are correctly reproduced at the atomic level.<sup>16</sup> Note that simulation of TCE is predominantly populated by the gauche conformer with an occupancy probability of 78%, as expected from its larger stability. The gauche conformer thus accounts for most of the nonbonded interactions resolved in RDF analysis.

**Solvation Properties.** TCE's solvation free energy into hexane (oil) was calculated via free energy perturbation (FEP)



**Figure 2.** Potential energy surfaces for torsional angles Cl–C2–C1–O (red) and H–O–C1–C2 (blue). Data is obtained from fully relaxed torsional scans at the MP2 level (solid lines) and empirical model (dashed lines). Note the two local minima for H–O–C1–C2 torsional angle at  $-70$  and  $+70^\circ$ , both gauche conformers.

**Table 2. Bulk Phase Properties of TCE**

	calculated	reference	deviation (%)
density (g/mL)	$1.59 \pm 0.013$	$1.49^a$	6.0
enthalpy of vaporization (kcal/mol)	$11.42 \pm 0.280$	$10.82^b$	5.5

<sup>a</sup>Experimental value available at ACS. <sup>b</sup>Theoretical value available at ACS. Calculated using Advanced Chemistry Development (ACD/Labs) Software V11.02 (1994–2018 ACD/Labs).

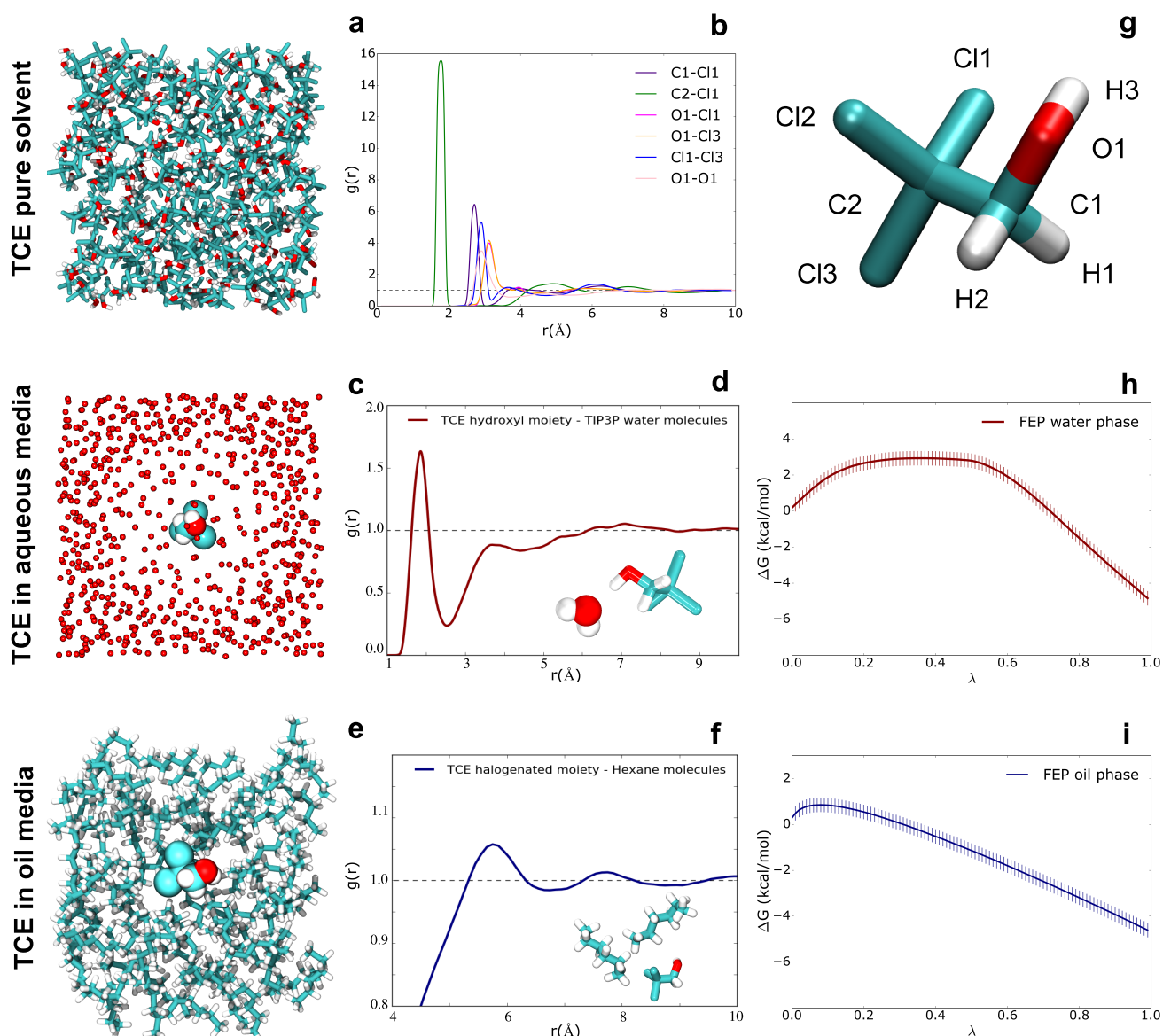
and amounts to  $\Delta G_{\text{oil}} = -4.60 \pm 0.001$  kcal/mol, in agreement with TCE's expected apolar character.<sup>17</sup> In spite of its favorable interaction with the apolar medium, TCE's solvation free energy in water is within the same range as that in oil ( $\Delta G_{\text{water}} = -4.83 \pm 0.001$  kcal/mol), suggesting that TCE is a lipophilic yet nonhydrophobic molecule. Figure 3 displays RDF profiles and solvation free energy for TCE in both water and oil (hexane).

**Transfer of TCE Across Membrane.** To characterize TCE's partition into the membrane, adaptive biasing force (ABF) simulations were employed to compute the potential of mean force (PMF) of a single TCE crossing a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) bilayer in the transmembrane  $z$  direction (Figure 4). Analysis of PMF evidences a clear energy minimum at the phospholipids' headgroup–tail interface region, with a stabilization of  $\Delta G_{\text{POPC}}^0 = -4.32 \pm 0.690$  kcal/mol relative to bulk. TCE behavior accompanies the trend of some anesthetics such as isoflurane and sevoflurane, which show a distinct preference for the membrane interface.<sup>20</sup> For comparison, Figure 4 also shows sevoflurane PMF, calculated under the same conditions as for TCE (anesthetic parameters come from Barber et al.<sup>21</sup>). PMF-derived atomic density for TCE is depicted in Figure 4c.

It is worth noting that TCE is a halogenated molecule containing a hydroxyl group that can act as a hydrogen bond

**Table 1. Electronic Properties of trans and gauche Conformers of TCE Calculated at the MP2 Level**

	$ \bar{\mu} $ (D)	$\bar{\alpha}$ (au)	$a_{xx}$	$a_{xy}$	$a_{yy}$	$a_{xz}$	$a_{yz}$	$a_{zz}$
trans	3.2	52.1	53.6	0.0	53.6	−4.2	0.0	49.1
gauche	1.6	53.0	55.5	−0.7	55.6	−5.9	0.3	47.7



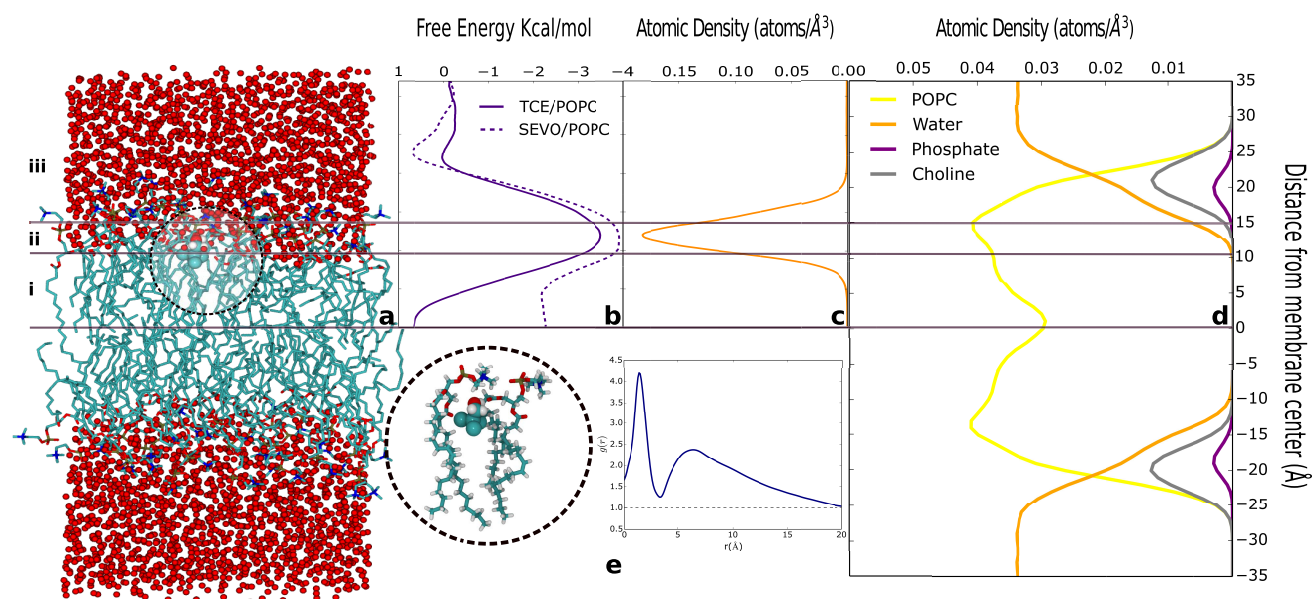
**Figure 3.** Solvation properties of TCE. First and second columns show respectively molecular systems and corresponding radial pair distribution functions (RDFs) for (a, b) pure TCE used in MD bulk phase simulations; (c, d) TCE in water; and (e, f) TCE in hexane. RDF of TCE's oxygen O1 relative to water hydrogen atoms evidences a first solvation layer at approximately 1.8 Å (d), whereas RDF of TCE in hexane shows a first solvation layer between 5 and 6 Å (f). (g) Licorice representation of the TCE gauche conformer. Free energy change as function of TCE's coupling/decoupling parameter  $\lambda$  in (h) bulk water phase and (i) in hexane. Error bars were estimated by the simple overlap sampling (SOS) method.<sup>18</sup> Molecular images were rendered using visual MD (VMD).<sup>19</sup>

donor. Indeed, TCE–membrane headgroup RDF indicates a first peak at about 1.5 Å, configuring a typical hydrogen bond between the ligand and the lipid (Figure 4e). The hydroxyl group observed in TCE is absent in the general anesthetic sevoflurane, thus rationalizing their differential stabilization within the membrane core. Although the former displays modest stabilization (less than 1 kcal/mol) in water, the energy difference in the latter favors membrane core by approximately 3 kcal/mol.

The PMF of TCE was integrated to estimate its POPC–water partition coefficient ( $\log K$ ), eq 4. The calculated and tabulated theoretical<sup>a</sup> values, respectively,  $\log K = 1.69 \pm 0.001$  and  $0.97 \pm 0.389$  at 298.15 K, are within the same order of magnitude, supporting model's ability to reproduce ligand's lipid solubility properties.

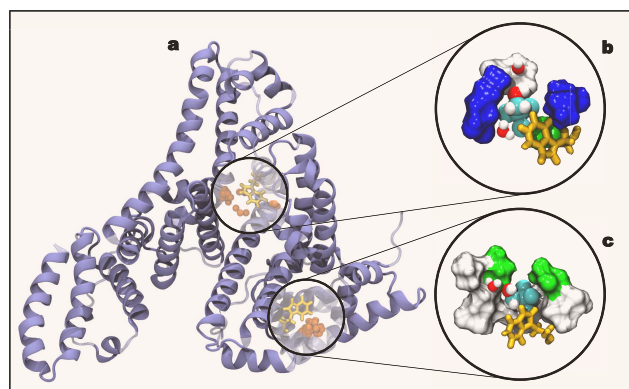
**Dissociation Constants for BSA Sites.** TCE is known to interact and bind to both human and bovine serum albumin (HSA and BSA, respectively). To further validate the proposed TCE parameters, model's affinity to these binding partners are confronted with ones determined from fluorescence quenching experiments.<sup>13</sup>

Briefly, in the experiment, BSA tryptophan residues W134 and W213 were effectively quenched by TCE at pH 7.0, allowing to pin down an apparent TCE dissociation constant of  $K_D = 3.3 \pm 0.3$  mmol/L. HSA contains a single tryptophan residue, namely, W214, which is analogous to BSA's W213. Thus, a combination of HSA and BSA quenching experiments meant individual dissociation constants of TCE from sites in the vicinity of BSA's W134 and W213 could be estimated, respectively, as  $K_D = 2.1 \pm 0.1$  and 12.0 mmol/L.<sup>13</sup>



**Figure 4.** Transfer of TCE across membrane. (a) TCE–membrane system considered in ABF simulations. Phospholipid tails (i), membrane headgroups (ii), and water (iii) are indicated. (b) Potential of mean force (PMF) of TCE along the transmembrane direction. Accumulated statistical error for the PMF was estimated following eq 3 (cf. Computational Methods) and amounts to 0.69 kcal/mol, indicating convergence of calculations. Sevoflurane PMF is shown for comparison. (c) Atomic density of TCE as computed from PMF in (b). (d) Atomic density profiles for water and membrane moieties. (e) Pair correlation between TCE and lipid headgroups. The first peak at 1.5 Å configures a typical hydrogen bond. The inset shows the molecular view of TCE binding to membrane headgroups via hydrogen bonds. Molecular images were rendered using VMD.<sup>19</sup>

Because quenching of W134 and W213 suggests two BSA binding sites for TCE, here labeled as s1 and s2, molecular docking and free energy perturbation (FEP) calculations were performed with the aim at reproducing experimental data (cf. Computational Methods) (Figure 5). All measures (calculated



**Figure 5.** TCE binding sites to BSA. (a) Ensemble-average structure of BSA (cartoon), along with tryptophan residues (yellow licorice) and set of TCE centroid configurations (orange points) determined from docking searches. (b, c) Atomistic details of TCE within binding sites s2 and s1, respectively, composed of residues K20, V40, K131, W134, G135 and L197, W213, S343, L346, S453, L480, V481. Site residues are represented by surfaces and colored by physical and chemical properties. Molecular images were rendered using VMD.<sup>19</sup>

and experimental) are reported in Table 3. The progression of free energy changes as a function of  $\lambda$  parameter for BSA binding sites during the course of five successive coupling/decoupling simulations is shown in Figure 6. Curves are well-behaved, which indicates that the time scale was small enough to correlate multiple homologous windows. As shown in Table 3, our predicted dissociation constants for ligand binding at

sites s1 and s2 are in the same order of magnitude as that of the experimental estimates. Besides, the predicted constants also agree qualitatively with measurements as the experimentally determined higher affinity of TCE to site s1 is recapitulated in the calculations.

Sites s1 and s2 are both buried in the protein and partially accessible to solvent. Independent MD simulations of the protein with bound ligands at sites s1 and s2 show that TCE remains confined to the binding sites, hydrated by few water molecules and in close contact with protein amino acids including W134 and W213 (Figure 7). Despite these structural similarities, TCE binds site s1 with a higher affinity. Given TCE's favorable interactions with lipids (Figure 4), the predominant hydrophobic nature of site s1, as highlighted in Figure 5, makes sense of the result.

According to eq 7, we also calculated the standard binding free energy corresponding to both simultaneously bounded states s1 and s2 (unique saturation):  $\Delta G_{s1,s2}^{\circ} = -5.45$  kcal/mol. This value, comparable to the standard binding free energy for POPC ( $\Delta G_{\text{POPC}}^{\circ} = -4.32 \pm 0.69$  kcal/mol), evidences TCE's greater affinity for the receptor rather than for the membrane. By these results, TCE is expected to bind protein targets with higher affinities, opening the possibility that it may bind ion channels when partitioning the membrane.

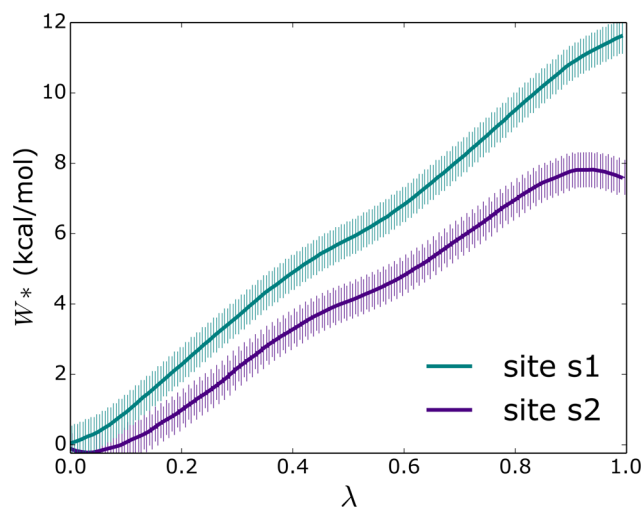
## CONCLUSIONS

A fine atomistic CHARMM36-compatible model for TCE, a sedative hypnotic drug, is presented in this article. Model development targets gas-phase conformations and molecular electrostatic potential with individual TIP3P water molecules via weak hydrogen bonding. Validation is ensured, as developed parameters appropriately reproduce ligand's physical and chemical properties, including liquid bulk properties, lipid partitioning, and interaction to protein targets.

**Table 3. Free Energy Changes for BSA Binding Sites s1 and s2 and Their Respective Dissociation Constants<sup>a</sup>**

	$k$	$W^*(\text{SOS})$	$\bar{\mu}(\text{SOS})$	$\Delta G_{\text{BSA}}^{\circ}$	$K_{\text{D}}(\text{calc})$	$K_{\text{D}}(\text{exp})$
s1	1.078	$-11.60 \pm 0.003$	$-4.83 \pm 0.004$	$-6.77 \pm 0.007$	$2.79^c$	$2.1 \pm 0.1^e$
s2	0.013	$-7.60 \pm 0.007$	$-4.83 \pm 0.004$	$-2.77 \pm 0.010$	$35.82^c$	$12.0^e$
s1, s2				$-5.78^b$	$99.94^d$	

<sup>a</sup>Units for  $k$ ,  $W^*$ ,  $\bar{\mu}$ ,  $\Delta G_{\text{BSA}}^{\circ}$ , and  $K_{\text{D}}$  are kcal/mol/Å<sup>2</sup>, kcal/mol, kcal/mol, kcal/mol, and mmol/L, respectively. <sup>b</sup>Calculated standard free energy of binding a single ligand to both binding sites (independent events). <sup>c</sup>Calculated dissociation constant values. <sup>d</sup>Calculated aggregate dissociation constant value. <sup>e</sup>Experimental dissociation constant values from ref 13.



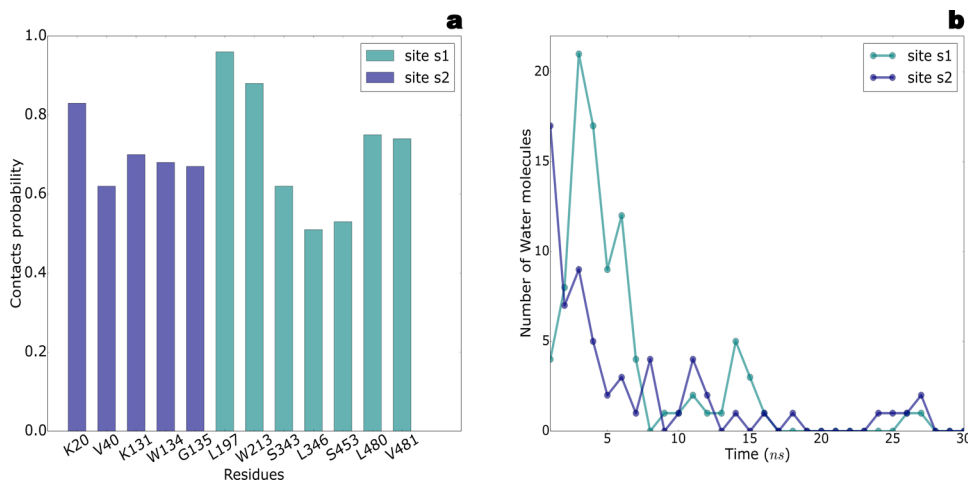
**Figure 6.** Computed free energy changes  $W^*$  as a function of decoupling parameter  $\lambda$  for BSA binding sites. Error bars were calculated using the simple overlap sampling (SOS) method.<sup>18</sup>

Overall, our results support that the presented TCE force-field parameters are robust and likely to be useful in a large series of *in silico* biological studies. TCE is known experimentally to interact favorably with albumin and phospholipids as judged respectively from its dissociation constant and water–oil partition coefficient. Figures 4 and 5 not only make sense of these experimental facts but also provide a detailed molecular model for ligand interaction to these substrates. Such a model is particularly useful to interpret and design novel experiments for characterization of ligand binding to membrane and albumin with potential consequen-

ces for our understanding of TCE action in membrane-embedded proteins. In this regard, our contribution opens the possibility to explore novel problems regarding the interaction of small ligands and ion channels, a field of immense interest.<sup>22</sup> Particular attention might be driven to K2P channels as the scientific literature still lacks information about their modulation mechanism by TCE and related molecules.

## COMPUTATIONAL METHODS

**Parametrization Adjustment.** The MATCH atom-typing toolset was used to set initial atom parameters for TCE.<sup>23</sup> The optimization protocol proceeded in stages: acquisition of quantum mechanics (QM) and molecular mechanics (MM) data, comparison of conformational properties, and refinements. Distributed as a VMD plugin,<sup>19</sup> the force field toolkit<sup>20</sup> contains helpful scoring algorithms to fit MM to QM data<sup>24</sup> that allowed determination of accurate TCE parameters. In detail, the TCE equilibrium geometry at the QM level was defined as the molecular target. Next, atomic charge distributions were determined on the basis of QM interactions with individual TIP3P water molecules. For each donor or acceptor hydrogen bond, a typical linear geometry was built and then bond distances were optimized keeping fixed all other degrees of freedom. Ligand–water interaction energies were scaled by a factor of 1.16, whereas hydrogen bond lengths were shifted by an offset of  $-0.2$  Å to yield appropriate bulk phase parameters.<sup>25,26</sup> QM partial atomic charges for the equilibrium geometry were used as a trial set for the charge optimization procedure. The objective function was then optimized on the basis of these QM interaction energies for TCE–water complexes until convergence was reached. During this iterative



**Figure 7.** Characterization of sites s1 and s2. Per-site amino acid contacts (a) and number of water molecules (b) within 3 Å of bound TCE. Data was computed over independent equilibrium MD simulations (30 ns) of the ligand–protein bound state.

process, the QM dipole was allowed to be overestimated by 20–50% to reach consistency with the bulk phase.<sup>27</sup>

Bond and valence angle parameters were optimized by computing the energetic perturbation of small distortions from the equilibrium QM geometry along redundant internal coordinates. Optimization proceeded until convergence of bond and angle objective functions. Torsional parameters were obtained by fitting QM potential energy surfaces (PESs) for TCE dihedrals Cl–C2–C1–O and H–O–C1–C2. Torsion scans at the QM level were conducted by rotating dihedrals from  $-180$  to  $+180^\circ$  with a step size of  $5^\circ$ , providing an energy function in the form

$$V(\phi) = \sum_{k=1}^N V_k [1 + \cos(k\phi - \delta_k)] \quad (1)$$

where  $V_k$  is the fitted barrier height (force constant), and  $k$  and  $\delta_k$  are dihedral multiplicity and phase, respectively. Parameters were adjusted to fit QM PES until convergence of the root-mean-square deviation. Finally, Lennard-Jones (LJ) parameters used to treat van der Waals interactions were assigned by transferability of similar chemical types available at the CHARMM general force field (CGenFF). After running a full optimization cycle, the geometry was minimized via a 1000-step gradient to compare the developed model to the QM target. As a standard practice, TCE–water interactions were reoptimized using this MM geometry, whereas Hessian and PES remained unmodified and a new round of parametrization was performed to ensure self-consistency. Further rounds were not required because no significant improvement was obtained. All MM calculations were performed using NAMD 2.10.<sup>28</sup>

**QM Calculations.** Molecular geometry was optimized at the MP2/6-31G(d) level, and initial partial atomic charges were derived from MP2/6-31G(d) Merz–Kollman charges.<sup>29</sup> Water interaction profiles were optimized at the HF/6-31G(d) level. Although higher levels of QM theory may lead to more accurate geometries and hydrogen bond energies, the chosen level of theory for water interaction profiles maintains consistency with the CHARMM additive force field.<sup>14,27</sup> The QM Hessian matrix and the relaxed PES were also calculated at the MP2/6-31G(d) level. Constraints were imposed to the scanned dihedrals during molecule minimization. All electronic structure calculations were performed in Gaussian 09.<sup>15</sup>

**MD Simulations.** Condensed phase simulations were performed in an NPT ensemble at 298.15 K and 1 bar using Langevin dynamics and the Langevin piston algorithm as implemented in NAMD 2.10.<sup>28</sup> Periodic boundary conditions were applied to all MD simulations. The particle mesh Ewald method was employed to evaluate full electrostatics using a real-space grid spacing of 1.2 Å or less. The multiple-time-step r-RESPA integrator was used, with a base time step of 2 fs for short-range nonbonded forces and extended time of 4 fs for soft, long-range interactions. Pairwise nonbonded interactions were truncated at a distance of 12 Å with a smooth switching function above (10 Å).

**Pure Solvent Properties.** The starting configuration for liquid phase simulations consisted of 216 TCE molecules placed with random orientation at the grid points of a cubic lattice ( $6.5 \times 6.5 \times 6.5 \text{ \AA}^3$ ). A 10 000-step minimization followed by gradual heating during 10 ps was applied to the whole system for equilibration. Equilibrium data was collected

for 0.4 ns to compute the average periodic cell volume to estimate TCE density.

Under the ideal-gas assumption and negligible mechanical work in the liquid phase,<sup>14</sup> the enthalpy of vaporization at a given temperature  $T$  can be written as

$$\Delta_{\text{vap}}H \approx -\frac{\langle U_{\text{liquid}} \rangle}{N} + \langle U_{\text{gas}} \rangle + RT \quad (2)$$

where  $\frac{\langle U_{\text{liquid}} \rangle}{N}$  and  $\langle U_{\text{gas}} \rangle$  correspond, respectively, to the average potential energy of the molecule in liquid and gas phases. First,  $\frac{\langle U_{\text{liquid}} \rangle}{N}$  was estimated on the basis of the liquid phase simulation of  $N = 216$  molecules of TCE. Then, each molecule's final configuration in the liquid phase was isolated and subsequently simulated in gas phase. These individual simulations comprised 50 ps of equilibration with a friction coefficient of  $5 \text{ ps}^{-1}$  followed by 50 ps of data collection. Finally,  $\langle U_{\text{gas}} \rangle$  was obtained by averaging over the potential energy relative to  $N = 216$  configurations.

**Solvation Properties.** TCE solvation free energies in bulk water and oil phase were evaluated using TIP3P water and hexane molecules, respectively.<sup>30</sup> Both solvent models are available in CHARMM36. Solvation free energies were computed via free energy perturbation calculations (FEP).<sup>31</sup> In these alchemical transformations, the solute is decoupled from the environment by turning its intermolecular interactions off, using a scalar parameter  $\lambda$ . A soft-core potential was adopted to scale the nonbonded pair potential of the perturbed system according to a dual coupling parameter (van der Waals and electrostatics).<sup>32</sup> Alchemical transformations were split into 100 “windows” and carried out explicitly in both directions, i.e., decoupling (integration over  $\lambda$  from 0 to 1) and recoupling simulations (integration over  $\lambda$  from 1 to 0). Each window contains 4 ps of relaxation followed by 66 ps of data collection. The soft-core potential shift distance was set to  $7.0 \text{ \AA}^2$  for water and  $5.5 \text{ \AA}^2$  for hexane, after several optimizations by trial and error. In total, five independent simulations (replicas) were performed with different initial velocities and seeds for the stochastic term in the Langevin thermostat. Statistical errors related to solvation free energies were computed with the simple overlap sampling (SOS) algorithm.<sup>18</sup>

**Partitioning to Lipid Bilayer.** A TCE molecule was inserted in a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) bilayer model, available in CHARMM36. A lamella with dimensions of approximately  $45.0 \times 43.0 \times 89.0 \text{ \AA}^3$  was built using 59 lipids. The POPC bilayer was hydrated by 3171 TIP3P water molecules. Free energy profile or potential of mean force (PMF) for TCE to cross the membrane was computed using the adaptive biasing force algorithm (ABF) as implemented in NAMD.<sup>33,34</sup> In the method, the reaction coordinate to be sampled  $\xi$  is discretized in  $i$  bins. PMF is then derived from a biasing force added to the system's equations of motion to counteract average forces acting on each bin. Here, the reaction coordinate  $\xi$ , with bin size  $0.1 \text{ \AA}$ , and defined as the distance  $z$  between TCE and the center of POPC lamella, was sampled over 300 ns simulation. The accumulated statistical error throughout the PMF was calculated according to the equation<sup>35</sup>

$$\text{Err}[\Delta G_{a \rightarrow b}] = \delta \xi \left( \sum_{i=i_a}^{i_b} \frac{\tau_i}{n_i \Delta t} \langle \Delta F_{\xi}^2 \rangle \right)^{1/2} \quad (3)$$

where  $i_a$  and  $i_b$  are bin indices delimiting the  $\xi$ -interval  $[a, b]$ ;  $\Delta t$  is the simulation time step; and, for each bin,  $n_i$ ,  $\tau_i$ , and  $\langle \Delta F_{\xi}^2 \rangle$ , respectively, stand for the number of samples accrued, autocorrelation time, and variance of the instantaneous force  $F_{\xi}$ .

The partition coefficient  $K$  was computed from the PMF following equation<sup>36</sup>

$$K = \frac{1}{|a - b|} \int_a^b e^{-\Delta G/RT} d\xi \quad (4)$$

where  $|a - b|$  is the system width from bulk to the membrane center, along  $z$ . Partition coefficient statistical uncertainty was estimated by propagating the average bin error  $\text{Err}[\Delta G_{a \rightarrow b}]/i$ .

**Binding Affinity to BSA.** Fluorescence quenching experiments suggest that TCE binds to bovine serum albumin (BSA) at two distinct sites, namely, in the vicinity of tryptophans 134 and 213.<sup>13</sup> To further characterize and validate the TCE model, molecular docking and free energy perturbation (FEP) calculations of TCE against BSA were performed and compared with experimental data.

AutoDock Vina software was used to dock the ligand into an ensemble of BSA equilibrium structures to allow for sampling of the protein's configurational space.<sup>37</sup> Binding sites were delimited from the ensemble of docking poses in such a way that each binding site was defined as the effective volume outlined by protein residues with the largest number of contacts to docking poses. In this manner, site s1 is defined by residues L197, W213, S343, L346, S453, L480, and V481 and site s2 by residues K20, V40, K131, W134, and G135 (Figure 5). Although docking was performed in vacuum, subsequent equilibrium and FEP simulations were conducted in the presence of explicit all-atom water molecules, thus taking environmental effects into account. Specifically, the simulated system comprised the protein receptor (BSA) embedded in a homogeneous reservoir, with diluted ligand (TCE) concentration. The protein is assumed to be in a well-defined conformational state in which it provides two distinct TCE binding sites, s1 and s2. Hence, the equilibrium constant for the process of bringing the ligand from the bulk into the bound state (s1, s2) can be solved by means of an MD/FEP approach.<sup>38</sup> The method requires a harmonic potential coupled to the ligand

$$u^*(\mathbf{R}) = \frac{1}{2} k [\mathbf{R} - \mathbf{R}^*]^2 \quad (5)$$

to guarantee that it is restrained to occupy the effective volume centered at the receptor binding site  $\mathbf{R}^*$ ; the restraint  $k$  exerts no force when the ligand center of mass is within a distance  $\mathbf{R}$  from the equilibrium position and exerts a harmonic restoring force when the ligand center of mass is out of this range. In our case, force constants were estimated from the root-mean-square fluctuation of docking solutions. The equilibrium binding constant for each individual site  $j$ ,  $j \in (s1, s2)$ , can be written as

$$K_j = \left( \frac{2\pi}{\beta k_j} \right)^{3/2} e^{-\beta[W_j^*(\mathbf{R}) - \bar{\mu}]} \quad (6)$$

By definition,  $W_j^*(\mathbf{R})$  is the reversible work to transfer a single ligand from the gas phase to the respective bound state and  $\bar{\mu}$  refers to its solvation free energy. Thereon, assuming that the binding sites are independent, an aggregate affinity constant, for both binding sites to be simultaneously occupied by a single ligand each, can be defined as the product of the individual constants<sup>38</sup>

$$K_{s1,s2} = K_{s1} \times K_{s2} \quad (7)$$

From eq 6, the standard binding free energy relative to bringing a single ligand from a reference standard reservoir concentration into the respective target site  $j$  can be defined as

$$\Delta G_j^\circ = -\beta^{-1} \ln[K_j \times (C^\circ)] \quad (8)$$

where  $C^\circ = 1$  M or in units of number density  $C^\circ = (1660 \text{ \AA}^3)^{-1}$ . Finally, from eq 7, we can derive the standard binding free energy relative to bringing two indistinguishable ligands from a reference standard reservoir concentration into simultaneously occupied binding sites s1 and s2

$$\Delta G_{s1,s2}^\circ = -\beta^{-1} \ln[K_{s1,s2} \times (C^\circ)^2] \quad (9)$$

To compute binding affinities described above, alchemical transformations were conducted by decoupling and recoupling the TCE gauche conformer from each BSA binding site separately. Starting from protein–TCE equilibrated systems as resolved from docking, decoupling/recoupling was carried out by varying the  $\lambda$  coupling parameter in steps of 0.01, amounting to 100 windows per transformation. Transformations in each direction spanned 2 ps of relaxation followed by 6.4 ns of data collection, thus totaling 12.8 ns of collection per replica, per site. For the purpose of improving statistics, site-specific FEP estimates and the associated statistical errors were determined from five independent decoupling/recoupling runs. Restraints on the ligand's center of the mass were included in the calculations to ensure appropriate sampling within the binding site so that TCE's chemical potential remains well defined during the final decoupling or initial recoupling stages.<sup>39,40</sup>

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.8b02017.

Empirical geometry data for the TCE gauche conformer as well as QM target deviations and reference data, partial atomic charges, LJ terms, and bonded parameters in the CHARMM force field format (PDF)

## ■ AUTHOR INFORMATION

### ✉ Corresponding Author

\*E-mail: treptow@unb.br.

### ORCID

Werner Treptow: 0000-0003-4564-3205

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The research was supported in part by the Brazilian Agencies CNPq, CAPES, and FAPDF under Grants 305008/2015-3, 23038.010052/2013-95, and 193.001.202/2016.



## ■ ADDITIONAL NOTE

<sup>a</sup>Theoretical value available at ACS. Calculated using Advanced Chemistry Development (ACD/Labs) software V11.02 (1994-2018 ACD/Labs).

## ■ REFERENCES

- (1) Liebig, J. Ueber die Zersetzung des Alkohols durch Chlor [On the degradation of alcohol by chlorine]. *Ann. Pharm.* **1832**, *1*, 31–32.
- (2) D'Agostino, J.; Terndrup, T. E. Chloral hydrate versus midazolam for sedation of children for neuroimaging: a randomized clinical trial. *Pediatr. Emerg. Care* **2000**, *16*, 1–4.
- (3) Roach, C. L.; Husain, N.; Zabinsky, J.; Welch, E.; Garg, R. Moderate sedation for echocardiography of preschoolers. *Pediatr. Cardiol.* **2010**, *31*, 469–473.
- (4) Schmalfuss, I. Oral sedation of pediatric patients for noninvasive radiological procedures: chloral hydrate versus midazolam. *J. Radiol. Nurs.* **2005**, *24*, 42–48.
- (5) Hare, M. Question 1. Chloral hydrate or midazolam: which is better for sedating children for painless diagnostic imaging? *Arch. Dis. Child.* **2012**, *97*, 750–752.
- (6) Hill, G. D.; Walbergh, D. B.; Frommelt, P. C. Efficacy of reconstituted oral chloral hydrate from crystals for echocardiography sedation. *J. Am. Soc. Echocardiogr.* **2016**, *29*, 337–340.
- (7) Avlonitou, E.; Balatsouras, D. G.; Margaritis, E.; Giannakopoulos, P.; Douniadakis, D.; Taskanikos, M. Use of chloral hydrate as a sedative for auditory brainstem response testing in a pediatric population. *Int. J. Pediatr. Otorhinolaryngol.* **2011**, *75*, 760–3.
- (8) Wheeler, D. S.; Jensen, R. A.; Poss, W. B. A randomized, blinded comparison of chloral hydrate and midazolam sedation in children undergoing echocardiography. *Clin. Pediatr.* **2001**, *40*, 381–387.
- (9) Meir, K.; Thomas, R.; Dement, W. *Principles and Practice of Sleep Medicine*, 5th ed.; Elsevier, 2010.
- (10) Lovinger, D. M.; Zimmerman, S. A.; Levitin, M.; Jones, M. V.; Harrison, N. L. Trichloroethanol potentiates synaptic transmission mediated by gamma-aminobutyric acidA receptors in hippocampal neurons. *J. Pharmacol. Exp. Ther.* **1993**, *264*, 1097–1103.
- (11) Peoples, R. W.; Weigh, F. F. Inhibition of excitatory amino acid-activated currents by trichloroethanol and trifluoroethanol in mouse hippocampal. *Br. J. Pharmacol.* **1998**, *124*, 1159–1164.
- (12) Harinath, S.; Sikdar, S. K. Trichloroethanol enhances the activity of recombinant human TREK-1 and TRAAK channels. *Neuropharmacology* **2004**, *46*, 750–760.
- (13) Solt, K.; Johansson, J. S. Binding of the Active Metabolite of Chloral Hydrate, 2,2,2-Trichloroethanol, to Serum Albumin Demonstrated Using Tryptophan Fluorescence Quenching. *Pharmacology* **2002**, *64*, 152–159.
- (14) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; S. Zhong, J. S.; Darian, E.; Guvench, O.; Lopes, P.; et al. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671–690.
- (15) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; R. C. J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision D.01; Gaussian Inc.: Wallingford, CT, 2009.
- (16) Takamuku, T.; Kumai, T.; Yoshida, K.; Otomo, T.; Yamaguchi, T. Structure and Dynamics of Halogenoethanol Water Mixtures Studied by Large-Angle X-ray Scattering, Small-Angle Neutron Scattering, and NMR Relaxation. *J. Phys. Chem. A.* **2005**, *109*, 7667–76.
- (17) Eftink, M. R.; Zajicek, J. L.; Ghiron, C. A. A hydrophobic quencher of protein fluorescence: 2,2,2-trichloroethanol. *Biochim. Biophys. Acta* **1977**, *491*, 473–481.
- (18) Lu, N.; Singh, J. K.; Kofke, D. A. Appropriate methods to combine forward and reverse free-energy perturbation averages. *J. Chem. Phys.* **2003**, *118*, 2977–2984.
- (19) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (20) Arcario, M. J.; Mayne, C. G.; Tajkhorshid, E. Atomistic Models of General Anesthetics for Use in Silico Biological Studies. *J. Phys. Chem. B* **2014**, *118*, 12075–12086.
- (21) Barber, A. F.; Carnevale, V.; Klein, M. L.; Eckenhoff, R. G.; Covarrubias, M. Modulation of a voltage-gated Na<sup>+</sup> channel by sevoflurane involves multiple sites and distinct mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 6726–6731.
- (22) Covarrubias, M.; Barber, A. F.; Carnevale, V.; Treptow, W.; Eckenhoff, R. G. Mechanistic Insights into Modulation of Voltage-Gated Ion Channels by Inhalational Anesthetics. *Biophys. J.* **2015**, *109*, 2003–2011.
- (23) Yesselman, J. D.; Price, D. J.; Knight, J. L.; Brooks, C. L., 3rd. MATCH: an atom-typing toolset for molecular mechanics force fields. *J. Comput. Chem.* **2012**, *33*, 189–202.
- (24) Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C. Rapid Parameterization of Small Molecules Using the Force Field Toolkit. *J. Comput. Chem.* **2013**, *34*, 2757–2770.
- (25) MacKerell, A. D.; Karplus, M. Importance of attractive van der Waals contribution in empirical energy function models for the heat of vaporization of polar liquids. *J. Phys. Chem.* **1991**, *95*, 10559–10560.
- (26) Jorgensen, W. L. Optimized intermolecular potential functions for liquid alcohols. *J. Phys. Chem.* **1986**, *90*, 1276–1284.
- (27) MacKerell, A. D.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (28) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (29) Singh, U. C.; Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, *5*, 129–145.
- (30) Pohorille, A. L.; Wilson, M. A.; Chipot, C. Interactions of alcohols and anesthetics with the water-hexane interface. *Prog. Colloid Polym. Sci.* **1997**, *103*, 29–40.
- (31) *Free Energy Calculations - Theory and Applications in Chemistry and Biology*; Chipot, C., Pohorille, A., Eds.; Springer Series in Chemical Physics; Springer-Verlag Berlin: Heidelberg, 2007; Vol. 86.
- (32) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J. Chem. Phys.* **1994**, *100*, 9025–9031.
- (33) Darve, E.; Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.* **2001**, *115*, 9169–9183.
- (34) Darve, E.; Rodríguez-Gómez, D.; Pohorille, A. Adaptive biasing force method for scalar and vector free energy calculations. *J. Chem. Phys.* **2008**, *128*, No. 144120.
- (35) Comer, J.; Gumbart, J. C.; Hémin, J.; Lelièvre, T.; Pohorille, A.; Chipot, C. The Adaptive Biasing Force Method: Everything You Always Wanted To Know but Were Afraid To Ask. *J. Phys. Chem. B* **2015**, *119*, 1129–1151.
- (36) MacCallum, J. L.; Tieleman, D. P. Computer simulation of distribution of hexane in a lipid bilayer: spatially resolved free energy, entropy and enthalpy profiles. *J. Am. Chem. Soc.* **2006**, *128*, 125–130.
- (37) Trott, O.; Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.
- (38) Stock, L.; Hosoume, J.; Treptow, W. Concentration-Dependent Binding of Small Ligands to Multiple Saturable Sites in Membrane Protein. *Sci. Rep.* **2017**, *7*, No. 5734.
- (39) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* **1997**, *72*, 1047–1069.
- (40) Roux, B.; Nina, M.; Pomès, R.; Smith, J. C. Thermodynamic stability for water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study. *Biophys. J.* **1996**, *71*, 670–681.

RESEARCH ARTICLE

# Binding of the general anesthetic sevoflurane to ion channels

Leticia Stock, Juliana Hosoume , Leonardo Cirqueira , Werner Treptow \*

Laboratório de Biologia Teórica e Computacional (LBTC), Universidade de Brasília DF, Brasil

\* [treptow@unb.br](mailto:treptow@unb.br)



 OPEN ACCESS

**Citation:** Stock L, Hosoume J, Cirqueira L, Treptow W (2018) Binding of the general anesthetic sevoflurane to ion channels. *PLoS Comput Biol* 14(11): e1006605. <https://doi.org/10.1371/journal.pcbi.1006605>

**Editor:** Rebecca C. Wade, Heidelberg Institute for Theoretical Studies (HITS gGmbH), GERMANY

**Received:** August 29, 2018

**Accepted:** October 26, 2018

**Published:** November 26, 2018

**Copyright:** © 2018 Stock et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** This work was supported by National Council for Scientific and Technological Development CNPq [grant number 305008/2015-3 (WT)], Coordenação de Aperfeiçoamento de Pessoal de Nível Superior CAPES [grant number 23038.010052/2013-95 (WT)], and Fundação de Apoio a Pesquisa do Distrito Federal FAPDF [grant number 193.001.202/2016 (WT)]. Doctoral fellowship was granted by National Council for

## Abstract

The direct-site hypothesis assumes general anesthetics bind ion channels to impact protein equilibrium and function, inducing anesthesia. Despite advancements in the field, a first principle all-atom demonstration of this structure-function premise is still missing. We focus on the clinically used sevoflurane interaction to anesthetic-sensitive Kv1.2 mammalian channel to resolve if sevoflurane binds protein's well-characterized open and closed structures in a conformation-dependent manner to shift channel equilibrium. We employ an innovative approach relying on extensive docking calculations and free-energy perturbation of all potential binding sites revealed by the latter, and find sevoflurane binds open and closed structures at multiple sites under complex saturation and concentration effects. Results point to a non-trivial interplay of site and conformation-dependent modes of action involving distinct binding sites that increase channel open-probability at diluted ligand concentrations. Given the challenge in exploring more complex processes potentially impacting channel-anesthetic interaction, the result is revealing as it demonstrates the process of multiple anesthetic binding events alone may account for open-probability shifts recorded in measurements.

## Author summary

General anesthetics are central to modern medicine, yet their microscopic mechanism of action is still unknown. Here, we demonstrate that a clinically used anesthetic, sevoflurane, binds the mammalian voltage-gated potassium channel Kv1.2 effecting a shift in its open probability, even at low concentrations. The results, supported by recent experimental measurements, are promising as they demonstrate that the molecular process of direct binding of anesthetic to ion channels play a relevant role in anesthesia.

## Introduction

Volatile and injected general anesthetics encompass a diverse array of small and uncharged chemotypes including haloalkanes, haloethers and alkylphenols. Despite efforts reaching back over a century, clarification of their microscopic mechanism in general anesthesia has proven difficult and wanting. A favored hypothesis proposes that ion channels in the brain are

Scientific and Technological Development CNPq [grant number 140845/2014-3 (LS)]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

implicated, among which members of ionotropic neurotransmitter receptors, voltage-gated and non-gated ion channels are best-known players [1–3]. Primary exemplars are the Cys-loop nicotinic acetylcholine and  $\gamma$ -aminobutyric acid class A receptors, the voltage-gated sodium and potassium channels, and the tandem pore potassium channels. An extensive series of electrophysiological studies corroborate the hypothesis by demonstrating a range of effects, from inhibition to potentiation, of general anesthetics on the various receptor targets. Beyond these electrophysiological studies of reductionist systems, the current view has gained additional support from gene knockout experiments demonstrating for some of these channels the *in vivo* role on a clinically-relevant anesthetic outcome. For instance, the knockout of the non-gated tandem pore potassium channel *trek-1* produces an animal model (*Trek1<sup>-/-</sup>*) resistant to anesthesia by inhalational anesthetics [4].

How general anesthetics modulate ion channels to account for endpoints of anesthesia must at some point build on understanding electrophysiological data in the context of ligand binding, a reasoning that has driven mounting efforts in the field. Currently, though not refuting other molecular processes likely to contribute to anesthetic action [5–7], crystallography and molecular dynamics studies support that anesthetics bind ion channels at clinical concentrations [8–16]. Binding interactions have been evidenced in anesthetic containing systems of mammalian voltage and ligand-gated channels, as well as bacterial channel analogs. Specifically, partitioning of anesthetics in the membrane core allows it to access and bind multiple transmembrane (TM) protein sites, featuring single or multiple occupancy states—a process that might depend further on chemotypes, channel types and conformations. Although some progress has been made in validating one or more aspects of the direct-site hypothesis, a first-principle demonstration that anesthetics bind ion channels to affect protein equilibrium and function as recorded in experiments is still unaccounted for.

Here, we focus our efforts on the haloether sevoflurane and its molecular interaction to Kv1.2, a mammalian voltage-gated potassium channel. Experimental work demonstrates that sevoflurane potentiates the channel in a dose-dependent manner [3,17,18]. Effects on current tracings include a leftward shift in the channel's conductance-voltage relationship and an increased maximum conductance. As extensively discussed in these past publications, at least two molecular mechanisms are expected to be involved in Kv channels potentiation by sevoflurane. One mechanism (i) might involve sites allosterically coupled to the electromechanical transduction directly responsible for controlling voltage-dependent gating. The other (ii) might involve distinct sites, which could modulate the channel's pore region and influence the stability of the conductive state and/or the unitary conductance. Here, we are interested in the investigation of mechanism (i) and its underlying structural hypothesis that sevoflurane binds the channel's open and closed states to impact protein equilibrium and therefore its voltage dependence. Among all other aspects that might impact channel-anesthetic interactions in general, our specific goal is to determine if sevoflurane binds the well-characterized open-conductive (O) and resting-closed (C) structures of Kv1.2 [19,20] in a conformation-dependent manner to impact its voltage-dependent open probability as recorded experimentally. Very recently, we have put forth an innovative structure-based study [21] dealing with the concentration-dependent binding of small ligands to multiple saturable sites in proteins to show that sevoflurane binds the open-pore structure of Kv1.2 at the S4S5 linker and the S6P-helix interface—a result largely supported by independent photolabeling experiments [22,23]. To our current goal, we aim therefore at extending these calculations to investigate sevoflurane interactions with the entire channel TM-domain and, more importantly, to resolve any conformational dependence in its binding process to channel structures C and O. In the following sections, we first provide the theoretical framework to study sevoflurane binding to a specific channel conformation under equilibrium conditions. A state-dependent strategy is put

forward to describe anesthetic binding in terms of occupancy states of all identified potential channel binding sites, embodying both concentration and multiple sites saturation effects. The strategy is then generalized to account for ligand impact in the C-O equilibrium, allowing for reconstruction of voltage-dependent open probabilities of the channel at various ligand concentrations. Anticipating our results, we find that sevoflurane binds Kv1.2 structures at multiple sites under saturation and concentration effects. Despite a similar pattern of molecular interactions, binding of sevoflurane is primarily driven towards the open-conductive state shifting leftward the open probability of the channel at diluted ligand concentrations.

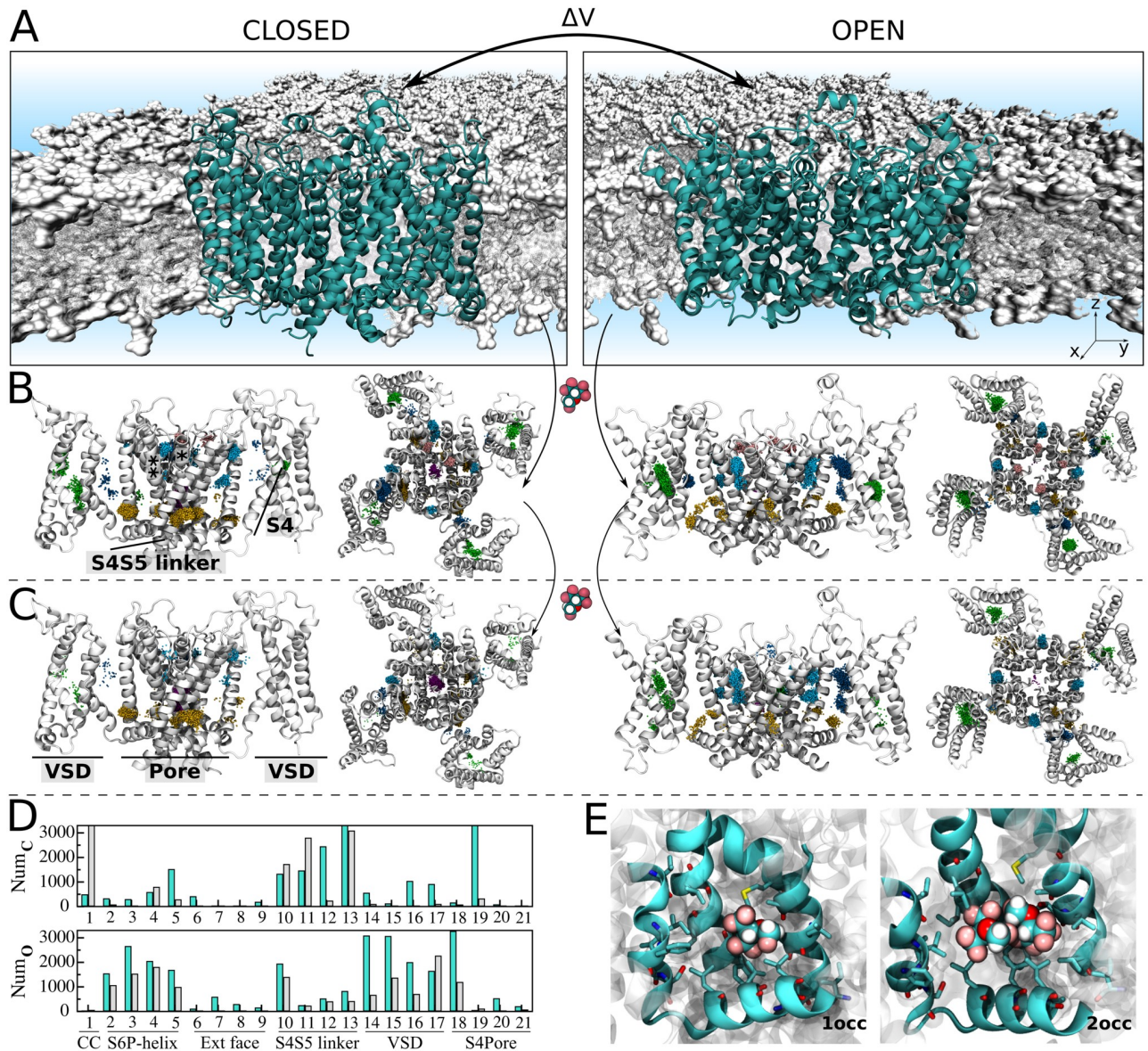
## Results

### Binding of anesthetics to multiple channel sites

We applied large-scale and flexible docking calculations to solve sevoflurane interactions to Kv1.2 structures  $X \equiv \{C, O\}$  (Fig 1). A total of ~ 6,000 docking solutions was generated per channel conformation and clustered into 21 ligand interaction sites. The interaction sites spread over the transmembrane region of the channel at the S4S5 linker, S6P-helix interface and at the extracellular face, next to the selectivity filter. Further docking sites were resolved within the voltage-sensor, at the S4Pore interface and within the channel central cavity. Redocking of sevoflurane generated in turn a total of ~ 13,000 solutions per channel conformation, solving the interaction of two ligands for all sites but the extracellular face.

From the docking ensembles, there are up to  $2 \times 3^{21}$  channel occupancy states that might contribute to sevoflurane binding and functional effects. To quantitatively evaluate this, we performed an extensive series of decoupling FEP calculations to estimate the per-site binding affinity for one and two bound ligands against the channel structures (*cf.* Computational Methods for details). Here, FEP calculations started from equilibrium ligand-bound channel structures, embedded in an explicit water-membrane environment (*cf.* RMSD analysis in S1 Fig). For the purpose of improving statistics, FEP estimates and the associated statistical errors were determined from at least two independent decoupling runs. Calculations were performed over ~ 7.0 ns per replica, per site, per conformation, to converge FEP estimates; in a total MD simulation time of ~ 2.0  $\mu$ s. S2 Fig shows the effectively sampled configuration space in FEP calculations for each of the channel structures. Systematic errors related to lack of site rehydration or relipidation during ligand decoupling were ruled out in S3 Fig showing equilibrium-like lipid or water coordination numbers of the channel structure at the final stages of FEP. Under these technical details, per-site equilibrium binding constants were quantified relative to a homogeneous and diluted aqueous solution occupied by ligands, with an excess chemical potential of  $\bar{\mu} = 0.10 \pm 0.09 \text{kcal.mol}^{-1}$ . As shown in S1 and S2 Tables, per-site binding constants are heterogeneous and take place over a diverse range, *i.e.*  $10^{-8} \text{(mM}^{-1})$  -  $10^{+2} \text{(mM}^{-2})$ . There is however a decreasing trend of affinities involving sites respectively at the S4S5 linker, S4Pore and S6P-helix interfaces, voltage sensor, central cavity and extracellular face.

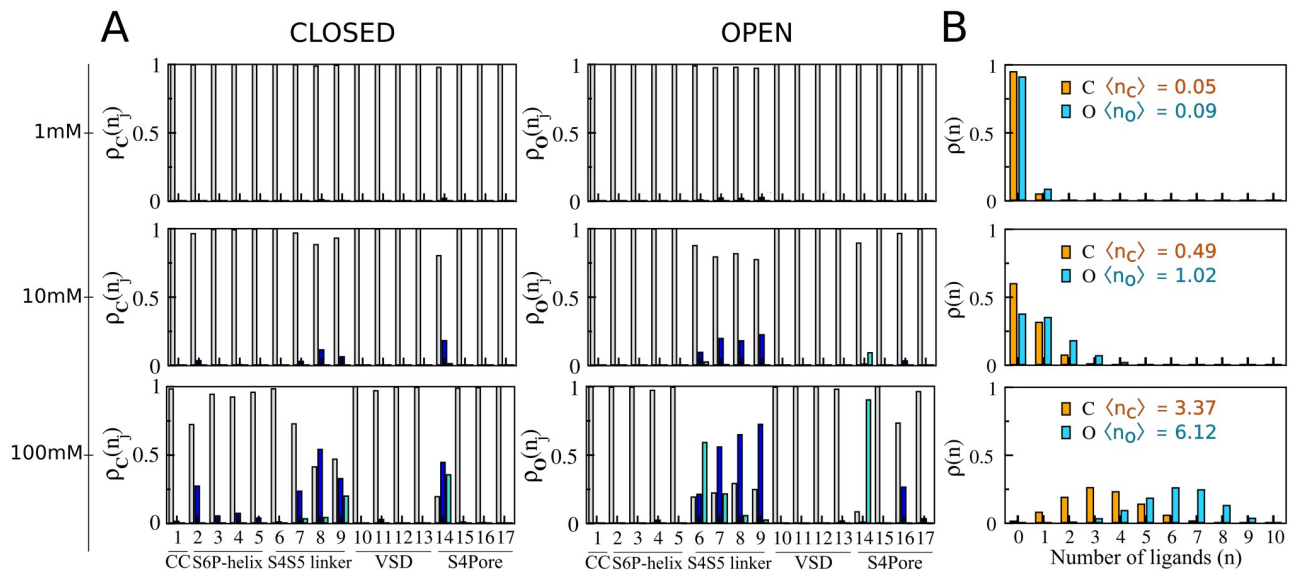
To determine if sevoflurane binds channel structures  $X \equiv \{C, O\}$  at clinically relevant concentrations, we computed binding probabilities  $\rho_X(n_1, \dots, n_s)$  for dilute concentrations of the ligand in solution, *i.e.* 1mM, 10mM and 100mM. Equilibrium constants  $K_X(n_1, \dots, n_s)$  for every occupancy state of the channel were then reconstructed from the per-site affinities to determine state probabilities via eq (2). Here, estimates of  $K_X(n_1, \dots, n_s)$  were determined for the condition of independent binding sites, as the minimum site-to-site distances of ~15 Å demonstrated their non-overlap distributions in each of the channel structures (*cf.* Computational Methods for details). At low 1mM concentration,  $\rho_X(n_1, \dots, n_s)$  is largely dominated by the empty state probability  $\rho_X(0_1, \dots, 0_s)$  indicating only a small fraction of bound states with non-negligible occurrences (S4 Fig). Within this fraction, the most likely states involve single



**Fig 1. Resolution of sevoflurane sites at the homotetrameric Kv1.2 structures C and O.** (A) Atomistic systems containing Kv1.2 structures (cyan) embedded in a fully-hydrated lipid bilayer (gray) were MD simulated to produce molecular ensembles considered for flexible docking calculations. (B) Docking poses for singly-occupied sites. Shown is the ensemble-average channel structures C and O, along with the set of centroid configurations of sevoflurane (points) determined from docking. Centroid configurations of sevoflurane were clustered as a function of their location on the channel structures, that is within the voltage-sensor (green), at the S4S5 linker (yellow), at the S4Pore (dark blue) and S6P-helix (light blue) interfaces, at the central cavity (violet) and extracellular face (pink). Each of these clusters was treated as an interaction site  $j$  for sevoflurane with volume  $\delta V_j$ . (C) Following another round of docking calculations starting from structures in (B), solutions for doubly occupied sites were resolved by determining if volumes  $\delta V_j$  could accommodate the centroid positions of two docked ligands at once. For better annotation, helices S4, S4S5 linker, S6 (\*\*) and P-helix (\*), as well as the pore and VSD domains are indicated in structures in (B) and (C). (D) Per site number of docking solutions for single (cyan) and double (gray) ligand occupancy. (E) Representative molecular structure resolved from docking. Voltage-sensor domains in two opposing channel subunits are not depicted for clarity in (B) and (C) lateral views.

<https://doi.org/10.1371/journal.pcbi.1006605.g001>

occupancy of the S4S5 linker or the S4Pore interface as shown by the marginal probabilities  $\rho_X(n_j)$  of individual sites (Fig 2). At higher concentrations, there is a clear shift of  $\rho_X(n_1, \dots, n_s)$  towards channel occupancy states that significantly enhance the average number of bound ligands. Careful inspection of  $\rho_X(n_j)$  confirms the major relevance of sites at the S4S5 linker



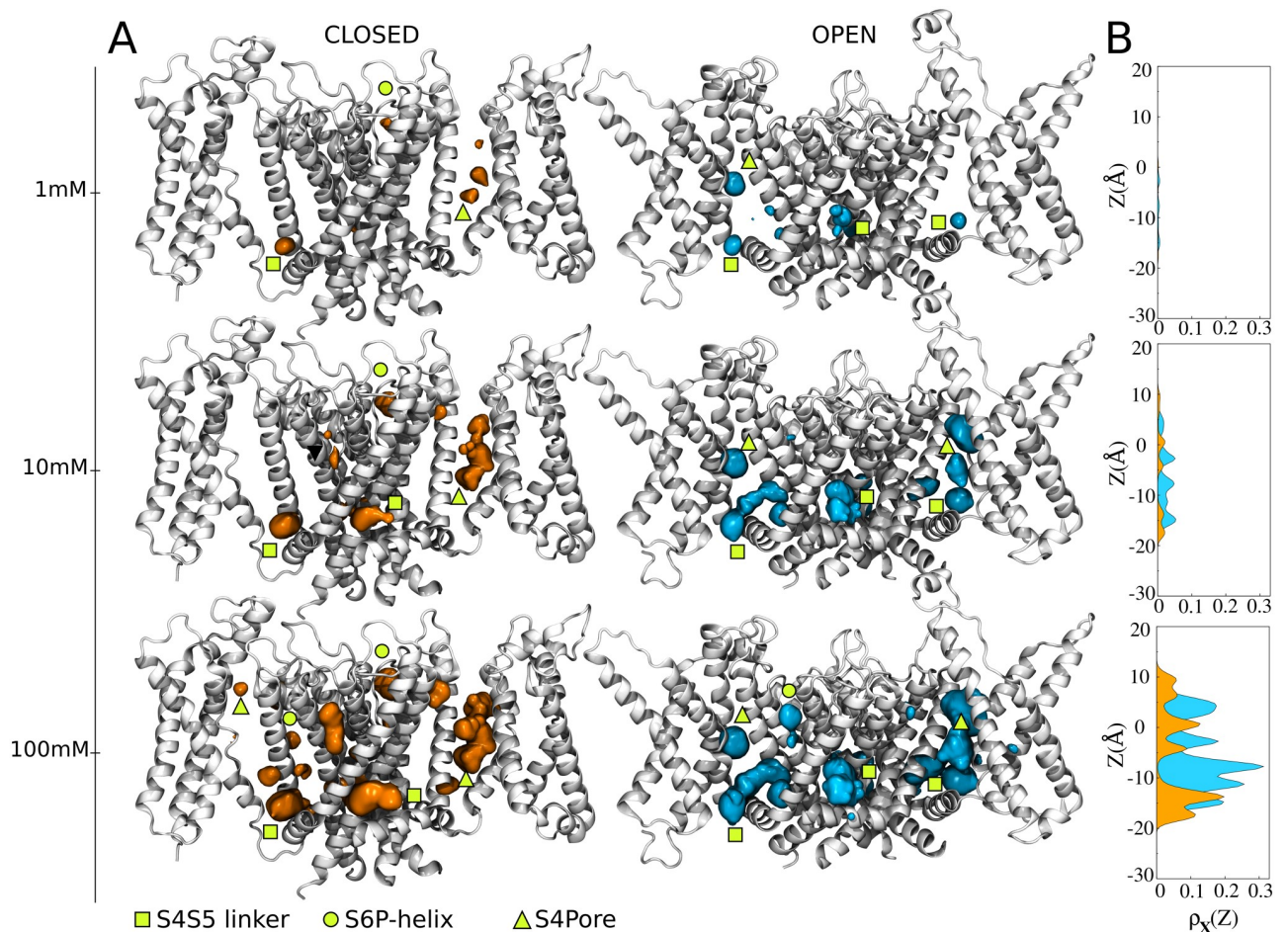
**Fig 2. C and O state-dependent binding probabilities for different concentrations of sevoflurane at the reservoir.** (A) Marginal probabilities  $\rho_X(n_j)$  of site  $j$ , for  $n_j = 0$  (gray),  $n_j = 1$  (blue) and  $n_j = 2$  (cyan). Marginals at the extracellular face of the channel are negligible for every structure/ concentration. (B) Probabilities  $\rho_X(n)$  for macrostates  $O_X^*(n)$  mapping an ensemble of accessible states  $O_X^*(n_1, \dots, n_s)$  in which  $n$  ligands bind the receptor regardless their specific distributions over the binding sites. Here,  $\rho_X(n_j)$  and  $\rho_X(n)$  were computed by coarse-graining over state probabilities in S4 Fig (cf. Computational Methods for details). Average number  $\langle n_X \rangle$  of bound ligands as a function of the reservoir concentration is indicated in (B).

<https://doi.org/10.1371/journal.pcbi.1006605.g002>

and S4Pore interface over the entire concentration range, accompanied by an increasing importance of binding regions at the S6P-helix interface. In contrast,  $\rho_X(n_j)$  for sites within the voltage-sensor, central cavity and nearby the extracellular face of the channel remain negligible over all concentrations. For completeness, note in S1 Table that equilibrium constants for doubly-occupied sites are comparable to or even higher than estimates for one-bound molecule thus revealing important saturation effects in which one or two sevoflurane molecules can stably bind the channel structures at individual sites. The result is especially true for spots at the S4S5 linker and S4Pore interface.

The complex distributions of the multiple occupied states of structures  $\mathbf{X} \equiv \{\mathbf{C}, \mathbf{O}\}$  were described in three dimensions by mapping  $\rho_X(n_1, \dots, n_s)$  into the position-dependent density  $\rho_X^j(\mathbf{R})$  of sevoflurane in each binding site  $j$  (cf. Computational Methods for details). As shown in Fig 3 and supplementary S1 and S2 Movies, the density of sevoflurane better convey the results by showing the spatially-mapped concentration dependent population of bound ligands. Projection of  $\rho_X^j(\mathbf{R})$  along the transmembrane direction  $z$  of the system,  $\rho_X^j(z)$ , stresses further the results. Note from  $\rho_X^j(\mathbf{R})$  that sevoflurane binds channel structures in a concentration dependent manner, binding preferentially the S4S5 linker and the interfaces S4Pore and S6P-helix over a range of concentrations.

So far, our calculations demonstrate that sevoflurane binds Kv1.2 structures over a spectrum of concentrations, preferentially at the linker S4S5 and at the segment interfaces S4Pore and S6P-helix. From a physical-chemical point of view, spots at these channel regions are primarily dehydrated, lipid accessible, amphiphilic pockets providing with favorable interaction sites for the polar lipophilic sevoflurane molecule (S5 Fig). It is worth mentioning that these findings recapitulate recent photolabeling experiments demonstrating that photoactive analogs of sevoflurane do interact to the S4S5 linker and at the S6P-helix interface of the open-



**Fig 3. C and O position-dependent binding probabilities for diluted concentrations of sevoflurane in the bulk.** (A) Shown is the ensemble average structure of the channel (white) along with the density  $\rho_X^j(\mathbf{R})$  of sevoflurane (orange and cyan) in each of the binding sites (isovalues of  $9 \times 10^{-5} \text{ \AA}^{-3}$ ). Densities pertaining to sites S4S5 linker, S6P-helix and S4Pore are indicated with yellow symbols. As described in Computational Methods, the determination of  $\rho_X^j(\mathbf{R})$  involved reweighing the marginal probability  $\rho_X(n_j)$  at the binding site  $j$  by the local equilibrium density of sevoflurane  $\rho_X(\mathbf{R}|n_j)$ . The marginal  $\rho_X(n_j)$  was computed by coarse-graining over state probabilities in S2 Fig whereas,  $\rho_X(\mathbf{R}|n_j)$  was calculated from the centroid distributions of docking solutions shown in Fig 1B and 1C. (B) Projection of  $\rho_X^j(\mathbf{R})$  along the transmembrane direction  $z$  of the system,  $\rho_X^j(z)$ .

<https://doi.org/10.1371/journal.pcbi.1006605.g003>

conductive Kv1.2 channel [22,23]. In detail, Leu317 and Thr384 were found to be protected from photoactive analogs, with the former being more protected than the latter. As shown in S6 Fig, atomic distances of bound sevoflurane to these amino-acid side chains are found here to be respectively  $7.28 \pm 2.5 \text{ \AA}$  and  $10.44 \pm 3.66 \text{ \AA}$ , in average more or less standard deviation. Such intermolecular distances are consistent with direct molecular interactions and therefore consistent with the measured protective reactions—similar conclusions hold for the closed channel as well. Besides that, our calculations recapitulate the stronger protection of Leu317 in the sense that, relative to sites at S6P-helix, the affinity of sevoflurane is found to be higher at the S4S5 linker considering its stable occupancy either by one or two ligands. The stable occupancy of the linker by one or two ligands as computed here, is also consistent with recent flooding-MD simulations of the homologous sodium channel NaChBac [14,24] and more importantly, with previous Ala/Val-scanning mutagenesis showing a significant impact of S4S5 mutations on the effect of general anesthetics on members of the  $\text{K}^+$  channel family [10].

In particular, a single residue (Gly329) at a critical pivot point between the S4S5 linker and the S5 segment underlies potentiation of Kv1.2 by sevoflurane [18]. Sevoflurane is found to be close to that amino acid when bound to the S4S5 linker.

In contrast to the aforementioned spots, sites within the voltage-sensor, within the main pore and nearby the extracellular face of the Kv1.2 structures are primarily hydrated, lipid-inaccessible, amphiphilic pockets (S5 Fig) that weaken sevoflurane interaction as reflected in the state- and space-dependent densities shown in Figs 2 and 3. The binding probabilities at these sites thus support that a non-negligible fraction of poses determined from docking (Fig 1D) corresponds to low affinity or false positives. In particular, because sevoflurane induces potentiation rather than blocking of Kv1.2 [17,18], we read the negligible or absent density of the ligand in the channel central-cavity as a self-consistent result of the study—especially for the open-conductive state. Supporting that conclusion, note that binding constants as computed here are upper bounds for the affinity of sevoflurane under ionic flux conditions in which potentiation takes place. Accordingly, as shown in S7 Fig, the binding affinity of a potassium ion at the central cavity overcomes that of sevoflurane due its binding and excess free-energies under applied voltages. Once bound, the ion destabilizes sevoflurane interactions and the molecule is not expected to bind the channel cavity at low concentrations. As also shown in S7 Fig and supplementary S3 Movie, even under the occurrence of rare binding events, sevoflurane appears unable to block the instantaneous conduction of potassium which is also consistent with its potentiating action.

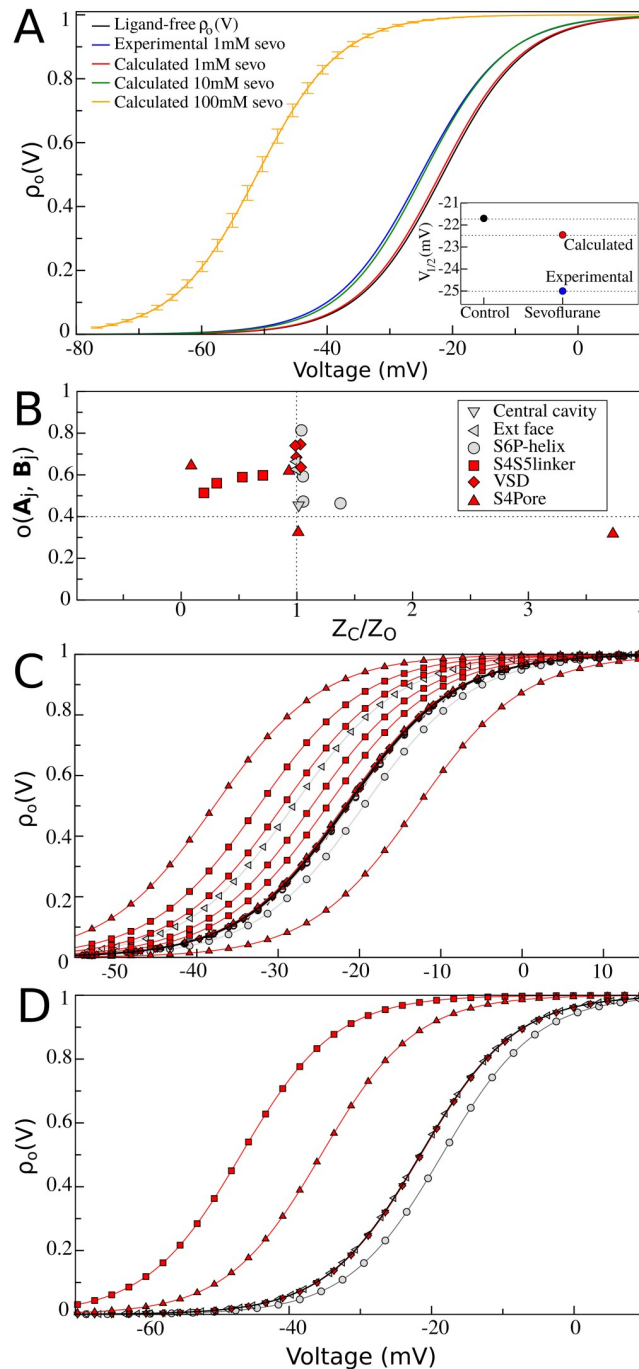
Weak interactions at the main pore and nearby the selectivity filter of Kv1.2 contrasts with sevoflurane binding at analogous regions of NaChBac [14,24], likely due major structural differences between Na<sup>+</sup> and K<sup>+</sup> channels. Specifically, the pore of potassium channels lacks lipid-accessible open-fenestrations of the sodium relatives and K<sup>+</sup>-selective filters are sharply distinct from Na<sup>+</sup>-selective ones.

### Anesthetic binding impacts channel energetics

Despite a comparable pattern of molecular interactions, careful inspection of  $\rho_X(n_j)$  or  $\rho_X^j(\mathbf{R})$  reveals that for most sites there is an obvious differential affinity of sevoflurane across Kv1.2 structures (Figs 2 and 3). The overall consequence for sevoflurane binding is then clear: the average number of ligands bound to the open-conductive channel systematically exceeds that for the resting-closed channel over the entire concentration range. There is therefore a remarkable conformational dependence for the anesthetic interaction, such that sevoflurane preferentially binds the open-conductive structure.

Implications for Kv1.2 energetics were then investigated by quantifying changes to the channel open probability  $\rho_O(V)$  induced by sevoflurane at concentrations of 1mM– 100mM (Fig 4). Specifically, from the partition functions  $Z_C(n_1, \dots, n_s)$  and  $Z_O(n_1, \dots, n_s)$  across the entire ensemble of occupancy states of the channel, solution of eqs (5) and (9) show that sevoflurane shifts leftwards the open probability of Kv1.2 in a concentration-dependent manner—voltage shifts amount from -1.0 mV to -30.0 mV with concentration increase of the ligand in solution. The result is particularly interesting, supporting that the approximately 3 mV probability shift recorded experimentally at 1 mM sevoflurane concentration can be explained by our structure-based probability predictions in the concentration range of 1–10 mM. The latter thus provides a theoretical basis to predict sevoflurane impact on channel energetics in a larger, not yet experimentally probed, range of concentrations. Additionally, for a fixed ligand concentration (100 mM), decomposition analysis reveals further that ratio values for the partition functions at individual sites  $j$  can be smaller, equal or larger than unity, implying a non-trivial interplay of conformation-dependent modes of action involving distinct sites (*cf.* Computational Methods for





**Fig 4. Sevoflurane binding effects on C-O equilibrium.** (A) Kv1.2 open probabilities for different sevoflurane concentrations in solution. Ligand-free and ligand-bound  $\rho_o(V)$  curves were respectively computed from eqs (5) and (9) by taking into consideration parameters,  $V_m = -21.9mV$  and  $\Delta Q = 3.85e_o$ , for best two-state Boltzmann fit of measured data for Kv1.2 free of ligands [18]. A reference experimental curve (blue) is shown for sevoflurane at 1 mM concentration, with best two-state Boltzmann parameters  $V_m = -25.1mV$  and  $\Delta Q = 4.00e_o$  [18]. Relative to the ligand-free channel, 1 mM of sevoflurane shifts the open probability of the channel leftward by  $\Delta V_m \approx -3mV$ . The inset explicitly shows both experimental (blue) and calculated (red)  $V_m$  shifts at 1mM of sevoflurane in solution. In black, ligand-free  $V_m$  is also shown for reference. For sevoflurane concentrations of 1, 10 and 100 mM, their respective open probabilities  $V_m$  are -22.3, -24.7 and -51.1 mV. Representative statistical errors for  $\rho_o(V)$  at 100mM sevoflurane concentration (error bars) were calculated by Monte Carlo bootstrap error analysis of the statistical uncertainty of independent FEP estimates considered in the calculations. (B) Decomposition analysis at 100mM ligand concentration. Shown is the FEP sampling overlap versus ratio values for C-O partition functions at the individual

binding sites  $j$ . Per-site ratio values can be equal, smaller or larger than unity meaning respectively that sevoflurane binding is not conformational dependent, stabilizes the open structure or stabilizes the closed structure. Binding sites located nearby flexible protein regions for which the root-mean-square deviation (RMSD) between channel structures is larger than 4.0 Å are highlighted in red (*cf.* Computational Methods and S2 Fig for details). (C) Decomposition analysis of  $\rho_O(V)$  curves in terms of partition ratio values showing in (B). (D) Same decomposition analysis in terms of an aggregate per-site contribution across channels subunits. At 100mM, binding of sevoflurane at the S4S5 linker and S4Pore interface significantly stabilizes the open structure of the channel which contrasts the mild stabilization of the closed structure due to ligand binding at the S6P-helix interface.

<https://doi.org/10.1371/journal.pcbi.1006605.g004>

details). In detail, binding of sevoflurane at low affinity sites within the voltage-sensor, central cavity and next to the extracellular face of the channel are mostly conformation-independent and do not impact open probability (ratio  $\approx 1$ ). On the other hand, conformation-dependent binding of sevoflurane to sites at the S4S5 linker and the S4Pore interface accounts for the overall stabilization of the open channel (ratio  $< 1$ ). That effect contrasts with the mild stabilization of the closed conformation of Kv1.2 induced by binding of sevoflurane at S6P-helix and reflected in rightward shifts of  $\rho_O(V)$  (ratio  $> 1$ ). The overall conformation-dependent binding process is therefore differentially encoded across distinct channel regions.

As extensively discussed in past publications [3,17,18], potentiation of Kv1.2 by sevoflurane has been attributed to stabilization of the open-conductive state of the channel via at least two molecular mechanisms. One mechanism (i) likely involving sites allosterically coupled to the electromechanical transduction responsible for controlling voltage-dependent gating; and another (ii) implicating distinct sites, which could influence the pore conductive state stability and/or unitary conductance. Here, our structural calculations based on the open-activated and resting-closed states of Kv1.2 were consistently designed to investigate the first (i) of these mechanisms. Given the critical role of S4 and S4S5 linker on the channel gating mechanism [19], it is reasonable that sevoflurane interactions with these segments, as found here, are at the origins of the experimentally measured voltage-dependent component of anesthetic action. While restricted to sevoflurane interactions with the resting-closed and open-conductive structures, the presented two-state binding model only embodies left or rightward shifts in the open probability of the channel, therefore it cannot clarify any molecular process accounting for the maximum conductance increase recorded experimentally. As supported by a recent kinetic modeling study [17], generalization of eqs (9) to include a third non-conducting open state yet structurally unknown is needed to account for such conductance effects and for that reason, the investigation of mechanism (ii) is beyond the scope of our study. We speculate however that binding of sevoflurane at the S4Pore and S6P-helix interfaces could allosterically interfere with pore domain operation, thus affecting channel's maximum conductance. A working hypothesis also raised in the context of anesthetic action on bacterial sodium channels [12,14], assumes indeed that non-conducting states of the selectivity filter are implicated. Corroboration of such an assumption from a molecular perspective is however not trivial and will necessarily involve further structural studies to demonstrate how ligand binding might impact non-conducting open states of the channel to affect maximum conductance.

## Discussion

Here, we carried out extensive structure-based calculations to study conformation-dependent binding of sevoflurane to multiple saturable sites of Kv1.2 structures  $X \equiv \{C, O\}$  under equilibrium conditions—the total MD simulation time was  $\sim 2.0 \mu\text{s}$ . Binding of sevoflurane was studied for ligand concentrations in the range of 1mM–100mM and saturation conditions up to  $n_j^{\text{max}} = 2$ . Our study relied on the assumption that molecular docking calculations performed

in vacuum can faithfully describe ligand interactions at protein sites. Specifically related to that assumption, we have considered the generated ensemble of docking solutions to estimate the location of binding sites  $\delta V_j$  and the local distribution of the ligand  $\rho_X(\mathbf{R}|n_j)$ . The generation of false positive hits is however a well documented drawback of docking algorithms as a result of limitations of the scoring function in describing ligand solvation energies and protein flexibility [25]. Given the same limitations of the scoring function, it is also not guaranteed that neither all binding hits, nor that  $\rho_X(\mathbf{R}|n_j)$  can be accurately known from docking. In this regard, although not considered here, it might be important to integrate docking results from different algorithms involving different scoring functions in order to characterize the bound ensemble. Still, thanks to the generality of the presented formulation, extension of the current investigation to sampling techniques other than docking, including all-atom flooding-MD simulations [9,11,12,14,16], might also be an important refinement in that direction.

Despite these sampling improvements that may eventually be obtained, the presented combination of extensive docking calculations against an ensemble of equilibrium receptor structures fit to handle protein flexibility, and FEP calculations based on fine force-fields to accurately estimate solvation energies are critical technical aspects of the applied methodology devised to minimize such drawbacks. Whereas docking was performed in vacuum, FEP calculations were carried out in presence of explicit all-atom lipids and water, therefore taking into consideration environmental effects in the estimated ligand binding free energies. In this regard, standard binding free-energies estimated from FEP (S1 and S2 Tables) are comparable or significantly smaller than the standard free energy for transferring the ligand from water to a pure lipid bilayer—supporting that sevoflurane is expected to partition preferentially into channel sites rather than the membrane. The partition coefficient ( $\log K$ ) and the related transfer free energy of sevoflurane between water and the lipid bilayer (POPC) amount respectively to  $2.64 \pm 0.96$  and  $-3.12 \pm 0.32$  kcal.mol<sup>-1</sup> as recently estimated by Tajkhorshid and coworkers [26]. Besides that, it is also important to note that the configuration space in FEP calculations overlap between channel structures at individual sites, i.e. sampling and binding affinities were evenly resolved between states (S2 Fig and Fig 4B)—meaning eventual biases or systematic errors were mitigated when comparing similar calculations between channel states according to the main goal here.

Under these technical considerations, we conclude that most of the identified binding sites are located nearby flexible protein regions for which the root-mean-square deviation between channel structures is larger than 4.0 Å. Then for the purpose of quantifying any direct ligand effect on channel energetics, the determined conformational dependence of sevoflurane binding to these gating-implicated protein regions appears robust and likely to impact function. Structural knowledge allied to solid electrophysiological data available for Kv1.2 make this channel an interesting model system for molecular-level studies of anesthetic action thereby justifying our choice. In detail, the atomistic structures account for most of the available experimental data characterizing closed and open conformations of the channel in the native membrane environment [20]. Previous findings support further that sevoflurane binds Kv1.2 to shift leftward its voltage-dependence and to increase its maximum conductance in a dose-dependent manner [18]. Despite a similar pattern of interactions, we found here a clear conformational dependence for sevoflurane binding at multiple channel sites. The ligand binds preferentially the open-conductive structure to impact the C-O energetics in a dose-dependent manner as dictated by the classical equilibrium theory for chemical reactions embodied in eq (9). Front of the difficulty in conceiving and characterizing other, still more complex molecular processes that might impact channel energetics under applied anesthetics [5–7], the result is revealing by showing that in principle the isolated process of sevoflurane binding to Kv1.2 accounts for open-probability shifts as recorded in experiments. Within this scenario, the

calculations reveal contrasting per-site contributions to the overall open probability of the channel. For instance, at 100mM concentration, binding of sevoflurane at the S4S5 linker and S4Pore interface significantly stabilizes the open structure of the channel overcoming the mild stabilization of the closed structure by ligand binding at the S6P-helix interface. By showing this non-trivial interplay of conformation-dependent modes of action involving distinct binding sites, the result is particularly insightful and should guide us to design novel site-specific mutagenesis and photolabeling experiments for further molecular characterization of anesthetic action.

Although not addressing the paucity of *in vivo* experimental evidences that a binding process to a specific molecular target as presented here is related to any clinically-relevant anesthetic outcome, our study adds support to the direct-site hypothesis by linking binding free-energy and protein energetics. As such, our study treats and reveals a new layer of complexity in the anesthetic problem that brings us novel paradigms to think their molecular action and to design/interpret research accordingly. To the best of our knowledge, the main-text Figs 3 and 4 represent in the context of structural studies, a deeper and first revealed view on the intricate mode of interactions that might take place between general anesthetics and ion channels to impact function in general.

## Methods

### Theory

**Anesthetic binding and channel energetics.** Consider the voltage-gated channel embedded in a large membrane-aqueous volume that contains  $N$  ligand molecules under dilution. The protein is assumed to remain in a well-defined conformational state  $X$ , in which it presents  $s$  distinct binding sites for ligands. For simplicity, we consider that ligands dissolve uniformly across the membrane-aqueous region of the system from where they can partition into the protein sites. The lipid and aqueous phases thus provide with a bulk volume  $V$  occupied by ligands at constant density  $\bar{\rho}$  and excess chemical potential  $\bar{\mu}$ . We consider further that every site  $j = 1, \dots, s$  corresponds to a discrete volume  $\delta V_j$  that can be populated by  $0 \leq n_j \leq n_j^{max}$  ligands. We denote by  $O_X^*(n_1, \dots, n_s)$  the specific occupancy state featuring  $n_j$  bound ligands at corresponding sites and by  $n = n_1 + \dots + n_s$  the total number of bound ligands in this state.

Under these considerations, solution of ligand binding to multiple receptor sites relies fundamentally in determining the equilibrium constant  $K_X(n_1, \dots, n_s)$  for the process  $O_X^*(0_1, \dots, 0_s) + nL \rightleftharpoons O_X^*(n_1, \dots, n_s)$  where,  $O_X^*(0_1, \dots, 0_s)$  is the empty receptor state with all ligands occupying the bulk. As shown in previous work [21], at a fixed temperature  $\beta = (k_B T)^{-1}$ , the binding constant  $K_X(n_1, \dots, n_s)$  can be evaluated from MD-based free-energy perturbation (FEP) calculations

$$K_X(n_1, \dots, n_s) = \frac{1}{n_1! \dots n_s!} \left[ \prod_{i=1}^n \left( \frac{2\pi}{\beta k_X(i)} \right)^{\frac{3}{2}} \right] e^{-\beta[W_X^*(n) - n\bar{\mu}]} \tag{1}$$

in which  $\bar{\mu}$  is the solvation free energy of the ligand in the bulk and  $W_X^*(n)$  corresponds to the free-energy of  $n$  site-specific bound ligands relative to a gas phase state given that ligands  $i = 1, \dots, n$  are restrained with force constants  $k_X(i)$  to occupy an effective site volume

$\left[ \prod_{i=1}^n \left( \frac{2\pi}{\beta k_X(i)} \right)^{\frac{3}{2}} \right]$  at structure  $X$ . Eq (1) is solved for the thermodynamic limit  $N \gg n$  and  $\frac{1}{n_1! \dots n_s!}$

corrects the binding constant for equivalent configurations of  $n_j$  indistinguishable ligands within the site volumes  $\delta V_j$ . Within this formulation, knowledge of  $K_X(n_1, \dots, n_s)$  ensures the

probability of any occupancy state

$$\rho_X(n_1, \dots, n_s) = \frac{\bar{p}^{(n_1+\dots+n_s)} K_X(n_1, \dots, n_s)}{\sum_{n'_1, \dots, n'_s} \bar{p}^{(n'_1+\dots+n'_s)} K_X(n'_1, \dots, n'_s)} \quad (2)$$

to be known in practice from free-energy calculations [27]. Here, the normalization condition appearing in the denominator of eq (2) runs over every occupancy state  $O(n'_1, \dots, n'_s)$  of the channel, ranging from  $O(0_1, \dots, 0_s)$  up to  $O(n_1^{max}, \dots, n_s^{max})$ . Note in eq (2)  $\rho_X(n_1, \dots, n_s)$  depends on the number density or concentration of the ligand at the reservoir thus providing a useful equation for investigation of concentration effects.

To investigate any conformational dependence on ligand binding, we consider eq (2) in the context of conformational equilibrium of the channel over a range of TM voltages. Specifically, we assume the very same microscopic system submitted to a Nernst potential induced by non-symmetrical electrolytes between membrane faces. The capacitive nature of the channel-membrane system ensures the Nernst potential accounts for a voltage difference  $V$  across the lipid bilayer. Accordingly, by denoting the entire set of instantaneous Cartesian coordinates of the channel as  $r^P$ , the free energy of the protein  $F_X(V)$  in a particular conformation  $X \equiv X(r^P)$

$$e^{-\beta F_X(V)} \propto \int dr^P \delta[X'(r^P) - X] e^{-\beta[U(r^P) + Q(r^P)V]} \quad (3)$$

can be written within an arbitrary constant, in terms of an effective potential energy of the protein  $U(r^P) + Q(r^P)V$  when coupled to the external voltage  $V$  with charge  $Q(r^P)$  [28]. Note the integral runs over the entire configurational space accessible to  $r^P$ , so long as it is compatible with conformation  $X$ , as indicated by the delta notation  $\delta[X'(r^P) - X]$ . From eq (3), the open probability of the channel then reduces to

$$\rho_O(V) = \frac{e^{-\beta F_O(V)}}{e^{-\beta F_C(V)} + e^{-\beta F_O(V)}} \quad (4)$$

for the case of a voltage-gated channel with two conformational states  $X \equiv \{C, O\}$  connected by the reaction process  $C \rightleftharpoons O$ . In terms of *chemical* free-energies of the receptor  $F_C(V=0)$  and  $F_O(V=0)$  and the corresponding *excess* free-energies  $\Delta F_C(V)$  and  $\Delta F_O(V)$ , eq (4) simplifies into the familiar two-state Boltzmann equation

$$\rho_O(V) = [1 + e^{+\beta \Delta Q (V_m - V)}]^{-1} \quad (5)$$

in which,

$$\Delta Q = -\frac{\Delta F_O(V) - \Delta F_C(V)}{V} \quad (6)$$

is the gating charge  $\Delta Q = Q_O - Q_C$  resulting from differences in the effective protein charge in each conformational state and

$$V_m = \frac{[F_O(V=0) - F_C(V=0)]}{\Delta Q} \quad (7)$$

is the midpoint voltage in which  $\rho_C(V) = \rho_O(V)$  [28]. From eq (5), the equilibrium constant between protein states  $C$  and  $O$  then writes as

$$K(V) = e^{-\beta \Delta Q [V_m - V]}$$

with  $K(0) = e^{-\beta V_m \Delta Q}$  determining their equilibrium at 0 mV. In eqs (6) and (7), the voltage-independent free energies account for the microscopic potential energy of the channel and its

solvation energy in each state whereas the corresponding voltage-dependent *excess* free energies are proportional to the applied voltage and associated protein charges.

By combining eqs (2) and (5) through a generalized thermodynamic-cycle analysis dealing with all possible states of the ligand-free and ligand-bound receptor, binding effects on the channel energetics can be then explicitly expressed over a range of membrane voltages

$$\rho_o(V) = \frac{K(V)Z_o(n_1, \dots, n_s)}{Z_c(n_1, \dots, n_s) + K(V)Z_o(n_1, \dots, n_s)} \quad (8)$$

in terms of the partition functions

$$\begin{cases} Z_c(n_1, \dots, n_s) = \sum_{n'_1, \dots, n'_s} \bar{p}^{(n'_1 + \dots + n'_s)} K_c(n'_1, \dots, n'_s) \\ Z_o(n_1, \dots, n_s) = \sum_{n'_1, \dots, n'_s} \bar{p}^{(n'_1 + \dots + n'_s)} K_o(n'_1, \dots, n'_s) \end{cases}$$

for the ensemble of occupancy states in each protein conformation. Eq (8) simplifies into

$$\rho_o(V) = \left[ 1 + \frac{Z_c(n_1, \dots, n_s)}{Z_o(n_1, \dots, n_s)} e^{+\beta \Delta Q(V_m - V)} \right]^{-1} \quad (9)$$

the two-state Boltzmann equation now embodying the free-energy contributions arising from ligand binding (*cf.* Computational Methods for details).

In this contribution, we consider eqs (1), (5) and (9) to investigate the molecular binding of sevoflurane to open and closed structures of Kv1.2, and its functional impact on channel energetics.

### Computational methods

A procedure was designed to solve the molecular binding of sevoflurane to the open-conductive (O) and resting-closed (C) structures of Kv1.2 for saturation conditions up to  $n_j^{max} = 2$ , under equilibrium conditions. For both channel structures, the procedure consisted of (i) an extensive production of docking solutions for the ligand-receptor interaction, (ii) clustering of docking solutions into binding sites along the receptor structure and (iii) estimation of binding affinities using the free-energy perturbation (FEP) method. First completion of steps (i) through (iii) solved the ligand channel interaction for singly-occupied binding sites. Double occupancy of the receptor sites was investigated by inputting the first generated ensemble of docked structures into another round of (i) through (iii) calculations. In detail, step (i) was accomplished by docking sevoflurane as a flexible ligand molecule against an MD-generated ensemble of membrane-equilibrated structures of the protein receptor. Docking calculations included the transmembrane domain of the channel, free from the membrane surroundings. Step (ii) provided the location of  $\delta V_j$  volumes lodging docking solutions for the ligand along the channel structures. Each of these volumes was treated as binding site regions in FEP calculations. Here, FEP calculations started from equilibrium structures of the ligand-bound channel embedded in a explicit water-membrane environment. For the purpose of improving statistics, FEP estimates and the associated statistical errors were determined from at least two independent decoupling runs. Calculations were performed over  $\sim 7.0$  ns per site, per conformation and per replica to converge FEP estimates, in a total MD simulation time of  $\sim 2.0$   $\mu$ s. Following this procedure, binding constants  $K_X(n_1, \dots, n_s)$  for channel structures  $X \equiv \{C, O\}$  were solved by inputting FEP estimates into eq (1), allowing for direct solution of state-dependent probability distributions  $\rho_X(n_1, \dots, n_s)$  via eq (2). Here, binding constants  $K_X(n_1, \dots, n_s)$

and the related standard binding free energies  $\Delta G_X^0(n_1, \dots, n_s)$  were solved for the condition of independent binding sites and relative to a homogeneous and diluted aqueous solution occupied by ligands at constant density  $\bar{\rho}$  and excess chemical potential  $\bar{\mu}$ . Ligand-free and ligand-bound open probability curves  $\rho_X(V)$  were respectively computed from eqs (5) and (9) by taking into consideration the previously determined mid-point voltage and steepness of the open probability curve of Kv1.2 free of ligands, *i.e.*  $V_m = -21.9 \text{ mV}$  and  $\Delta Q = 3.85e_0$  as determined from best two-state Boltzmann fit of the measured conductance-voltage data of the channel [18]. Note that any ligand-induced shift in eq (9) is determined by the partition function ratio between open and closed structures and not by the choice of these reference parameters. Probabilities  $\rho_X(n_1, \dots, n_s)$  and  $\rho_X(V)$  were determined for sevoflurane concentrations in the range of 1mM–100mM (or in density units,  $6.02 \times 10^{-7} \text{ \AA}^{-3}$ – $6.02 \times 10^{-5} \text{ \AA}^{-3}$ ).

**Membrane equilibrated channel structures.** The Kv1.2 structure in the open-conductive (O) state was obtained from Treptow and Tarek [29]. The construct was previously acquired via molecular dynamics (MD) simulations of the published x-ray crystal structure [19]. The resting-closed (C) structure of Kv1.2 was obtained from Delemotte *et al.* [30]. Modeling details and validation can be found in the original papers. It is important to clarify that besides the structure by Delemotte *et al.*, other resting-state models have been proposed for Kv1.2 [31–34]. Given that these structures were proven to provide with a consensus model for the resting state of the channel [35], we focus our investigation on the former model.

Structures C and O were embedded in the lipid bilayer for Molecular Dynamics (MD) relaxation and subsequent molecular docking of sevoflurane. Specifically, each structure featuring three  $\text{K}^+$  ions (s4s2s0) at the selectivity filter was inserted in a fully hydrated and zwitterionic all atom palmitoyloleoylphosphatidylcholine (POPC) phospholipid bilayer. After assembled, each macromolecular system was simulated over a MD simulation spanning  $\sim 20$  ns, at constant temperature (300 K) and pressure (1 atm), neutral pH, and with no applied TM electrostatic potential. The channel structures remained stable in their starting conformations throughout the simulations. Indeed, the RMSD profile of the channel in each simulation converges to a plateau value of approximately 4.0 Å, indicative of structural stability of the constructs [29,30]. Structures sampled in the steady phase of the trajectories were considered in subsequent docking and FEP calculations.

**Molecular Docking.** We used *AutoDock Vina* [36] to dock sevoflurane against the MD-generated ensemble of channel structures C and O. Each ensemble included 120 independent channel configurations at least. Docking solutions were resolved with an exhaustiveness parameter of 200, by searching a box volume of  $100 \times 100 \times 100 \text{ \AA}^3$  containing the transmembrane domain of the protein receptor. Sevoflurane was allowed to have flexible bonds for all calculations. Clustering of docking solutions was carried out following a maximum neighborhood approach.

**Molecular dynamics.** All MD simulations were carried out using the program NAMD 2.9 [37] under Periodic Boundary Conditions. Langevin dynamics and Langevin piston methods were applied to keep the temperature (300 K) and the pressure (1 atm) of the system fixed. The equations of motion were integrated using a multiple time-step algorithm [38]. Short- and long-range forces were calculated every 1 and 2 time-steps respectively, with a time step of 2.0 fs. Chemical bonds between hydrogen and heavy atoms were constrained to their equilibrium value. Long-range electrostatic forces were taken into account using the Particle Mesh Ewald (PME) approach [39]. The CHARMM36 force field [40] was applied and water molecules were described by the TIP3P model [41]. All the protein charged amino acids were simulated in their full-ionized state (pH = 7.0). All MD simulations, including FEP and voltage-driven simulations (see next), were performed on local HPC facility at LBTC.

**Free-energy perturbation (FEP).** Eq (1) was simplified here for the condition of ligand interactions to multiple independent sites—a condition that appears to be fulfilled at the channel structures featuring sparse binding sites for sevoflurane. Within this scenario, binding constants for structures  $X \equiv \{C, O\}$  were factorized as the product of independent equilibrium constants

$$K_X(n_1, \dots, n_s) = K_X(n_1, 0_2, \dots, 0_s) \times \dots \times K_X(0_1, \dots, 0_{s-1}, n_s)$$

where,

$$K_X(n_1, 0_2, \dots, 0_s) = \frac{1}{n_1!} \left[ \prod_{i=1}^{n_1} \left( \frac{2\pi}{\beta k_X(i)} \right)^{\frac{3}{2}} \right] e^{-\beta[W_X^*(n_1) - n_1\bar{\mu}]}$$

...

$$K_X(0_1, \dots, 0_{s-1}, n_s) = \frac{1}{n_s!} \left[ \prod_{i=1}^{n_s} \left( \frac{2\pi}{\beta k_X(i)} \right)^{\frac{3}{2}} \right] e^{-\beta[W_X^*(n_s) - n_s\bar{\mu}]}$$

denote respectively the binding constant of  $n_j$  ligands to each of the  $j$  sites at structure  $X$ .

Accordingly, the *excess* chemical potential  $\bar{\mu}$  associated with coupling of the ligand from gas phase to bulk water and  $W_X^*(n_j)$  associated with coupling of  $n_j$  ligands from gas phase to site  $j$  under restraints were quantified via FEP. Because computation of  $\bar{\mu}$  does not depend on the choice of concentration, so long as the same thermodynamic state is used for the solution and gas phases, we estimated the *excess* potential by considering one sevoflurane molecule embedded into a water box of 60 x 60 x 60 Å<sup>3</sup>.  $W_X^*(n_j)$  was computed by taking into consideration the whole ligand-channel-membrane system.

All FEP calculations were performed in NAMD 2.9 [37] by considering the Charmm-based parameters for sevoflurane as devised by Barber *et al.* [42]. Starting from channel-membrane equilibrated systems containing bound sevoflurane as resolved from docking, forward transformation were carried out by varying the coupling parameter in steps of 0.01. Each transformation then involved a total of 100 windows, each spanning over 31800 steps of simulation. For the purpose of improving statistics, free-energy estimates and associated statistical errors were determined using the simple overlap sampling (SOS) formula [43] based on at least two independent FEP runs.

Specifically for ligand-protein calculations, the free-energy change  $W_X^*(1_j)$  for singly-occupied sites  $j$  was computed as a FEP process that involves ligand coupling to a vacant site. Differently, for doubly-occupied sites,  $W_X^*(2_j)$  was computed as a two-step FEP process involving ligand coupling to a vacant site  $W_X^*(1_j)$  followed by binding of a second ligand at the preoccupied site  $W_X^*(2_j|1_j)$ . Because  $W_X^*(2_j)$  is a state function, the stepwise approach is equivalent to a single-step process involving simultaneous coupling of two ligands to the protein site that is,  $W_X^*(2_j) = W_X^*(1_j) + W_X^*(2_j|1_j)$ . The colvars module [44] in NAMD 2.9 was used to apply the harmonic restraint potentials when computing these quantities.

The value of  $W_X^*(n_j)$  depends on the parameters of the restraint potential adopted in the FEP calculation, *i.e.* the reference positions of the ligands in the bound state  $\mathbf{R}_X^*(1_j), \dots, \mathbf{R}_X^*(n_j)$  and the magnitude of force constants  $\{k_X(1_j), \dots, k_X(n_j)\}$ . By minimizing the contribution of the restraint potential to the binding free-energy  $W_X^*(n_j)$ , Roux and coworkers [45] devised optimum choices for the parameters

$$\{\mathbf{R}_X^*(1_j) = \langle \mathbf{R}_X(1_j) \rangle, \dots, \mathbf{R}_X^*(n_j) = \langle \mathbf{R}_X(n_j) \rangle\}$$



and

$$\left\{ k_x(1_j) = \frac{3\beta^{-1}}{\langle [\delta \mathbf{R}_x(1_j)]^2 \rangle}, \dots, k_x(n_j) = \frac{3\beta^{-1}}{\langle [\delta \mathbf{R}_x(n_j)]^2 \rangle} \right\}$$

in which,  $\langle \mathbf{R}_x(1_j) \rangle, \dots, \langle \mathbf{R}_x(n_j) \rangle$  and  $\langle [\delta \mathbf{R}_x(1_j)]^2 \rangle, \dots, \langle [\delta \mathbf{R}_x(n_j)]^2 \rangle$  are respectively the equilibrium average positions for each of the  $n_j$  bound ligands at site  $j$  and their corresponding mean-square fluctuations when interacting to structure  $\mathbf{X}$ . Here, these parameters were estimated from the space of docking solutions and the resulting force constants, in the range of 0.03 to 1.35 kcal/mol/Å<sup>2</sup>, were considered for computations of the bound state.

**Sampling overlap.** Here, a per-site measure of sampling overlap  $o(A_j, B_j)$  between FEP configurations in structures  $\mathbf{C}$  and  $\mathbf{O}$

$$o(A_j, B_j) = 1 - \frac{\sqrt{\text{tr}((A_j^{1/2} - B_j^{1/2})^2)}}{\sqrt{\text{tr}A_j + \text{tr}B_j}}$$

was determined [46] from the square root of the covariance matrices  $A_j$  and  $B_j$  associated respectively to  $\mathbf{C}$  and  $\mathbf{O}$  samples at site  $j$ . Specifically,  $A_j$  and  $B_j$  were computed as symmetric  $3 \times 3$  covariance matrices for centroid positions  $\mathbf{R}_j$  of the ligand at site  $j$

$$X_j = \langle (\mathbf{R}_j - \langle \mathbf{R}_j \rangle) \cdot (\mathbf{R}_j - \langle \mathbf{R}_j \rangle)^T \rangle$$

and their square roots

$$X^{1/2} = R \text{diag}(\lambda_1^{1/2}, \lambda_2^{1/2}, \lambda_3^{1/2}) R^T$$

were solved from the column major eigenvectors  $\{\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3\}$  of the rotation matrix  $\mathbf{R}$  and the associated eigenvalues  $\{\lambda_1, \lambda_2, \lambda_3\}$ . Note that overlap is expectedly 1 for identical samplings and 0 for orthogonal configuration spaces.

**Absolute binding free energy and ensemble averages.** An absolute binding free-energy  $\Delta G_x^o(n_1, \dots, n_s)$  [47] associated with state  $O_x^*(n'_1, \dots, n'_s)$  can be defined as

$$\Delta G_x^o(n_1, \dots, n_s) = -\beta^{-1} \ln [K_x(n_1, \dots, n_s) \times (C^o)^n]$$

where it is understood that this refers to the free energy of binding  $n$  ligands to the protein structure  $\mathbf{X} \equiv \{\mathbf{C}, \mathbf{O}\}$  from a reference standard reservoir concentration  $C^o = 1\text{M}$  or in units of number density  $C^o = (1,660 \text{Å}^3)^{-1}$ . Still, the ensemble average of any thermodynamic property of the system  $A_x(n'_1, \dots, n'_s)$  for state  $O_x^*(n'_1, \dots, n'_s)$  can be known

$$\langle A_x \rangle = \sum_{n'_1, \dots, n'_s} \langle A_x \rangle_{(n'_1, \dots, n'_s)} \rho_x(n'_1, \dots, n'_s)$$

from the state probability in eq (2).

**Position-dependent probability densities.** As demonstrated in reference [21], state-dependent probabilities  $\rho_x(n_1, \dots, n_s)$  for channel structures  $\mathbf{X} \equiv \{\mathbf{C}, \mathbf{O}\}$  can be mapped into the probability density  $\rho_x(\mathbf{R})$  of any given ligand  $i$  to occupy position  $\mathbf{R}$  in the system (regardless the position of the remaining  $N - 1$  ligands). Given our original consideration that the reservoir is a homogeneous volume occupied by ligands with position-independent density  $\bar{\rho}$ , the probability  $\rho_x(\mathbf{R})$  simplifies to

$$\rho_x(\mathbf{R}) = \begin{cases} \rho_x^j(\mathbf{R}), \forall \mathbf{R} \in \delta V_j \\ \bar{\rho}, \text{reservoir} \end{cases}$$

for every protein site  $j = 1, \dots, s$ . The determination of  $\rho_X(\mathbf{R})$  thus reduces in practice to knowledge of the per-site density  $\rho_X^j(\mathbf{R})$

$$\rho_X^j(\mathbf{R}) = \sum_{n_j=0}^{n_j^{\max}} \rho_X(n_j) \times \rho_X(\mathbf{R}|n_j)$$

where,  $\rho_X(\mathbf{R}|n_j)$  is the local density at site  $j$  when occupied exactly by  $n_j$  molecules and  $\rho_X(n_j)$  is the probability for this occupancy state. The probability  $\rho_X(\mathbf{R}|n_j)$  describes the local equilibrium density of the ligand, conditional to a specific number of bound molecules that satisfies  $\int_{\delta V_j} d\mathbf{R} \rho_X(\mathbf{R}|n_j) = n_j$ . In contrast,

$$\rho_X(n_j) = \sum_{n'_1, \dots, n'_s} \delta_{n'_j, n_j} \rho_X(n'_1, \dots, n'_s)$$

denotes the marginal probability of site  $j$  to be occupied by  $n_j$  ligands regardless the occupancy of the other sites.

The formulation above establishes a formal relation between space-dependent and state-dependent densities of the system. At a fine level, this relation involves the set of equilibrium constants  $K_X(n_1, \dots, n_s)$  satisfying  $\rho_X(n_j)$ . From  $\rho_X^j(\mathbf{R})$ , spatial projections of  $\rho_X(\mathbf{R})$  along the transmembrane  $z$  direction of the system can be determined as

$$\rho_X(z) = \bar{\rho} \times A(z) + \sum_{j=1}^s \rho_X^j(z)$$

where,  $A(z) = \Delta x \Delta y$  is the total area of the membrane-aqueous region along the Cartesian  $x$  and  $y$  directions.

**Coarse-graining over states.** Consider any macrostate  $O_X^*(n)$  of the system mapping an ensemble of accessible states  $O_X^*(n_1, \dots, n_s)$  in which  $n$  ligands bind the receptor regardless their specific distributions over the binding sites. Because  $O_X^*(n)$  is degenerate, the probability density of the macrostate

$$\rho_X(n) = \sum_{n'_1, \dots, n'_s} \delta_{n', n} \rho_X(n'_1, \dots, n'_s)$$

can be determined by coarse-graining over the receptor states  $O_X^*(n_1, \dots, n_s)$  featuring exactly  $n = n_1 + \dots + n_s$  bound ligands. Here, the Kronecker delta function  $\delta_{n', n}$  ensures summation over states accessible to  $O_X^*(n)$  only.

**Binding of potassium and sevoflurane at the main-pore of Kv1.2.** FEP calculations to quantify the binding free-energy of sevoflurane against a preoccupied central cavity of Kv1.2 with bound potassium was computed as described in the Free-Energy Perturbation (FEP) section. Specifically, the free-energy change  $W_o^*(2_j)$  for double occupancy of the central-cavity by potassium and sevoflurane was computed as a two-step FEP process involving coupling of the ion to the central cavity  $W_o^*(1_j)$  followed by binding of the anesthetic at the preoccupied cavity  $W_o^*(2_j|1_j)$ , that is,  $W_o^*(2_j) = W_o^*(1_j) + W_o^*(2_j|1_j)$ . Absolute binding free energies  $\Delta G_o^o(0_1, \dots, 1_j, \dots, 0_s)$  and  $\Delta G_o^o(0_1, \dots, 2_j, \dots, 0_s)$  were then computed from the respective per-site binding constants  $K_o(0_1, \dots, 1_j, \dots, 0_s)$  and  $K_o(0_1, \dots, 2_j, \dots, 0_s)$ . An in-water excess potential of  $-65.10 \text{ kcal.mol}^{-1}$  was estimated for potassium. Specifically for  $K^+$ , a total binding free-energy was obtained by summing up its absolute binding free energy with its charge ( $q$ ) excess free energy ( $q\phi V$ ) under an applied external voltage  $V$  [48,49]. The voltage coupling  $\phi$

was determined in the form of the “electrical distance”

$$\delta_\epsilon = \frac{\partial}{\partial V} \Phi(V)|_{V=0}$$

where,  $\Phi(V)$  is the local-electrostatic potential of the ion at the central cavity of the open channel. In practice, we solve  $\delta_\epsilon$  from two independent 2ns-long simulations at voltages  $V = 0mV$  and  $V = 600mV$ . For both runs,  $\Phi(V)$  was estimated from the electrostatic potential map of the system and subsequently considered to solve  $\delta_\epsilon$  for  $\delta V = 600mV$ .

To investigate the conduction properties of Kv1.2 with sevoflurane bound to the main pore, the open channel structure was simulated under depolarized-membrane conditions using a charge-imbalance protocol [50].

**Partition function decomposition.** In the limit of  $s$  independent sites, binding constants can be factorized as the product of independent equilibrium constants then ensuring the associated partition function to be factorized in terms

$$Z_X(n_1, \dots, n_s) = Z_X(n_1, 0_2, \dots, 0_s) \times \dots \times Z_X(0_1, \dots, 0_{s-1}, n_s)$$

of per-site contributions. That decomposition is particularly useful to estimate the per-site contributions impacting the open probability of the channel as defined in eq (9). For any given site  $j$ , ratio values

$$\frac{Z_C(0_1, \dots, n_j, \dots, 0_s)}{Z_O(0_1, \dots, n_j, \dots, 0_s)} \begin{cases} = 1 \\ < 1 \\ > 1 \end{cases}$$

mean respectively that ligand binding is not conformational dependent, stabilizes the open structure or stabilizes the closed structure.

**Derivation of main-text Eq (8).** The voltage-dependent open probability for a two-state channel writes according to

$$\rho_o(V) = \frac{\rho_o(V)}{\rho_c(V) + \rho_o(V)}$$

where,

$$\begin{cases} \rho_c(V) = \rho_c(0_1, \dots, 0_s, V) + \sum_{n'_1, \dots, n'_s \neq 0_1, \dots, 0_s} \rho_c(n'_1, \dots, n'_s) \\ \rho_o(V) = \rho_o(0_1, \dots, 0_s, V) + \sum_{n'_1, \dots, n'_s \neq 0_1, \dots, 0_s} \rho_o(n'_1, \dots, n'_s) \end{cases}$$

embodies respectively every occupancy state  $O(n'_1, \dots, n'_s)$  of the channel conformations  $X \equiv \{C, O\}$ . The state probabilities rewrite according to

$$\begin{cases} \rho_c(V) = \rho_c(0_1, \dots, 0_s, V) \times \sum_{n'_1, \dots, n'_s} \bar{p}^{(n'_1 + \dots + n'_s)} K_C(n'_1, \dots, n'_s) \\ \rho_o(V) = \rho_o(0_1, \dots, 0_s, V) K(V) \times \sum_{n'_1, \dots, n'_s} \bar{p}^{(n'_1 + \dots + n'_s)} K_O(n'_1, \dots, n'_s) \end{cases}$$

by noting that the voltage-dependent equilibrium constant between channel conformations is given by

$$\rho_o(0_1, \dots, 0_s, V) = \rho_c(0_1, \dots, 0_s, V) K(V)$$

and that every occupancy probability

$$\begin{cases} \rho_C(n_1, \dots, n_s) = \rho_C(0_1, \dots, 0_s, V) \times \bar{p}^{(n_1+\dots+n_s)} K_C(n_1, \dots, n_s) \\ \rho_O(n_1, \dots, n_s) = \rho_O(0_1, \dots, 0_s, V) \times \bar{p}^{(n_1+\dots+n_s)} K_O(n_1, \dots, n_s) \end{cases}$$

derives from the respective binding constant and voltage-dependent probability of the ligand-free reference state. From above, we then conclude that the voltage-dependent open probability of the channel can be expressed in terms of the partition functions

$$\begin{cases} Z_C(n_1, \dots, n_s) = \sum_{n'_1, \dots, n'_s} \bar{p}^{(n'_1+\dots+n'_s)} K_C(n'_1, \dots, n'_s) \\ Z_O(n_1, \dots, n_s) = \sum_{n'_1, \dots, n'_s} \bar{p}^{(n'_1+\dots+n'_s)} K_O(n'_1, \dots, n'_s) \end{cases}$$

such that,

$$\rho_O(V) = \frac{K(V)Z_O(n_1, \dots, n_s)}{Z_C(n_1, \dots, n_s) + K(V)Z_O(n_1, \dots, n_s)}$$

## Supporting information

**S1 Fig. Root mean square deviation (RMSD) profiles of channel structures C and O along equilibrium MD simulations.** Heavy TM domain atoms of the channel were included in the calculation, considering the starting conformation (simulation time  $t = 0$  ns) as the reference structure. Channel structures remained stable throughout the simulations. RMSD profiles converge to a plateau value of approximately 4.0 Å, indicative of structural stability of the constructs. Equilibrium structures sampled in the steady phase of the trajectories were used in subsequent docking and FEP calculations.

(TIFF)

**S2 Fig. Overlap analysis.** (A) Per-site distributions of one and two sevoflurane molecules bound to channel structures C and O. Distributions were sampled in FEP calculations based on reference positions  $R_X^*(1_j)$ ,  $R_X^*(2_j)$  and force constants  $\{k_X(1_j), k_X(2_j)\}$  known from docking (cf. Computational Methods). Only centroid positions of the ligand are shown (dots). All binding sites but spots at the S6P-helix and the extracellular face of the channel are located nearby flexible protein regions (light to dark red) for which the root-mean-square deviation (RMSD) between channel structures is larger than 4.0 Å. (B) Sampling overlap  $O(A_j, B_j)$  between ligand distributions in (A) (cf. Computational Methods). Overlap is larger than 0.4 for the majority of binding sites implying a similar set of configurations effectively sampled for closed and open states. Here, RMSD and overlap were computed after elimination of overall structural rotation and translation by fitting channel structures at segments S1, S2, S3 and P-helix.

(TIFF)

**S3 Fig. Re-hydration and re-lipidation.** Site-specific lipid or water coordination number difference ( $\Delta n$ ) between final FEP configuration and equilibrium trajectories. For both closed (C) and open (O) channel structures,  $\Delta n$  is a function of lipid or water distances from individual binding site's geometric center.  $\Delta n$  is computed considering the average number of water/lipid molecules within throughout equilibrium trajectories (cf. S1 Fig), as well as the average coordination number in the same binding sites in all four channel subunits at the end of the FEP calculation. Averaging statistical uncertainty is propagated and shown as error bars. Note that coordination number at the final ligand-decoupled stage of FEP is very similar ( $\Delta n \approx 0$ ) to

equilibrium reference values determined for membrane-embedded, ligand-free channel structures. Note as well in [S5 Fig](#) that sites S4S5 linker and S6P-helix interface are lipid exposed, whereas sites within the voltage-sensor, S4Pore interface and extracellular face are water accessible.

(TIFF)

**S4 Fig. State-dependent binding probabilities for different concentrations of sevoflurane in the bulk.** Shown are sorted values of  $\rho(n_1, \dots, n_s)$  over the occupancy states of channel structures **C** and **O**. Strings for the four most likely states are stated in the center of the plots—the first line corresponding to the most likely, and the last to the fourth most likely state.

(TIFF)

**S5 Fig. Local molecular interactions and physical chemical environment of sevoflurane within Kv1.2 binding sites.** The first column of each conformation (**C** and **O**) displays a surface representation of the amino acids that compose the respective binding site, colored by their physical chemical character—white: apolar, green: polar non-charged, red: negatively charged and blue: positively charged. The second column shows the water molecules within a 5 Å radius from the site's geometrical center. Of note, sites S4S5 linker and S6P-helix are predominantly dehydrated and lipid accessible.

(TIFF)

**S6 Fig. Average atomic distances between the centroids of sevoflurane and photolabeled Kv1.2 residues in structures C and O.** Shown are average distances between sevoflurane when bound to S6P-helix site and Thr384, and average distances of the ligand when bound to S4S5-linker site to residues Leu317 and G329. Distances were measured by considering both ensembles of equilibrium protein structures inputted into docking searches, and sevoflurane docking poses pertaining to a given binding site.

(TIFF)

**S7 Fig. Ion-sevoflurane equilibrium.** (A) Absolute and excess free-energies (kcal.mol<sup>-1</sup>) for binding potassium (yellow) and/or sevoflurane (blue) at the main pore of Kv1.2 (white). Excess free energies at 100mV are shown in parentheses. More favorable absolute and excess free-energies ensure single occupancy by potassium to be more likely than that by sevoflurane. In contrast, double occupancy by potassium and sevoflurane is unfavorable due to a positive absolute free energy for binding the molecule at the ion occupied cavity. (B) Shown are time-dependent trajectories of potassium ions diffusing through the open pore of Kv1.2 despite one bound sevoflurane molecule at the central cavity. The voltage-driven MD simulations were carried out at a depolarized potential of 600 mV to increase the rate of sampling of conduction events. Simulations spanned a total of ~ 30 ns.

(TIFF)

**S1 Movie. Three-dimensional visualization of position-dependent densities of sevoflurane bound to the closed Kv1.2, at 100 mM.** Kv1.2 channel and ligand densities are depicted in white and orange, respectively. Video displays 360° side view, followed by an extracellular view. Relates to main text [Fig 3](#).

(MPG)

**S2 Movie. Three-dimensional visualization of position-dependent densities of sevoflurane bound to the open Kv1.2, at 100 mM.** Kv1.2 channel and ligand densities are depicted in white and blue, respectively. Video displays 360° side view, followed by an extracellular view. Relates to main text [Fig 3](#).

(MPG)

**S3 Movie. Ion conduction simulation in the presence of sevoflurane bound to channel central cavity.** Molecular dynamics simulation with 600mV depolarizing potential was carried out to enhance sampling of conduction events. Kv1.2 channel, sevoflurane and K<sup>+</sup> ions can be seen in white, blue and yellow, respectively. Water molecules nearby potassium ions and sevoflurane are also shown. Note that the presence of sevoflurane does not hinder ion conduction. (MPG)

**S1 Table. FEP calculations and equilibrium binding constants for singly- and doubly-occupied sites *j* of the closed channel structure<sup>#</sup>.** (PDF)

**S2 Table. FEP calculations and equilibrium binding constants for singly- and doubly-occupied sites *j* of the open channel structure<sup>#</sup>.** (PDF)

## Author Contributions

**Conceptualization:** Werner Treptow.

**Formal analysis:** Leticia Stock, Juliana Hosoume, Leonardo Cirqueira.

**Funding acquisition:** Werner Treptow.

**Writing – original draft:** Werner Treptow.

**Writing – review & editing:** Leticia Stock, Werner Treptow.

## References

1. Franks NP, Honoré E. The TREK K2P channels and their role in general anaesthesia and neuroprotection. *Trends Pharmacol Sci*. 2004 Nov; 25(11):601–8. <https://doi.org/10.1016/j.tips.2004.09.003> PMID: 15491783
2. Franks NP. General anaesthesia: from molecular targets to neuronal pathways of sleep and arousal. *Nat Rev Neurosci*. 2008; 9:370–386. <https://doi.org/10.1038/nrn2372> PMID: 18425091
3. Covarrubias M, Barber AF, Carnevale V, Treptow W, Eckenhoff RG. Mechanistic Insights into the Modulation of Voltage-Gated Ion Channels by Inhalational Anesthetics. *Biophys J*. 2015 Nov 17; 109(10):2003–11. <https://doi.org/10.1016/j.bpj.2015.09.032> PMID: 26588560
4. Heurteaux C, Guy N, Laigle C, Blondeau N, Duprat F, Mazzuca M, et al. TREK-1, a K<sup>+</sup> channel involved in neuroprotection and general anesthesia. *EMBO J*. 2004; 23(13):2684–95. <https://doi.org/10.1038/sj.emboj.7600234> PMID: 15175651
5. Cantor RS. The Lateral Pressure Profile in Membranes: A Physical Mechanism of General Anesthesia. *Biochemistry*. 1997 Mar 1; 36(9):2339–44. <https://doi.org/10.1021/bi9627323> PMID: 9054538
6. Finol-Urdaneta RK, McArthur JR, Juranka PF, French RJ, Morris CE. Modulation of KvAP Unitary Conductance and Gating by 1-Alkanols and Other Surface Active Agents. *Biophys J*. 2010 Mar 3; 98(5):762–72. <https://doi.org/10.1016/j.bpj.2009.10.053> PMID: 20197029
7. Roth R, Gillespie D, Nonner W, Eisenberg RE. Bubbles, Gating, and Anesthetics in Ion Channels. *Biophys J*. 2008 Jun 1; 94(11):4282–98. <https://doi.org/10.1529/biophysj.107.120493> PMID: 18234836
8. Nury H, Van Renterghem C, Weng Y, Tran A, Baaden M, Dufresne V, et al. X-ray structures of general anaesthetics bound to a pentameric ligand-gated ion channel. *Nature*. 2011 Jan 20; 469(7330):428–31. <https://doi.org/10.1038/nature09647> PMID: 21248852
9. Brannigan G, LeBard DN, Hénin J, Eckenhoff RG, Klein ML. Multiple binding sites for the general anesthetic isoflurane identified in the nicotinic acetylcholine receptor transmembrane domain. *Proc Natl Acad Sci*. 2010 Aug 10; 107(32):14122–7. <https://doi.org/10.1073/pnas.1008534107> PMID: 20660787
10. Barber AF, Liang Q, Amaral C, Treptow W, Covarrubias M. Molecular mapping of general anesthetic sites in a voltage-gated ion channel. *Biophys J*. 2011 Oct 5; 101(7):1613–22. <https://doi.org/10.1016/j.bpj.2011.08.026> PMID: 21961587

11. LeBard DN, Hénin J, Eckenhoff RG, Klein ML, Brannigan G. General Anesthetics Predicted to Block the GLIC Pore with Micromolar Affinity. *PLoS Comput Biol* [Internet]. 2012 May [cited 2015 Jul 12]; 8(5). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3364936/>
12. Raju SG, Barber AF, LeBard DN, Klein ML, Carnevale V. Exploring Volatile General Anesthetic Binding to a Closed Membrane-Bound Bacterial Voltage-Gated Sodium Channel via Computation. *PLoS Comput Biol*. 2013 Jun 13; 9(6):e1003090. <https://doi.org/10.1371/journal.pcbi.1003090> PMID: 23785267
13. Jayakar SS, Dailey WP, Eckenhoff RG, Cohen JB. Identification of Propofol Binding Sites in a Nicotinic Acetylcholine Receptor with a Photoreactive Propofol Analog. *J Biol Chem*. 2013 Mar 1; 288(9):6178–89. <https://doi.org/10.1074/jbc.M112.435909> PMID: 23300078
14. Barber AF, Carnevale V, Klein ML, Eckenhoff RG, Covarrubias M. Modulation of a voltage-gated Na<sup>+</sup> channel by sevoflurane involves multiple sites and distinct mechanisms. *Proc Natl Acad Sci*. 2014 May 6; 111(18):6726–31. <https://doi.org/10.1073/pnas.1405768111> PMID: 24753583
15. Kinde MN, Bondarenko V, Granata D, Bu W, Grasty KC, Loll PJ, et al. Fluorine-19 NMR and computational quantification of isoflurane binding to the voltage-gated sodium channel NaChBac. *Proc Natl Acad Sci*. 2016 Nov 29; 113(48):13762–7. <https://doi.org/10.1073/pnas.1609939113> PMID: 27856739
16. Arcario MJ, Mayne CG, Tajkhorshid E. A membrane-embedded pathway delivers general anesthetics to two interacting binding sites in the *Gloeobacter violaceus* ion channel. *J Biol Chem*. 2017 Jun 9; 292(23):9480–92. <https://doi.org/10.1074/jbc.M117.780197> PMID: 28420728
17. Barber AF, Liang Q, Covarrubias M. Novel activation of voltage-gated K<sup>+</sup> channels by sevoflurane. *J Biol Chem*. 2012 Oct 4; 287:40425–32. <https://doi.org/10.1074/jbc.M112.405787> PMID: 23038249
18. Liang Q, Anderson WD, Jones ST, Souza CS, Hosoume JM, Treptow W, et al. Positive Allosteric Modulation of Kv Channels by Sevoflurane: Insights into the Structural Basis of Inhaled Anesthetic Action. *PLoS ONE*. 2015 Nov 24; 10(11):e0143363. <https://doi.org/10.1371/journal.pone.0143363> PMID: 26599217
19. Long SB, Campbell EB, MacKinnon R. Crystal structure of a mammalian voltage-dependent Shaker family K<sup>+</sup> channel. *Science*. 2005; 309:897–903. <https://doi.org/10.1126/science.1116269> PMID: 16002581
20. Stock L, Souza C, Treptow W. Structural Basis for Activation of Voltage-Gated Cation Channels. *Biochemistry*. 2013 Mar 5; 52(9):1501–13. <https://doi.org/10.1021/bi3013017> PMID: 23317082
21. Stock L, Hosoume J, Treptow W. Concentration-Dependent Binding of Small Ligands to Multiple Saturable Sites in Membrane Proteins. *Sci Rep*. 2017 Jul 18; 7(1):5734. <https://doi.org/10.1038/s41598-017-05896-8> PMID: 28720769
22. Bu W, Liang Q, Zhi L, Maciunas L, Loll PJ, Eckenhoff RG, et al. Sites and Functional Consequence of Alkylphenol Anesthetic Binding to Kv1.2 Channels. *Mol Neurobiol*. 2017 Feb 15; 1–11.
23. Woll KA, Peng W, Liang Q, Zhi L, Jacobs JA, Maciunas L, et al. Photoaffinity Ligand for the Inhalational Anesthetic Sevoflurane Allows Mechanistic Insight into Potassium Channel Modulation. *ACS Chem Biol*. 2017 May 19; 12(5):1353–62. <https://doi.org/10.1021/acscchembio.7b00222> PMID: 28333442
24. Barber AF, Carnevale V, Raju SG, Amaral C, Treptow W, Klein ML. Hinge-bending motions in the pore domain of a bacterial voltage-gated sodium channel. *Biochim Biophys Acta*. 2012 May 9; 1818(9):2120–5. <https://doi.org/10.1016/j.bbame.2012.05.002> PMID: 22579978
25. Deng N, Forli S, He P, Perryman A, Wickstrom L, Vijayan RSK, et al. Distinguishing Binders from False Positives by Free Energy Calculations: Fragment Screening Against the Flap Site of HIV Protease. *J Phys Chem B*. 2015 Jan 22; 119(3):976–88. <https://doi.org/10.1021/jp506376z> PMID: 25189630
26. Arcario MJ, Mayne CG, Tajkhorshid E. Atomistic Models of General Anesthetics for Use in *in Silico* Biological Studies. *J Phys Chem B*. 2014 Oct 23; 118(42):12075–86. <https://doi.org/10.1021/jp502716m> PMID: 25303275
27. Chipot C, Pohorille A, editors. *Free Energy Calculations: Theory and Applications in Chemistry and Biology*. Springer; 2007. 536 p.
28. Roux B. The membrane potential and its representation by a constant electric field in computer simulations. *Biophys J*. 2008 Nov 1; 95(9):4205–16. <https://doi.org/10.1529/biophysj.108.136499> PMID: 18641071
29. Treptow W, Tarek M. Environment of the gating charges in the Kv1.2 Shaker potassium channel. *Biophys J*. 2006 May 1; 90(9):L64–66. <https://doi.org/10.1529/biophysj.106.080754> PMID: 16533847
30. Delemotte L, Tarek M, Klein ML, Amaral C, Treptow W. Intermediate states of the Kv1.2 voltage sensor from atomistic molecular dynamics simulations. *Proc Natl Acad Sci U S A*. 2011 Apr 12; 108(15):6109–14. <https://doi.org/10.1073/pnas.1102724108> PMID: 21444776
31. Henrion U, Renhorn J, Börjesson SI, Nelson EM, Schwaiger CS, Bjelkmar P, et al. Tracking a complete voltage-sensor cycle with metal-ion bridges. *Proc Natl Acad Sci*. 2012 May 29; 109(22):8552–7. <https://doi.org/10.1073/pnas.1116938109> PMID: 22538811

32. Jensen MØ, Jogini V, Borhani DW, Leffler AE, Dror RO, Shaw DE. Mechanism of Voltage Gating in Potassium Channels. *Science*. 2012 Apr 13; 336(6078):229–33. <https://doi.org/10.1126/science.1216533> PMID: 22499946
33. Pathak MM, Yarov-Yarovoy V, Agarwal G, Roux B, Barth P, Kohout S, et al. Closing in on the resting state of the shaker K<sup>+</sup> channel. *Neuron*. 2007; 56:124–40. <https://doi.org/10.1016/j.neuron.2007.09.023> PMID: 17920020
34. Vargas E, Bezanilla F, Roux B. In search of a consensus model of the resting state of a voltage-sensing domain. *Neuron*. 2011 Dec 8; 72(5):713–20. <https://doi.org/10.1016/j.neuron.2011.09.024> PMID: 22153369
35. Vargas E, Yarov-Yarovoy V, Khalili-Araghi F, Catterall WA, Klein ML, Tarek M, et al. An emerging consensus on voltage-dependent gating from computational modeling and molecular dynamics simulations. *J Gen Physiol*. 2012 Dec 1; 140(6):587–94. <https://doi.org/10.1085/jgp.201210873> PMID: 23183694
36. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010; 31(2):455–461. <https://doi.org/10.1002/jcc.21334> PMID: 19499576
37. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005; 26(16):1781–1802. <https://doi.org/10.1002/jcc.20289> PMID: 16222654
38. Izaguirre JA, Reich S, Skeel RD. Longer time steps for molecular dynamics. *J Chem Phys*. 1999 May 22; 110(20):9853–64.
39. Darden T, York D, Pedersen L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J Chem Phys*. 1993 Jun 15; 98(12):10089–92.
40. Huang J, MacKerell AD. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J Comput Chem*. 2013 Sep 30; 34(25):2135–45. <https://doi.org/10.1002/jcc.23354> PMID: 23832629
41. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983 Jul 15; 79(2):926–35.
42. Barber AF, Carnevale V, Klein ML, Eckenhoff RG, Covarrubias M. Modulation of a voltage-gated Na<sup>+</sup> channel by sevoflurane involves multiple sites and distinct mechanisms. *Proc Natl Acad Sci*. 2014 May 6; 111(18):6726–31. <https://doi.org/10.1073/pnas.1405768111> PMID: 24753583
43. Lu N, Kofke DA, Woolf TB. Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods. *J Comput Chem*. 2004 Jan 15; 25(1):28–40. <https://doi.org/10.1002/jcc.10369> PMID: 14634991
44. Fiorin G, Klein ML, Hénin J. Using collective variables to drive molecular dynamics simulations. *Mol Phys*. 2013; 111(22–23):3345–62.
45. Roux B, Nina M, Pomès R, Smith JC. Thermodynamic stability of water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study. *Biophys J*. 1996 Aug; 71(2):670–81. [https://doi.org/10.1016/S0006-3495\(96\)79267-6](https://doi.org/10.1016/S0006-3495(96)79267-6) PMID: 8842206
46. Hess B. Convergence of sampling in protein simulations. *Phys Rev E*. 2002 Mar 1; 65(3):031910.
47. Gilson MK, Given JA, Bush BL, McCammon JA. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys J*. 1997 Mar; 72(3):1047–69. [https://doi.org/10.1016/S0006-3495\(97\)78756-3](https://doi.org/10.1016/S0006-3495(97)78756-3) PMID: 9138555
48. Dong H, Fiorin G, Carnevale V, Treptow W, Klein ML. Pore waters regulate ion permeation in a calcium release-activated calcium channel. *Proc Natl Acad Sci*. 2013 Oct 7; 201316969.
49. Souza CS, Amaral C, Treptow W. Electric fingerprint of voltage sensor domains. *Proc Natl Acad Sci*. 2014 Dec 9; 111(49):17510–5. <https://doi.org/10.1073/pnas.1413971111> PMID: 25422443
50. Delemotte L, Dehez F, Treptow W, Tarek M. Modeling membranes under a transmembrane potential. *J Chem Phys B*. 2008; 112:5547–50.