



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA

Lucas Stefan Minuzzi Neumann

Estudo In Silico de Aquaporinas
de *Leishmania major*

CAMPINAS

2020

Lucas Stefan Minuzzi Neumann

Estudo In Silico de Aquaporinas de *Leishmania major*

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

Orientador: Prof. Dr. Munir Salomão Skaf

**O ARQUIVO DIGITAL CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO
DEFENDIDA PELO ALUNO LUCAS STEFAN MINUZZI NEUMANN, E
ORIENTADA PELO PROF. DR. MUNIR SALOMÃO SKAF**

CAMPINAS

2020

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

N397e Neumann, Lucas Stefan Minuzzi, 1995-
Estudos *in silico* de aquaporinas de *leishmania major* / Lucas Stefan Minuzzi Neumann. – Campinas, SP : [s.n.], 2020.

Orientador: Munir Salomão Skaf.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Química.

1. Dinâmica molecular. 2. Teoria de equações integrais. 3. Aquaporinas. I. Skaf, Munir Salomão, 1963-. II. Universidade Estadual de Campinas. Instituto de Química. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: *In silico* study of *leishmania major* aquaporins

Palavras-chave em inglês:

Molecular dynamics

Integral equation theory

Aquaporins

Área de concentração: Físico-Química

Titulação: Mestre em Química na área de Físico-Química

Banca examinadora:

Munir Salomão Skaf [Orientador]

Leandro Martínez

Rodrigo Leandro Silveira

Data de defesa: 20-02-2020

Programa de Pós-Graduação: Química

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

- ORCID do autor: <https://orcid.org/0000-0002-7333-1850>

- Currículo Lattes do autor: <http://lattes.cnpq.br/7946343600006341>

BANCA EXAMINADORA

Prof. Dr. Munir Salomão Skaf (Orientador)

Prof. Dr. Leandro Martínez (UNICAMP)

Dr. Rodrigo Leandro Silveira

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

O arquivo digital corresponde à versão final da Dissertação defendida pelo aluno Lucas Stefan Minuzzi Neumann, aprovada pela Comissão Julgadora em 20 de Fevereiro de 2020.

“Theoretical chemists tend to use the word ‘prediction’ rather loosely to refer to any calculation that agrees with experiment, even when the latter was done before the former” (KARPLUS, M., 2003).

Ao **software livre** e ao **software de código aberto** sem os quais a realização de trabalhos como este não seriam possíveis em países periféricos como o nosso.

Agradecimentos

Cabe aqui reconhecer que, qualquer que possa ser o mérito do trabalho realizado nestes últimos dois anos, este é, em sua maior parte, resultado não somente de meus esforços, mas do trabalho e do apoio de muitos. A todos, os meus mais sinceros agradecimentos. Faço menção especial àqueles que a contribuição foi (e É) essencial durante o período:

- À minha família, sobretudo a minha mãe, pelo amor, apoio e (principalmente) paciência incondicionais. Agradeço também ao meu pai por me convencer, no dia anterior à prova, a deixar de fazer seleção de mestrado da UFSM para ir à Florianópolis realizar a seleção da UNICAMP;
- Ao Prof. Dr. Munir, pela oportunidade, pela orientação durante o período e, igualmente, pelo trabalho frente a PRP em benefício de toda comunidade acadêmica – mesmo que, ao custo de alguns cabelos brancos;
- Aos meus colegas de laboratório, certo de que o convívio sempre tornou a rotina mais leve. Em especial, agradeço à Caroline e ao Rodrigo (a despeito desse compor a banca de avaliação) pela amizade e pela ajuda no desenvolvimento do projeto;
- A todos os amigos que fiz dentro e fora do IQ: obrigado por terem me feito realmente sentir parte da comunidade acadêmica da UNICAMP. Em especial ao Adriano por toda ajuda prestada (que não foi pouca). Ao Vinicius por ter me acompanhado nas aulas, seminários e congressos (e pelas inúmeras folhas e canetas emprestadas). E aos meus amigos de República (e agregados): Rogério, Igor, Tiago, Antônio, Luccas e Iago, Nilson e Alan;
- Aos meus amigos de longa data que, mesmo longe e espalhados “mundo afora”, sempre se fizeram presentes;
- Ao Instituto de Química e à UNICAMP por oferecerem uma estrutura ímpar aos alunos e pós-graduandos;
- Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico/CNPq (131181/2018-1) pela bolsa de estudos;
- À Fundação de Pesquisa do Estado de São Paulo/FAPESP (2013/08293-7) pelo financiamento e recursos computacionais utilizados no trabalho;
- O presente trabalho foi realizado com apoio da Coordenação de Perfeccionamento de Pessoal de Nível Superior – Brasil/CAPES (Código de Financiamento 001).

Resumo

De acordo com a Organização Mundial da Saúde, a leishmaniose, doença causada pelo protozoário *Leishmania spp.*, afeta anualmente mais de 12 milhões de pessoas e está presente em 88 países. Entretanto, ainda hoje, fármacos à base de antimônio constituem a principal forma de tratamento da doença, apesar de exibirem alta toxicidade e de os protocolos de tratamento existentes serem inadequados à realidade das populações de risco – de baixa renda e com pouco ou nenhum acesso ao saneamento básico e à saúde pública. Soma-se a isso o crescente número de relatos de cepas selvagens de *Leishmania spp.* que apresentam resistência a esta classe de drogas. Logo, é clara a necessidade do desenvolvimento de novas vias de combate à doença. Neste sentido, a literatura especializada tem sugerido que as aquaporinas do protozoário seriam um alvo farmacológico interessante a ser estudado, uma vez que estas proteínas, responsáveis por facilitar o transporte passivo de água entre o meio intra e extracelular, também agem como a principal via de entrada do compostos de antimônio no protozoário. Ainda assim, apenas uma das cinco aquaporinas expressas em *Leishmania major* já foi caracterizada quanto sua função celular – ainda que sua estrutura não seja conhecida. Neste contexto, o presente trabalho buscou estudar quatro aquaporinas expressas em *Leishmania major* e uma aquaporina presente em *Plasmodium falciparum* com uso de metodologias *In Silico* de dinâmica molecular (DM) e de teoria molecular de solvatação via equações integrais 3D-RISM – *Reference Site Interaction Model*. Simulações de DM foram empregadas para avaliar propriedades estruturais dos canais e caracterizá-los quanto ao processo de transporte de água no poro. Como técnicas baseadas em DM, aplicadas ao estudo da permeabilidade de espécies diversas, são computacionalmente muito custosas e talvez não se justificam dada à qualidade das estruturas, o 3D-RISM foi empregado como alternativa para computar os perfis energéticos dos processos de permeação da amônia, uréia, metanol, dióxido de carbono, cloreto, íon sódio e hidrônio nos canais das proteínas. A análise das simulações sugere similaridades estruturais entre LmAQP1 e PfAQP1 e entre LmAQP α e LmAQP γ , sendo essas refletidas no transporte de água. Cálculos com 3D-RISM levantam a hipótese que LmAQP1 e PfAQP1 exibam permeabilidade frente a ânions e LmAQP β exiba permeabilidade à amônia. Demais resultados para PfAQP1 e LmAQP1 concordam qualitativamente com resultados experimentais reportados na literatura.

Abstract

According to the World Health Organization, leishmaniasis, a disease caused by the protozoan *Leishmania spp.*, affects more than 12 million people annually and is present in 88 countries. However, even today, antimony-based drugs are the main form of treatment for the disease, despite exhibiting high toxicity and the existing treatment protocols are inadequate to the reality of at-risk populations - low income with little or no access to basic sanitation and public health. Added to this is the growing number of reports of wild *Leishmania spp.* that are resistant to this class of drugs. Therefore, the need to develop new ways to combat the disease is clear. In this sense, the specialized literature has suggested that protozoan aquaporins would be an interesting pharmacological target to be studied, since these proteins, responsible for the transport of water in the body, also act as the main pathway for antimony to enter the protozoan. Even so, only one of the five aquaporins expressed in *Leishmania major* is characterized in terms of its cellular function - although its structure and mechanisms of regulation means are not known. In this context, the present work sought to study four aquaporins expressed in *Leishmania major* and one aquaporin present in *Plasmodium falciparum* using *In Silico* methodologies of molecular dynamics (MD) and molecular solvation theory via integral 3D-RISM equations - reference site interaction model . MD simulations were used to evaluate structural properties of the channels and to characterize them regarding the water transport process in the pore. As MD-based techniques applied to the study of the permeability of other species are computationally expensive, 3D-RISM was used as an alternative to compute the energy profiles of the permeation processes in the protein channels of ammonia, urea, methanol, carbon dioxide, chloride, sodium and hydronium ion. The analysis of the simulations suggests structural similarities between LmAQP1 and PfAQP1 and between LmAQP α and LmAQP γ and these are reflected in the transport of water in the pores. Calculations with 3D-RISM raise the hypothesis that LmAQP1 and PfAQP1 exhibit permeability against anions and LmAQP β exhibits ammonia permeability. Other results for PfAQP1 and LmAQP1 agree qualitatively with experimental results reported in the literature.

Lista de Abreviaturas e Siglas

LmAQP1 – Aquagliceroporina da *Leishmania major*, código no banco de dados

UniProt: Q4Q6R2

PfAQP1 – Aquagliceroporina da *Plasmodium falciparum*, código no PDB: 3C02

LmAQP α – Aquaporina clássica da *Leishmania major*, código no banco de dados

UniProt: Q4Q2F7

LmAQP β – Aquaporina clássica da *Leishmania major*, código no banco de dados

UniProt: Q4QBK5

LmAQP γ – Aquaporina clássica da *Leishmania major*, código no banco de dados

UniProt: Q4Q537

4NEF – Aquaporina 2 Humana

5I32 – *Arabidopsis thaliana* aquaporina TIP2

AQPs – Aquaporinas

DM – Dinâmica Molecular

KH – Relação de Fechamento de Kovalenko-Hirata

MC – Monte Carlo

MDIIS – Algoritmo Modificado da Inversão Direta do Espaço Interativo

PME – *Particle Mesh Ewald*

POPC – 1-Palmitoil-2-oleoilfosfatidilcolina

RISM – *Reference Interaction Site Model*

Sumário

1 Introdução.....	13
1.1 Leishmanias e Leishmaniose	13
1.2 Aquaporinas.....	15
1.2.1 Estrutura.....	16
1.2.2 Permeabilidade das Aquaporinas	19
1.2.3 Aquaporinas da <i>Leishmania Major</i>	19
2 Dinâmica molecular.....	21
2.1 Campos de força.....	22
2.2 Equações do Movimento e Métodos de Integração.....	25
2.3 Controle de Temperatura e Pressão.....	25
2.4 Condições Periódicas de Contorno.....	26
2.5 Análises de Trajetórias.....	27
2.5.1 Flutuações Estruturais.....	27
2.5.2 Energia Livre.....	28
2.5.3 Cálculos de Permeabilidade.....	29
3 <i>Reference Site Interaction Model</i>	32
3.1 Funções de Correlação.....	32
3.2 1D-RISM.....	34
3.3 3D-RISM.....	35
4 Métodos.....	37
4.1 Modelagem Estrutural.....	37
4.2 Construção do Sistema.....	37
4.3 Simulações de Dinâmica Molecular.....	38
4.4 Análises de Trajetória.....	38
4.5 RISM.....	39
4.5.1 1D-RISM.....	39

4.5.2 3D-RISM.....	40
5 Estudo Comparativo de Aquagliceroporinas de <i>Leishmania major</i> e <i>Plasmodium falciparum</i>	41
5.1 Descrição e Análise Estrutural.....	41
5.2 Dinâmica Molecular.....	45
5.2.1 Estabilidade.....	45
5.2.2 Permeabilidade à Água.....	49
5.3 3D-RISM.....	52
5.4 Conclusões.....	57
6 Aquaporinas clássicas da <i>Leishmania major</i>	59
6.1 Descrição e Análise Estrutural.....	60
6.2 Dinâmica Molecular.....	62
6.2.1 Estabilidade.....	62
6.2.2 Permeabilidade à Água.....	66
6.3 3D-RISM.....	69
6.4 Conclusões.....	73
7 Considerações Finais.....	74
Referências Bibliográficas.....	76

Capítulo 1

Introdução

1.1 Leishmanias e Leishmaniose

De acordo com a Organização Mundial da Saúde, a leishmaniose afeta anualmente mais de 12 milhões de pessoas e está presente em 88 países (PONTE-SUCRE, 2017). Apesar disso, formas de tratamento da doença pouco mudaram nos últimos 50 anos, sendo que ainda hoje medicamentos à base de antimônio são a principal via de tratamento da doença. Estes, por possuírem fortes efeitos colaterais e necessitarem de diversas aplicações por via parental e por longos períodos (20 a 30 dias), são inadequados à realidade das principais populações de risco – classes menos abastadas e com acesso precário ao atendimento médico. Igualmente, é preocupante a existência de cepas selvagens de *Leishmania spp.* resistentes a esta classe de drogas. Diante disso, é clara a necessidade de estudos que busquem novas vias de combate a doença (PONTE-SUCRE, 2017).

Todavia, a busca de novas formas de tratamentos constitui um desafio bastante complexo. Uma vez que a forma virulenta do parasita no ser humano – forma amastigota – aloja-se no fagossomo de células macrófago. Portanto, qualquer droga que tenha como alvo o protozoário deve: (1) ser absorvida na corrente sanguínea do paciente; (2) atravessar a membrana celular do macrófago; (3) atravessar a membrana do fagossomo; (4) não ser desativada pelo meio de baixo pH do fagossomo; (5) e, finalmente, atravessar a membrana da *Leishmania spp.* e ligar-se seletivamente ao alvo. Ainda assim, há na literatura um crescente interesse no estudo de proteínas de membrana de parasitas como um alvo farmacológico potencial. Uma vez que, além da importância biológica desempenhada por proteínas, o desenvolvimento de resistência à drogas deste tipo é muito mais lento (KUN, 2009).

Entre as proteínas que tem despertado particular interesse como alvo destacam-se a família das aquaporinas, proteínas de membrana que facilitam o transporte passivo da água, visto que uma das aquaporinas da *Leishmania major*,

LmAQP1, constitui a principal via de entrada dos medicamentos antimoniais na célula (como ilustrado na Figura 1.1). Além disso, a LmAQP1 também participa no transporte de outras espécies químicas e influi na osmotaxia (FRÉZARD, 2014; FIGARELLA, 2007).

Sabe-se que aquaporinas de protistas realizam um número maior de funções que as expressas em mamíferos, o que, possivelmente, indica a existência de diferenças estruturais significativas que poderiam ser exploradas. No entanto, pouco é conhecido sobre a estrutura e função das demais aquaporinas expressas por *Leishmania spp.* (FIELD, 2017; SONG, 2014).

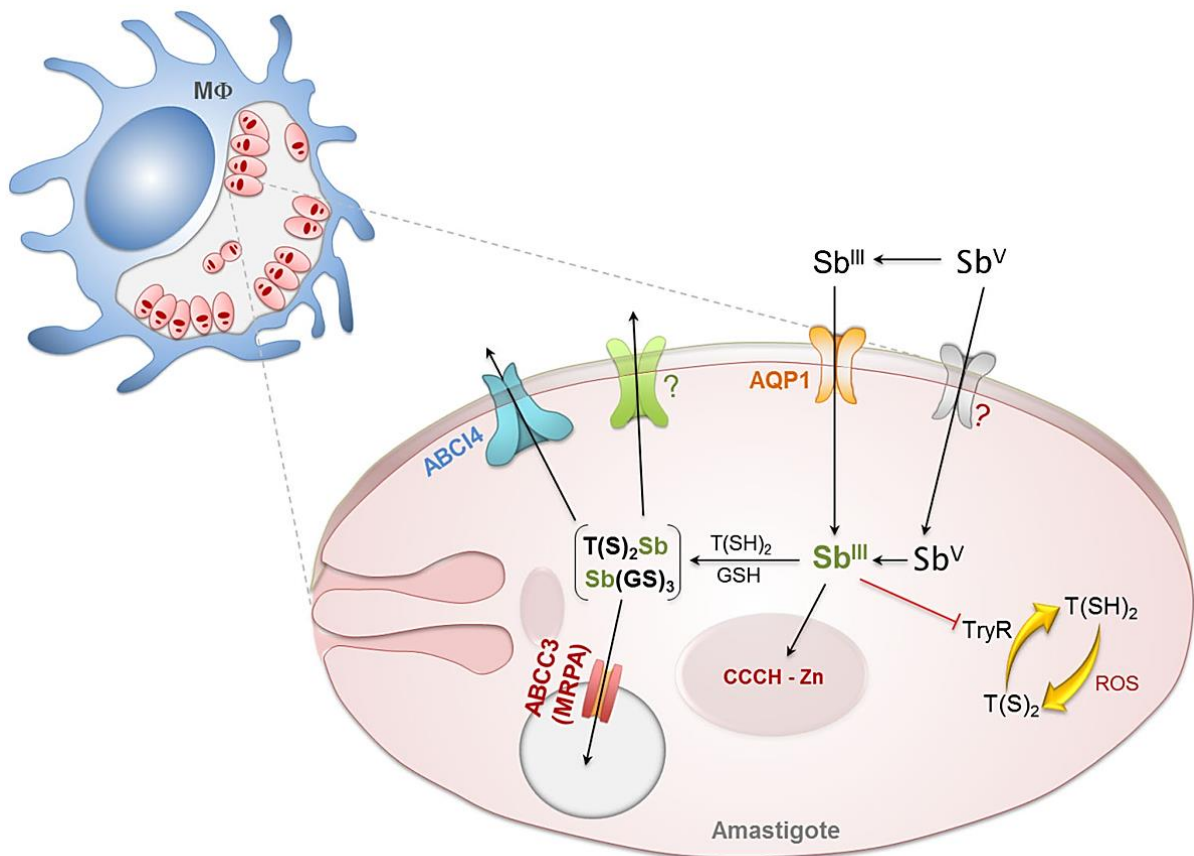


Figura 1.1 Ilustração das propriedades farmacológicas das drogas antimoniais, incluindo metabolismo e rota de transporte no organismo do parasita na sua forma amastigota. Retirado de Frézard, 2014.

1.2 Aquaporinas

Aquaporinas constituem uma família de proteínas integrais de membrana (MIPs) presentes em virtualmente todos os organismos, sendo característica comum a todos os membros da família o transporte passivo de água no canal proteico concomitante ao bloqueio da passagem de prótons. Ademais, mesmo que a estrutura secundária seja altamente conservada, sabe-se que membros da família podem realizar outras funções, como o transporte de glicerol, peróxido de hidrogênio, amônia, entre outros, e possuir mecanismos de regulação diversos (MAUREL, 2009). É também comum que um organismo expresse diversas aquaporinas, sendo 13 presentes em mamíferos e mais de 30 em algumas espécies de plantas (KALDENHOFF, 2006).

A análise filogenética da família aponta uma divergência evolucionária antiga, separando as aquaporinas em duas grandes subfamílias: aquaporinas clássicas, permeáveis exclusivamente à água, e aquagliceroporinas, permeáveis a glicerol e outros solutos de pequena massa molecular. Embora estruturalmente semelhantes, membros dessas famílias exibem uma identidade de aminoácidos relativamente baixa entre si, normalmente em torno de 20% (Tabela 1). Existem na literatura alguns modelos que buscam explicar a baixa semelhança e a grande diversidade da família: um deles sendo o surgimento das aquaporinas pela evolução convergente conjunta de diversos linhagens originais de diferentes proteínas (GIORNO, 2014); outra possível explicação seria que a baixa identidade observada é resultado da identificação errônea de algumas sequências de aminoácidos nos bancos de dados (o que não é um algo incomum) (ALLEVA, 2012).

As aquaporinas clássicas ainda podem ser divididas em subgrupos (o conjunto é por vezes chamado de super-aquaporinas). A exemplo: TIPs (*tonoplast intrinsic proteins*), PIPs (*plasma membrane intrinsic proteins*) e NIPs (*nodulin26-like intrinsic protein*) (ABASCAL, 2013).

1.2.1 Estrutura

Estruturalmente, a família é caracterizada pela presença de seis α -hélice e dois *loops* altamente conservados. Na Figura 1.2 é possível distinguir as seis α -hélices trans-membranares e os dois *loops* que visualmente formam a 'sétima α -hélice' da estrutura. A estrutura desses *loops*, metade α -hélice e metade sem estrutura bem definida (*random coil*), é central para a ação do canal, atuando como uma barreira à passagem de prótons.

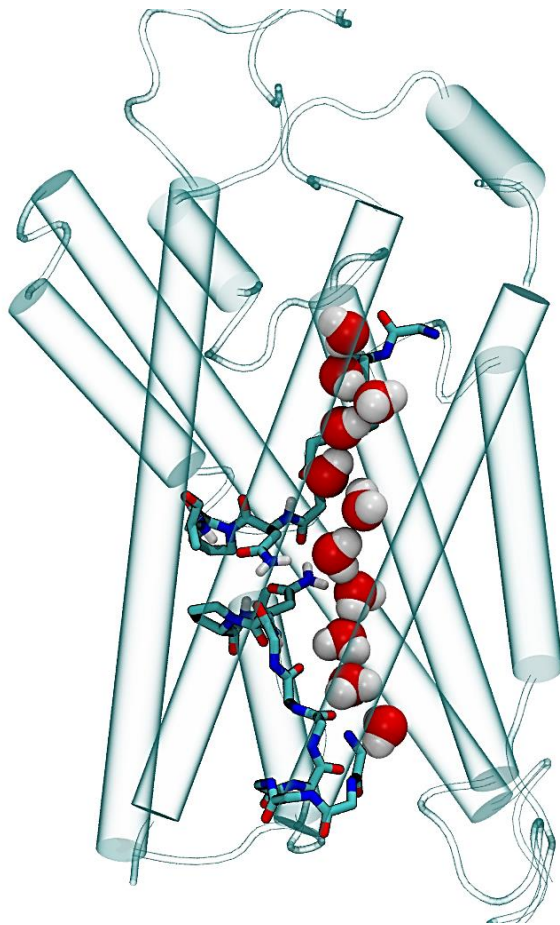


Figura 1.2 Representação estrutural da Aquagliceroporina do *Plasmodium falciparum* (PfAQP1) com os dois motivos NPA e parte da cadeia principal em destaque.

O enovelamento global da proteína produz uma estrutura que se assemelha a uma ampulheta com um canal que perpassa a proteína e tem porção mais estreita no ponto médio. Ambos C e N terminais encontram-se sempre no lado intracelular do poro. O canal em si é constituído em sua maior parte por resíduos hidrofóbicos, com a exceção de duas regiões:

- **Motivo NPA (asparagina - prolina - alanina):** talvez a característica estrutural mais marcante da família seja a presença de dois motivos NPA altamente conservados. A região é responsável pela exclusão da passagem de prótons pelo canal. Dois são os fatores estruturais que contribuem para esse efeito: (1) a barreira eletrostática criada pelo dipolo das α hélices dos dois *loops*, (2) a posição das asparaginas, que força as moléculas do canal a uma orientação que desfavorece mecanismo de *Grotthus* de transporte de prótons em meio aquoso (BURYKIN, 2003) (Figura 1.3).
- **Filtro seletivo (SF):** está localizado próximo à porção extracelular e compõe a região mais estreita do poro. A região SF é caracterizada pela presença de resíduos carregados, na maioria dos casos uma arginina (Figura 1.3). O grau de variação apresentado na região é atribuído ao fato desta ser determinante para o controle da seletividade no poro frente a outras pequenas moléculas que não a água. A exemplo, aquagliceroporinas tem seu SF composto de arginina e resíduos aromáticos; já aquaporinas clássicas possuem a mesma arginina associada a uma histidina, o que torna o filtro muito mais estreito e, portanto, restritivo à passagem do esqueleto carbônico do glicerol, por exemplo (HUB, 2008).

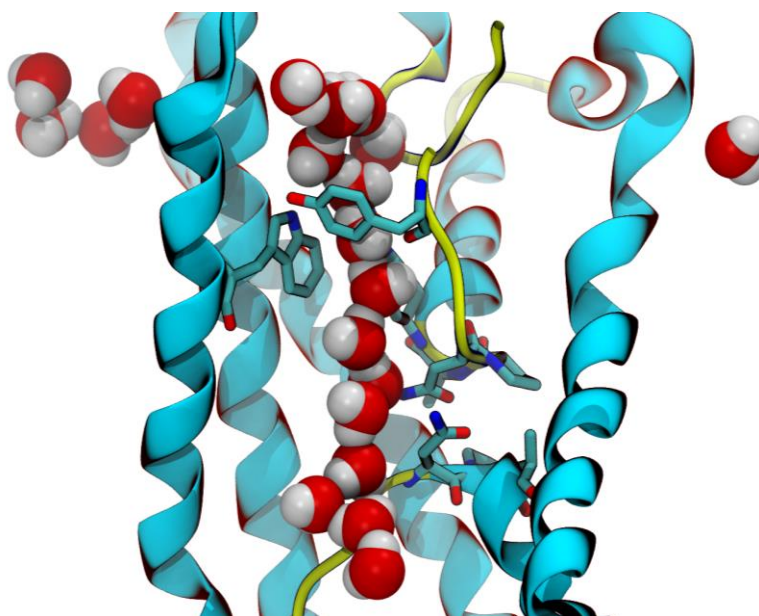


Figura 1.3 Resíduos do SF acima e dois motivos NPA abaixo. Orientação dos dipolos das moléculas de água dentro do canal de uma aquaporina. No mecanismo de Grotthuss os dipolos da molécula de água devem estar alinhados. Dentro do poro da AQP a condição não é satisfeita devido à conformação das asparaginas dos NPAs, que orienta a molécula de água de forma a impedir o 'salto' dos prótons.

Aquaporinas podem ainda exibir uma histidina conservada próxima à porção citosólica. Esse resíduo exibe uma mobilidade considerável de cadeia lateral, e a literatura indica que o estado de protonação do resíduo está associado ao mecanismo de oclusão do canal, em resposta a mudanças de pH citosólico. Sendo a oclusão do canal ocasionada pela ancoramento de um *loop* do N-terminal da proteína na forma protonada da histidina em questão (SACHDEVA, 2014).

Além do controle por pH, alguns membros da família ainda exibem alterações conformacionais em resposta à concentração citosólica de Ca^{+2} . Sabe-se também que motivos de fosforilação estão associados a uma transição entre os estados aberto e fechado do poro (FREITES, 2019; TORNROTH-HORSEFIELD, 2006).

A unidade funcional dessa família de proteínas é a de um tetrâmero (Figura 1.4). Sendo que o tetrâmero pode ser composto de diferentes aquaporinas, podendo a estequiometria de conjunto influir ou não nas propriedades do canal (JOZEFKOWICZ, 2016). Especula-se também que o poro interno formado pelo tetrâmero possa possuir função biológica (JOZEFKOWICZ, 2016).

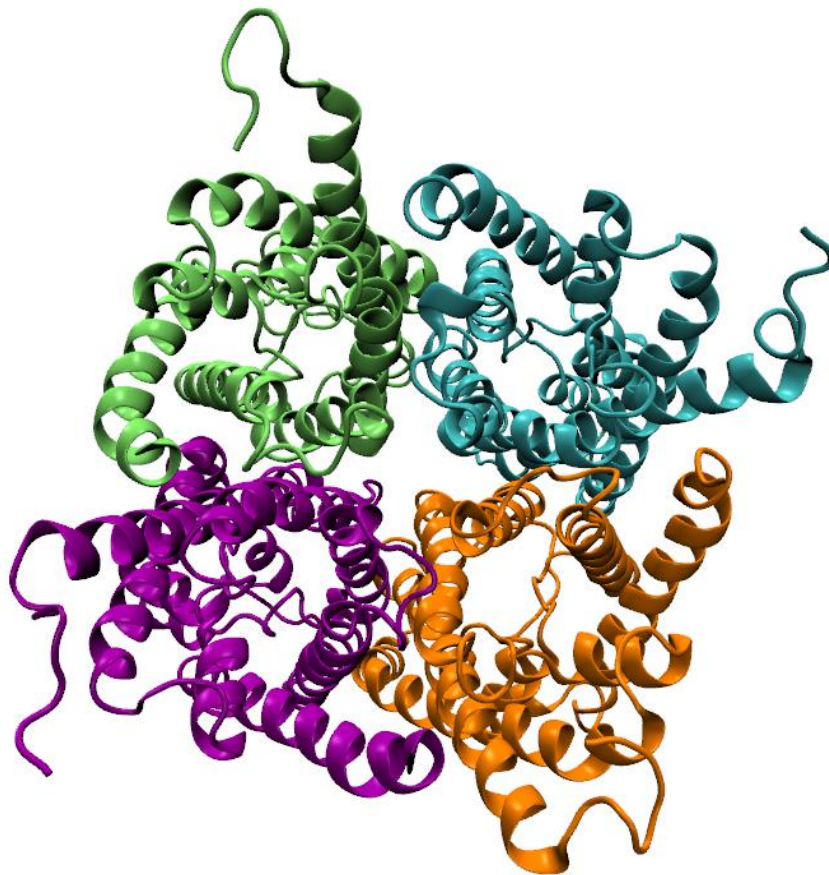


Figura 1.4 Representação do tetrâmero da Aquaporina LmAQP1.

1.2.2 Permeabilidade das Aquaporinas

O canal das AQP's pode ser permeável a diversas espécies químicas, entre elas citam-se: amônia, uréia, peróxido de hidrogênio, óxido nitroso, dióxido de carbono, oxigênio, $\text{As}(\text{OH})_3$, $\text{Sb}(\text{OH})_3$, boratos, alguns ânions (nitrato, cloreto, etc.), ácidos orgânicos, álcoois e açúcares pequenos (GIORGIO, 2014). Destas, amônia, ureia e glicerol parecem possuir maior relevância biológica. Por outro lado, o transporte de dióxido de carbono, ânions e boratos é característico de aquaporinas expressas em plantas (KALDENHOFF, 2006). Entende-se que a permeabilidade de espécies de arsênio e antimônio estejam relacionadas com a permeabilidade do canal ao glicerol, o que provavelmente decorre destas espécies apresentarem volumes semelhantes (FRÉZARD, 2014).

1.2.3 Aquaporinas da *Leishmania major*

Quanto às aquaporinas expressas por *Leishmania major*, sabe-se que a LmAQP1 é permeável a glicerol, $\text{As}(\text{OH})_3$ e $\text{Sb}(\text{OH})_3$ e tem um papel importante na osmotaxia (FIGARELLA, 2007). No tocante às demais (LmAQP α - δ), a literatura apenas nota que essas se assemelham às aquaporinas expressas em plantas (5132 na Tabela 2). A aquaporina LmAQP δ não sendo expressa em níveis significativos na forma promastigota do parasita, e as aquaporinas LmAQP α e LmAQP γ parecem apresentar motivos de fosforilação em duas tirosinas (BIYANI, 2011). Ao que indica a literatura, somente a LmAQP1 é presente na membrana celular do protistas; as demais aquaporinas são subcelulares (BIYANI, 2011).

LmAQP1 por sua vez possui um elevado grau de semelhança com PfAQP1, uma aquagliceroporina expressa em *Plasmodium falciparum*. A PfAQP1 apresenta permeabilidade a diversas espécies. Estudos mostram que o canal é capaz de transportar água, glicerol e amônia, $\text{As}(\text{OH})_3$, $\text{Sb}(\text{OH})_3$ e, em menor grau, ureia (APONTE-SANTAMARÍA, 2010). O grau de identidade é significativamente maior para com a LmAQP1 (Tabela 1) do que com as demais aquaporinas de *Leishmania*, visto que ambas são aquagliceroporinas.

Grau de Identidade das Aquaporinas Estudadas								
(%)	LmAQP 1	LmAQP α	LmAQP β	LmAQP γ	LmAQP δ	3C02	4NEF	5I32
LmAQP1	314 a.a.	22.27	20.77	18.87	17.54	32.24	28.32	24.15
LmAQPα	22.27	230 a.a.	40.72	41.82	43.61	21.93	21.10	26.48
LmAQPβ	20.77	40.72	278 a.a.	36.82	33.21	22.03	22.07	24.90
LmAQPγ	18.87	41.82	36.82	301 a.a.	32.99	18.86	27.36	24.90
LmAQPδ	17.54	43.61	33.21	32.99	593 a.a.	20.34	22.81	23.57
PfAQP1	32.24	21.93	22.03	18.86	20.34	258 a.a.	24.05	21.79
4NEF	28.32	21.20	22.07	27.36	22.81	24.05	242 a.a.	37.50
5I32	24.15	26.48	24.90	24.90	23.57	21.79	37.50	275 a.a.

Tabela 1 Grau de identidade (%) entre aquaporinas presentes em *Leishmania major* e três outras aquaporinas com estrutura cristalográfica resolvida no PDB (3C02(PfAQP1), 4NEF e 5I32). Em destaque estão as estruturas cristalográficas com maior grau de identidade para com as AQPs da *Leishmania major*.

O conhecimento a nível molecular das características estruturais, do funcionamento e da regulação de proteínas é, hoje, indispensável quando se almeja investigar problemas complexos como o desenvolvimento de novos fármacos ou mesmo no estudo de mecanismos fisiológicos. Neste contexto, buscou-se caracterizar quatro aquaporinas presentes em *Leishmania major* – LmAQP1, LmAQP α - γ – quanto a sua funcionalidade e características estruturais. Embora os resultados aqui apresentados tenham cunho apenas qualitativo, espera-se que este primeiro "screening" proporcione indícios quanto à função biológica exercida por estas proteínas no protozoário.

Capítulo 2

Dinâmica molecular

Dinâmica molecular (DM) é, em essência, uma técnica computacional que propaga iterativamente as equações do movimento de um sistema clássico de n -corpos por meio de integração numérica, dado que se saiba computar as forças entre as partículas. Técnicas de DM são amplamente empregadas no estudo de biomoléculas em escala atômica. Além disso, contrário a métodos baseados em Monte Carlo, a DM se presta particularmente bem ao estudo de propriedades dependentes do tempo – e.g. difusão.

Em 1975, Levitt e Warshel publicaram a primeira simulação de DM de uma proteína: a *Bovine Pancreatic Trypsin Inhibitor* – um sistema de ~700 átomos simulado por alguns picosegundos e com uso de solvente implícito. Desde então, a lei de Moore tem beneficiado enormemente a técnica, de modo que hoje são comuns na literatura estudos com mais de milhão de átomos e/ou na escala dos milissegundos (GOOSSENS, 2019; DE VIVO, 2016) – o que talvez fosse impensável em 1957, quando Alder e Wainwright conceberam a DM para calcular o coeficiente virial de um sistema de 108 esferas rígidas.

Esses primeiros estudos de DM de proteínas detiam-se em questões sobre seu comportamento dinâmico e estabilidade estrutural (mobilidade, enovelamento, alosteria, etc.), sendo, já nesta época, formulados muitos dos conceitos de termodinâmica de proteínas modernos (COOPER, 1976; BROOKS, 1983). Ainda assim, nos primeiros anos, a técnica já mostrava sua versatilidade quando, a exemplo, em 1979 foi utilizada para propor dois novos mecanismos para a ligação do oxigênio na mioglobina (CASE, 1979). Todavia, no contexto do presente trabalho, talvez as conquistas mais relevantes da DM sejam nos estudos com aquaporinas do início dos anos 2000, que esclareceram a dinâmica do poro e como se dava a exclusão da passagem de prótons no canal (SCHULTEN, 2004; GROOT, 2005). A família foi identificada em 1992 e a primeira estrutura cristalográfica data de 1999; na época o tetrâmero da aquaporina era o maior sistema biológico a ser simulado com resolução atômica.

Mesmo com os avanços recentes na técnica e a crescente capacidade computacional, a DM ainda possui severas limitações: uma das principais sendo as das escalas de tempo simuladas que são ordens de magnitudes aquém das necessárias para observar boa parte dos processos bioquímicos/biológicos (KARPLUS, 2003; VAN GUNSTEREN, 2018).

Existem excelentes textos que descrevem as mais variadas técnicas de DM de modo abrangente (LEACH, 2001; FRENKEL, 2001; GRIEBEL, 2007). Aqui se faz apenas uma breve exposição sobre os fundamentos das simulações de DM. Organiza-se o restante do capítulo nos seguintes tópicos: (1) campos de força; (2) equações do movimento e métodos de integração; (4) controle de temperatura e pressão; (5) condições de contorno periódicas e (6) análises das trajetórias.

2.1 Campos de Força

Existe um grande número de formas de representar sistemas de químicos/biofísicos (*coarse-grain*, *united atoms*, *all atoms*, etc.) e de descrever os potenciais de interação entre as partículas de tais sistemas (*classical FF*, *polarizable FF*, *reaxFF*, etc.); aqui, descreve-se somente o *all-atoms force field* na forma como está implementado no campo de força CHARMM36m (HUANG, 2017).

Nota-se que os parâmetros utilizados nos cálculos das forças entre pares de partículas (Figura 2.1) são derivados empiricamente, normalmente esses são ajustados de maneira a reproduzir um conjunto de dados experimentais ou cálculos *ab initio* (MACKERELL, 2004).

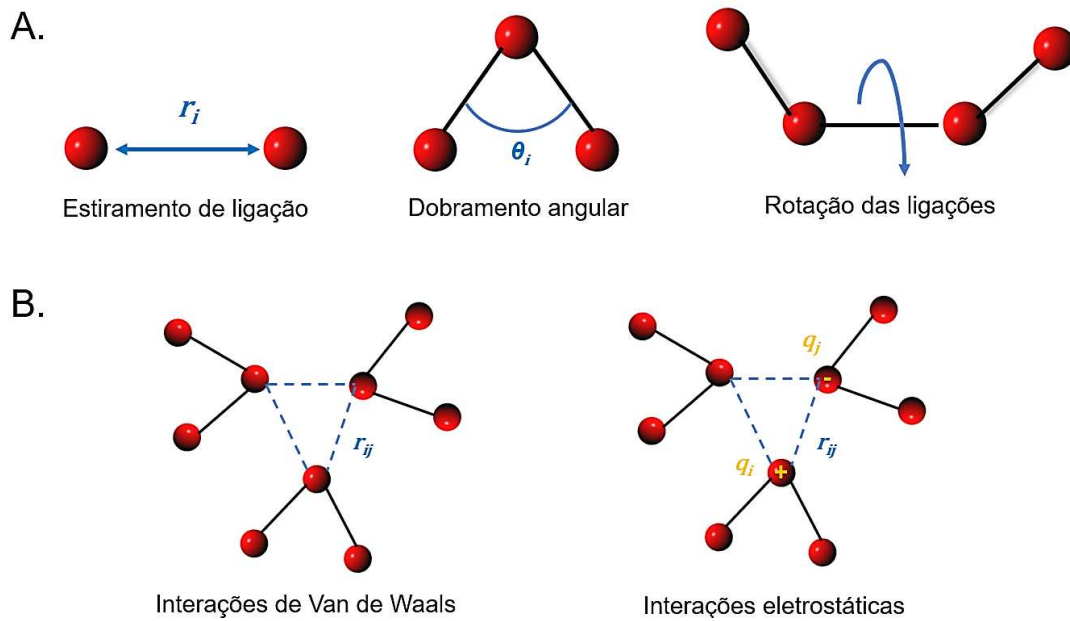


Figura 2.1. Representação gráfica das cinco principais contribuições para o campo de força, sendo: (A) interações covalentes e (B) interações não-covalentes.

Considera-se a energia potencial (V_{total}) do sistema como composta por dois termos:

$$V_{\text{total}}(r^n) = V_{\text{ligado}}(r^n) + V_{\text{não-ligado}}(r^n)$$

sendo r^n o vetor que representa a posição de todos os átomos do sistema. Como a técnica de DM utilizada aqui não considera quebra de ligações, a informação sobre a conectividade (topologia) dos átomos deve ser fornecida de antemão. Na Figura 2.1 são representados os termos dos potenciais ligados e não-ligados. Os termos ligados podem ser separados em:

$$V_{\text{ligado}}(r^n) = V_{\text{vibracional}}(r^n) + V_{\text{angular}}(r^n) + V_{\text{torcional}}(r^n)$$

onde cada potencial, definido para pares de átomos, é:

$$V_{\text{vibracional}}(r) = k_r (r - r_o)^2$$

$$V_{\text{angular}}(\theta) = k_\theta (\theta - \theta_o)^2$$

$$V_{\text{torcional}}(\phi) = k_\phi (1 + \cos(n\phi - \delta))$$

Onde k_r e k_θ são as constantes harmônicas de estiramento e deformação angular, r_0 o valor de equilíbrio do comprimento da ligação, θ_0 ângulo de equilíbrio da ligação, k_ϕ é o valor associado à altura das barreiras rotacionais, n indica a periodicidade das barreiras e δ é o ângulo de fase. Também são definidos potenciais torcionais impróprios para impor planaridade a determinados tipos de diedros (e.g. carbonilas, anéis aromáticos etc).

Por sua vez os termos não-ligados são separados em duas contribuições:

$$V_{\text{não-ligado}}(r^n) = V_{LJ}(r^n) + V_{\text{coulomb}}(r^n)$$

$$V_{LJ}(r_{ij}) = \epsilon \left[\left(\frac{R_{\text{mim}}}{r_{ij}} \right)^{12} - \left(\frac{R_{\text{min}}}{r_{ij}} \right)^6 \right]$$

$$V_{\text{coulomb}}(r_{ij}) = \frac{q_i q_j}{\epsilon} \left(\frac{1}{r_{ij}} \right)$$

As interações de Lennard-Jones (LJ) são interações de curto alcance e não são computadas para distâncias maiores que um valor de corte pré-estabelecido (*cutoff*). Já o tratamento das interações eletrostáticas requer métodos específicos devido ao longo alcance espacial das forças de Coulomb relativamente às forças dispersivas. Para tal, um dos métodos mais utilizados (e adotado neste trabalho) é o *Particle Mesh Ewald* (PME), que decompõe as interações eletrostáticas em um termo no espaço direto de curto alcance e uma série de mais rápida convergência no espaço recíproco (DARDEN, 1993). A principal vantagem deste método é possuir uma complexidade computacional de $O(n) = n \log(n)$.

Interações de três ou mais corpos, polarizabilidade e outros efeitos eletrônicos não são representados no campo de força CHARMM36m. Há campos de força que introduzem termos para tratar tais forças, mas isso tende a comprometer sua escalabilidade com o tamanho do sistema. Ressalta-se, todavia, que nenhum campo de força é uma representação fidedigna da natureza e, portanto, o mesmo deve ser escolhido baseado não na quantidade de termos utilizados, mas em quais termos (e escalas de tempo) são relevantes para o fenômeno a ser estudado.

2.2 Equações do Movimento e Métodos de Integração

O método de integração numérica *Velocity-Verlet* é um dos métodos comumente empregados para computar trajetórias integrando as equações do movimento de Newton. Dado que se pode determinar a aceleração de partículas no instante t pela equação:

$$M a(r^n) = -\nabla V(r^n)$$

sendo as posições e velocidades conhecidas no instante t , o algoritmo de *verlet-velocity* consiste em:

- (1) calcular a velocidade no instante $v(t + \frac{1}{2}\Delta t) = v(t) + \frac{1}{2}a(t)\Delta t$;
- (2) calcular as posições $r(t + \Delta t) = r(t) + v(t + \frac{1}{2}\Delta t)\Delta t$;
- (3) calcular $a(t + \Delta t)$ partindo das posições calculadas na etapa anterior – $r(t + \Delta t)$;
- (4) finalmente, calcular $v(t + \Delta t) = v(t + \frac{1}{2}\Delta t) + \frac{1}{2}a(t + \Delta t)\Delta t$.

Esse método é preferível a outros esquemas de integração (como o método de Euler) por ser numericamente estável e o erro na posição ser da ordem de $(\Delta t)^4$ e erro associado às velocidades de $(\Delta t)^2$ (LEACH, 2002).

2.3 Controle de Temperatura e Pressão

Trajетórias obtidas por simulação de DM convencional correspondem a um *ensemble microcanônico (NVE)*. No entanto, para sistemas de interesse na modelagem de biomoléculas é mais conveniente trabalhar no *ensemble NTP*. Portanto, para gerar trajetórias com temperatura e pressão constante, é necessário introduzir formas de controle dessas grandezas ao computar a trajetória. Para o controle de temperatura uma das alternativas é modificar as equações do movimento de Newton:

$$Mr'' = -\nabla V(r) - \gamma r'(t) + \sqrt{2\gamma k_b T} f(t)$$

Esta equação diferencial corresponde à dinâmica de Langevin, onde r , r' e r'' são o vetor da posição dos átomos do sistema e sua primeira e segunda derivada em relação ao tempo, respectivamente. $V(r)$ é a energia potencial, γ é o coeficiente de amortecimento – o segundo termo é, portanto, uma força de amortecimento – e $f(t)$ uma função randômica com distribuição de probabilidade de uma gaussiana centrada na origem – $\langle f(t) \rangle = 0$. Assim, a introdução do termo estocástico resulta em um ensemble de velocidades que flutua próximo do valor médio para a temperatura T .

Para controle da pressão, emprega-se o pistão de Langevin que introduz um grau de liberdade adicional ao sistema correspondente ao volume da caixa. Este grau de liberdade funciona efetivamente como um ‘pistão’ que varia as dimensões do sistema e rescala a posição dos átomos a fim de a manter a pressão flutuando próximos a um valor preestabelecido. De forma similar ao controle de temperatura, introduz-se às equações de movimento do ‘pistão’ um termo de amortecimento e um termo estocástico.

Uma descrição mais rigorosa de ambos métodos pode ser encontrada nas referências (BERENDSEN, 1976; BROOKS, 1995).

2.4 Condições de Contorno Periódicas

Dado que os sistemas de interesse contêm na ordem de 10^4 - 10^6 átomos, a imposição de bordas a fim de restringir o movimento das partículas / fixar o volume (seja por ou uma parede rígida ou um potencial harmônico) introduziria artefatos sem sentido físico na simulação. Isso é consequência de que, nestes sistemas, uma porção significativa dos átomos está em contato com as bordas da caixa e, portanto, não representam adequadamente o ‘bulk’, porção que de fato interessa ser simulada (ou, em outras palavras, a razão $\frac{\text{área de superfície}}{\text{volume do sistema}}$ é alta). Uma solução elegante para o problema foi introduzida já em 1955 por Alder e Lewison: aplicar condições periódicas de contorno ao sistema. O método consiste em, quando uma molécula ultrapassa os limites da caixa, ‘transportá-la’ para o outro lado da caixa.

Condições de contorno periódicas são trivialmente implementadas atualizando a posição das partículas com as seguintes equações:

$$\begin{aligned}x_{new} &= x - l_x \left[\frac{x}{l_x} \right] \\y_{new} &= y - l_y \left[\frac{y}{l_y} \right] \\z_{new} &= z - l_z \left[\frac{z}{l_z} \right]\end{aligned}$$

Para uma caixa com origem em (0, 0, 0) e sendo l_x , l_y , l_z o comprimento da caixa nas direções x , y e z , respectivamente. Como consequência, ao se computar a força entre pares de partículas, deve-se calcular as distâncias pelo método de mínima imagem (I_{ij}), que considera a mínima distância entre as partículas virtuais e reais:

$$\begin{aligned}x_{ij} &= x_i - x_j \\I_{ij} &= x_{ij} - l_x \left[\frac{x_{ij}}{l_x} \right]\end{aligned}$$

Onde I_{ij} a distância de imagem mínima entre os vetores x_i e x_j .

2.5 Análises das Trajetórias

2.5.1 Flutuações Estruturais

Root Mean Square Deviation (RMSD) é uma medida do quanto um conjunto Ω de átomos no tempo t desvia das posições correspondentes de uma estrutura de referência. O valor é computado por:

$$RMSD_{\Omega}(t) = \left(\frac{1}{N_{\Omega}} \sum_{i \in \Omega} [r_i(t) - r_i^{ref}]^2 \right)^{\frac{1}{2}}$$

Sendo N_{Ω} o número de átomos no conjunto de interesse, $r_i(t)$ o vetor da posição do átomo i no instante t e r_i^{ref} a posição do átomo i na estrutura de referência. A estrutura no tempo t e a estrutura de referência devem ser alinhadas a fim de

eliminar possíveis contribuições provenientes dos graus de liberdade rotacionais e translacionais. A *Root Mean Square Fluctuation* (RMSF), por sua vez, é definida como a média, sobre as diferentes configurações, do RMSD de cada resíduo:

$$RMSF_{res} = \langle RMSD_{res} \rangle$$

Em simulações de proteínas, o RMSD é útil na medida em que mede o quanto a estrutura da proteína se modifica ao longo da simulação. O RMSF, por sua vez, indica em quais regiões da proteína esses movimentos ocorrem.

2.5.2 Energia Livre

De acordo com a mecânica estatística, a população em um microestado de dado sistema está relacionada com a energia deste microestado por meio da relação:

$$p_i \propto \exp(-e_i / k_B T)$$

Logo, a diferença de energia livre entre dois microestados A e B pode ser escrita em função de suas populações:

$$\frac{p_A}{p_B} = \exp\left(\frac{e_B - e_A}{k_B T}\right)$$

$$e_B - e_A = -k_B T \ln\left(\frac{p_A}{p_B}\right)$$

Enquanto o microestado é definido pela a distribuição de todas as partículas nos níveis de energia $e_A, e_B, \dots, e_N$, o macroestado de um sistema é definido pelo conjunto de suas propriedades macroscópicas (temperatura, pressão, densidade etc). Para sistemas com um grande número de partículas a probabilidade de observar macroestados diferentes que o macroestado mais provável tende a zero (a distribuição torna-se mais estreita com o aumento do número de partículas).

Dessa forma, o único macroestado de interesse no equilíbrio é macroestado estado mais provável.

Para um ensemble NTp, o macroestado é aquele que maximiza a função de distribuição no número de partículas nos níveis de energia e satisfaz as restrições de número de partículas, temperatura e pressão constantes. Assim, a energia livre associada e esse é dada pela relação:

$$G = -k_B T \ln \Delta$$

Onde Δ é a função de partição do ensemble NTp. Dessa forma, a variação de energia livre entre dois estados é calculada simplesmente contando quantas partículas foram observadas em cada estado durante a simulação, ou seja, obtém-se as populações:

$$\Delta G_{A,B} = -k_B T \ln\left(\frac{p_A}{p_b}\right)$$

Assim, caso a amostragem do espaço de fases gerada pela DM se aproxime da distribuição de populações em um limite de tempo infinito de simulação, a relação acima pode ser aplicada.

Como tais relações termodinâmicas são abordadas em detalhe em qualquer curso introdutório de mecânica estatística, optou-se por não mostrar a derivação formal dessas expressões neste texto.

2.5.3 Cálculos de Permeabilidade

Para soluções diluídas e baixos gradientes de concentração, pode-se aproximar que o fluxo entre dois reservatórios (j_w) como linearmente proporcional a diferença de concentrações (ΔC):

$$j_w = p_f \Delta C$$

Onde $p_f[\text{cm}^3\text{s}^{-1}]$ é a coeficiente de permeabilidade osmótica do sistema. Considere-se dois reservatórios unidos por um canal e tome-se a variável coletiva n em sua forma diferencial dn como:

$$dn = \sum_{i \in C(t)} dz_i / L ,$$

Onde $C(t)$ é o conjunto de moléculas contidos no canal no instante t , L o comprimento do canal e z_i a posição no eixo z da partícula i . A definição de dn é tal que, quando uma molécula atravessa o canal na direção $+z$, n sofrerá um acréscimo de $+1$, da mesma forma, ao atravessar na direção $-z$, $\Delta n = -1$. Portanto, $n(t)$ quantifica o fluxo de partículas líquido/efetivo entre os reservatórios. Desse modo, podemos escrever o fluxo como:

$$j_w = \frac{1}{N_A} \langle n(t) \rangle / t$$

Ou seja, o fluxo no equilíbrio $\langle n(t) \rangle = 0$ e, como neste cenário as flutuações no valor de $n(t)$ são aleatórias, poderíamos descrever $n(t)$ por um caminho aleatório (randômico) em uma dimensão, de tal modo que vale a relação de Einstein:

$$\langle n^2(t) \rangle = 2D_n t$$

Onde D_n é coeficiente de difusão associado à variável n . Combinando a primeira equação com a relação (Lei de Ficks):

$$j_w = \frac{-\Delta\mu}{k_B T} \frac{1}{N_A} D_n ,$$

Obtém-se:

$$p_f = v_w D_n$$

Onde v_w é o volume de uma molécula do solvente (nos sistemas estudados, a água). Assim, pode-se obter o valor de p_f computando o coeficiente angular de $\langle n^2(t) \rangle$ em simulações no equilíbrio – portanto sem a necessidade de aplicar um gradiente ou de pressão ou de concentração.

De maneira similar, o fluxo de autodifusão pode ser aproximado como variando linearmente com ΔC_{tr}

$$j_{tr} = p_d \Delta C_{tr}$$

Onde o subscrito tr indica a molécula traço e p_d [cm^3s^{-1}] é chamado de coeficiente de permeabilidade difusiva (ou autodifusão). Assumindo reservatórios 1 e 2 com $C_{tr}^{(1)} = 0$ e $C_{tr}^{(2)} = C$ (logo $\Delta C_{tr} = C$), o fluxo de autodifusão é apenas o número médio de moléculas tr que cruzam o canal em um intervalo de tempo vezes o respectivo volume de cada molécula. Como as propriedades das moléculas tr e das demais moléculas são idênticas, o número médio de moléculas tr que cruzam o canal deve ser proporcional ao número total de moléculas que cruzam o canal vezes a concentração do traço. Logo:

$$q_{tr} v_w = q_0 \Delta C_{tr}$$

ou:

$$p_d = v_w q_0$$

sendo q_0 o número de moléculas que atravessaram o canal em um intervalo de tempo e q_{tr} o número de moléculas traço que atravessa o canal. Como o número total de eventos de permeação não depende de C_{tr} , a relação vale para $C_{tr}^{(1)} \neq 0$. Desse modo, basta contar os eventos unidirecionais de permeação ocorridos ao longo de uma simulação para obter p_d . Pode-se ainda relacionar p_f e p_d pela razão:

$$\frac{p_f}{p_d} = N + 1$$

Onde N é uma constante. Para um transporte de modo “fila-indiana” (*single-file*), p_d representa a taxa de evento de permeação, em que uma molécula atravessa todo o canal, e p_f a taxa do movimento das moléculas dentro do canal (taxa de ‘pulos’ / ‘hops’). Assim, $(N + 1)$ pode ser interpretado como o número médio de ‘pulos’ por evento de permeação. Nota-se também que ambos p_d e p_f podem ser medidos experimentalmente, sendo que valores para diversas aquaporinas, tanto *in cito* como *in silico*, estão tabelados na literatura.

Uma descrição mais aprofundada de ambos métodos pode ser encontrada nas referências (ZHU, 2004).

Capítulo 3

Reference Interaction Site Model

3.1 Funções de Correlação

Para a descrição mecânico-estatística de líquidos é útil a definição da função de correlação – $g^{(2)}(r_1, r_2)$ ou simplesmente $g(r_{12})$ – entre duas partículas 1 e 2 que forneça a probabilidade de 2 estar a uma distância r_{12} de 1. Como $g(r)\rho 4\pi r^2 dr$ é o número de partículas em uma calota esférica de espessura dr (sendo ρ a densidade numérica do bulk), pode-se interpretar $g(r_{12})$ como o fator de conversão de densidade numérica global à densidade local centrada em 1. É possível demonstrar também que, caso a energia potencial do sistema puder ser escrita como a soma de potenciais de pares, todas as relações termodinâmicas do sistema podem ser expressas em função de $g(r_{12})$. Disso, obtém-se uma relação semelhante à da secção (2.5.2):

$$g(r) = h(r) - 1 = \exp\left(-\frac{\omega(r)}{k_b T}\right),$$

onde $\omega(r)$ é o potencial de força média (PMF). Logo, da mesma forma que a população está associada com sua energia, também a função de distribuição de determinado estado está relacionada com uma função de energia que, neste caso, se expressa em função de uma coordenada espacial. Efetivamente, toma-se $\omega(r)$ como equivalente à energia livre na coordenada na coordenada r .

Ademais, $g(r_{12})$ é uma quantidade que pode ser medida experimentalmente – fator de estrutura – ou calculado por simulações numéricas de DM e MC. Para alguns sistemas simples, é possível também derivar $g(r_{12})$ analiticamente. Um dos métodos empregados para tal é através da solução da equação de Ornstein-Zernik (OZ), que divide a correlação total em dois termos, correlação direta e indireta:

$$h(r_{12}) = c(r_{12}) + \rho \int c(r_{13})h(r_{23})dr_3 ,$$

onde $h(r_{12})$ é a função de correlação total de 1 em 2, $c(r_{12})$ a correlação direta e ρ a densidade numérica. O segundo termo, a correlação indireta, poderia ser interpretado como a influência propagada diretamente da molécula 1 para 3 – $c(r_{13})$ – que exerce influência na molécula 2 através de todas outras partículas – já que $h(r_{23})$ pode ser expresso em termo de suas correlações diretas e indiretas:

$$\begin{aligned} h(r_{12}) = c(r_{12}) + \rho \int c(r_{13})c(r_{23})dr_3 \\ + \rho^2 \int \int c(r_{13})c(r_{34})h(r_{24})dr_3dr_4 \end{aligned}$$

A integral dupla pode ainda ser escrita em função de suas correlações diretas e indiretas e assim por diante, gerando uma série de integrais múltiplas.

Para resolver a equação acima é necessário ainda uma segunda expressão que relacione $h(r)$ e $c(r)$ – chamada de relação de fechamento. A derivação dessas expressões está fora do escopo deste estudo, mas pode ser encontrada na referência (HUANG, 2013). A relação de fechamento utilizada neste trabalho foi a de Kovalenko-Hirata (KH), que assume a forma de:

$$\begin{aligned} h_\gamma(r) &= \exp(t(r)) - 1 \text{ se } g(r) \leq 0 \\ h_\gamma(r) &= t(r) \text{ se } g(r) > 0 \end{aligned}$$

$$t(r) = -u_\gamma(r) / k_b T + h_\gamma(r) - c_\gamma(r) + B(r)$$

Na relação de fechamento de KH o termo $B(r)$ é considerado como sendo igual a zero.

Somente as variáveis $r_{12}, r_{13}, r_{23} \dots r_{(n-1)n}$ não são suficientes para descrever líquidos moleculares. Para tal, é preciso introduzir a orientação espacial das moléculas (por meio de seus ângulos de Euler $\Omega(\theta, \phi, \rho)$ na equação).

Nesta forma a equação de OZ pode ser escrita como:

$$h(r_{12}, \Omega_1, \Omega_2) = c(r_{12}, \Omega_1, \Omega_2) + \frac{\rho}{\Omega} \int c(r_{13}, \Omega_1, \Omega_3) h(r_{23}, \Omega_3, \Omega_2) dr_3 d\Omega_3$$

Onde Ω é um fator de normalização.

3.2 1D-RISM

Devido ao fato de a equação integral para líquidos moleculares ser função de 6 variáveis, sua resolução é computacionalmente muito trabalhosa – o que também limita suas aplicações. Todavia, a equação pode ser reduzida de 6 à 1 variável, aplicando a aproximação de sítios de interação (RISM) introduzida por Chandler e Anderson em 1972 (CHANDLER, 1972). Essa assume que o potencial de interação entre duas moléculas pode ser expressa como:

$$u_{12} = \sum_{\alpha, \gamma} u_{\alpha\gamma}(r)$$

E de forma similar:

$$c(r_{12}) = \sum_{\alpha, \gamma} c_{\alpha, \gamma}(r_{\alpha 1} - r_{\gamma 2})$$

Onde α e γ são os sítios da molécula 1 e 2, respectivamente. Dessa maneira é possível escrever uma equação de OZ para cada par de sítios fixando a distância entre os sítios – $(r_{\alpha n} - r_{\gamma m})$ – e fazendo a média das funções de correlação sobre todas as possíveis orientações espaciais. Assim, obtém-se:

$$h_{\alpha\gamma}(r) = \sum_{\mu, \nu} \omega_{\alpha\mu}(r) \otimes c_{\mu\nu}(r) \otimes [\omega_{\nu\gamma}(r) + \rho h_{\nu\gamma}(r)]$$

Onde $\otimes$ indicam integrais de convolução e $\omega_{\alpha\gamma}(r)$, a função de correlação intramolecular, é definida como:

$$\omega_{\alpha\gamma}(r) = \delta_{\alpha\gamma}\delta(r) + (1 - \delta_{\alpha\gamma})\frac{\delta(r - l_{\alpha\gamma})}{4\pi l_{\alpha\gamma}^2}$$

Onde $\delta_{\alpha\gamma}$ é o delta de Kronecker, $\delta(r)$ é a função delta de Dirac e $l_{\alpha\gamma}$ a distância entre os sítios. Devido às integrais de convolução, a equação é mais facilmente resolvida no espaço recíproco. Numericamente, esses sistemas de equações são resolvidos com o algoritmo de transformada rápida de Fourier (FFT).

3.3 3D-RISM

Para sistemas solvente-soluto, no limite da diluição infinita do soluto, podemos escrever as equações de OZ para o sistema da seguinte forma:

$$\begin{aligned} h^{VV}(r_{12}, \Omega_1, \Omega_2) &= c^{VV}(r_{12}, \Omega_1, \Omega_2) + \rho^V \int c^{VV}(r_{13}, \Omega_1, \Omega_3) h^{VV}(r_{23}, \Omega_3, \Omega_2) dr_3 d\Omega_3 \\ h^{UV}(r_{12}, \Omega_1, \Omega_2) &= c^{UV}(r_{12}, \Omega_1, \Omega_2) + \rho^V \int c^{UV}(r_{13}, \Omega_1, \Omega_3) h^{VV}(r_{23}, \Omega_3, \Omega_2) dr_3 d\Omega_3 \\ h^{UU}(r_{12}, \Omega_1, \Omega_2) &= c^{UU}(r_{12}, \Omega_1, \Omega_2) + \rho^V \int c^{UV}(r_{13}, \Omega_1, \Omega_3) h^{UU}(r_{23}, \Omega_3, \Omega_2) dr_3 d\Omega_3 \end{aligned}$$

Sendo o superíndice U e V referentes ao soluto e ao solvente respectivamente. A correlação total $h^{UV}(r)$ é de particular interesse, uma vez que descreve as correlações solvente solutos – termo relevante ao se estudar processos de solvatação. Nota-se também que esta é função de $h^{VV}(r_{23}, \Omega_2, \Omega_3)$ – correlação solvente-solvente – que pode ser obtida pelo 1D-RISM.

A mesma aproximação utilizada no 1D-RISM como forma de diminuir a dimensionalidade pode ser aplicada aqui. No entanto, é mais útil somente integrar com respeito aos grau de liberdade do solvente, pois desse modo se obtém a distribuição 3D do solvente em torno do soluto, que é mais fácil de ser interpretada. Aplicando a aproximação temos as equações da teoria 3D-RISM:

$$h_{\gamma}^{UV} = \sum_{\alpha} \int c_{\alpha}^{UV}(r-r') \chi_{\alpha\gamma}^{VV}(r') dr'$$

Onde:

$$\chi_{\alpha\gamma}^{VV}(r) = \omega_{\alpha\gamma}^{VV}(r) + \rho h_{\alpha\gamma}^{VV}(r)$$

Em que r e r' são os vetores das posições dos sítios. Pode-se, assim, obter as funções de correlação soluto-solvente entre sítios do solvente e o soluto. A solução é computada em pontos em uma malha tridimensional.

Devido ao método proporcionar uma maneira computacionalmente barata de obter propriedades no equilíbrio de sistemas complexos, o 3D-RISM tem sido utilizado amplamente em estudo computacionais de macromoléculas. É possível citar como exemplos: afinidade e seletividade de canais transportadores de íons (PHONGPHANPHANEE, 2010), predição de propriedades químicas diversas de potenciais fármacos (ANSARI, 2015) e no estudo da dimerização de proteínas (SRIVASTAVA, 2020).

No entanto, ressalta-se que o método sofre de uma série de limitações, tais como: a não garantia da convergência das soluções para solventes com mais de 12 sítios e, mais fundamentalmente, a da aproximação de sítios tratar as espécies como conjunto de esferas fusionadas, logo não considera nenhum dos possíveis confôrmeros das espécies do solvente/soluto. Por usar os mesmos potenciais dos campos de força da DM, também é limitado pela capacidade desses de descreverem as interações do sistema.

Capítulo 4

Métodos

4.1 Modelagem Estrutural

As sequências de aminoácidos para LmAQP1, LmAQP α , LmAQP β e LmAQP γ foram retiradas do banco de dados UniProt (UNIPROT CONSORTIUM, 2019). Partindo das sequências, empregou-se a modelagem por homologia, utilizando o servidor *online SwissModel* (WATERHOUSE, 2018), para obter as estruturas 3D. Sendo as estruturas de modelo (*template*) escolhidas com base em suas identidades de sequência das estruturas cristalográficas disponíveis no PDB para com as AQPs de *Leishmania*.

Alguns dos resíduos nas porções iniciais e finais das sequências não foram considerados na modelagem, devido à baixa similaridade com as estruturas de referência. A baixa similaridade sugere que esses segmentos não fazem parte da estrutura canônica do AQP (não fazem parte das seis hélices alfa e dos dois *loops* altamente conservados). Além disso, sua exclusão melhorou significativamente a qualidade do modelo e reduziu o custo computacional das simulações.

Os tetrâmetros AQP (estrutura quaternária) também foram modelados no servidor *SwissModel*. Posteriormente, todos os modelos obtidos foram verificados quanto a possíveis rotâmetros de cadeia lateral nas regiões NPA e SF.

A estrutura cristalográfica para PfAQP1 pode ser encontrada no PDB, essa possui resolução de 2 Å (NEWBY, 2008).

4.2 Construção do Sistema

O estado de protonação para cada aminoácido foi atribuído usando o servidor H++ (GORDON, 2005). Em que o pH para LmAQP1 foi considerado igual a 5, devido a esta estar em contato direto com o fagossomo do macrófago, meio de pH baixo (VANNIER-SANTOS, 2002). As demais AQPs de *Leishmania* são subcelulares

(BIYANI, 2011), portanto o pH foi definido como sendo igual a 7. Para PfAQP1, o pH também foi considerado igual a 7.

Cada tetrâmero AQP foi incorporado em uma bicamada de lipídica de POPC, composta de aproximadamente 200 moléculas, usando o conjunto de ferramentas do servidor *online* CHARMM-GUI (JO, 2009). O sistema foi solvatado usando o modelo de água TIP3P e as cargas neutralizadas com a adição de íons KCl na concentração de 0,15M. Os sistemas resultantes variaram em tamanho entre 80-100 mil átomos.

4.3 Dinâmica Molecular

Todas as simulações de dinâmica molecular foram executadas no *software* NAMD 2.13 (PHILLIPS, 2005) usando o campo de força CHARMM36m (HUANG, 2017) no *ensemble* isotérmico-isobárico (NPT) com condições periódicas de contorno. A temperatura foi mantida constante a 310 K com o termostato Langevin e a pressão mantida em 1 bar com o pistão de Langevin (RYCKAERT, 1976). Todas as ligações envolvendo átomos de hidrogênio foram mantidas fixas em seus comprimentos de equilíbrio usando o algoritmo RATTLE (MIYAMOTO, 1992), de modo a permitir o uso de um passo de integração de 2 fs. Foi utilizado um *cutoff* de 12Å para as interações de curto alcance e as interações de longo alcance foram tratadas com o algoritmo de *Particle Mesh Ewald* (PME) (DARDEN, 1993).

Adotou-se o esquema de equilibração sugerido pelo CHARMM-GUI, esse, em suma, consistia de 7 etapas: uma minimização de 10000 passos, seguida por 6 etapas de equilíbrio com um passo de 1 fs, onde restrições harmônicas foram aplicadas aos diédricos do sistema e gradualmente relaxadas em cada uma das etapas subsequentes. O tempo total de equilibração somou 10 ns. Após as etapas de equilibração, simulações de 300 ns com um passo de tempo de 2 fs foram realizadas para todos os cinco sistemas.

4.4 Análises de Trajetória

Todas as análises de trajetória foram realizadas com a biblioteca MDAnalysis (MICHAUD-AGRAWAL, 2011) da linguagem Python. As demais figuras foram renderizadas no *software* VMD (HUMPHREY, 1996).

O raio médio do canal foi calculado com o *software* HOLE 2.2 (SMART, 1996) utilizando 100 *frames* retirados da simulação.

MDLovoFit foi usado para identificar as porções mais e menos móveis da das proteínas e calcular seus respectivos RMSD (MARTÍNEZ, 2015).

O perfil de energia livre para a água no interior do canal foi obtido dividindo-se o volume do poro em *bins* e calculando-se o logaritmo do número de água em cada *bins* durante a simulação.

A permeabilidade osmótica (p_f) foi calculada usando o método de difusão coletiva (ZHU, 2004), em que $\langle n(t)^2 \rangle$ foi calculado computando todos os pares $(n(t + \Delta t) - n(t))^2$ para intervalos de tempo $t - \Delta t < 100ps$. A trajetória foi dividida em três segmentos de 100 ns e esses foram analisados separadamente. p_f foi então obtido ajustando uma reta à curva $\langle n(t) \rangle^2$.

A permeabilidade da autodifusão (p_d) foi calculada contando-se os eventos de permeação ocorridos durante a simulação.

4.5 RISM

Os cálculos de 1D-RISM e 3D-RISM foram feitos com as implementações encontradas no pacote AmberTools19 (MIYAMOTO, 2019).

4.5.1 1D-RISM

Os cálculos foram realizados em seis soluções diferentes: amônia 0,1M, uréia 0,1 M, dióxido de carbono 1mM, metanol 0,1 M e uma solução de HCl com pH = 5. O modelo de água utilizado foi o TIP3P à concentração de 55,5 M. Íons NaCl à concentração de 0,1 M foram considerados em todas soluções. Diferentes concentrações de soluto e NaCl foram testadas, mas não houve impacto perceptível nos resultados, a parte o maior tempo de execução e pior convergência.

O parâmetro do campo de força para os cálculos de 1D-RISM foram retirados

do campo de força CHARMM36m, exceto pelos parâmetros para o hidrônio que foram retirados da literatura (BONTHUIS, 2016). A relação de fechamento utilizada foi a de KH e as soluções computadas em uma malha 16384 pontos com espaçamento de 0,025 Å. O algoritmo modificado da inversão direta do espaço iterativo (MDIIS) foi empregado com tolerância residual de 10^{-12} . Todos sistemas foram considerados na temperatura de 310 K.

4.5.2 3D-RISM

Para cada sistema solvente-proteína, retirou-se 100 *frames* igualmente espaçados da trajetória da DM e um dos monômeros em cada frame foi usado como *input* para o cálculo do 3D-RISM (representando o soluto no cálculo). Uma caixa de 64 Å x 64 Å x 80 Å e uma malha de 256 x 256 x 320 pontos foram empregadas. O refinamento adicional da malha não alterou os resultados do cálculo. O algoritmo MDIIS foi utilizado com tolerância de 10^{-3} e KH tomada como a condição de contorno.

O potencial da força média (PMF) foi calculado com a média de $g_\alpha(z)$ em 100 estruturas, tomando como o sítio de interação de referência (α) aquele que mais se aproxima do centro de massa da molécula de solvente (por exemplo, no caso da molécula de uréia, o carbono central foi usado para calcular o PMF). O $\langle g_\alpha(z) \rangle$ foi normalizado com um fator proporcional a área do canal em função de z .

O $\langle g_\alpha(z) \rangle$ foi então, em uma área de 8 Å x 8 Å, integrado a coordenadas x e y – área que encapsula todo o poro da proteína no plano. Para obter o PMF aplicou-se a relação:

$$PMF = -k_B T \ln \frac{\langle g_\alpha(z) \rangle}{g_\alpha(z_{bulk})}$$

A metodologia descrita em mais detalhe pode ser encontrada em outros trabalhos na literatura (ARIZ-EXTREME, 2017).

Capítulo 5

Estudo Comparativo de Aquagliceroporinas de *Leishmania major* e *Plasmodium falciparum*

O conhecimento a nível molecular das características estruturais, do funcionamento e da regulação de proteínas é, hoje, indispensável quando se almeja investigar problemas complexos como o desenvolvimento de novos fármacos ou mesmo no estudo de mecanismos fisiológicos. Deste modo, busca-se, neste trabalho, por meio de simulação de dinâmica molecular, modelar e caracterizar a aquagliceroporina da *Leishmania major* quanto a sua estrutura e permeabilidade à água. Realizam-se estudos do mesmo tipo para a aquagliceroporina do *Plasmodium falciparum*, análoga da LmAQP1 e com estrutura cristalográfica no PDB. Procura-se, também, avaliar a permeabilidade dos canais frente à amônia, uréia, peróxido de hidrogênio, dióxido de carbono, íons hidrônio, Na⁺ e Cl⁻. Como análises deste tipo, empregando técnicas baseadas em DM, são computacionalmente muito custosas e talvez não se justifiquem, dada a qualidade da estrutura da LmAQP1, opta-se por usar o *Reference Interaction Site Model* (RISM), que permite calcular a função de correlação entre a solvente e proteína, para obter o perfil de energia livre do processo de transporte.

Embora os resultados aqui apresentados tenham cunho apenas qualitativo, espera-se que este primeiro "*screening*" possa fornecer indícios das funções biológicas exercidas por estas proteínas no protozoário.

5.1 Descrição e Análise Estrutural

Ambas proteínas possuem o SF do tipo ARG/Ar, característico das aquagliceroporinas. No entanto, a região aromática é composta por triptofano e tirosina em LmAQP1 (Figura 5.1) e triptofano e fenilalanina em PfAQP1 (Figura 5.2). Há também, próxima à região aromática da PfAQP1, uma cisteína conservada que é responsável por tornar o canal sensível a mercúrio (Hg^{+2}) – o cátion é um inibidor clássico de AQPs, a oclusão do canal ocorre pela ligação do Hg^{+2} à cisteína (Tabela 5.1). Na LmAQP1 esse resíduo não está presente, o que possivelmente é a razão de, experimentalmente, não se observar a inibição (GOURBAL, 2004). A histidina (H126 e H69), responsável em algumas aquaporinas pela mudança de conformação do canal em resposta ao pH citosólico, é conservada em ambas estruturas. A região do canal composta majoritariamente de resíduos hidrofóbicos (*hydrophobic rings*) é também similar em ambas estruturas.

O motivo NPA, curiosamente, não é conservado na PfAQP1 (Figura 5.1), mas a modificação – de NPA para NLA e NPS – não afeta em grau significativo a permeabilidade à água no canal (APONTE-SANTAMARÍA, 2010).

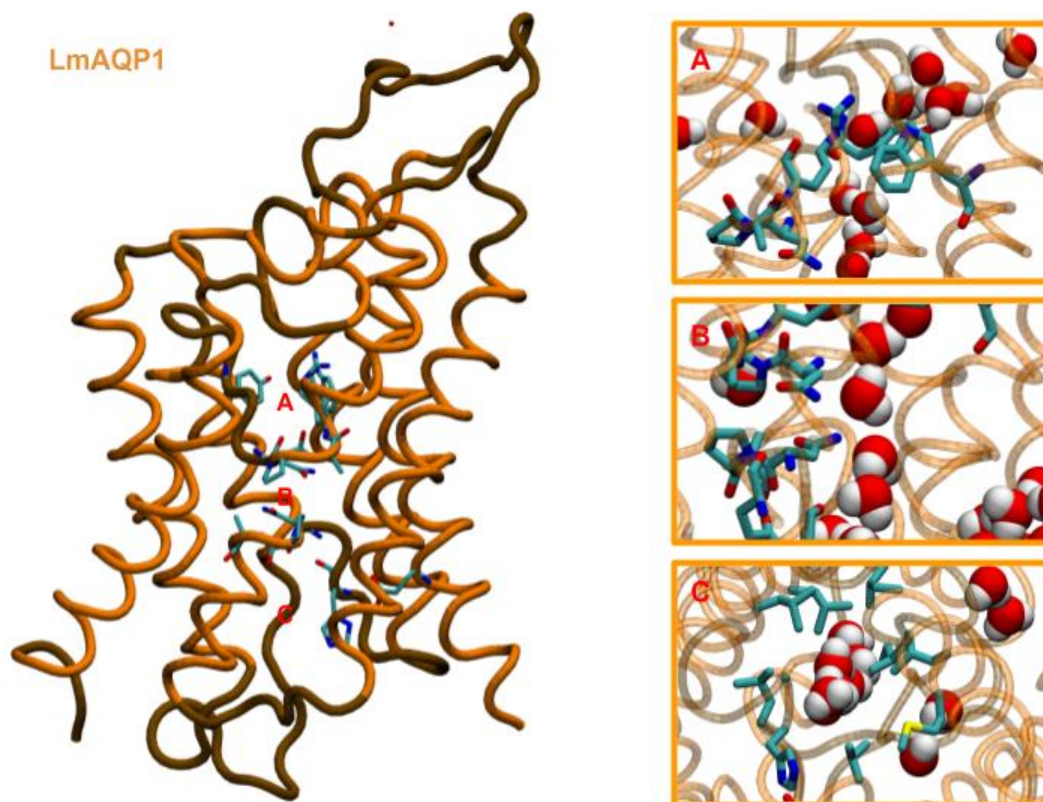


Figura 5.1 Representação estrutural da LmAQP1. (A) filtro seletivo; (B) motivo NPA; (C) região hidrofóbica e histidina H126.

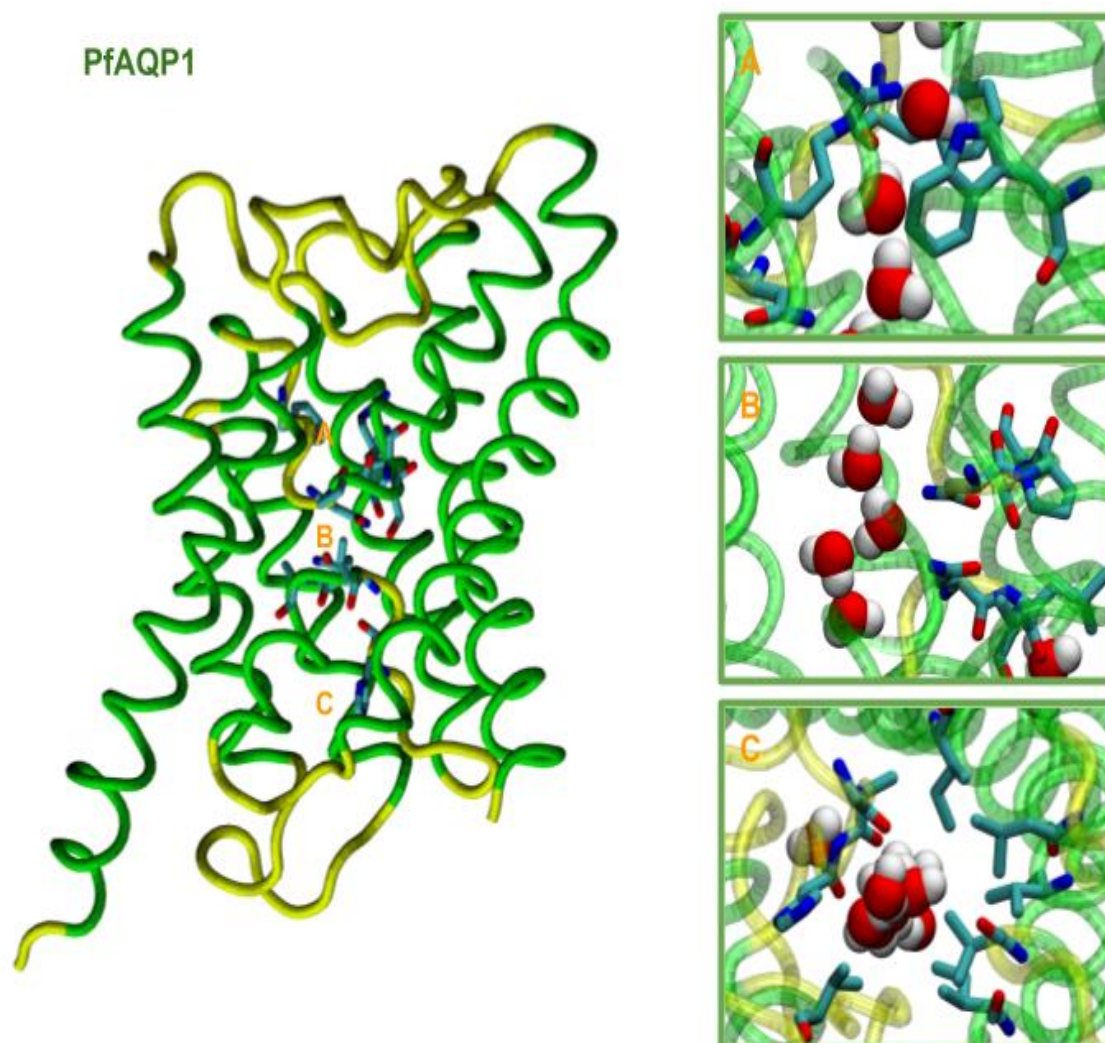


Figura 5.2 Representação estrutural da PfAQP1. (A) filtro seletivo; (B) motivo NPA não conservado; (C) região hidrofóbica e histidina H69

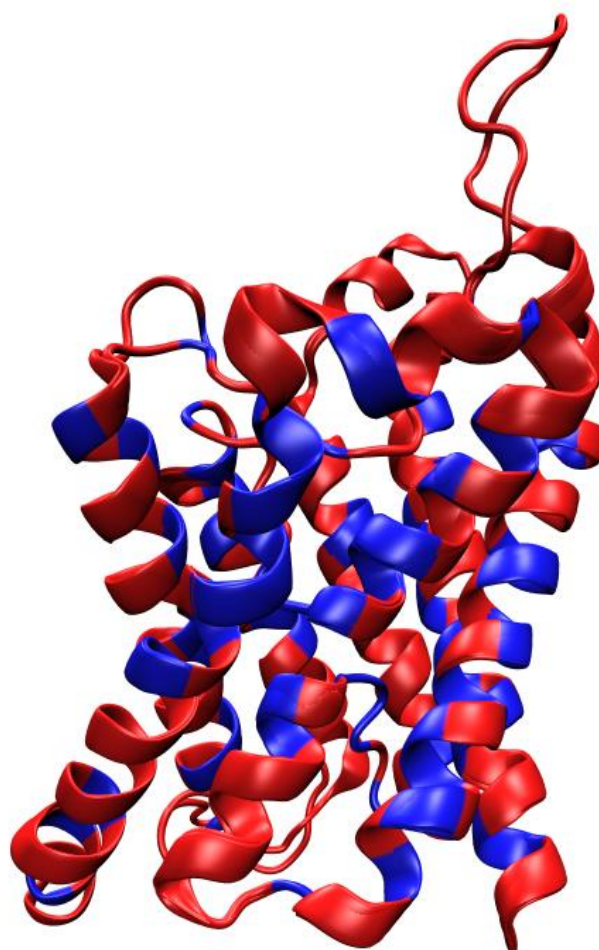


Figura 5.3 Representação das estruturas alinhadas da LmAQP1 e PfAQP1. Cor indica os resíduos conservados (azul) e não conservados (vermelho)

O alinhamento das sequências (Tabela 5.1) mostra uma porção significativa (50-70 resíduos) na LmAQP1 que não possui análoga PfAQP1. Essa região corresponde a um *loop* N-terminal que não compõe a estrutura canônica do canal, mas parece estar envolvido em mecanismos de regulação/degradação (UZCATEGUI, 2008). O mesmo *loop* varia de tamanho em diferentes cepas de *Leishmania major*. Com a retirada da região a identidade para com a PfAQP1 chega a 36-37%. Ademais, nota-se que, em sua maior parte, os resíduos conservados encontram-se nas alfa-hélices e porções centrais da proteína (Figura 5.3). As regiões não ordenadas possuem baixa identidade (e similaridade) e variam significativamente em tamanho (o que é comum em proteínas da família).

LmAQP1 & PfAQP1 - Alinhamento				
EstSec	1	-----		7
PfAQP1	1	MH-----MLFYK		7
		M+ + Y+		
LmAQP1	1	MNSPTTMPLQCHDAETQLYVDKESPEGLPIGNQMHEEEDQHESKRNFMSQNRNPLYRYR		60
EstSec	8	H1 -- -- H2		60
PfAQP1	8	SYVREFIGFLGTFVLMFLGEG--ATANFH--TTGL---SGDNVYKLCGLGLAVFFGILV		60
		+RE++ EF GTF L+ G G AT FH TT + + + + G + + +		
LmAQP1	61	WRLREYVAEFFGTFFLVTFGTGVVATTVFHGGTTAMYQSNSSYLAITFGAFLAISLFL		120
EstSec	61	H3		120
PfAQP1	61	SAKLSGAHLNLAIVSIGLSSINKFDLKKIPVYFFAQLLGAFVGTSTVYGLYHGFISNSKIP		120
		S +SG HLN AV++ F K+P YF AQ LG FVG + Y L+ ++		
LmAQP1	121	SMAVSGGHLNPAVTLANCVFGTFPWNKLPGYFLAQFLGGFVGAANTYVLFKSHFDEAEKR		180
EstSec	121	H4		171
PfAQP1	121	QFANETSRL-----PSISLTGAFFNELILTGILLVILVVVDENICGKFHILKLSS		171
		ET + P+++ T A ++E+ T L++ IL + D + ++		
LmAQP1	181	LLLNETMASKYGGIFATYPNVANTYAVNSEVNTMALNMGILAITDARMTPAVDYKPKVA-		239
EstSec	172	H5		231
PfAQP1	172	VVGLIILCIGITFGGNTGFALMPSRDLGSRFLSLIAYGKDTFTKDNFYFNVPLVAPCVGS		231
		+GL++ IGI G N+ + LNP+RDL R LS + +G + FT ++YFN+PLV P VG+		
LmAQP1	240	-IGLLLFVIGIASGINSSYGLMPARDLSPRILSAMLWGSEPFTLHSYYFNIPLVVPFVGA		298
EstSec	232	H6		258
PfAQP1	232	VVFCQFYDKVICPLVDLANNEKDGVDL		258
		+ Y I P		
LmAQP1	299	LFGMFLYVFFIIP-----PSC		314

Tabela 5.1 Alinhamento das sequências de LmAQP1 e PfAQP1. Motivo NPA em vermelho (R é considerado pertencente ao motivo NPARD nas aquagliceroporinas), SF em laranja, histidina em verde e cisteína, responsável pela sensibilidade ao Hg⁺², em amarelo. Linhas 'EstSec' indicam a estrutura secundária adotada pela PfAQP1, em que 'H1', 'H2' etc correspondem as alfa-hélices da proteína.

5.2 Dinâmica Molecular

5.2.1 Estabilidade

Em sua maioria, as cadeias laterais de resíduos pertencentes ao poro se mantiveram relativamente rígidas em ambos os canais. No entanto, a histidina (H69 e H129) de ambas aquaporinas apresentou significativa mobilidade. Na PfAQP1 a cadeia lateral da arginina (R197) do SF se mostrou também flexível, a ponto de ocluir o canal momentaneamente durante a simulação. A tirosina (Y257) do SF da LmAQP1 alterna entre duas conformações durante algumas das simulações, o que

também levou à perturbação do fluxo de água durante a simulação – talvez artefato da modelagem, já que o resíduo não é conservado na estrutura de referência.

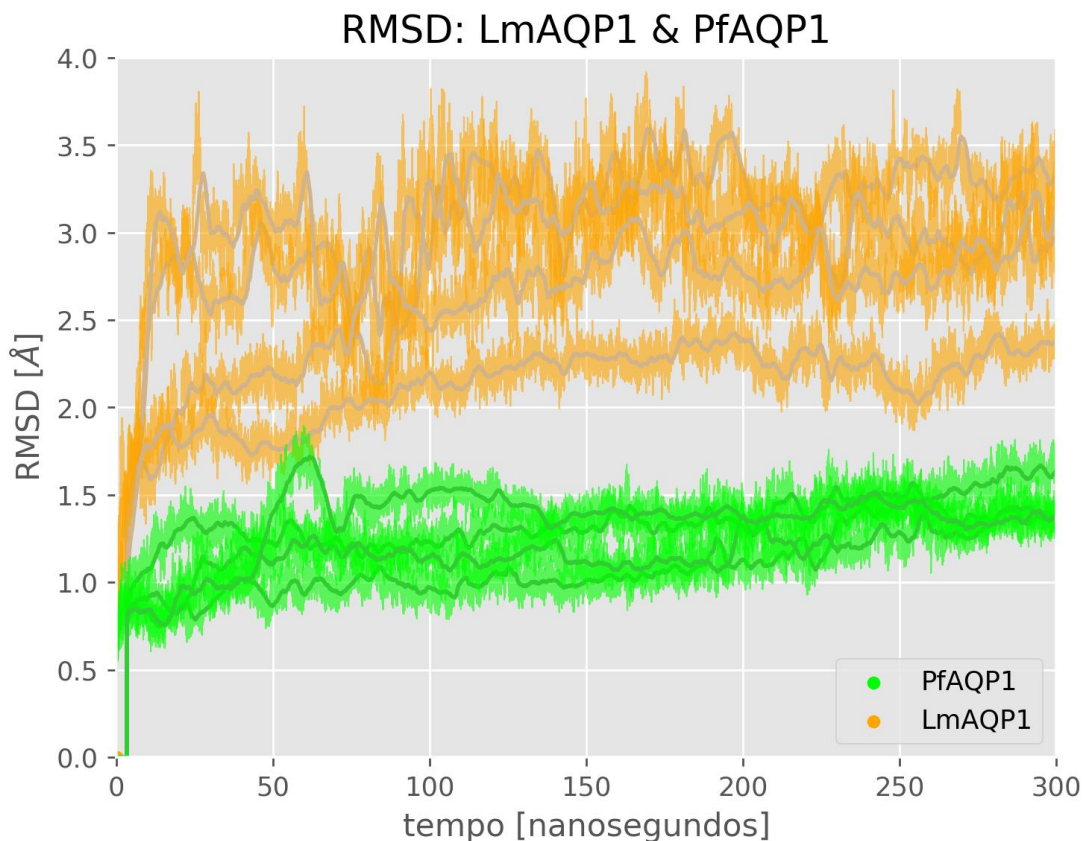


Figura 5.4 RMSD dos carbonos alfa da LmAQP1 (laranja) e PfAQP1 (verde) para cada monômero ao longo dos 300 ns de simulação.

O RMSD da estrutura modelada ficou próximo de 2.5-3.0 Å, enquanto a estrutura cristalográfica apresentou valores expressivamente menores – próximos de 1.5 Å (Figura 5.4). Isso pode ser resultado da diferença de qualidade entre as duas estruturas. No entanto, o RMSF da LmAQP1 indica que a mobilidade apresentada pela estrutura é resultante, em grande parte, da contribuição de um conjunto pequeno de resíduos muito móveis (Figura 5.5). E, como mostrado pelo alinhamento, a LmAQP1 possui um conjunto de *loops* não conservados e significativamente maiores que os de PfAQP1. Esses ainda estão em contato com o meio intra/extracelular, o que também lhes confere uma maior mobilidade que os demais resíduos que se encontram em contato com a membrana. Assim, se for excluída a contribuição dos 20% de carbonos alfa mais móveis e se alinhar o restante da estrutura, o RMSD permanece em torno de 1.0 Å (Figura 5.5). Desse modo, mesmo a que esses resíduos estejam mal enovelados e/ou trate-se de um

domínio intrinsecamente desordenado, o canal em si, restante da proteína, se mantém relativamente estável e preserva sua estrutura ao longo da simulação.

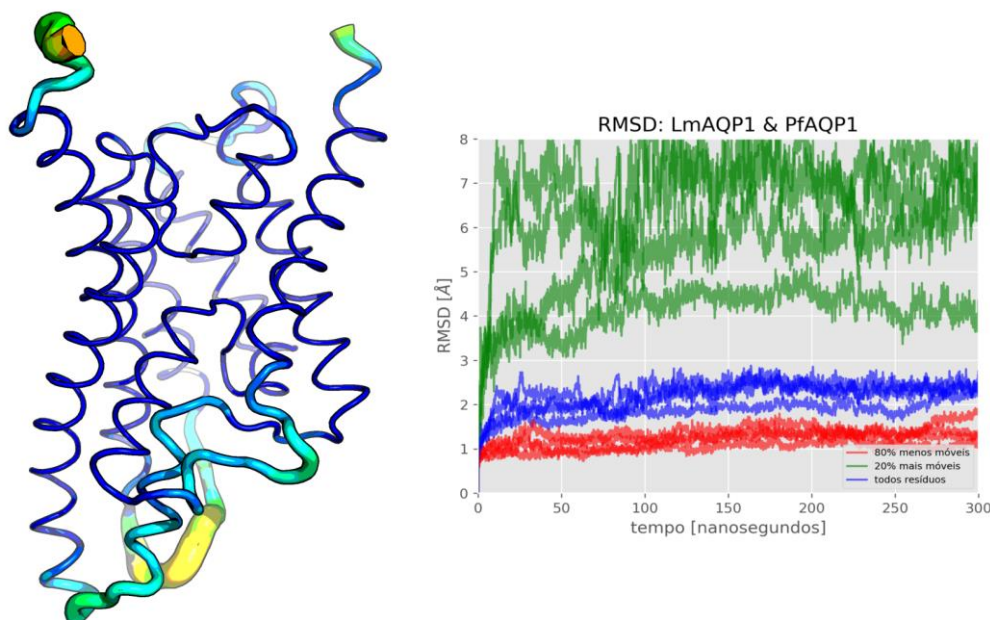


Figura 5.5 RMSF por resíduo da LmAQP1 representada estruturalmente (à esquerda) e RMSD ao longo da simulação dos carbonos alfa de diferentes porções da LmAQP1 (à direita): 80% menos móveis (vermelho), 20% mais móveis (verde) e toda a proteína (azul).

O raio médio do canal (Figura 5.6) com a média tomada sobre 100 frames apresentou pequena variância em PfAQP1 – o que condiz com o baixo RMSD da estrutura. No entanto, o SF apresentou maior desvio do valor médio, devido à alta mobilidade da cadeia lateral da arginina da região. Observa-se também um estreitamento no poro proveniente da região do NPA (próximo a $\zeta = 6$ Å) e, em maior grau, na região do SF ($\zeta = 15$ Å). Esse é o ponto mais estreito do canal, com raio próximo a 1,5 Å, e é o que confere à estrutura o aspecto de ampulheta descrito na literatura. A região hidrofóbica ($\zeta = 0$ Å), muito mais larga que as demais, também é relativamente estreita, possuindo raio próximo a 2 Å, mas nota-se que a região também apresenta grau significativo de variação.

O poro de LmAQP1 aparenta ser levemente mais estreito, no entanto, ambos canais tem seu ponto de afunilamento em 13 Å, sendo também o aspecto geral/perfil muito semelhante nos dois. O motivo NPA é um pouco mais pronunciado e está mais distante do SF (Figura 5.6). A divergência mais significativa entre os dois, todavia, parece ser a região hidrofóbica (entre $\zeta > -10$ e $\zeta < 0$). Essa região

apresenta uma variabilidade muito mais pronunciada que as demais e apresenta um mínimo local que não é observado na PfAQP1.

Nota-se que flutuações em $\zeta > 15$ Å devem ser entendidas como um artefato do programa e não são representativas do comportamento real da proteína.

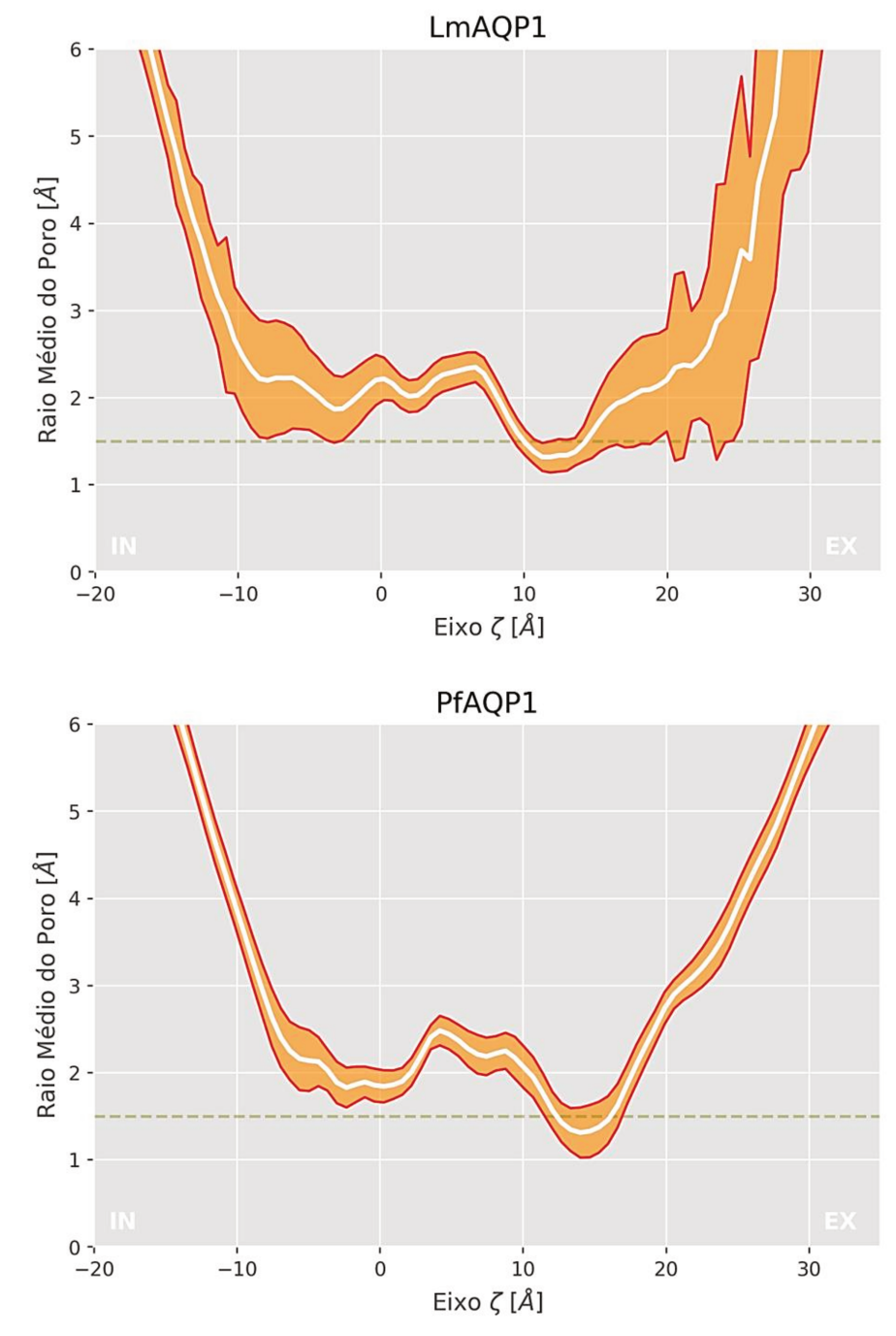


Figura 5.6 Raio médio do canal ao longo da coordenada do poro ζ centrada no centro geométrico da proteína. Região em laranja marca um desvio padrão acima e abaixo do raio médio. 'IN' e 'EX' denotam as porções intracelulares e extracelulares respectivamente. Linha pontilhada marca raio igual a 1,5 Å.

5.2.2 Permeabilidade à água

Os valores de p_d e p_f para a LmAQP1 e PfAQP1 indicam que o canal é capaz de transportar água e concordam qualitativamente com valores de outros estudos tanto *in silico* com *in situ* (APONTE-SANTAMARÍA, 2010) (Tabela 5.2). A razão entre p_f / p_d está abaixo do esperado – o valor deveria ser representativo, em algum grau, do número de mínimos energéticos dentro do canal.

	p_f [$10^{-14}\text{cm}^3\text{s}^{-1}$]	p_d [$10^{-14}\text{cm}^3\text{s}^{-1}$]	p_f / p_d
LmAQP1	7.06 ± 0.48	2.40 ± 0.16	2.74 ± 0.27
PfAQP1	10.67 ± 0.45	3.89 ± 0.14	2.74 ± 0.15

Tabela 5.2 Permeabilidade osmótica (p_f) e permeabilidade auto difusiva (p_d) em LmAQP1 e PfAQP1.

Embora a análise visual das trajetórias deixe claro que há a passagem de águas nos poros da proteína, é difícil afirmar que PfAQP1 é a mais permeável somente com os valores de p_f e p_d . Isso porque: (1) a qualidade dos modelos das estruturas não é a mesma; (2) sabe-se que o método não produz uma boa concordância com a medida experimental; (3) a variância entre monômeros de uma mesma proteína é significativa (valores para PfAQP1 estão entre $8.80 - 15.6 \times 10^{-14}\text{cm}^3\text{s}^{-1}$, e para LmAQP1 entre $6.10 - 9.12 \times 10^{-14}\text{cm}^3\text{s}^{-1}$).

No entanto, é interessante notar as populações de ligação de hidrogênio formadas entre água-água e água-proteína dentro do canal (Figura 5.7). Um número maior de ligações proteína-água indica que a água encontra-se mais fortemente ligada à proteína, enquanto um número menor de interações água-água indica que o ambiente do poro perturba as redes de ligação de hidrogênio entre as moléculas. Na Figura 5.7 observa-se que o número médio de ligações água-proteína similar entre LmAQP1 e PfAQP1, ainda que a distribuição da última seja mais larga no último. No entanto o número de interação água-água é significativamente menor na LmAQP1 o que aponta uma maior afinidade do solvente pela proteína do que por si próprio. Esses resultados são consistentes com o fato da PfAQP1 transportar água mais rapidamente.

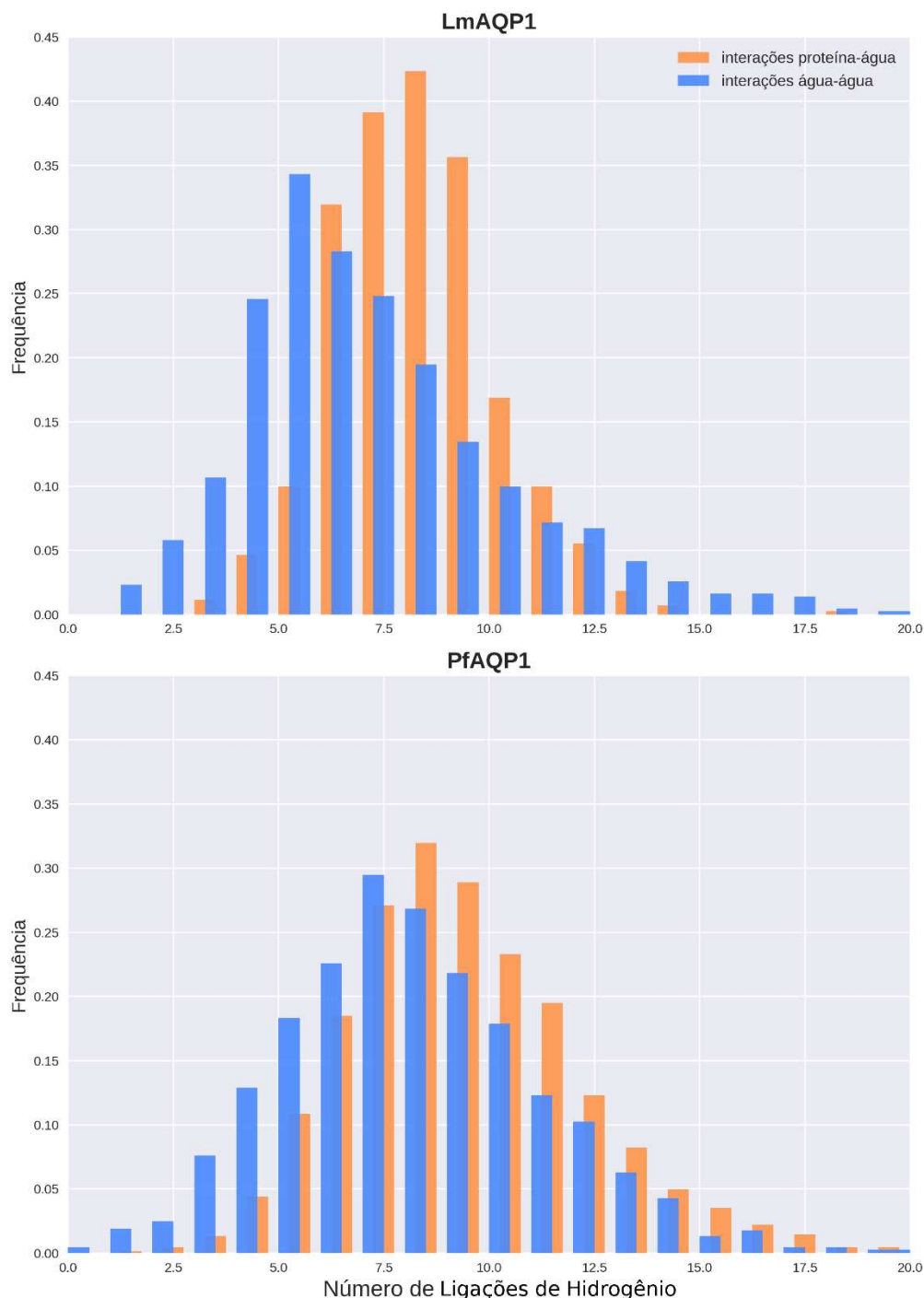


Figura 5.7 Número médio de ligação de hidrogênio dentro do poro de LmAQP1 e PfAQP1.

O simples registro da posição das moléculas de água dentro do canal fornece uma primeira aproximação de superfície de energia livre para o processo de transporte (ou pode ser interpretado como uma mapa de densidade) (Figura 5.8). Sendo que os mínimos de energia bem definidos no gráfico correspondem aos sítios onde a água realiza ligação de hidrogênio com o esqueleto da proteína. A figura, portanto, ilustra o modelo de *'collect diffusion model'* – em que as moléculas se

movimentam por pulos discretos (“hops”) concertados entre os sítios de ligação. As regiões mais estreitas do poro (SF, NPA) também podem ser facilmente distinguidas.

O perfil para LmAQP1 mostra mínimos menos definidos do que o PfAQP1, mas ocupando posições similares. Uma vez que o tamanho do poro na porção intracelular é similar para as duas proteínas, pode-se atribuir a diferença observada à maior flexibilidade da estrutura da LmAQP1. Também observa-se claramente a redução do raio do canal próximo ao SF da LmAQP1. Em PfAQP1, a redução acontece devido a uma região próxima do SF mais larga e populada por diversas moléculas de água ($\zeta \in [5,10]$ e $r \in [4,8]$). No entanto, observa-se o máximo de energia na região.

Superfície de Energia Livre da Água

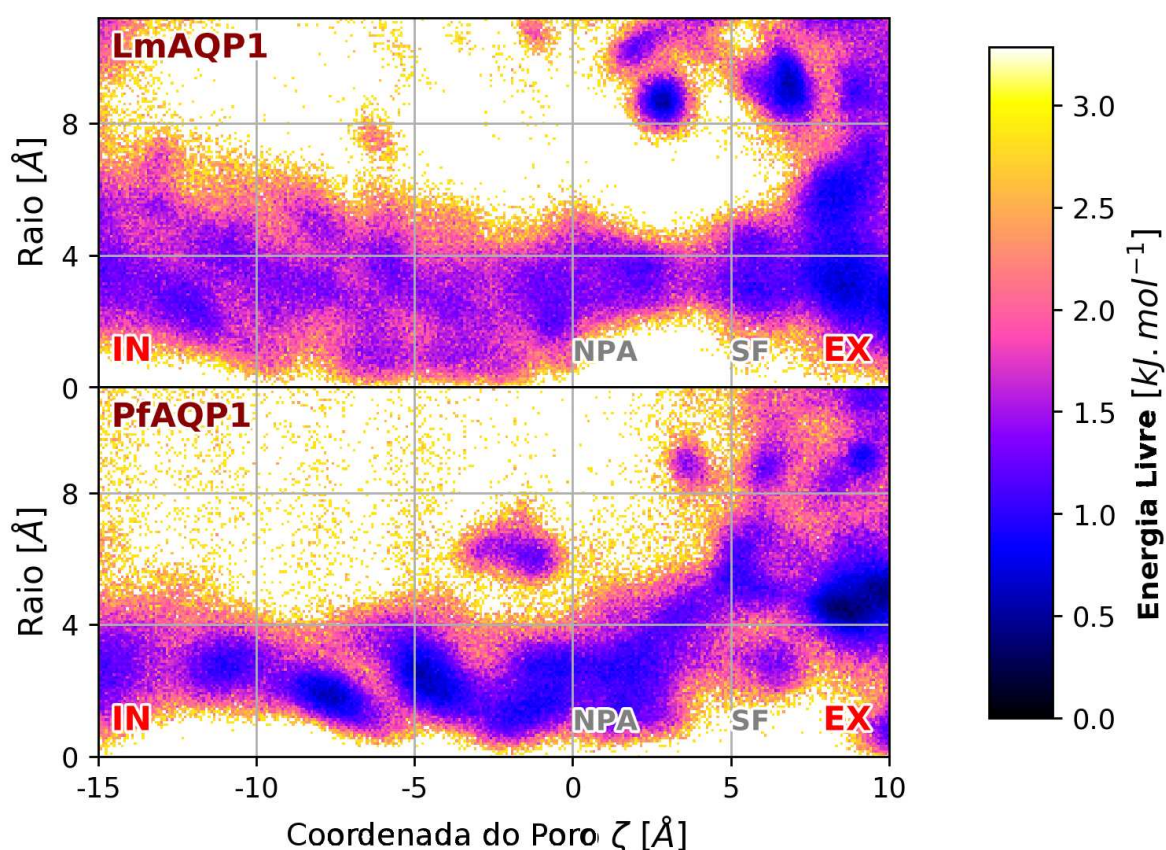


Figura 5.8 Superfície de energia livre da molécula de água dentro do poro da LmAQP1 (acima) e PfAQP1 (abaixo) em função da distância para o centro geométrico da proteína no plano XY (eixo y no gráfico) e no eixo ζ (eixo x).

5.3 3D-RISM

Aplicando a equação do 3D-RISM aos sistemas, obtemos o $g_{\alpha}(r)$ para todos os sítios α do solvente definidos em uma malha tridimensional. Um ponto com $g_{\alpha}(r) = 2$, indica uma probabilidade 2 vezes maior de encontrar o sítio α neste ponto do que no *bulk*. Qualitativamente, é possível observar a afinidade dos diferentes solventes com a proteína (Figura 5.9) somente pela inspeção da malha. Espécies como o Na^+ e o H_3O^+ exibem região no poro com $g_{\alpha}(r)$ próximo a zero, o que corresponde às barreiras de energia do processo de transporte. Da mesma forma, a inspeção do $g_{\alpha}(r)$ dos hidrogênios da molécula de água indicam os sítios onde essas realização ligação de hidrogênio e os ângulos em que estas interação é acontece.

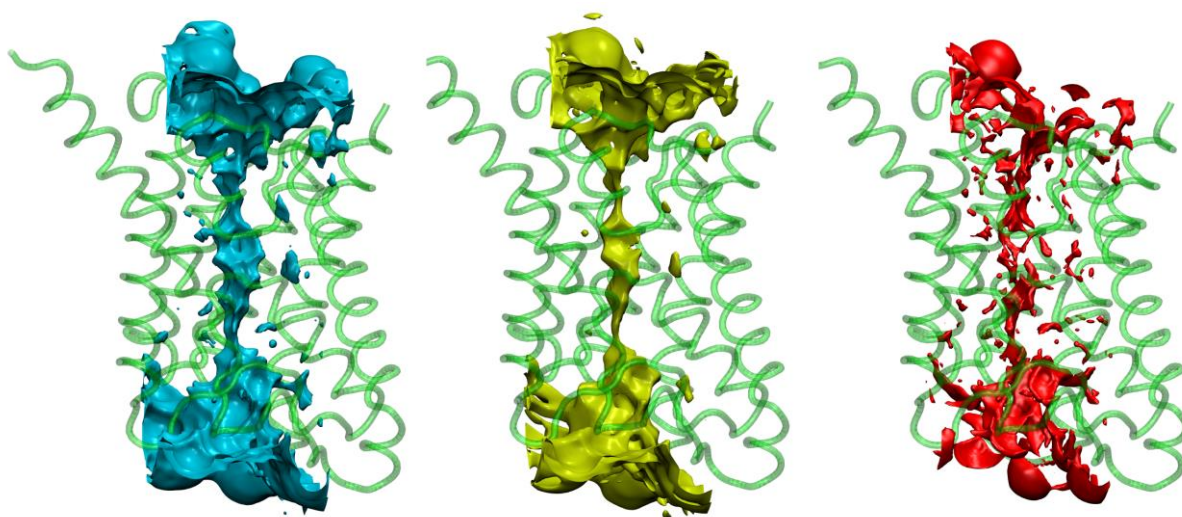


Figura 5.9 Da esquerda para a direita, $g(r)$ da água, amônia e do hidrônio no canal da Pf AQP1. Na imagem está representado o região na qual o $g(r) > 1.5$.

As barreiras de energia ao longo do tempo do PMF (Figura 5.10) concordam com os valores de raio médio do canal e, da mesma forma, os mínimos locais no PMF correspondem, em posição, aos apresentados na Figura 5.8. Nota-se, ainda, a barreira do NPA composta por dois picos e um mínimo, sendo este último correspondente ao estado da molécula de água, ligando-se às asparaginas dos motivos.

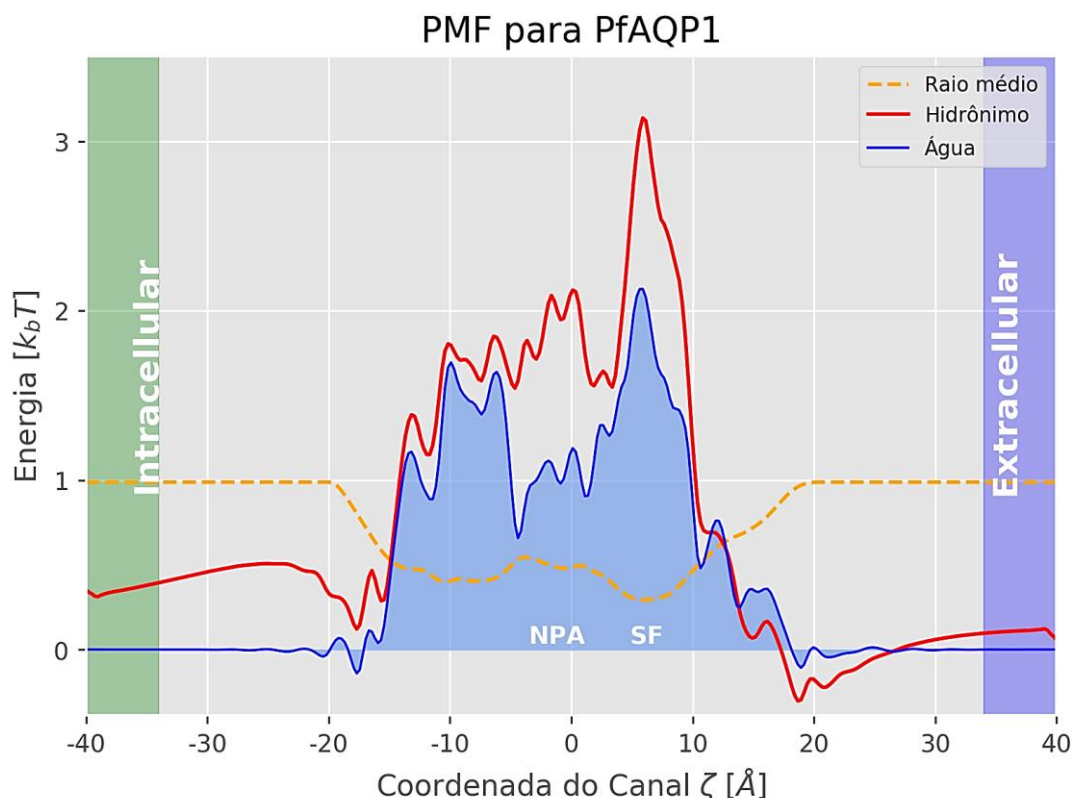


Figura 5.10 PMF médio de 100 estruturas da PfAQP1 para o hidrônio e para a água. Eixo x é a coordenada z do canal centrado no centro de massa da proteína, e o eixo y a energia em unidades de $k_B T$. Raio médio do canal em relação ao raio máximo (~ 8 Å) está representado na linha tracejada. O cátion hidrônio aqui é usado como uma forma de controle positivo, uma vez que todas as proteínas da família das aquaporinas não permitem sua passagem.

Para o hidrônio, mesmo este não sendo adequadamente representado pelos modelo de sítios/parâmetros de campo de força, o PMF obtido é coerente. O perfil de energia é mais acentuado que o da água em todos os pontos, sendo que nas regiões do NPA e SF as barreiras de energia são muito maiores que as da água.

Ainda assim, ambos PMF apresentam barreiras mais altas que as esperadas – água em torno de $2k_B T$, principalmente. Como alternativa, pode-se normalizar o PMF em função do raio do canal. O procedimento consiste de computar a razão entre a densidade numérica e a área do canal e a densidade numérica e área do *bulk* – obtendo, efetivamente, a densidade de moléculas relativa à densidade do *bulk*.

Com normalização (Figura 5.11), o perfil de energia livre do canal muda de uma barreira para o de um fosso potencial. Também é interessante notar que o PMF normalizado para a água indica que o transporte é processo espontâneo e, ao

mesmo tempo, a normalização conserva a barreira energética para a passagem do hidrogênio.

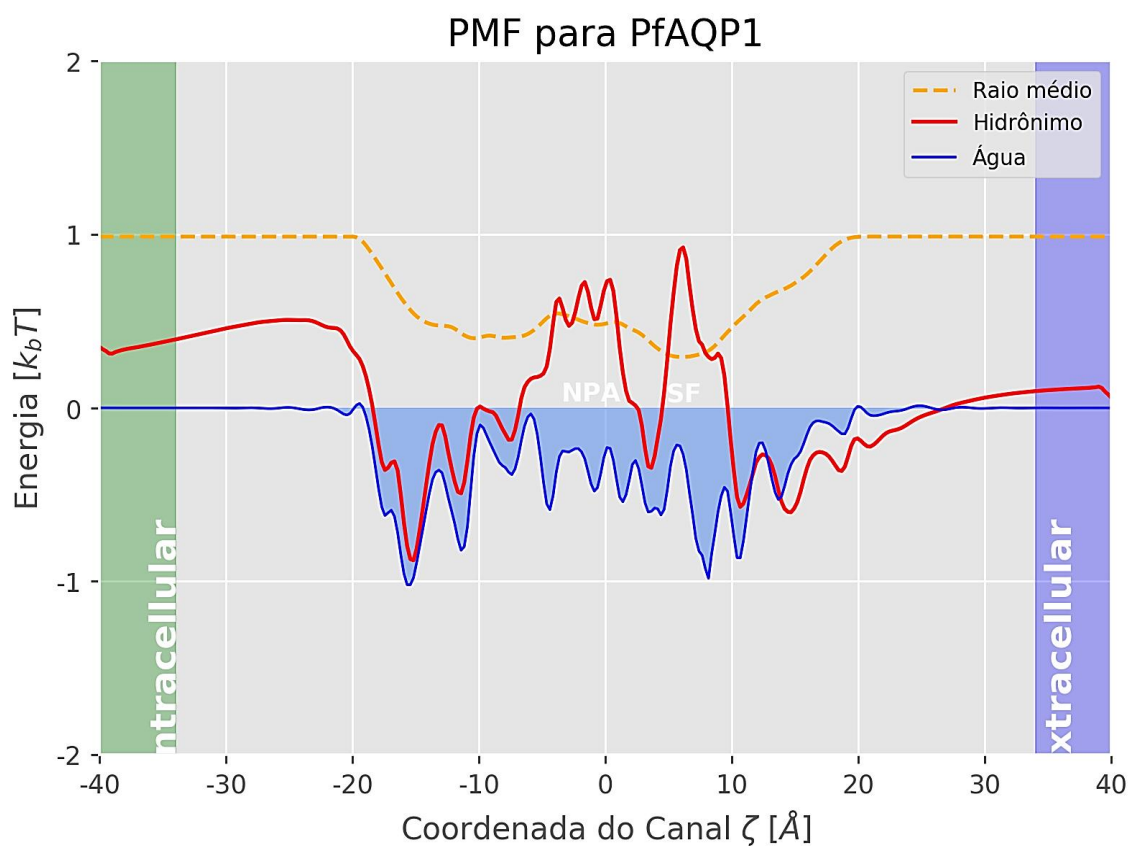


Figura 5.11 PMF médio normalizado pela área do canal da PfAQP1 para o hidrônio e para a água. Eixo x é a coordenada ζ do canal centrado no centro de massa da proteína, e o eixo y a energia em unidades de $k_B T$. Raio médio do canal em relação ao ratio máximo (~ 8 Å) está representado na linha tracejada.

A seguir, encontram-se os PMFs obtidos para os diferentes solventes utilizados no estudo:

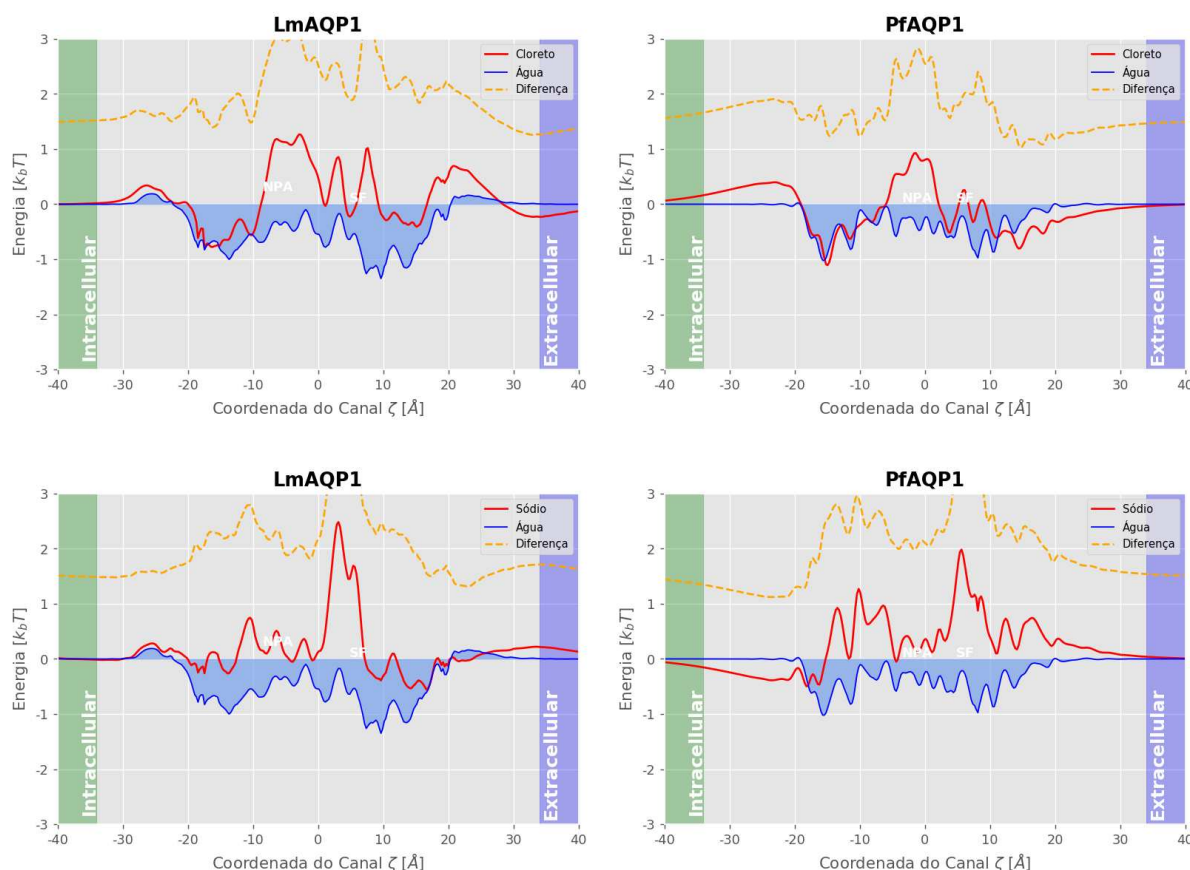


Figura 5.12 PMF normalizado pela área obtido para cloreto (acima) e sódio (abaixo).

Íons cloreto e sódio: O sódio apresenta barreiras bastante acentuadas em ambas as proteínas, o que é esperado, visto que as regiões do SF e NPA estão sob pressão evolucionária para impedir a passagem de cátions de tamanho similar. No entanto, mesmo nas regiões hidrofóbicas se observam barreiras elevadas. Já o cloreto apresenta um perfil energético mais suave, mesmo sendo uma espécie carregada (nota-se que algumas AQPs são permeáveis a cloreto, nitrato e outros ânions). Ademais, a região mais significativa da exclusão parece ser o NPA ($\zeta = 5 \text{ \AA}$) e não SF ($\zeta = 10 \text{ \AA}$). A LmAQP1 apresenta um PMF mais elevada e, portanto, não favorável a passagem do cloreto. Em PfAQP1, o cloreto apresenta uma barreira de energia livre menor ao evento de transporte, o que pode indicar uma possível capacidade do canal de transportar ânions.

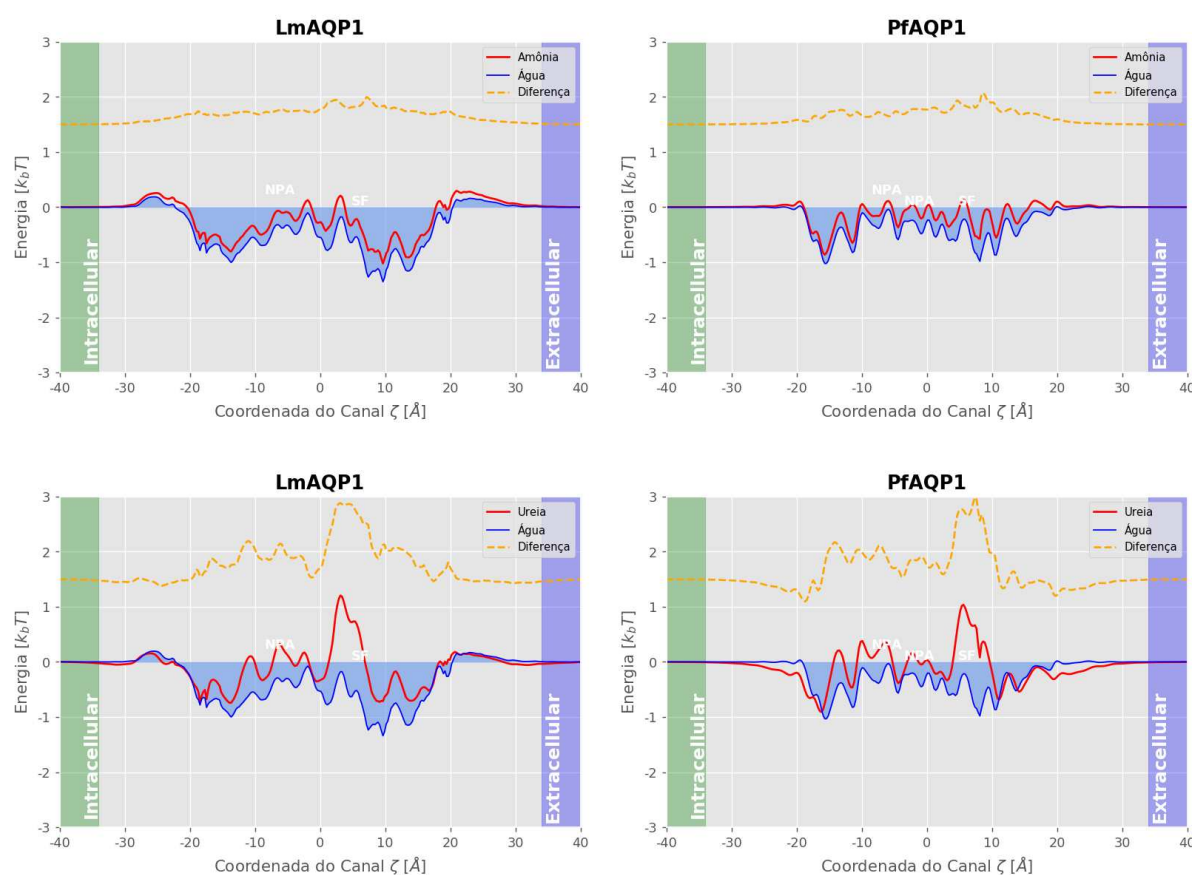


Figura 5.13 PMF normalizado pela área obtido para amônia(acima) e ureia (abaixo).

Amônia e uréia: sabe-se que a PfAQP1 é uma aquaporina permeável à amônia e em menor grau à ureia (10x menos permeável). Isso é aparente nos PMF: barreira próxima ao zero na amônia e de ~ 1 $k_B T$ para ureia. Para LmAQP1, a tendência se repete, ainda que as barreiras sejam levemente maiores.

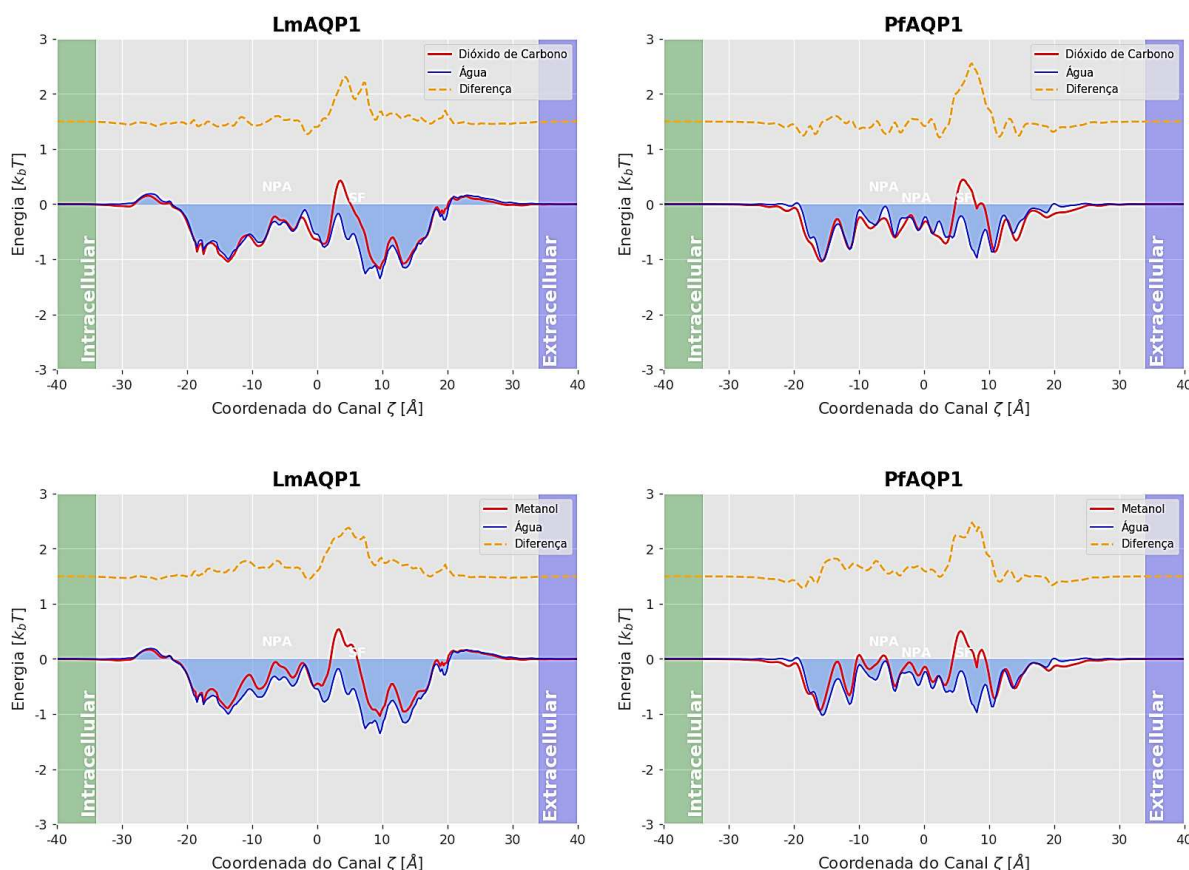


Figura 5.14 PMF normalizado pela área obtido para dióxido de carbono (acima) e metanol (abaixo).

Dióxido de carbono e metanol - devido ao CO_2 ser uma molécula pequena e apolar é esperado que seu transporte pelo canal ocorra. Ainda assim, uma barreira significativa, próxima a $0.5 k_B T$, é observada para a LmAQP1 e PfAQP1. Surpreendentemente, o PMF só diverge significativamente ao da água na região do SF – isso talvez indique que o fator dominante no processo seja simplesmente o tamanho da molécula no caso do CO_2 . O metanol, ainda que mais similar à água, apresentou um perfil semelhante, com ambas as barreiras não ultrapassando $0.5 k_B T$ e o máximo localizado na região do SF.

5.4 Conclusões

- A estrutura da LmAQP1 foi modelada por homologia e, apesar de apresentar um RMSD mais elevado, a região central do canal foi preservada e se manteve estável no decorrer das simulações;

- Ambas LmAQP1 e PfAQP1 possuem um perfil de poro similar, conforme indicado pelo raio médio dos canais – o que atesta a semelhança das estruturas, sendo que a LmAQP1 possui um canal levemente mais estreito (próximo a 1 Å no SF) e também uma maior variabilidade do raio médio;
- Foi observada, em ambas aquaporinas, a passagem de água no canal durante as simulações, sendo que a PfAQP1 aparenta ser a mais permeável. Padrões de ligação de hidrogênio e populações de água no poro foram calculados. Interação água-água mostram-se reduzidas na LmAQP1 de modo a aumentar a afinidade entre a espécie e a proteína;
- Cálculos com a teoria 3D-RISM apontam que a LmAQP1 é permeável à amônia, dióxido de carbono, metanol e, de forma bastante reduzida, à ureia. PfAQP1 parece também ter capacidade de transporte de cloreto e dióxido de carbono.

Capítulo 6

Aquaporinas clássicas de *Leishmania major*

Pouco se sabe sobre as demais aquaporinas da *Leishmania major*, LmAQP α - δ , sendo seus papéis biológicos e funções na célula desconhecidas, à parte que facilitam o transporte passivo da água. A literatura nota apenas que, pela análise filogenética, essas são similares às aquaporinas de plantas e são subcelulares (BIYANI, 2011). Além disso, foram identificados motivos de fosforilação na sequência de LmAQP α (Y185 e Y192). Desse modo, o presente estudo busca explorar as características estruturais e de permeabilidade frente à água das LmAQP α - γ , utilizando em conjunto o 3D-RISM, a fim de investigar a permeabilidade frente a diferentes espécies químicas. Excluí-se a LmAQP δ do estudo por esta não ser expressa na forma amastigota e possuir baixa similaridade com as demais e proteínas do PDB (a sequência de aminoácido desta tem duas vezes o número de aminoácidos das demais).

6.1 Descrição e Análise estrutural

Alinhamento - Estrutura de Referência e AQPs estudadas					
5I32	1	MSHHHHHHHHHDSNGI	PTENLYFQGAGVAFSGSFDSSFLASLRAYLAEFISTLLFVF	-AGVGSAlAYAKLTSDAA	75
LmAQPα	1	MLSEF-----	-----LSQLVAEFVGTFLVLTLIALASV-----		28
LmAQPγ	1	MATKANQMLTGANDDQV [32]	PAPVLNTNTVIQKFSCLQSLRFDALMSAAVVEFIGTFLLVLTVP	LSII-----	97
LmAQPβ	1	MCALQLDAAMGAEEKPY	----LARQIPLIQLSRYQ-----	AKFVCEGIGTFIFLMTTSLAEM-NCGLHAVDGK	63
5I32	76	LDTpGLVAIA-VCHGFALFVAVAIGANISGGHV	NPAVTFGLAV	GGQITVITGVFYWIAQLLGSTAACFLLKYVTGGL	151
LmAQPα	29	---GVGALAPIPIGFMLAAMCFTFGYISGAHF	NPAISFAVFI	NRKMTLRRTVMYIVVQLAGSFCASLYASAIIG-L	100
LmAQPγ	98	---QNAEMAPISIGFLLMSLVFSFGYLSGGHI	NPMVTFSVWL [7]	YATFNKRRLIMYVVAQMAGGVAAAFYCMMINGRD	177
LmAQPβ	64	TRT---RNLAPIAEGFMLCVLIFMFGYISGGHF	NPAVTFGVVM	IRGMRVEEAISYVVAQVVGALVGAGLSILVHGST	137
5I32	152	A-VPTHSVAAGLGS--IEGVVMEIIITFALVYTVYATAADPKKGS	gTIAPLAIGLIVGANILAAGPFSGGSM	NPAARSFG	228
LmAQPα	101	Q-IPAPVAKGDLAN-TWQVLLCELVYTFALTSVVLHVCFSRQRSN	--DFYGFAGMTLMAAGFSVGGFTSGAF	NPAVATG	176
LmAQPγ	178	F--PVPNMGRSFNR-MLRGFLAESVFTFVLCVVLHVAISQQRNN	--NFYGFAGIFSMLCGGLTCGSISSGVF	NPAVATP	252
LmAQPβ	138	QHLPAPQVVQNTAEcIFSAFAEAVFTMLLVTVVLHAAYSQQRNN	--DFYGLAVGMCLLASQ	YAVGVSGGAF	NPAVATG 215
5I32	229	PAV----AAGDFS-----GHVVYVWGPLIGGGLAGLIYGNVF	MG----	SSEHVPLAS-ADF	275
LmAQPα	177	TQL----VLCVYNCDPLFYFWVYWIAPICGAFIASVIYQLLD	TH----	ESVPVVLGKEAVY	230
LmAQPγ	253	LIVVRCLFSFNEAACTPMASLWVYWASEAVGAVCASIMYLAQ	NLGDEV-----		301
LmAQPβ	216	LQVTKFIAAGYFTQ---LLYLWLYWAAPACGAIAAFLFMLTH [4]	DEDGELQQQRVARHLYSSF		278

Tabela 6.1 Alinhamento das sequências de 5I32, LmAQPα, LmAQPβ e LmAQPγ. Motivo NPA em vermelho, SF em laranja, histidina em azul motivo de fosforilação da LmAQPα.

5I32, uma TIP AQP permeável à amônia com estrutura disponível no PDB, exibe maior similaridade com a LmAQPα e LmAQPγ (ainda assim a similaridade é baixa, ~20%).

No banco de dados UniProt, a LmAQPγ é identificada como uma aquaporina putativa, por não apresentar motivo NPA canônico (Figura 6.2). Ainda assim, o restante da estrutura se alinha razoavelmente com as LmAQPα e LmAQPβ. Demais AQPs apresentam NPA conservado (Tabela 6.1), mas não exibem os resíduos do filtro seletivo (SF), que na 5I32 são uma arginina e uma histidina (como é característico de aquaporinas clássicas). Dessa forma, as AQPs LmAQPα e LmAQPγ apresentam um canal essencialmente hidrofóbico; já LmAQPβ apresenta um glutamato e uma glicina (E52 e Q197) formando um SF (Figura 6.1), mas este não parece formar uma região bem definida dentro do poro, uma vez que a região não corresponde a um estreitamento significativo do canal.

Nota-se, também, um *loop* na porção citosólica em todas estruturas, que localiza-se próximo à entrada do canal na LmAQPα. Motivos de fosforilação próximos indicam que este possa estar ligado a mecanismos de regulação do canal (na literatura é reportado que aquaporinas de planta exibem mecanismos similares). A histidina, por vezes associada a regulação por pH do canal, encontra-se

conservada em todas as estruturas. Ademais, o alinhamento mostra que LmAQP α exibe maior similaridade com LmAQP γ do que com LmAQP β .

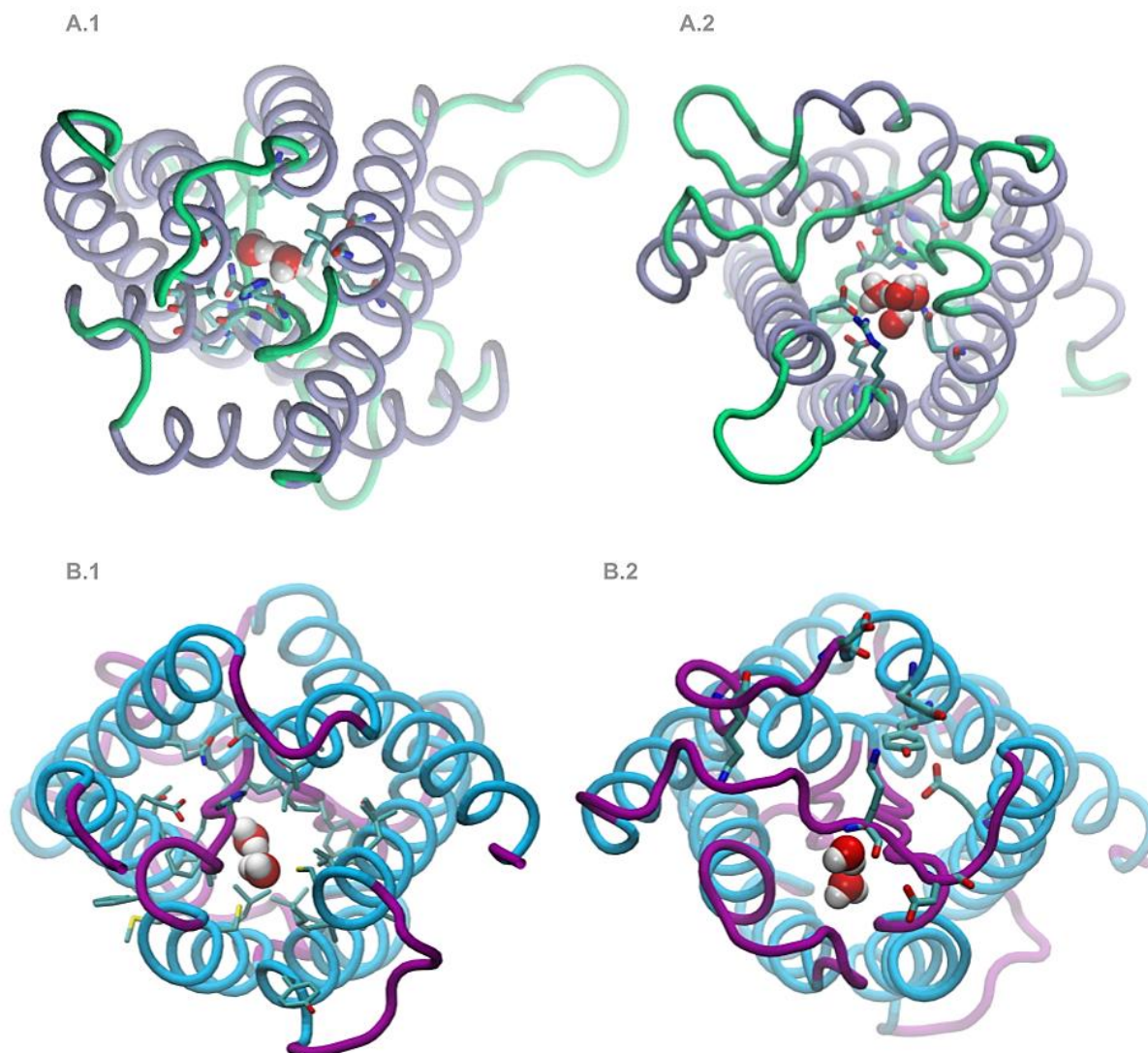


Figura 6.1 Representação do poro da LmAQP α e LmAQP β visto do lado intra e extracelular. (A.1) porção próxima ao lado intracelular de LmAQP β , histidina (H92) e resíduos hidrofóbicos do poro de LmAQP β ; (A.2) porção próxima ao lado extracelular de LmAQP β , resíduos do SF (E52 e Q197) e resíduos carregados próximos; (B.1) porção próxima ao lado intracelular de LmAQP α , histidina (H56) e resíduos hidrofóbicos; (B.2) porção próxima ao lado extracelular de LmAQP α , motivos de fosforilações (Y185 e Y192) e resíduos carregados próximos.

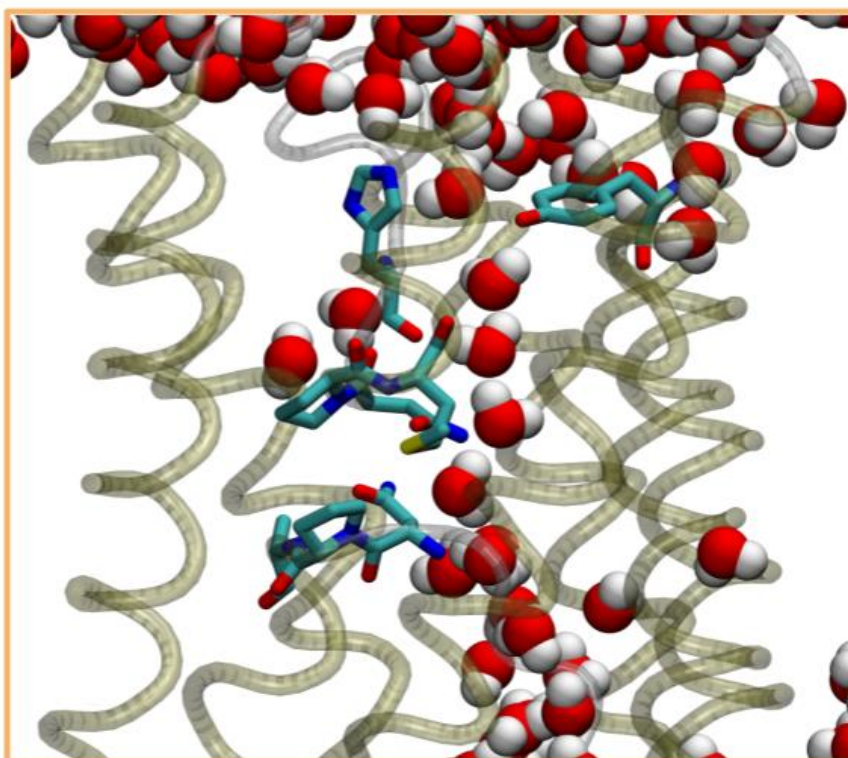


Figura 6.2 Poro da LmAQP γ : NPA não conservado com a presença de uma Metionina e histidina (H125) interagindo com uma tirosina (Y222) bloqueando o movimento das moléculas de água no canal.

6.2 Dinâmica Molecular

6.2.1 Estabilidade

Devido ao grau de identidade das proteínas de referência utilizadas, as estruturas obtidas por homologia possuem qualidade significativamente pior do que a LmAQP1. Assim, os valores de RMSD medidos ao longo da simulação foram bastante elevados. No entanto, como na LmAQP1, ao se separar a contribuição dos carbonos α mais móveis e realinhar a estrutura, os RMSD diminuíram consideravelmente, próximos de 2 Å para todos os átomos e entre 1 e 1,5 Å para os 80% menos móveis (Figura 6.2).

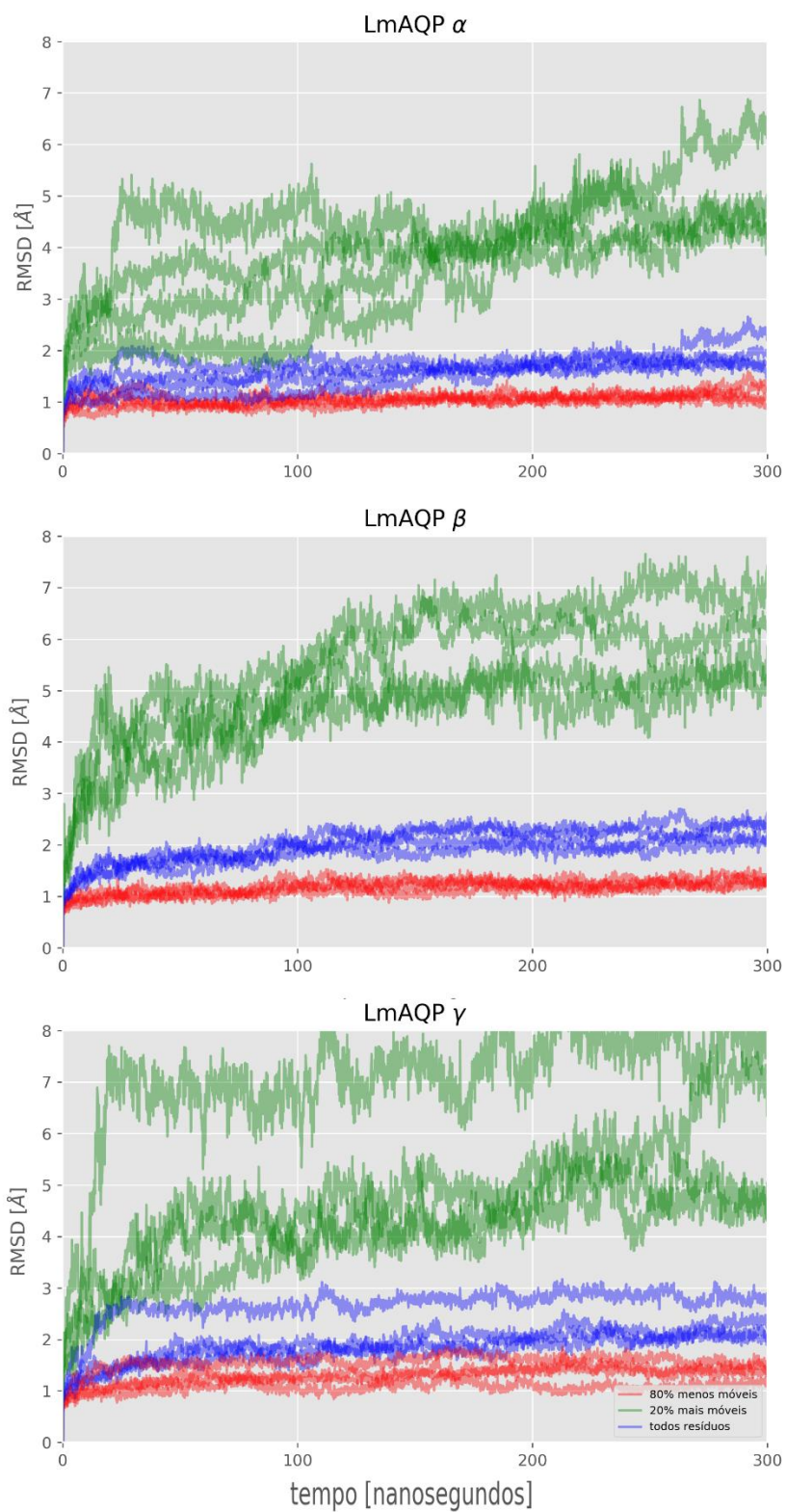


Figura 6.3 RMSD dos carbonos α de diferentes porções da LmAQP α , LmAQP β , LmAQP γ ao longo da simulação. Em vermelho, o RMSD dos 80% menos móveis, em verde dos 20% mais móveis e em azul o RMSD de toda a proteína.

O raio médio do canal das três AQPs apresentou perfil contrastante com os de LmAQ1 e PfAQP1 (Figura 5.6) (Figura 6.4), sendo que na LmAQP α a ausência do SF canônico parece tornar o ambiente do poro mais homogêneo (sem mínimos bem definidos). Já o SF de LmAQP β parece não restringir o raio do canal significativamente – pelo contrário, a região possui raio em torno de 2.5 Å com sua porção mais estreita a próxima à H92 ($\zeta = 0$) e composta majoritariamente por resíduos hidrofóbicos.

A LmAQP γ , por sua vez, apresenta valores menores de raio na porção citosólica em decorrência de Y222, pertencente a um loop intracelular, que interage com a H125 (Fig 6.2). Devido a sua alta mobilidade, este parece ocluir o canal momentaneamente. Nota-se, ainda, que ambas LmAQP α e LmAQP γ possuem o poro com formato irregular, de maneira que este se desvia de uma trajetória retilínea (isto sendo melhor observado na representação tridimensional – Figura 6.5). Todas estruturas apresentam uma variação significativa no raio médio. No entanto, ressalta-se que, para $\zeta > 15$, os valores são em grande parte resultado de artefatos do algoritmo, e não representativos do comportamento da proteína.

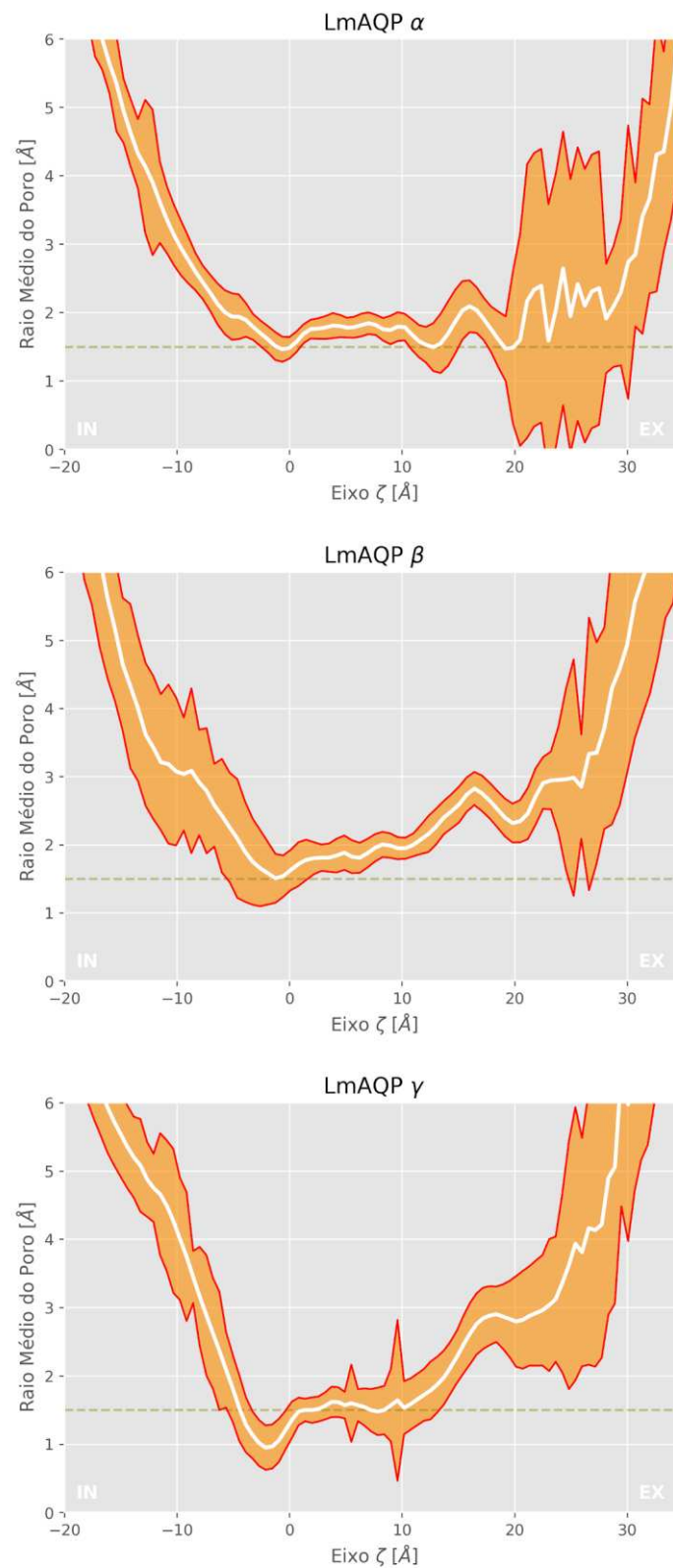


Figura 6.4 Raio médio do canal ao longo da coordenada do poro ζ centrada no centro geométrico da proteína. Região em laranja marca um desvio padrão acima e abaixo do raio médio. 'IN' e 'EX' denotam as porções intracelulares e extracelulares respectivamente. Linha pontilhada marca raio igual a 1,5 Å.

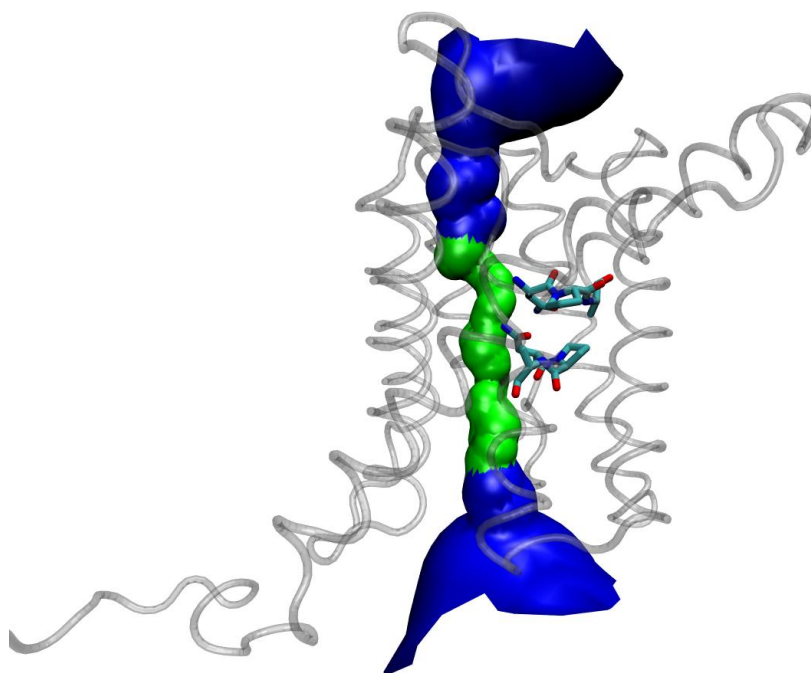


Figura 6.5 Poro da LmAQP α calculado com o programa HOLE. Nota-se a curva no canal acima do motivo NPA.

6.2.2 Permeabilidade à Água

Observa-se a passagem de água no canal em todas as simulações, ainda que em menor grau que nas aquagliceroporinas, uma vez que o fluxo é frequentemente interrompido devido ao movimento de alguns resíduos (vide Y222 e H125 em LmAQP γ). Ainda assim, p_f e p_d (Tabela 6.2) atestam que há transporte de moléculas de água no sistema e este é comparável com os valores reportados para outras AQPs da literatura.

A razão p_f / p_d parece concordar com as medidas experimentais para proteínas da família (entre 7-14), também havendo certa relação com o número de mínimos observados na Figura 6.6.

	p_f [$10^{-14}\text{cm}^3\text{s}^{-1}$]	p_d [$10^{-14}\text{cm}^3\text{s}^{-1}$]	p_f / p_d
LmAQPα	4.71 ± 0.46	1.37 ± 0.34	3.43 ± 0.91
LmAQPβ	7.42 ± 0.91	1.02 ± 0.24	7.27 ± 1.93
LmAQPγ	5.17 ± 0.36	0.99 ± 0.18	5.22 ± 1.01

Tabela 6.2 Permeabilidade osmótica (p_f) e permeabilidade auto difusiva (p_d) em LmAQP α LmAQP β e LmAQP γ .

O mapa de densidade (Figura 6.6) chama a atenção para a similaridade entre os poros de LmAQP α e LmAQP γ . Estes apresentaram um número similar de mínimos e em posições análogas, onde a correspondência para valores de $\zeta < 0$, porção próxima ao citosol, é quase exata. Ambos os poros estão centrados em valores de raio próximos a $4.0 \text{ \AA}$, o que pode indicar a irregularidade do trajeto do poro (Figura 6.5). No mapa de densidade, a tirosina Y222 na LmAQP γ está localizada próximo de $\zeta = -5$, a região que apresenta uma pequena barreira entre dois mínimos. Logo, então é razoável assumir que o efeito da movimentação da cadeia lateral é significativo.

A LmAQP β , por sua vez, apresenta mínimos mais esparsos e uma população significativamente maior de moléculas de água. Ademais, uma região de baixa densidade/barreira não é perceptível na região correspondente ao SF (em $\zeta = 5$), o que atesta a pouca importância estrutural do mesmo. Na porção mais estreita do canal (Figura 6.4) próximo a $\zeta = -5$, ocorre a maior barreira de energia à passagem de moléculas de água do raio do canal, como indicado pela significativa redução da densidade local.

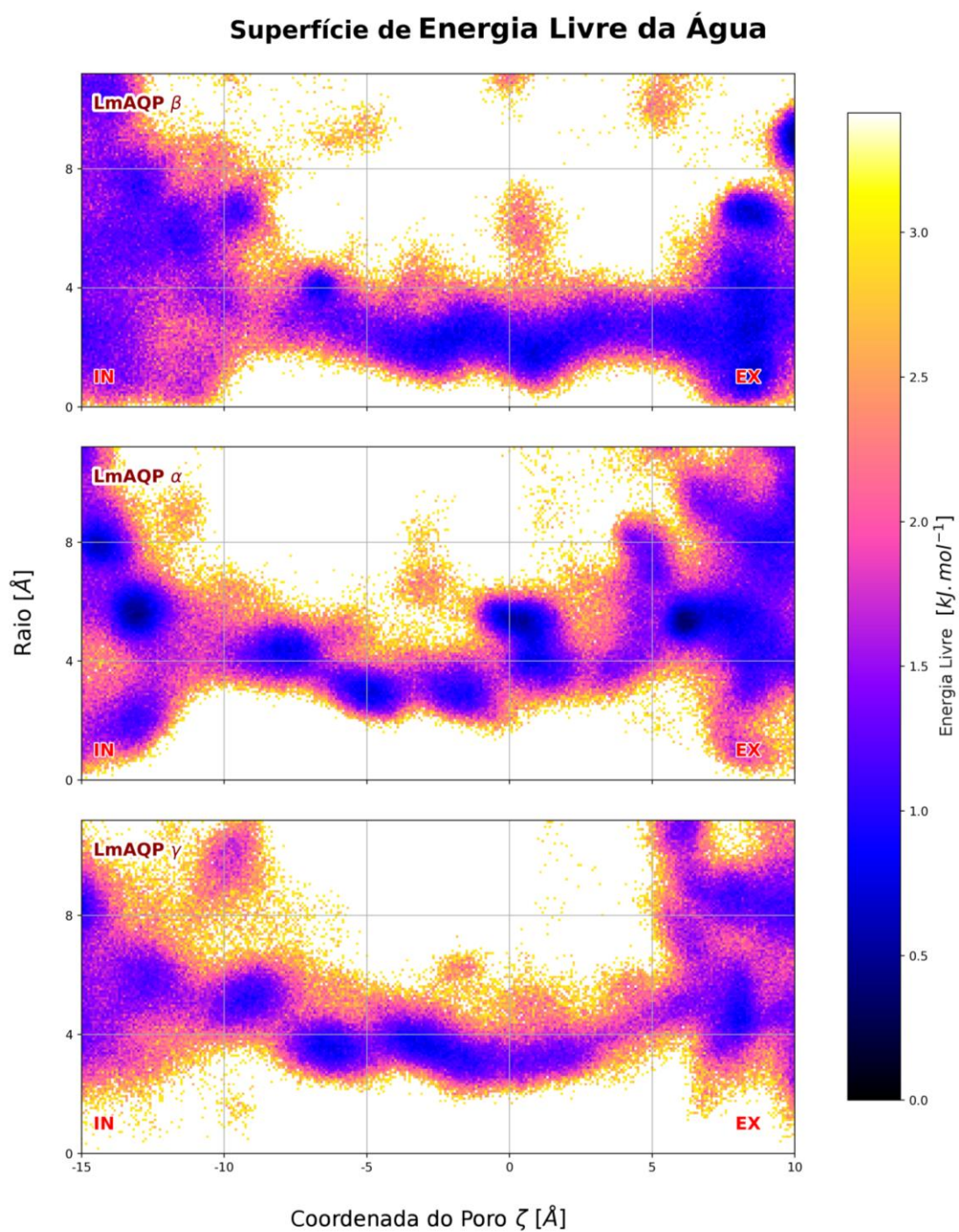


Figura 6.6 Mapa de densidade da água dentro do canal da LmAQP α , LmAQP β e LmAQP γ . 'EX' indica a porção do canal em contato com o meio extracelular e 'IN' com o meio intracelular;

6.3 3D-RISM

Em um primeiro momento, a análise dos PMFs do canal leva a concluir que o comprimento dos canais são significativamente maiores que os apresentados em PfAQP1 (Figura 6.7). No entanto, o perfil energético nas extremidades do canal é afetado pela presença dos *loops* extracelulares da proteína próximos a região, assim, mesmo que já no ‘*bulk*’ o PMFs apresenta barreira energéticas e fossos que não correspondem ao sistema real. Deve-se entender estas regiões enquanto um artefato computacional resultante da integração da função de correlação se dar apenas em uma área retangular fixa. Assim, o PMF para LmAQP α deve ser considerado apenas no intervalo de $\zeta \in [-10, 20]$, para LmAQP β no intervalo $\zeta \in [-10, 15]$ e para LmAQP γ $\zeta \in [-10, 15]$.

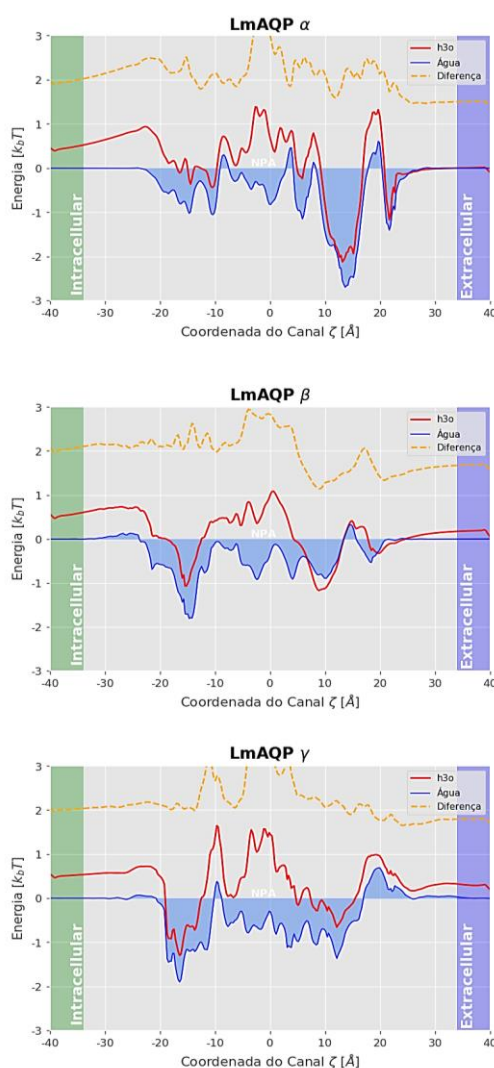


Figura 6.7 PMF normalizado pela área obtido para água e hidrônio.

Íons cloreto e sódio: a região responsável pela seletividade contra ânions parece ser o NPA visto que estas regiões apresentam barreiras de entre 2-3 $k_B T$. Essas também são expressivamente maiores que as em LmAQP1 e PfAQP1 o que sugere que as aquaporinas clássicas não exibem permeabilidade significativa à ânions. Barreiras para o Na^+ , como esperado, indicam que não há passagem do cátion.

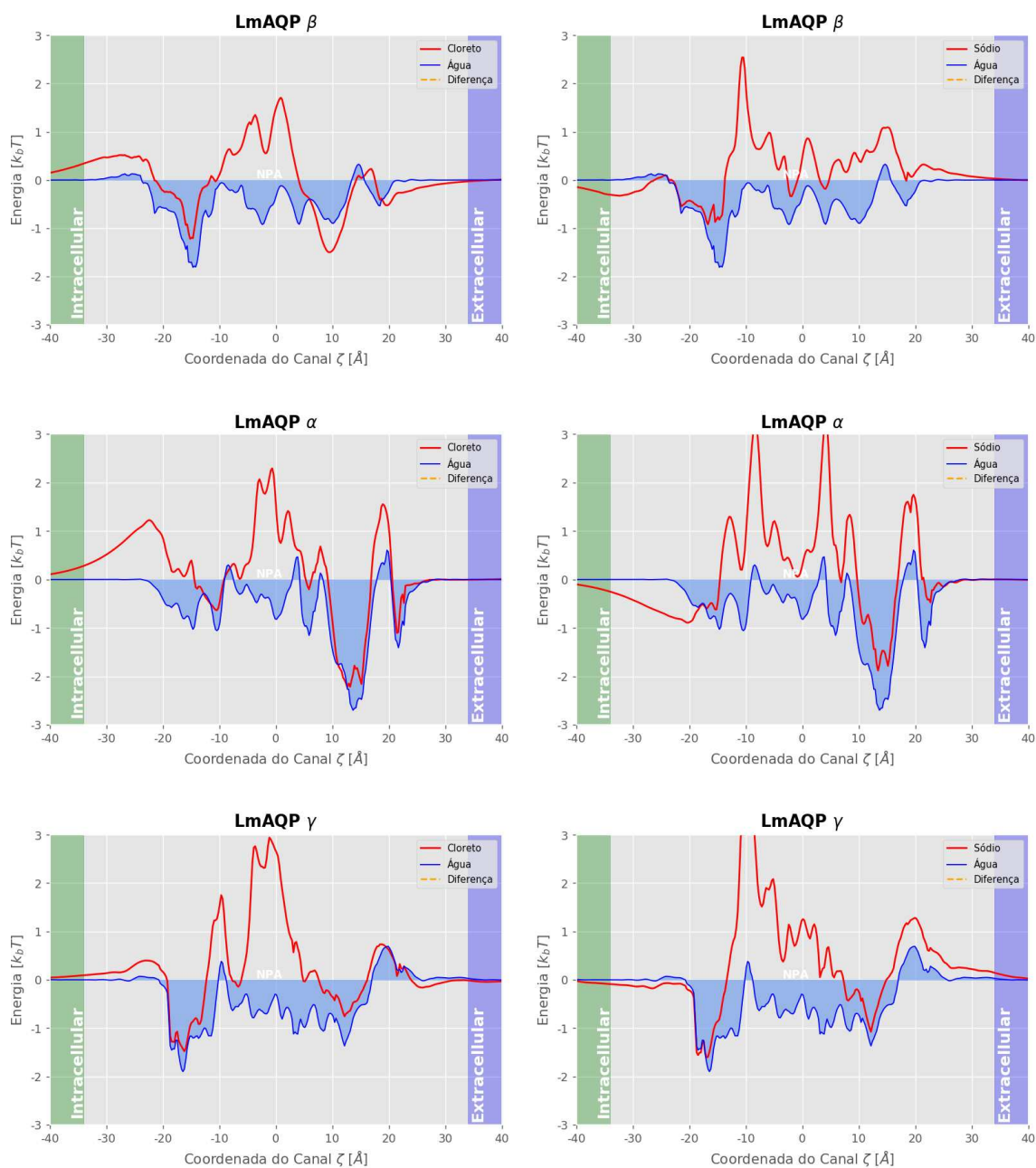


Figura 6.8 PMF normalizado pela área obtido para o cloreto e o sódio.

Amônia e ureia: LmAQP β possui barreiras relativamente pequenas para amônia, enquanto em LmAQP α e LmAQP β estas chegam a pouco mais de 1 $k_B T$. Para a ureia, todas as barreiras são bastante elevadas, principalmente nas porções mais estreias dos canais, o que indica que predomina o efeito estérico neste caso.

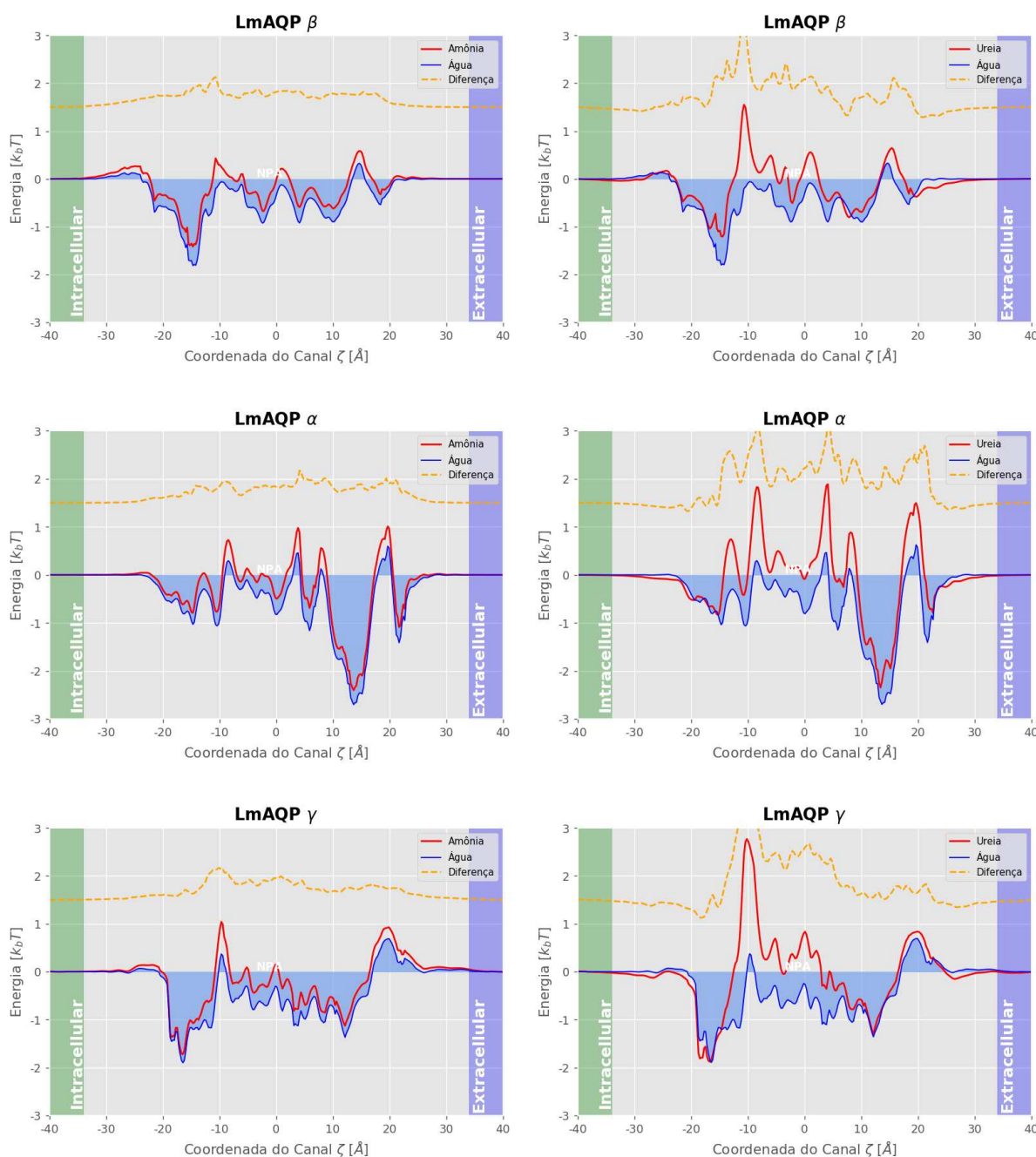


Figura 6.9 PMF normalizado pela área obtido para amônia e ureia.

Dióxido de carbono e metanol: o perfil energético do dióxido de carbono diverge significativamente do perfil para a molécula água em LmAQP γ . Uma barreira acentuada próxima aos resíduos Y222 e H125 ($\zeta = -10$), na região do NPA, foram os principais contribuintes para tal. Já o metanol difere em maior grau do perfil energético exibido para a água. Barreiras em LmAQP α e LmAQP γ são significativas, de 1 k_bT e 2 k_bT , respectivamente. Já em LmAQP β , a barreira é menos restritiva ($\sim 0.8 k_bT$), ainda que seja relevante quando comparada aos demais solventes.

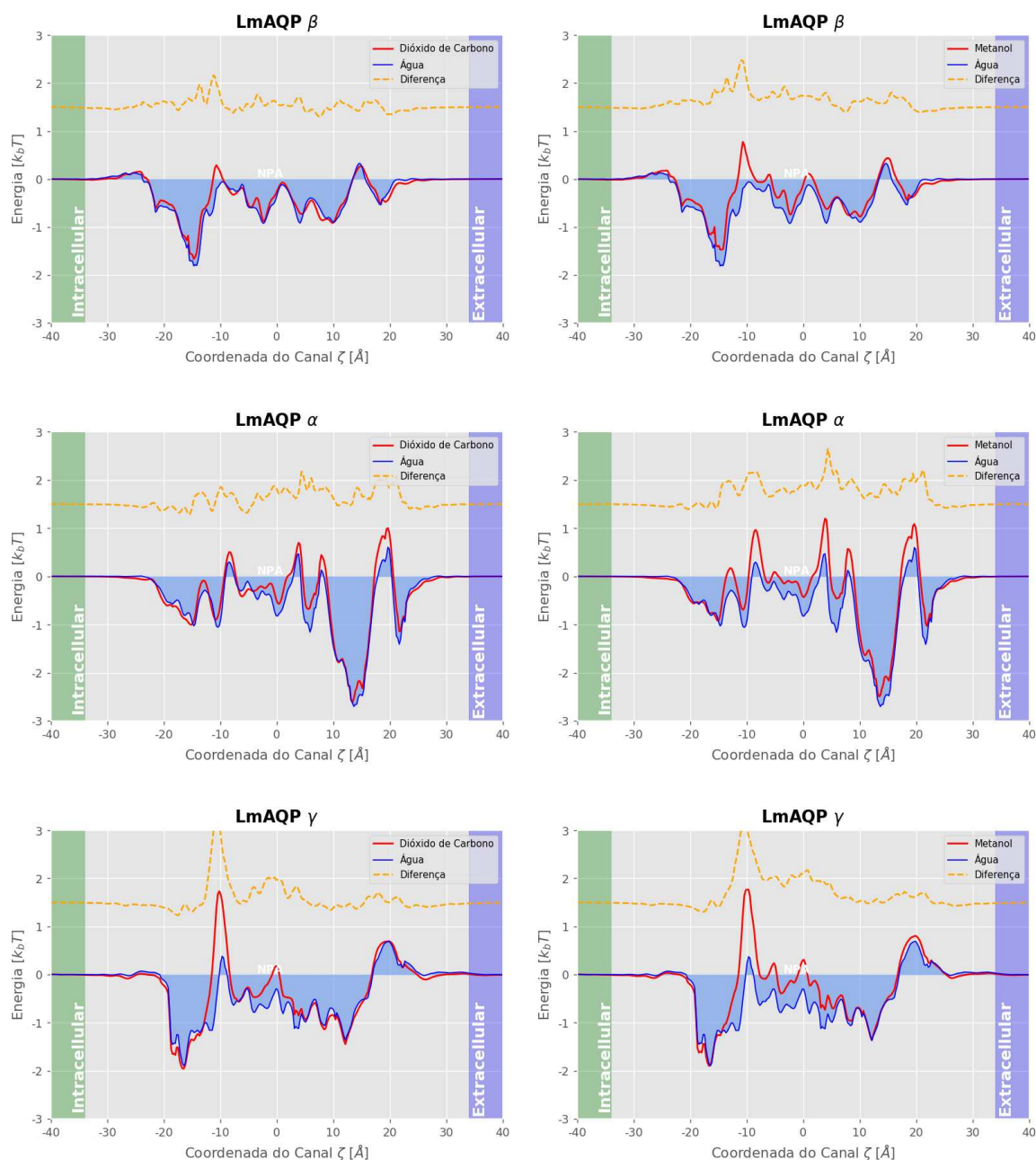


Figura 6.10 PMF normalizado pela área obtido para dióxido de carbono e metanol.

6.4 Conclusões

- As estruturas da LmAQP α , LmAQP β e LmAQP γ foram obtidas por via de modelagem por homologia. LmAQP α e LmAQP γ apresentam estruturas semelhantes, sendo que nenhuma delas exibe a região do SF, de modo que o poro dessas é muito mais hidrofóbico que as demais aquaporinas estudadas. LmAQP β parece possuir um SF não canônico. No lugar de resíduos carregados positivamente (histidina e arginina), como é comum em aquaporinas clássicas, a proteína apresenta duas glutaminas e uma glicina próximos à porção extracelular. No entanto, a importância da região não é clara, uma vez que não é observado um estreitamento do raio próximo ao SF e também não se notam modificações do mapa de densidade.
- Ainda que os modelos gerados apresentem qualidade significativamente pior que a estrutura da LmAQP1, foi observada, nas simulações, a preservação da estrutura do canal e passagem de água no poro, ainda que interrompida por vezes;
- Canais da LmAQP α e LmAQP γ caracterizam-se principalmente por apresentar um perfil irregular, ainda que os sítios de ligação da água sejam bem definidos nestes. É possível que LmAQP γ esteja em uma configuração não produtiva, dada a interação da histidina 125 com a tirosina 222 no poro;
- LmAQP β apresenta SF e raio de poro relativamente largos. O valor de p_f sugere que esta característica favorece o transporte de água;
- Nenhuma das aquaporinas exibe transporte de ânions, mas a região parece ser de crucial importância no processo. LmAQP α e LmAQP γ podem exibir permeabilidade reduzida à amônia. Em LmAQP β , no entanto, o perfil é muito mais similar ao da água, o que indica que esta seja capaz de transportar a espécie.

Capítulo 7

Considerações finais

Dado o exposto, o uso do 3D-RISM para estudo da superfície de energia livre para diferentes moléculas ao longo do canal de aquaporinas (e indiretamente dos processos de transporte destas moléculas através do canal) parece ser uma alternativa ao uso de métodos mais caros de amostragem ampliada. No entanto, este encontra suas limitações em sistemas onde os graus de liberdade da proteína e diferentes conformações do solvente sejam importantes, ainda que algumas dessas limitações possam ser mitigadas computando a média de múltiplas estruturas. O método parece concordar com os resultados experimentais para a maioria dos sistemas utilizados (e.g. hidrônio e cátions em todas aquaporinas e moléculas orgânicas - metanol - pequenas em LmAQP1 e PfAQP1).

Neste estudo, foi possível observar que similaridades estruturais entre LmAQP1 e PfAQP1 e entre LmAQP α e LmAQP γ se refletem nos perfis energéticos calculados pelo 3D-RISM. Destacam-se as seguintes observações: (1) PfAQP1 e LmAQP1 apresentam barreiras energéticas reduzidas para o transporte de ânions e CO₂. A região de maior importância no transporte de ânions parece ser a do motivo NPA; (2) de maneira geral, PMFs para a LmAQP1 apresentaram barreiras de energia maiores, mesmo sendo similares aos de PfAQP1; (3) metanol e CO₂ possuem perfis de energia similares indicando talvez que o efeito estérico do raio do poro seja o dominante neste caso; (4) O perfil energético da amônia indica que esta é permeável no poro da LmAQP1, LmAQP β e PfAQP1; (5) Canais da LmAQP α e LmAQP γ são essencialmente hidrofóbicos e apresentam irregularidades no trajeto; (6) Motivo de fosforilação na LmAQP α próximos ao poro e o bloqueio parcial do canal da LmAQP γ por Y222 indicam que a permeabilidade à água do canal pode ser modulada por modificações nestes resíduos; (6) Resíduos carregados no ambiente do poro, o que normalmente caracteriza a região do SF, parecem não afetar os perfis de permeabilidade em LmAQP β .

Todavia, ressalta-se que um estudo mais detalhado dos sistemas será necessário, a fim de elucidar suas funções na célula. Seriam de grande interesse, principalmente, a obtenção de estruturas cristalográficas e o estudo de diferentes estados de protonação e fosforilação. Ao tocante do RISM, julga-se que seria produtivo realizar análises similares com outras espécies (e.g. ácidos orgânicos, cetonas, boratos etc) e com as formas iônicas das substâncias já estudadas (e.g. amônio, bicarbonato).

Referências Bibliográficas

1. ABASCAL, F., IRISARRI, I. & ZARDOYA, R. Diversity and evolution of membrane intrinsic proteins. *Biochim. Biophys. Acta - Gen. Subj.* 1840, 1468–1481 (2014).
2. ALDER, B. J., FRANKEL, S. P. & LEWINSON, V. A. Radial distribution function calculated by the Monte-Carlo method for a hard sphere fluid. *J. Chem. Phys.* 23, 417–419 (1955).
3. ALLEVA, K., CHARA, O. & AMODEO, G. Aquaporins: Another piece in the osmotic puzzle. *FEBS Lett.* 586, 2991–2999 (2012).
4. APONTE-SANTAMARÍA, C., HUB, J. S. & DE GROOT, B. L. Dynamics and energetics of solute permeation through the *Plasmodium falciparum* aquaglyceroporin. *Phys. Chem. Chem. Phys.* 12, 10246–10254 (2010).
5. ARIZ-EXTREME, I. & HUB, J. S. Potential of Mean Force Calculations of Solute Permeation Across UT-B and AQP1: A Comparison between Molecular Dynamics and 3D-RISM. *J. Phys. Chem. B* 121, 1506–1519 (2017).
6. BATEMAN, A. UNIPROT: A worldwide hub of protein knowledge. *Nucleic Acids Res.* 47, D506–D515 (2019).
7. BIYANI, N. et al. Characterization of *leishmania donovani* aquaporins shows presence of subcellular aquaporins similar to tonoplast intrinsic proteins of plants. *PLoS One* 6, (2011).
8. BONTHUIS, D. J., MAMATKULOV, S. I. & NETZ, R. R. Optimization of classical nonpolarizable force fields for OH⁻ and H₃O⁺. *J. Chem. Phys.* 144, (2016).
9. BROOKS, B. & KARPLUS, M. Harmonic dynamics. *Biophysics (Oxf.)* 80, 6571–6575 (1983).
10. CASE, D. A. & KARPLUS, M. Dynamics of ligand binding to heme proteins. *J. Mol. Biol.* 132, 343–368 (1979).
11. CHANDLER, D. & ANDERSEN, H. C. Optimized cluster expansions for classical fluids. II. Theory of molecular liquids. *J. Chem. Phys.* 57, 1918–1929 (1972).
12. COOPER, A. *Biochemistry : Cooper.* 73, 2740–2741 (1976).
13. DARDEN, T., YORK, D. & PEDERSEN, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089–10092 (1993).
14. DE VIVO, M., MASETTI, M., BOTTEGONI, G. & CAVALLI, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* 59, 4035–4061 (2016).
15. FELLER, S. E., ZHANG, Y., PASTOR, R. W. & BROOKS, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* 103, 4613–4621 (1995).

16. FIELD, M. C. et al. Anti-trypanosomatid drug discovery: An ongoing challenge and a continuing need. *Nat. Rev. Microbiol.* 15, 217–231 (2017).
17. FIGARELLA, K. et al. Biochemical characterization of *Leishmania major* aquaglyceroporin LmAQP1: Possible role in volume regulation and osmotaxis. *Mol. Microbiol.* 65, 1006–1017 (2007).
18. FREITES, J. A., NÉMETH-CAHALAN, K. L., HALL, J. E. & TOBIAS, D. J. Cooperativity and allostery in aquaporin 0 regulation by Ca^{2+} . *Biochim. Biophys. Acta - Biomembr.* 1861, 988–996 (2019).
19. FRENKEL, D., SMIT, B., *Understanding Molecular Simulation*, Academic Press, Inc, 2nd, (2001).
20. FRÉZARD, F., MONTE-NETO, R. & REIS, P. G. Antimony transport mechanisms in resistant *leishmania* parasites. *Biophys. Rev.* 6, 119–132 (2014).
21. GRIEBEL, M., KNAPEK, S., ZUMBUSCH, G., *Numerical Simulation in Molecular Dynamics: Numerics, Algorithms, Parallelization, Applications*, Springer Publishing Company, Incorporated, 1st, (2007).
22. GOOSSENS, K. & DE WINTER, H. Molecular Dynamics Simulations of Membrane Proteins: An Overview. *J. Chem. Inf. Model.* 58, 2193–2202 (2018).
23. GORDON, J. C. et al. H⁺⁺: A server for estimating pK_as and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* 33, 368–371 (2005).
24. HASHMI, S. & AL-SALAM, S. Acute myocardial infarction and myocardial ischemia-reperfusion injury: A comparison. *Int. J. Clin. Exp. Pathol.* 8, 8786–8796 (2015).
25. HUANG, J. et al. CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* 14, 71–73 (2016).
26. HUB, J. S. & DE GROOT, B. L. Mechanism of selectivity in aquaporins and aquaglyceroporins. *Proc. Natl. Acad. Sci.* 105, 1198–1203 (2008).
27. HUMPHREY, W., DALKE, A. & SCHULTEN, K., *J. Mol. Graph.* 14, 33–38 (1996).
28. JO, S., LIM, J. B., KLAUDA, J. B. & IM, W. CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* 97, 50–58 (2009).
29. KALDENHOFF, R. & FISCHER, M. Aquaporins in plants. 169–176 (2006). doi:10.1111/j.1748-1716.2006.01563.x
30. KARPLUS, M. Molecular dynamics of biological macromolecules: A brief history and perspective. *Biopolymers* 68, 350–358 (2003).
31. KUN, J. F. & DE CARVALHO, E. G. Novel therapeutic targets in *Plasmodium falciparum*: Aquaglyceroporins. *Expert Opin. Ther. Targets* 13, 385–394 (2009).
32. LEACH, A. R. *Molecular modelling: principles and application*. Harlow: Pearson Prentice Hall (2001)

33. LINDAHL, V., GOURDON, P., ANDERSSON, M. & HESS, B. Permeability and ammonia selectivity in aquaporin TIP2;1: Linking structure to function. *Sci. Rep.* 8, 1–13 (2018).
34. MACKERELL, A. D. Empirical force fields for biological macromolecules: Overview and issues. *J. Comput. Chem.* 25, 1584–1604 (2004).
35. MARTÍNEZ, L. Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. *PLoS One* 10, 1–10 (2015).
36. MIYAMOTO, S. & KOLLMAN, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* 13, 952–962 (1992).
37. NEWBY, Z. E. R. et al. Crystal structure of the aquaglyceroporin PfAQP from the malarial parasite *Plasmodium falciparum*. *Nat. Struct. Mol. Biol.* 15, 619–625 (2008).
38. PALMER, D. S., MIŠIN, M., FEDOROV, M. V. & LLINAS, A. Fast and General Method to Predict the Physicochemical Properties of Druglike Molecules Using the Integral Equation Theory of Molecular Liquids. *Mol. Pharm.* 12, 3420–3432 (2015).
39. PEREZ DI GIORGIO, J. et al. Prediction of aquaporin function by integrating evolutionary and functional analyses. *J. Membr. Biol.* 247, 107–125 (2014).
40. PERKYN, J. & PETTITT, B. M. A site-site theory for finite concentration saline solutions. *J. Chem. Phys.* 97, 7656–7666 (1992).
41. PHILLIPS, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781–1802 (2005).
42. PONTE-SUCRE, A. et al. Drug resistance and treatment failure in leishmaniasis: A 21st century challenge. *PLoS Negl. Trop. Dis.* 11, 1–24 (2017).
43. ROY, A., KUCUKURAL, A. & ZHANG, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738 (2010).
44. RYCKAERT, J. P., CICCOTTI, G. & BERENDSEN, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23, 327–341 (1977).
45. SALOMON-FERRER, R., CASE, D. A. & WALKER, R. C. An overview of the Amber biomolecular simulation package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 3, 198–210 (2013).
46. SHARMA, M., MANDAL, G., MANDAL, S., BHATTACHARJEE, H. & MUKHOPADHYAY, R. Functional role of lysine 12 in *Leishmania major* AQP1. *Mol. Biochem. Parasitol.* 201, 139–145 (2015).
47. SMART, O. S., NEDUVELIL, J. G., WANG, X., WALLACE, B. A. & SANSOM, M. S. P. HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.* 14, 354–360 (1996).

48. SONG, J., MAK, E., WU, B. & BEITZ, E. Parasite aquaporins: Current developments in drug facilitation and resistance. *Biochim. Biophys. Acta - Gen. Subj.* 1840, 1566–1573 (2014).
49. SRIVASTAVA, R., CHATTOPADHYAYA, M. & BANDYOPADHYAY, P. Calculation of salt-dependent free energy of binding of β -lactoglobulin homodimer formation and mechanism of dimer formation using molecular dynamics simulation and three-dimensional reference interaction site model (3D-RISM): diffuse salt ions and non-po. *Phys. Chem. Chem. Phys.* (2020). doi:10.1039/c9cp05578a
50. TEIXEIRA, M. H. & ARANTES, G. M. Balanced internal hydration discriminates substrate binding to respiratory complex I. *Biochim. Biophys. Acta - Bioenerg.* 1860, 541–548 (2019).
51. TÖRNROTH-HORSEFIELD, S. et al. Structural mechanism of plant aquaporin gating. *Nature* 439, 688–694 (2006).
52. UZCATEGUI, N. L. et al. Alteration in glycerol and metalloid permeability by a single mutation in the extracellular C-loop of *Leishmania major* aquaglyceroporin LmAQP1. *Mol. Microbiol.* 70, 1477–1486 (2008).
53. VAN GUNSTEREN, W. F. et al. Validation of Molecular Simulation: An Overview of Issues. *Angew. Chemie - Int. Ed.* 57, 884–902 (2018).
54. WARSHEL, M. & LEVITT, A. Computer Simulation Protein Folding_warshel.pdf. *Nature* 253, 694–698 (1975).
55. WATERHOUSE, A. et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303 (2018).
56. YANEF, A. et al. Loop B serine of a plasma membrane aquaporin type PIP2 but not PIP1 plays a key role in pH sensing. *Biochim. Biophys. Acta - Biomembr.* 1858, 2778–2787 (2016).
57. YOSHIDA, N., PHONGPHANPHANEE, S., MARUYAMA, Y., IMAI, T. & HIRATA, F. Selective ion-binding by protein probed with the 3D-RISM theory. *J. Am. Chem. Soc.* 128, 12042–12043 (2006).
58. ZHU, F., TAJKHORSHID, E. & SCHULTEN, K. Collective diffusion model for water permeation through microscopic channels. *Phys. Rev. Lett.* 93, 1–4 (2004).
59. ZHU, F., TAJKHORSHID, E. & SCHULTEN, K. Theory and simulation of water permeation in Aquaporin-1. *Biophys. J.* 86, 50–57 (2004)