UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

ÉRIKA CHANG DE AZEVEDO

UDP-N-acetilglicosamina 2-epimerase de *Staphylococcus aureus*: estrutura, dinâmica e
prospecção de novos ligantes

São Carlos
2020

ÉRIKA CHANG DE AZEVEDO

UDP-N-acetilglicosamina 2-epimerase de *Staphylococcus aureus*: estrutura, dinâmica e
prospecção de novos ligantes

Tese apresentada ao Programa de Pós-
Graduação em Física do Instituto de Física de
São Carlos da Universidade de São Paulo, para
obtenção do título de Doutora em Ciências.

Área de concentração: Física Aplicada
Opção: Física Biomolecular
Orientador: Prof. Dr. Alessandro Silva
Nascimento.

Versão Corrigida

(versão original disponível na Unidade que aloja o Programa)

São Carlos

2020

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Azevedo, Érika Chang de
UDP-N-acetilglicosamina 2-epimerase de *Staphylococcus aureus*: estrutura, dinâmica e prospecção de novos ligantes / Érika Chang de Azevedo; orientador Alessandro Silva Nascimento - versão corrigida -- São Carlos, 2020.
137 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Aplicada Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2020.

1. *Staphylococcus aureus*. 2. Ácidos teicóicos de parede. 3. MRSA. 4. UDP-N-acetilglicosamina 2-epimerase. 5. Antibióticos beta-lactâmicos. I. Silva Nascimento, Alessandro, orient. II. Título.

Agradecimentos

Gostaria de agradecer à CAPES e à Fapesp, pelo auxílio financeiro.

Agradeço também ao Instituto de Física de São Carlos, IFSC-USP, ao Grupo de Biotecnologia Molecular (GBM) e ao Grupo de Biofísica Molecular “Sérgio Mascarenhas” pela infraestrutura.

No decorrer do meu doutoramento, tive a felicidade de contar com diversas pessoas, sem as quais este trabalho não seria possível. A primeira pessoa a quem dedico meu trabalho e agradeço é meu orientador, Alessandro Silva Nascimento. Agradeço pela paciência e pela dedicação. Por ter acreditado em mim e por ter me apoiado nos meus piores momentos, que não foram poucos. Pelas correções, pelos conselhos, pelos cafés, por todos esses anos.

Gostaria de agradecer também à minha família, que compreendeu todas as vezes que não pude estar presente. Mas gostaria de agradecer em especial à minha mãe, Suely, que partiu muito cedo deste mundo, mas que me deu bases para ser quem eu sou. Minha mãe, que apesar de não ter concluído seu curso superior, sempre valorizou minha educação, me ensinando que este é meu bem mais precioso, e que ninguém nunca poderá toma-lo de mim. Agradeço também a minha avó materna, Wang Lian Tje, que partiu durante meus anos de doutorado, e a qual não pude acompanhar durante seus últimos dias terrenos. Esta mulher forte, severa, mas que se mostrou doce e amável em seus últimos anos de vida. Sinto-as presentes em tudo o que faço e acredito que ambas estariam muito orgulhosas da mulher e pesquisadora que me tornei.

Agradeço a meus colegas de grupo, Heloísa, João Victor, Erick, Karina, Milton, Geraldo, Victor, Stheffany, Livinha, Raíssa, Luana e Angélica, os quais sempre me ofereceram apoio, amizade e cafés. Muitos deles já não são parte do nosso grupo, seguindo outros rumos e outras carreiras, mas ainda assim muito presentes nos resultados obtidos neste trabalho. Gostaria de agradecer em especial ao Milton, que se foi devido a uma situação trágica, mas não sem antes deixar ao mundo seu exemplo de serenidade e comprometimento. Passamos algum tempo juntos, jogamos nos finais de semana, conversamos sobre o trabalho e sobre a vida. Sua partida foi muito sentida, nunca te esqueceremos.

Além disso, diversos técnicos foram essenciais para a realização deste trabalho, portanto gostaria de agradecer à Maria, Livia, Josimar e João Possato pela constante disposição em ajudar a todos do GBM. Durante o período em que realizei experimentos no Grupo de Biofísica Molecular “Sérgio Mascarenhas”, também contei com a ajuda das técnicas Andressa e Isabel (“Belezinha”). Muito obrigada.

Não poderia deixar também de agradecer aos meus amigos que me acompanharam nesta jornada, alguns desde o mestrado, desde a graduação, outros da vida. Agradeço à Ana Junger, Cássia, Hellen, Stephânia (“Ste”), Everton (“Tonzinho”), Ana Laura, Laís Brazaca, Francisco, Tiago Tolve, Marcos, Murilo (“Muddylow”) entre tantos outros que caminharam junto comigo por todos esses anos.

Meu último, e mais especial, agradecimento é dedicado ao meu noivo, Paulo, o qual estive ao meu lado “*through thick and thin*” durante os últimos 11 anos, que me apoiou, brigou comigo, me ofereceu colo (e chocolate), que sempre tentou me fazer mais feliz. E conseguiu. Palavras não são suficientes para te agradecer, amor.

Enfim, agradeço a todos aqueles que direta ou indiretamente contribuíram para a realização deste trabalho.

“Até mais, e obrigado pelos peixes!”

Douglas Adams - O Guia do Mochileiro das Galáxias

RESUMO

AZEVEDO, E. C. **UDP-N-acetilglicosamina 2-epimerase de *Staphylococcus aureus*: estrutura, dinâmica e prospecção de novos ligantes.** 2020. 137p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2020.

A bactéria Gram-positiva *Staphylococcus aureus* é uma causadora importante de infecções resistentes a antibióticos nos sistemas de saúde em todo o mundo. Tem sido demonstrado que os ácidos teicóicos de parede (WTA) pode ser um alvo importante para o desenvolvimento de fármacos que combatem infecções resistentes, em especial aos antibióticos beta-lactâmicos. A enzima UDP-N-acetilglicosamina (UDP-GlcNac) 2-epimerase ou MnaA de *S. aureus*, é uma das primeiras enzimas que compõe a via de biossíntese da WTA. Neste trabalho, simulações de dinâmica molecular detalhadas utilizando a estrutura cristalográfica da enzima foram realizadas para caracterizar as mudanças conformacionais que ocorrem devido à presença ou ausência do cofator UDP e do substrato UDP-GlcNac. Utilizando diferentes técnicas de simulação, como dinâmicas moleculares aceleradas pela adição de um potencial enviesado (ABMD) e pela adição de potenciais gaussianos ao potencial total ou de diedro do sistema (GAMD), foi possível acessar o perfil energético para a proteína com e sem ligantes. Os dados obtidos em mais de 32 μ s de dinâmica molecular indicam que os aminoácidos presentes no sítio ativo e carregados positivamente têm papel fundamental no movimento de fechamento entre os domínios e na manutenção do substrato no sítio catalítico. Além disso, foi possível concluir que a proteína apresenta um perfil energético dinâmico, que é modificado pela presença das moléculas de UDP/UDP-GlcNac e que fornece indícios sobre o mecanismo de entrada do substrato e saída do produto. Para a busca de ligantes e possíveis inibidores da enzima, foram realizados experimentos de docagem molecular, onde foram encontrados 17 possíveis ligantes. Para a realização dos testes *in vitro*, a proteína recombinante foi produzida e testada em ensaios de DSF e DSF quantitativo com a adição do substrato e das moléculas selecionadas no processo de docagem. Foi identificada a interação entre a enzima MnaA e o beta-lactâmico ticarcilina, abrindo a possibilidade de uma nova farmacodinâmica associada a este antibiótico, o qual já é utilizado comercialmente.

Palavras chaves: *Staphylococcus aureus*. Ácidos teicóicos de parede. MRSA. UDP-N-acetilglicosamina 2-epimerase. Antibióticos beta-lactâmicos.

ABSTRACT

AZEVEDO, E. C. *Staphylococcus aureus* UDP-N-acetylglucosamine 2-epimerase: structure, dynamics and ligand prospection. 2020. 137p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2020.

Staphylococcus aureus (a Gram-positive bacteria) is a major cause of antibiotic-resistant infections in healthcare systems worldwide. It has been shown that the wall teichoic acid (WTA) presents an important target for the development of drugs that fight resistant infections, especially beta-lactam antibiotics. An UDP-N-acetylglucosamine (UDP-GlcNac) 2-epimerase enzyme (or MnaA) from *S. aureus*, is one of the first enzymes of the WTA biosynthesis pathway. In this work, simulations of molecular dynamics were performed using the crystallographic structure of the enzyme to characterize conformational changes as they occur due to the presence or absence of the UDP (cofactor) and the UDP-GlcNac (substrate). Using different simulation techniques, such as accelerated molecular dynamics by adding a biased potential (ABMD) and adding Gaussian potentials to the total or dihedral potential of the system (GAMD), it was possible to access the energy profile of the systems with and without ligands. The data collected in more than 32 μ s of simulation allowed us to analyze the role of the positively charged amino acids present in the active in the movement between the domains and in the maintenance of the substrate in the catalytic site. Also, the protein has an active energy profile which is modified by the presence of UDP / UDP-GlcNac molecules and could lead to understanding the mechanisms in which the substrate enters and/or exits the active site. A search for ligands and possible inhibitors was carried out using the LiBELA docking software, where 17 possible ligands were found. In vitro tests were made using the recombinant protein in differential scanning fluorimetry (DSF) and quantitative DSF (qDSF) assays. An interaction between MnaA and the beta-lactam antibiotic, ticarcillin, was found, leading to the possibility of a new pharmacodynamics associated with this drug, that is already commercially available.

Keywords: *Staphylococcus aureus*. Wall teichoic acid. MRSA. UDP-N-acetylglucosamine 2-epimerase. Beta-lactam antibiotics.

LISTA DE FIGURAS

- Figura 1 – Mecanismos de transferência de DNA empregados por bactérias. A transformação ocorre quando há a lise de uma célula e a posterior transferência do DNA livre para uma célula receptora. A transdução é o processo no qual o DNA é transferido por intermédio de um vírus bacteriófago. O processo de conjugação ocorre pela junção de duas células onde é formado um canal para a transferência do DNA..... 28
- Figura 2 – Linha temporal indicando os períodos de introdução dos antibióticos no tratamento de doenças e a identificação da resistência aos mesmos em bactérias. O intervalo de tempo entre a introdução de novos antibióticos e a identificação de resistência tem sido cada vez menor, tornando a demanda pelo desenvolvimento de novos fármacos cada vez mais urgente..... 29
- Figura 3 – Infográfico fornecido pela Organização Mundial da Saúde destacando as principais causas da resistência a antibióticos. Dentre elas, se encontra o mal-uso de antibióticos para tratamentos de doenças em humanos e na pecuária, além da precarização dos procedimentos de higiene em ambientes hospitalares e a falta de novos antibióticos para tratamento de infecções bacterianas. 30
- Figura 4 – Infográfico fornecido pela Organização Mundial da Saúde sobre as duas principais vias de propagação de bactérias resistentes a antibióticos: pelo uso inadequado dos medicamentos e através da alimentação. 31
- Figura 5 – Número de novos agentes antibacterianos desenvolvidos desde o período de 1983-1987, que era de 16 novas moléculas, até 2008-2012, que foi de apenas duas. 32
- Figura 6 – Estruturas dos antibióticos beta-lactâmicos. 36
- Figura 7 – Distribuição dos antibióticos mais produzidos e consumidos no mundo, os beta-lactâmicos cefalosporina e penicilina somados representam 54% do total. 38
- Figura 8 – Espectro de ação de antibióticos utilizados in vivo. Em destaque, o aminoglicosídeo tobramicina e o beta-lactâmico penicilina, os quais são utilizados em conjunto para o combate de infecções causadas por agentes desconhecidos. 39
- Figura 9 – Composição da parede celular em bactérias Gram-positivas, contendo a membrana celular, a parede de peptidoglicanos e a ácidos teicóicos de parede..... 41
- Figura 10 - Via de síntese dos ácidos teicóicos de parede(WTA). A síntese se inicia na região interna da membrana plasmática, com a ação da proteína TarO, terminando ligada à porção externa da parede de peptidoglicanos. 41

Figura 11 –O “Paradoxo dos genes essenciais”, onde a deleção das enzimas TarGH causa a morte celular (A), porém a deleção conjunta de um dos primeiros passos (TarO ou TarA) permite a viabilidade do microrganismo (D).	42
Figura 12 –Via de síntese dos ácidos teicóicos de parede. Em destaque, a proteína MnaA, a qual fornece substrato para as enzimas TarA e TarO.....	45
Figura 13 –Estrutura cristalográfica da cadeia A da proteína UDP-GlcNac 2-epimerase de <i>S. aureus</i> (PDB: 5ENZ), com uma molécula de UDP em sítio ativo, vista de frente (A) e rotacionada de 90° (B), como indicado na figura. Em vermelho, o domínio C-terminal e em azul, o domínio N-terminal.	46
Figura 14 –Mecanismo de interconversão entre os compostos UDP-GlcNac e UDP-ManAc, para a enzima MnaA de <i>S. aureus</i> , que é não-hidrolisante.	47
Figura 15 –Estruturas cristalográficas das enzimas UDP-GlcNac 2- epimerase de <i>S. aureus</i> (5ENZ, em verde)) e <i>M. jannaschii</i> , resolvida sem (4NEQ, em vermelho) e com ligantes (4NES, em azul). As estruturas foram alinhadas por seu domínio N-terminal, de modo a evidenciar as diferenças de ângulo de abertura entre os domínios. Os ligantes das estruturas foram removidos da figura para facilitar a visualização.....	48
Figura 16 –Representações atomísticas e SIRAH de aminoácidos, ácidos nucleicos, solvente e íons.	55
Figura 17 - Construção para produção da proteína produzida, a qual possui uma cauda de hexa-histidinas em sua região N- terminal, seguida pela proteína carreadora tiorreoxina, um sítio para clivagem pela enzima TEV protease e a proteína de interesse.	59
Figura 18 –Estrutura cristalográfica da cadeia A da proteína MnaA de <i>S. aureus</i> . Em vermelho, os aminoácidos S128, G349 e P213, que compõe o ângulo de abertura medido (CV).....	65
Figura 19 –Simulações de equilíbrio para a proteína SaMnaA. (A) Ângulo de abertura observado durante a simulação de 1 μ s do sistema APO. (B) Histograma dos ângulos de abertura observados durante a simulação de 1 μ s do sistema APO. (C) Ângulo de abertura observado durante a simulação do sistema HOLO. (D) Histograma dos ângulos de abertura observados durante a simulação do sistema HOLO. (E) Ângulo de abertura observado durante a simulação de 5 μ s do sistema APO. (F) Histograma dos ângulos de abertura observados durante a simulação de 5 μ s do sistema APO.	68
Figura 20 –Análise das componentes principais da trajetória do sistema APO durante a simulação de equilíbrio. Projeções da primeira (A) e da segunda (B) componentes da análise.....	69
Figura 21 –Métodos para a redução de dimensionalidade. Evolução temporal das técnicas utilizadas para a decomposição das posições atômicas provenientes	

- de uma simulação de dinâmica molecular em conjuntos de dados representativos menores e posterior análise por MSM..... 71
- Figura 22 – Processo de análise de dados de uma simulação por MSM. (a) Uma ou mais trajetórias de dinâmica molecular, onde cada frame é separado por um intervalo de tempo da simulação. (b) É realizada uma decomposição discreta dos estados encontrados na trajetória. (c) É construída uma matriz de contagem das transições observadas, a qual registra quantas vezes o sistema transiciona de uma posição na linha para uma posição na coluna. (d) A partir da matriz anterior, é calculada uma matriz de probabilidade de transições reversíveis, que é a transposição da anterior, de modo a garantir que as transições forward ocorram um número igual de vezes que as transições reverse. (e) São calculadas as populações ajustadas que são mais termodinamicamente estáveis. (f) Representação esquemática dos autoprocesso dominantes da trajetória, evidenciando os autofluxos entre os estados..... 72
- Figura 23 – Análise de MSM da trajetória da simulação de 5 μ s para o sistema APO da proteína SaMnaA. (A) Comparação entre os ângulos CV medidos (gráfico superior) e a tIC 1 da análise tICA (gráfico inferior). (B) Mapa da tIC 1 versus tIC 2, onde os estados encontrados estão representados por três agrupamentos de pontos vermelhos, verdes e azuis. (C) Estruturas representativas dos três estados encontrados para o sistema. As esferas representam os resíduos utilizados para o cálculo de CV. 73
- Figura 24 – Ângulos acessados durante a simulação coarse-grained. O gráfico à esquerda mostra os ângulos CV durante os 25 μ s de simulação. À direita, a distribuição de ângulos. É possível notar que o pico de distribuição se encontra em aproximadamente 75°, configurando o estado “aberto”, esperado para a proteína APO. 74
- Figura 25 - Análise das componentes principais da trajetória do sistema HOLO durante a simulação de equilíbrio. Os ligantes estão representados em azul. Projeções da primeira (A) e da segunda (B) componentes da análise..... 75
- Figura 26 – Análise da interação entre UDP-GlcNac, UDP e SaMnaA. Mapa das interações entre UDP (A) e UDP-GlcNac (B) com a proteína SaMnaA. Contribuição da energia potencial total de interação entre UDP (C) e UDP-GlcNac (D) e os os principais aminoácidos interagentes da proteína SaMnaA..... 76
- Figura 27 – Simulações de GAMD para o sistema APO da estrutura SaMnaA. Os resultados das três simulações estão mostrados: GAMD-APO-I em preto, GAMD-APO-II em vermelho e GAMD-APO-III em azul. (A) Ângulos CV observados em função do tempo de simulação. (B) Distribuição dos ângulos CV durante a simulação. (C) PMF computada após o rebalanceamento. 79
- Figura 28 – Dados da simulação GAMD para o sistema HOLO. Os resultados das três simulações independentes estão mostrados: GAMD-HOLO I em preto, GAMD-HOLO II em vermelho GAMD-HOLO III em azul. (A) Ângulos CV observados durante a simulação. (B) Histograma dos ângulos

observados em função do tempo de simulação. (C) PMFs calculadas após o rebalanceamento. (D) RMSD calculado para as moléculas de UDP (preto) e UDP-GlcNac (vermelho). (E) Energias totais de interação computadas entre a proteína e as moléculas de UDP (preto) e UDP-GlcNac (vermelho). 80

Figura 29 – Simulações de ABMD para o sistema APO. Gráficos representativos das simulações denominadas ABMD-APO-I (linha superior), ABMD-APO-II (linha média) e ABMD-APO-III (linha inferior). (A) Evolução temporal do ângulo CV. (B) Distribuição dos ângulos CV. (C) Potencial de força média. As linhas em cinza indicam a evolução temporal e convergência da PMF em 10 simulações de 25 ns cada. 82

Figura 30 - Simulações de ABMD para o sistema HOLO. Gráficos representativos das simulações denominadas ABMD-HOLO-I (linha superior), ABMD-HOLO-II (linha média) e ABMD-HOLO-III (linha inferior). (A) Evolução temporal do ângulo CV. (B) Distribuição dos ângulos CV. (C) Potencial de força média. As linhas em cinza indicam a evolução temporal e convergência da PMF em 10 simulações de 25 ns cada. 83

Figura 31 – Energia total de interação potencial computada entre UDP (preto) e UDP-GlcNac (vermelho) com a proteína SaMnaA para as três simulações ABMD. 84

Figura 32 – Forças eletrostáticas nas estruturas de SaMnaA. A soma das forças eletrostáticas de interação entre os domínios C-terminal e N-terminal para os estados aberto (laranja), intermediário (cinza) e fechado (roxo), com os ligantes presentes no sítio ativo (ciano). 85

Figura 33 – Distribuição de ângulos CV acessados pelos mutantes construídos para o sistema APO, no tempo de 350 ns de simulação para cada mutante. No gráfico à esquerda se encontram em destaque as distribuições para a proteína sem mutações (curva em preto), para o mutante M62E-Q66E-L68D (curva em verde) e para o mutante K39E-M62E-Q66E-L68D-M43E (curva em vermelho). É possível observar que as mutações resultaram em um deslocamento das distribuições de ângulos, indicando uma estabilização da proteína em seus estados “intermediário” e “fechado”. 85

Figura 34 – Interação prevista pelo programa de docking LiBELA para a molécula de ticarcilina e a proteína MnaA. (A) Pose obtida para o composto ticarcilina (em amarelo) e UDP (em azul) no sítio ativo; (B) Interações de hidrogênio (em azul) calculadas pelo programa de visualização Chimera entre a ticarcilina e a proteína. 88

Figura 35 - Expressão e purificação da proteína MnaA. Gel de poliacrilamida 15%, onde 1) Marcador de massa molecular; 2) Amostra antes da indução por 0,1 mM de IPTG; 3) Amostra após a indução por 0,1 mM IPTG; 4) Sobrenadante após lise por ultra-sonicação; 5) Pellet após lise por ultra-sonicação; 6) Amostra não ligada à coluna; 7 e 8) Lavagem da coluna de níquel com tampão de ligação; 9) Eluição com tampão adicionado de 500 mM de imidazol; 10) Proteína após a clivagem. 96

Figura 36 – Temperatura de desenovelamento para a MnaA ligada às moléculas selecionadas durante a docagem.	98
Figura 37 – Dados de DSF quantitativo para a enzima UDP-GlcNac 2-epimerase ligada ao seu substrato. (A) Variação da temperature em função do aumento da concentração de ligante. (B) Fluorescência medida para cada concentração de ligante. (C) Curvas isotérmicas no intervalo de temperaturas de 50 a 64 °C, onde a fração de proteína desenovelada está descrita em função da concentração de ligante, em μM . (D) Sobreposição das curvas isotérmicas para todo o interval de temperaturas definido anteriormente.	101
Figura 38 – Relação de van't Hoff com relação ao K_d calculado para os ligantes UDP-GlcNac e ticarcilina.	102
Figura 39 – Variação de T_m em função do aumento da concetração para os compostos testados.	103
Figura 40 - Dados de DSF quantitativos para a enzima UDP-GlcNac 2-epimerase ligada à ticarcilina. (A) Variação da temperature em função do aumento da concentração de ligante. (B) Fluorescência medida para cada concentração de ligante. (C) Curvas isotérmicas no intervalo de temperaturas de 50 a 64 °C, onde a fração de proteína desenovelada está descrita em função da concentração de ligante, em μM . (D) Sobreposição das curvas isotérmicas para todo o interval de temperaturas definido anteriormente.	105
Figura 41 - Teste de solubilidade da proteína TarA. Gel de poliacrilamida 15%, no qual 1) Marcador de massa molecular; 2) Amostra antes da indução por 0,1 mM de IPTG; 3) Amostra após a indução por 0,1 mM de IPTG; 4) Pellet após lise sem detergentes; 5) Sobrenadante após lise sem detergentes; 6) Pellet após lise utilizando o detergente CHAPS 20 mM; 7) Sobrenadante após lise utilizando o detergente CHAPS 20 mM, 8) Pellet após lise utilizando o detergente Sarcosyl 0,1%; 9) Sobrenadante após lise utilizando o detergente Sarcosyl 0,1%; 10) Marcador de massa molecular; 11) Pellet após lise utilizando o detergente Triton X-100 1%; 12) Sobrenadante após lise utilizando o detergente Triton X-100 1%. A seta indica a proteína de interesse.	107
Figura 42 – Purificação por cromatografia de afinidade da proteína TarA. Dados obtidos com o auxílio do equipamento ÄKTA Purifier (GE Healthcare Life Sciences), o qual mostra a absorção em 280 nm da amostra em função do volume de eluição.	108
Figura 43 - Purificação da proteína TarA por cromatografia de afinidade. Gel de poliacrilamida 15%, onde 1) Proteína TarA purificada; 2) Amostra após a indução por 0,1 mM de IPTG; 3) Sobrenadante após lise com tampão de lise e detergente CHAPS 40 mM; 4) Pellet após lise com tampão de lise e detergente CHAPS 40 mM; 5 e 6) Amostra não ligada à coluna; 7) Lavagem da coluna de níquel com tampao de ligação; 8) a 13) Primeiro pico de absorção observado durante o gradiente de imidazol; 14) a 22) Segundo pico de absorção observado durante o gradiente de imidazol.	108

Figura 44 - Diagrama de fases para a cristalização de proteínas. Baseado na variação da concentração de proteína e de precipitante.	109
Figura 45 - Gota na condição de cristalização com tampão HEPES 0,1 M pH 7,5, citrato de sódio 1,4 M, proporção de amostra:tampão 2:1.	110
Figura 46 – (A) Cristal obtido na condição 1 M citrato de sódio e 0,1 M HEPES pH 7,8, proporção de amostra:tampão 2:1. (B) Difração de raios-X dos cristais obtidos na condição tampão HEPES 0,1 M pH 7,5, citrato de sódio 1,4 M, onde o círculo cinza marca a resolução de 7,7 Å e o círculo vermelho marca pontos de difração observados.	111
Figura 47 -Imagem da intensidade de emissão de radiação pela amostra, diminuída da intensidade de emissão do tampão. Em vermelho, emissão da amostra antes da desnaturação por calor e em preto, emissão da amostra após a desnaturação. Linhas verticais demarcam os máximos de intensidade de emissão de cada uma das amostras. A amostra foi excitada em 280 nm.	135
Figura 48 –Espectroscopia de dicroísmo circular para a enzima <i>SaMnaA</i> . (A) Gráfico dos resultados obtidos durante os experimentos. As setas indicam regiões de estrutura secundária. (B) Gráfico de absorção de diversas estruturas secundárias conhecidas quando submetidas à luz circularmente polarizada, onde o espectro para alfa-hélice está representado pela linha sólida, folhas-beta por pontos e traços, voltas-beta pela linha pontilhada e as regiões sem estrutura (<i>random coil</i>) pelos traços.....	137
Figura 49 –Espectroscopia de dicroísmo circular para a enzima <i>SaMnaA</i> . (A) Gráfico dos resultados obtidos durante os experimentos. As setas indicam regiões de estrutura secundária. (B) Gráfico de absorção de diversas estruturas secundárias conhecidas quando submetidas à luz circularmente polarizada, onde o espectro para alfa-hélice está representado pela linha sólida, folhas-beta por pontos e traços, voltas-beta pela linha pontilhada e as regiões sem estrutura (<i>random coil</i>) pelos traços.....	137

LISTA DE TABELAS

Tabela 1 –	Classes de antibióticos descobertos durante a “Era de Ouro”. As classes sublinhadas correspondem a moléculas sintéticas. Os demais são compostos naturais.	33
Tabela 2 –	Mecanismos de ação adquiridos pelas bactérias para evitar os efeitos de antibióticos.	35
Tabela 3 –	Compostos inibitórios das enzimas que fazem parte da síntese dos ácidos teicóicos de parede.	44
Tabela 4 –	Testes de re-sensibilização de bactérias cepa MRSA COL através da mutação e deleção dos genes que codificam para as proteínas Cap5 e MnaA.....	46
Tabela 5 –	Conjunto total de simulações de dinâmica molecular realizados no trabalho, divididos por método, sistema simulado e tempo de simulação.	57
Tabela 6 –	Primers utilizados para a clonagem dos genes que codificam para as proteínas TarA, TarB, TarS, TarL, TarO, TarI, TarJ e MnaA de S. aureus SA16.....	59
Tabela 7 –	Estado de abertura das estruturas cristalográficas conhecidas das enzimas UDP-GlcNac 2-epimerase. *cadeia sem ligantes.....	66
Tabela 8 –	Ligantes selecionados durante o processo de inspeção visual dos compostos após o processo de docagem.	89
Tabela 9 –	Compostos e concentrações utilizados para a realização dos experimentos de DSF quantitativo.....	100

LISTA DE ABREVIATURAS E SIGLAS

ABMD	Dinâmica molecular adaptativa
CD	Dicroísmo circular
CDC	Center for disease control and prevention
DNA	Ácido desoxirribonucleico
DSF	Fluorimetria de varredura diferencial
EARSS	European Antimicrobial Resistance Surveillance System
GAMD	Dinâmica molecular acelerada por gaussianas
GlcNac	N-acetilglicosamina
ManNac	N-acetilmanosamina
MRSA	<i>Staphylococcus aureus</i> resistente a meticilina
OMS	Organização mundial de saúde
qDSF	Fluorimetria de varredura diferencial quantitativa
Tar	Teichoic acid ribitol
UDP	Uridina difosfato
WTA	Ácidos teicóicos de parede

LISTA DE SÍMBOLOS

°C	graus Celsius
K	graus Kelvin
M	molar

SUMÁRIO

1	INTRODUÇÃO	27
1.1	Resistência a antibióticos.....	27
1.2	Antibióticos beta-lactâmicos e MRSA.....	36
1.3	Staphylococcus aureus e seus ácidos teicóicos de parede (WTA).....	39
1.4	Os ácidos teicóicos de parede como alvo para o desenvolvimento de novos fármacos	42
1.5	A enzima UDP-N-acetilglicosamina de S. aureus (SaMnaA).....	44
2	OBJETIVOS.....	51
3	MATERIAIS E MÉTODOS.....	53
3.1	Metodologias computacionais	53
3.1.1	Parametrização da proteína e dos ligantes	53
3.1.2	Dinâmica molecular em equilíbrio	54
3.1.3	Dinâmica molecular acelerada por gaussianas.....	55
3.1.4	Dinâmica molecular adaptativa	56
3.1.5	Análise estrutural.....	57
3.1.6	Docagem de compostos	57
3.2	Metodologias experimentais	58
3.2.1	Clonagem e expressão das enzimas envolvidas na síntese da WTA de S. aureus	58
3.2.2	Testes de solubilidade	60
3.2.3	Purificação por afinidade	60
3.2.4	Clivagem pela enzima TEV protease.....	61
3.2.5	Fluorescência intrínseca	61
3.2.6	Espectroscopia de dicroísmo circular (CD)	62
3.2.7	Experimentos de Cristalização.....	62
3.2.8	Fluorimetria de varredura diferencial (DSF) e Fluorimetria de varredura diferencial quantitativa (qDSF)	63
4	RESULTADOS E DISCUSSÃO	65

4.1	Análise das estruturas cristalográficas existentes de MnaA.....	65
4.2	Simulações de dinâmica molecular	67
4.2.1	Simulações de equilíbrio.....	67
4.2.2	Simulações de dinâmica molecular aceleradas	78
4.2.3	Conclusões parciais.....	87
4.3	Docagem e seleção de compostos	87
4.4	Clonagem, expressão e purificação da enzima MnaA.....	95
4.5	Triagem experimental dos candidatos	97
4.5.1	Conclusões parciais.....	106
4.6	Outras enzimas da via de biossíntese de WTA	106
5	Conclusões	113
REFERÊNCIAS.....		117
APÊNDICES.....		125

1 INTRODUÇÃO

1.1 Resistência a antibióticos

A descoberta dos antibióticos foi crucial para a medicina moderna, viabilizando diversos procedimentos que anteriormente não eram possíveis devido à grande incidência de infecções por microrganismos, como por exemplo o transplante de órgãos e a quimioterapia para o tratamento de câncer.¹ Os antibióticos antibacterianos perturbam processos bioquímicos importantes que resultam na inibição do crescimento celular, no caso dos agentes bacteriostáticos, e morte celular, no caso dos agentes bactericidas.

Alguns microrganismos podem ser intrinsecamente resistentes a certos antibióticos.² Além disso, bactérias são amplamente conhecidas por sua facilidade em adquirir genes de resistência a antibióticos por sua mobilização e transferência horizontal.^{1,3-4} O termo “resistência” se refere à habilidade dos microrganismos não somente de sobreviver, mas de se replicar na presença de antibióticos.⁵ Esta habilidade de se adaptar às mais diversas condições, seja com a adição de componentes tóxicos para as células, ou pelas mudanças climáticas e condições extremas de temperatura foi adquirida por meio da evolução em aproximadamente 3,5 bilhões de anos. A maior parte dos antibióticos utilizados atualmente são moléculas produzidas por microrganismos como forma de proteção de seu ambiente e fontes tróficas.³ Apesar disso, a presença de genes que favorecem a sobrevivência e multiplicação de microrganismos não pode ser considerada inofensiva, visto que estes representam um reservatório de genes de resistência que podem ser transferidos para organismos que são patógenos humanos.²

As modificações que causam a resistência a antibióticos ocorrem de forma randômica, sendo selecionadas aquelas que conferem vantagens evolutivas aos microrganismos. Podem ocorrer mutações que alteram o DNA da célula (mudanças de bases, deleções e inversões do DNA) e aquisição de material genético, na forma de plasmídeos (pequenas moléculas de DNA, que podem ser lineares ou circulares), de outras bactérias por transformação, transdução ou conjugação.³ Os processos de transferência de DNA se diferem pela forma com que o material genético é transferido do organismo doador para o receptor, como mostra a Figura 1.

Figura 1 – Mecanismos de transferência de DNA empregados por bactérias. A transformação ocorre quando há a lise de uma célula e a posterior transferência do DNA livre para uma célula receptora. A transdução é o processo no qual o DNA é transferido por intermédio de um vírus bacteriófago. O processo de conjugação ocorre pela junção de duas células onde é formado um canal para a transferência do DNA.

Fonte: MADIGAN.⁶

A transformação consiste na aquisição do DNA livre de uma célula sendo captado por outra. Já a transdução é a transferência de DNA mediada por um vírus. O método para transferência horizontal de material genético entre bactérias é a conjugação, a qual é caracterizada pela transferência de DNA no contato célula-célula, codificado por um plasmídeo conjugativo na célula doadora.⁶

Contudo, o uso extensivo de antibióticos no controle de infecções em humanos, em animais e na agricultura causou uma mobilização de genes para resistência em populações bacterianas sem precedentes.¹ A aquisição de resistência por microrganismos importantes para a saúde humana tem sido observada sistematicamente desde a descoberta dos primeiros

antibióticos, e em períodos cada vez mais curtos em relação à sua introdução no mercado, como mostra a Figura 2.⁷

Figura 2 – Linha temporal indicando os períodos de introdução dos antibióticos no tratamento de doenças e a identificação da resistência aos mesmos em bactérias. O intervalo de tempo entre a introdução de novos antibióticos e a identificação de resistência tem sido cada vez menor, tornando a demanda pelo desenvolvimento de novos fármacos cada vez mais urgente.

Fonte: Adaptada de CDC.⁸

Segundo a Organização Mundial de Saúde (OMS), as principais causas de infecções resistentes são a prescrição excessiva de antibióticos, pacientes que terminam seus tratamentos precocemente, uso excessivo de antibióticos na pecuária e na criação de peixes, falta de controle das infecções em hospitais e clínicas e falta de higiene e saneamento precário, como mostra a Figura 3.⁹

Figura 3 – Infográfico fornecido pela Organização Mundial da Saúde destacando as principais causas da resistência a antibióticos. Dentre elas, se encontra o mal-uso de antibióticos para tratamentos de doenças em humanos e na pecuária, além da precarização dos procedimentos de higiene em ambientes hospitalares e a falta de novos antibióticos para tratamento de infecções bacterianas.

Fonte: Adaptada de OMS.⁹

Ainda segundo a OMS, a resistência a antibióticos pode se propagar através de duas principais vias: durante o tratamento de pacientes ou durante o tratamento de animais para consumo, como mostra a Figura 4.⁹

Figura 4 – Infográfico fornecido pela Organização Mundial da Saúde sobre as duas principais vias de propagação de bactérias resistentes a antibióticos: pelo uso inadequado dos medicamentos e através da alimentação.

Fonte: Adaptada de WHO.⁹

Novos mecanismos de resistência surgem e se difundem globalmente com frequência, caracterizando um risco para a saúde humana e para o de tratamento de doenças causadas por infecções comuns. Os patógenos denominados pela sigla ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter sp.*) representam a maior parte das infecções hospitalares resistentes nos Estados Unidos, o que os torna importantes alvos para o desenvolvimento de novos fármacos.¹⁰ O desenvolvimento de novos antibióticos será sempre necessário para combater a aquisição de resistência por microrganismos, visto que a aquisição de resistência é um processo que ocorre naturalmente. A descoberta de novas moléculas precisa ocorrer na mesma taxa em que os microrganismos se modificam, o que não vem acontecendo nos últimos anos, como mostra a Figura 5.

Figura 5 – Número de novos agentes antibacterianos desenvolvidos desde o período de 1983-1987, que era de 16 novas moléculas, até 2008-2012, que foi de apenas duas.

Fonte: Adaptada de CDC. ¹¹

A pesquisa e o desenvolvimento de antibióticos podem ser divididos em diferentes épocas, de acordo com a estratégia utilizada pelos pesquisadores. A chamada “Era de Ouro” é caracterizada pela utilização da técnica de triagem adotada por Selman Waksman, um dos pesquisadores envolvidos na descoberta da estreptomicina. Esta envolvia a triagem de bactérias do solo que produziam metabólitos os quais impediam o crescimento de patógenos. A maior parte dos antibióticos desenvolvidos nesta época era, portanto, de origem natural, como mostra a Tabela 1.

Tabela 1 – Classes de antibióticos descobertos durante a “Era de Ouro”. As classes sublinhadas correspondem a moléculas sintéticas. Os demais são compostos naturais.

Classe química	Alvo	Modo de ação	Exemplos
Sulfonamidas	Síntese de folato	Bacteriostático	Sulfanilamida
Aminoglicosídeos	Síntese de proteínas	Bactericida	Espectinomomicina, canamicina, neomicina
Beta-lactâmicos	Síntese da parede celular	Bactericida	Penicilinas, cefalosporinas, carbapenemos
Tetraciclínas	Síntese de proteínas	Bacteriostático	Tetraciclina, doxiciclina
Cloranfenicóis	Síntese de proteínas	Bacteriostático	Cloranfenicol
Macrolídeos	Síntese de proteínas	Bacteriostático	Eritromicina, claritromicina
Glicopeptídeos	Síntese da parede celular	Bactericida	Vancomicina, teicoplanina
Oxazolidionas	Síntese de proteínas	Bacteriostático	Linezolida
Ansamícinas	Síntese de RNA	Bactericida	Rifamicina
Quinolonas	Síntese de DNA	Bactericida	Ciprofloxacina
Streptograminas	Síntese de proteínas	Bactericida	Pristinamicina

Fonte: Adaptada de BROWN.¹

A partir da metade da década de 1960, com a dificuldade de serem encontrados novos fármacos eficazes e com baixa toxicidade utilizando a plataforma de Waksman, surgiu a chamada “Era da Química Medicinal”. O desenvolvimento de antibióticos neste período foi dominado por ciclos de inovação que tinham como foco a criação de versões sintéticas das moléculas descobertas na era anterior, com modificações que aumentassem sua eficácia ou que evadissem dos efeitos da resistência desenvolvida pelos microrganismos. O desenvolvimento nas tecnologias de *high-throughput*, sequenciamento de genomas e tratamento otimizado de estruturas proteicas trouxe novas possibilidades para a descoberta de novos fármacos. Este modelo ficou conhecido como “*genes-to-drugs*” (genes-a-fármacos, em português), sendo aplicado com bastante sucesso em diversas áreas. Do início dos anos 1960 até o final dos anos 1990, porém, uma nova onda de resistência surgiu, caracterizando o período como a “Era da resistência”, onde os esforços para a descoberta de novos fármacos, apesar de promissores, não geravam moléculas viáveis para o tratamento de doenças *in vivo*. Atualmente, continuamos a enfrentar o problema da rápida aquisição de resistência por parte

das bactérias. Porém, se iniciou uma busca por novos alvos essenciais nos organismos-alvo, além da realização de modificações nas moléculas-líder (*lead compounds*) encontradas.¹

A resistência a antibióticos pode ser caracterizada pelos diferentes mecanismos pelos quais as bactérias evadem aos efeitos dos fármacos. Dentre os mecanismos, podem ser apontados aqueles em que a bactéria previne o acesso da molécula ao alvo farmacológico, por conta de uma diminuição na permeabilidade da célula ou pelo aumento no efluxo, eliminando o antibiótico. As células Gram-positivas e Gram-negativas, por exemplo, possuem membranas externas compostas de forma diferente, fazendo com que as bactérias Gram-negativas possuam menor permeabilidade. Além disso, bactérias podem realizar modificações ou adições para proteção nos alvos moleculares, além de modificações na estrutura da molécula do antibiótico, por hidrólise ou pela transferência de um grupamento químico.² As estratégias de defesa utilizadas pelas bactérias descritas de forma mais detalhada e exemplos se encontram na Tabela 2.

Antibióticos podem ter como alvos diferentes vias essenciais, além de poder agir em mais de uma espécie bacteriana. Dentre os principais mecanismos de ação dos antibióticos, podemos citar a inibição da síntese da parede celular, ruptura da membrana celular, inibição da estrutura e função de ácidos nucleicos, inibição da síntese proteica e o bloqueio de vias metabólicas essenciais para a viabilidade das células.¹²

Tabela 2 – Mecanismos de ação adquiridos pelas bactérias para evitar os efeitos de antibióticos.

Mecanismo de ação (estratégia de defesa)	Descrição
Restringindo o acesso do antibiótico	<p>Restringindo o número ou modificando o tamanho das entradas na parede celular, bactérias resistentes podem manter antibióticos fora do interior da célula.</p> <p>Exemplo: Bactérias Gram-negativas possuem uma membrana externa que as protege do ambiente. Essas bactérias podem utilizar esta membrana para impedir a entrada de fármacos de forma seletiva.</p>
Eliminando o antibiótico	<p>Bactérias resistentes podem utilizar proteínas presentes em suas paredes celulares para bombear antibióticos que entrem na célula.</p> <p>Exemplo: Algumas bactérias <i>Pseudomonas aeruginosa</i> podem produzir bombas que eliminam vários antibióticos importantes, incluindo fluoroquinolonas, beta-lactâmicos, cloranfenicol e trimetoprima.</p>
Destruindo o antibiótico	<p>Algumas bactérias resistentes utilizam enzimas que digerem as moléculas dos antibióticos e as tornam ineficazes.</p> <p>Exemplo: Bactérias <i>Klebsiella pneumoniae</i> produzem enzimas chamadas carbapenemases, que quebram o fármaco carbapenemo e a maioria dos beta-lactâmicos.</p>
Modificando o antibiótico	<p>Outras bactérias resistentes utilizam de enzimas que alteram as moléculas dos antibióticos de modo que estas têm sua eficácia diminuída.</p> <p>Exemplo: bactérias <i>Staphylococcus aureus</i> adicionam compostos aos fármacos aminoglicosídicos para modificar sua função.</p>
Modificando os alvos do antibiótico	<p>Muitos antibióticos são planejados para identificar e destruir partes específicas (ou alvos) de uma bactéria. Bactérias resistentes podem modificar a estrutura destes alvos de tal forma que o antibiótico não os reconheça, permitindo a sobrevivência das bactérias.</p> <p>Exemplo: Bactérias <i>E. coli</i> com podem adicionar um componente à porção externa da parede celular através do gene mcr-1, de tal forma que o antibiótico colistina não consegue se ligar.</p>
Evitando os efeitos do antibiótico	<p>Alguns antibióticos são planejados para causar a disfunção de processos críticos para a sobrevivência das bactérias, como o processo de síntese de nutrientes. Se for bem-sucedido, o antibiótico impede os microrganismos de realizar todos os passos necessários para o processo. Algumas bactérias resistentes, no entanto, desenvolveram processos novos e diferentes para contornar os efeitos dos fármacos. O novo processo pode ser mais lento, mas evita os efeitos dos antibióticos.</p> <p>Exemplo: Algumas bactérias <i>Staphylococcus aureus</i> podem evitar os efeitos do antibiótico trimetoprima.</p>

Fonte: Adaptada de CDC.¹⁸

1.2 Antibióticos beta-lactâmicos e MRSA

Os antibióticos beta-lactâmicos, que incluem as penicilinas, cefalosporinas e cefamicinas (Figura 6), são fármacos de primeira escolha para o tratamento de infecções causadas por diversos organismos Gram-positivos, incluindo *S. aureus*, que é uma das principais causas de infecções hospitalares em todo o mundo.¹³ Esta bactéria pode ser encontrada na pele e mucosa de humanos e outros animais¹⁴ e, apesar de inofensiva em condições normais, pode causar uma variedade de infecções em humanos, tais como foliculite, furúnculos e carbúnculos, impetigo, mastite, infecções em feridas e síndrome da pele escaldada estafilocócica, além de condições mais perigosas, como bacteremia, pneumonia, endocardite, infecções ósseas e articulares e síndrome do choque tóxico.¹⁵ De acordo com dados coletados pelo CDC entre 2011 e 2014, esse microrganismo é responsável por aproximadamente 10% das infecções relacionadas ao sistema de saúde analisadas nos Estados Unidos, ocupando o segundo lugar na classificação de patógenos mais frequentemente reportados ao NHSN (do inglês, *National Healthcare Safety Network* ou Rede Nacional de Segurança da Saúde). Além disso, *S. aureus* é responsável por cerca de 20% das infecções associadas a procedimentos cirúrgicos, onde cerca de 40% das amostras testadas apresentaram resistência a antibióticos. As infecções que acometem a corrente sanguínea provenientes do uso de cateter, as infecções urinárias causadas pelo uso de sonda e as pneumonias associadas ao uso de ventilador mecânico, possuem cerca de 50% dos isolados de resistentes a oxacilina, meticilina e cefoxitina.¹⁶

Figura 6 – Estruturas dos antibióticos beta-lactâmicos.

Fonte: Adaptada de WILLIAMS.¹⁷

Os beta-lactâmicos agem pela inibição da síntese e da ligação transversal entre polímeros específicos da parede celular de bactérias, os peptidoglicanos, ligando-se e acilando transpeptidases (que realizam a reação de ligação cruzada entre duas cadeias peptídicas ao glicano), o que resulta no enfraquecimento da parede celular e lise da célula.¹⁹ A resistência a estes antibióticos em *S. aureus* é mediada por dois principais mecanismos, ambos adquiridos pela transferência horizontal de genes. O primeiro confere à bactéria resistência à penicilina, que foi o primeiro fármaco beta-lactâmico a ser introduzido para o tratamento de doenças.²⁰ Para evitar os efeitos do antibiótico, o microrganismo secreta uma enzima beta-lactamase, a qual cliva o anel beta-lactâmico, presente neste grupo de antibióticos. A beta-lactamase é codificada pelo gene *blaZ*, mantido em um plasmídeo. A resistência mediada por beta-lactamases está presente na maioria das cepas de *S. aureus* fazendo com que a penicilina seja raramente utilizada para o tratamento de infecções causadas pela bactéria.²⁰

O segundo mecanismo para resistência é adquirido pela transferência lateral do gene *mecA*, que codifica uma nova transpeptidase chamada *penicillin binding protein 2A* ou PBP2a, a qual também se liga a fármacos beta-lactâmicos, como a meticilina, porém com baixa afinidade, e viabiliza células em concentrações de antibióticos que deveriam ser letais às mesmas em condições normais.¹⁹ Outra enzima que participa na resistência de *S. aureus* a antibióticos beta-lactâmicos é a PBP4, codificada pelo gene *pbp4*, e que também possui função de transpeptidase, realizando reparos em defeitos encontrados na parede de peptidoglicanos. Foram encontradas em cepas resistentes, mutações do tipo *missense*, que adicionam serinas ao sítio ativo da proteína, e mutações pontuais no promotor do gene *pbp4*. As mutações encontradas no promotor levam a célula a apresentar dois fenótipos diferentes: uma expressão elevada da proteína PBP4 e mais ligações de *crosslink* entre os peptidoglicanos. A ação das duas mutações leva a um efeito sinérgico, elevando a resistência das células aos antibióticos beta-lactâmicos.²⁰ As bactérias que apresentam genes que apresentam este tipo de resistência à meticilina são chamados de MRSA (do inglês, *Methicillin Resistant S. aureus*, ou *S. aureus* resistente a meticilina).

No ano de 2014, bactérias MRSA foram responsáveis 90% das infecções por *S. aureus* na região da América, e 60% na região da Europa. Nos Estados Unidos, infecções por bactérias resistentes representam mais de 50% de todos os isolados de *S. aureus*, resultando em uma taxa de mortalidade de 20%.^{19, 21} Esta alta taxa de mortalidade está associada em parte ao estado de saúde anterior do paciente, tais como aqueles que passaram por cirurgias ou em pacientes imunocomprometidos. No Brasil, em 2006, um estudo em um hospital no estado de

Pernambuco avaliou pacientes infectados por *S. aureus* e concluiu que 34,5% possuíam MRSA.²² Em 2017, a OMS emitiu um comunicado listando 12 famílias bacterianas resistentes a antibióticos que representam grande risco para a saúde humana e reforçando a importância do desenvolvimento de novos fármacos para combater estas populações. Dentre os organismos listados, MRSA está classificado como alta prioridade para pesquisa e desenvolvimento de novos medicamentos.²³

A resistência a antibióticos apresentada pelos organismos MRSA possui um impacto direto na saúde humana e nos sistemas de saúde. Dados obtidos em 2007 a partir de um pesquisa em 31 países que participam do Sistema Europeu de Vigilância à Resistência Antimicrobiana (do inglês, *European Antimicrobial Resistance Surveillance System* ou EARSS), mostrou que 27.771 infecções da corrente sanguínea causadas por MRSA resultaram em 5.503 mortes e 255.683 dias de internação em excesso, gerando um débito de 44 milhões de euros (63,1 milhões de dólares).²⁴

Tendo em vista a dificuldade em combater as infecções causadas por bactérias, a busca por novas estratégias se faz necessária. Dentre elas se encontram o desenvolvimento de novos fármacos (que inibem o crescimento ou eliminam as populações bacterianas), desenvolvimento de vacinas utilizando anticorpos ou anticorpos conjugados a fármacos, utilização de fagos ou proteínas derivadas de fagos, realização de terapias moduladoras da microbiota, entre outros.²⁵

Apesar da grande incidência de resistência bacteriana a estes antibióticos, beta-lactâmicos continuam sendo umas das melhores alternativas para o tratamento de infecções por patógenos bacterianos, devido ao seu perfil seguro e baixo custo. Em conjunto, as penicilinas e cefalosporinas correspondem a mais da metade de todos os antibióticos produzidos e consumidos no mundo (Figura 7).⁶

Figura 7 – Distribuição dos antibióticos mais produzidos e consumidos no mundo, os beta-lactâmicos cefalosporina e penicilina somados representam 54% do total.

Fonte: MADIGAN.⁶

A combinação de agentes para re-sensibilização a esses fármacos é uma opção promissora que implica na necessidade do entendimento de mecanismos e vias metabólicas envolvidas na resistência.^{19, 26} De fato, a combinação de fármacos tem sido utilizada com sucesso em diversos casos, como a combinação de trimetoprim e sulfametoxazol, que afetam o metabolismo de folato de forma sinérgica.¹ Outra combinação conhecida é a administração conjunta de clavulanato, que inibe beta-lactamases, comum em bactérias Gram-negativas, e amoxicilina, que é um antibiótico que possui o grupamento beta-lactâmico em sua estrutura química. O tratamento de infecções as quais não se conhece o patógeno também é comumente realizado com a combinação entre beta-lactâmicos e aminoglicosídeos para a cobertura de um amplo espectro de organismos (Figura 8).¹

Figura 8 – Espectro de ação de antibióticos utilizados *in vivo*. Em destaque, o aminoglicosídeo tobramicina e o beta-lactâmico penicilina, os quais são utilizados em conjunto para o combate de infecções causadas por agentes desconhecidos.

Fonte: MADIGAN.⁶

1.3 *Staphylococcus aureus* e seus ácidos teicóicos de parede (WTA)

S. aureus é uma bactéria Gram-positiva, aeróbio facultativo, que forma cachos celulares com pigmentação amarela. Esta espécie bacteriana tem sua parede celular formada por camadas de peptidoglicano envoltas por uma camada de ácidos teicóicos de parede (do inglês, *Wall Teichoic Acid* ou WTA), externas à membrana plasmática, como mostra a Figura 9.

A WTA é uma parede celular formada por glicofosfatos aniônicos, não essencial para a viabilidade das células.¹³ Além disso, sabe-se que os ácidos teicóicos estão envolvidos na manutenção da forma celular, resistência a peptídeos antimicrobianos, formação de biofilmes, tolerância a ácidos e eficiência de secreção de proteínas em meio de cultura.²⁷ Em *S. aureus*, a ausência dos ácidos teicóicos de parede mostrou que as células, apesar de viáveis em

condições laboratoriais, apresentam sua capacidade colonizadora e infectiva comprometida.²⁸ A síntese dos ácidos teicóicos de parede em *S. aureus* é realizada pela ação sequencial de enzimas denominadas Tar (do inglês, *teichoic acid ribitol*), e é iniciada na região citoplasmática da célula (Figura 10).

O polímero de ácidos teicóicos em *S. aureus* é composto por cadeias de poli(ribitol-fosfato) e inicia-se pela ação da enzima TarO, a qual catalisa a transferência de uma molécula de UDP-N-acetilglicosamina para o lipídeo undecaprenil-fosfato. Em seguida, a proteína TarA realiza a transferência de um resíduo de N-acetilmanosamina ao lipídio. A unidade de ligação à membrana plasmática interna é finalizada pela enzima TarB com a adição de um *sn*-glicerol-3-fosfato. Após a síntese do resíduo de ligação, passos para completar o polímero de poli(ribitol-fosfato) são realizados. Primeiramente, uma molécula de glicerol-fosfato é adicionada ao glicopeptídeo pela ação da proteína TarF, seguida pela polimerização de ribitol-fosfato pela enzima TarL. A decoração intracelular do polímero em *S. aureus* inclui modificações de ambos N-acetilglicosamina alfa e beta-ligados, catalisados pelas enzimas TarM e TarS, respectivamente. A cadeia de poli(ribitol-fosfato) é então exportada para a superfície extracelular da membrana plasmática pelo sistema transportador do tipo ABC composto pelo complexo das enzimas TarG e TarH. Uma vez exportado, este polímero ligado ao lipídio é D-anilado pelos produtos gênicos do *operon* *dltABCD* e transferido do lipídio carreador undecaprenol para o peptideoglicano.²⁹

A ausência das enzimas TarA ou TarO na via de síntese dos ácidos teicóicos de parede torna as demais proteínas da via dispensáveis. Na presença das mesmas, as proteínas que constituem os últimos passos da síntese do polímero são essenciais para a viabilidade das células (Figura 11). Este fenômeno é conhecido como "paradoxo dos genes essenciais" e, alguns dos motivos possíveis para que isto ocorra são o acúmulo de precursores tóxicos nas últimas etapas do processo e/ou o sequestro do precursor biosintético essencial, bactoprenol, que leva à depleção de PG, pois sua síntese também depende de bactoprenol como carreador de lipídeos.¹³

Figura 9 – Composição da parede celular em bactérias Gram-positivas, contendo a membrana celular, a parede de peptidoglicanos e a ácidos teicóicos de parede.

Fonte: Adaptada de SEWELL.²⁹

Figura 10 - Via de síntese dos ácidos teicóicos de parede(WTA). A síntese se inicia na região interna da membrana plasmática, com a ação da proteína TarO, terminando ligada à porção externa da parede de peptidoglicanos.

Fonte: Adaptada de SEWELL.²⁹

Figura 11 – O “Paradoxo dos genes essenciais”, onde a deleção das enzimas TarGH causa a morte celular (A), porém a deleção conjunta de um dos primeiros passos (TarO ou TarA) permite a viabilidade do microrganismo (D).

Fonte: Adaptada de SEWELL.²⁹

1.4 Os ácidos teicóicos de parede como alvo para o desenvolvimento de novos fármacos

Diversas enzimas que participam da via de síntese dos ácidos teicóicos de parede se mostraram bons possíveis alvos para a re-sensibilização de MRSA. Esforços recentes da indústria farmacêutica resultaram na triagem de aproximadamente 20.000 compostos com efeitos inibitórios da síntese dos ácidos teicóicos de parede, dentre eles o targocil, ácido indol tríclicos, N-aril-triazóis e C-aril-triazóis (Tabela 3).²⁹ A molécula targocil possui como alvo a proteína TarG da via de biossíntese da WTA. A inibição somente da TarG, porém, não se mostrou eficaz para a eliminação da bactéria. A administração *in vivo* inibidor conjuntamente com o antibiótico beta-lactâmico imipenem em concentrações sub-inibitórias, por outro lado, mostrou um efeito de adição ou sinergia, demonstrando eficácia no tratamento de infecções por MRSA em animais. A proteína TarG, portanto, pode ser considerada um alvo para o desenvolvimento de fármacos bactericidas.¹³

Estudos em que foi realizada a deleção da enzima TarA em *S. aureus* resistentes mostraram a re-sensibilização das células a antibióticos carregados negativamente (ácido fusídico e fosfomicina) e a diminuição de sua infectividade em ratos. Um ponto interessante neste caso, é a carga elétrica líquida negativa dos fármacos encontrados. A WTA também é carregada negativamente, indicando que o aumento da sensibilidade é devido à ausência da mesma, não mais repelindo as moléculas dos fármacos e aumentando a acessibilidade dos compostos aos alvos intracelulares.³⁰

A proteína TarO também foi amplamente estudada devido ao seu potencial como sítio para ligação de inibidores do crescimento bacteriano em *S. aureus*. Esta possui papel semelhante àquele da TarA na via de síntese da WTA. Sua deleção mostrou diminuir a infectividade e a propagação do microrganismo em modelos vivos.²⁹ Além disso, a inibição da proteína utilizando a molécula tunicamicina mostrou re-sensibilizar cepas de *S. aureus* resistentes à ação do antibiótico beta-lactâmico imipenemo.¹³ A inibição desta proteína está ligada a defeitos na septação e separação celular. Beta-lactâmicos específicos para PBP2 inibem a atividade de transpeptidação da proteína, resultando em baixos níveis de ligações transversais na parede celular primária. A inibição da síntese dos ácidos teicóicos de parede impede a localização e função apropriada da PBP4, evitando, portanto, a transpeptidação secundária da parede celular. O baixo nível de transpeptidação da parede celular nestas circunstâncias se torna insuficiente, levando à morte celular.²⁹

A adição de açúcares à WTA também se mostrou importante para a resistência a antibióticos beta-lactâmicos. Como mencionado anteriormente, são adicionados ambos alfa e beta-N-acetilglucosamina à cadeia crescente de ácidos teicóicos. Estudos realizados por Sobhanifar *et al.*^{26, 28} demonstraram que somente os beta-açúcares influenciam a suscetibilidade de cepas MRSA a antibióticos beta-lactâmicos. A proteína responsável pela adição destes açúcares é a TarS, uma beta-glicosiltransferase. Foi observado em experimentos *in vitro* a ligação da proteína PBP2a à WTA indicando uma relação entre a parede e a biossíntese da camada de peptidoglicanos. A sensibilização das bactérias a meticilina ocorre pois os beta-O-GlcNac servem como estrutura para o recrutamento da proteína PBP2a. A deleção da proteína TarS evitaria o mecanismo de resistência causado pela proteína insensível a beta-lactâmicos, pois a mesma se encontraria deslocada de sua posição usual na célula.²⁶

Tabela 3 - Compostos inibitórios das enzimas que fazem parte da síntese dos ácidos teicóicos de parede.

Nome químico ou série	Estrutura	Alvo
Targocil		TarG
Ácido indol tricíclico		TarG
N-aril triazol		TarG
C-aril triazol		TarG
Ticlopidina		TarO

Fonte: SEWELL.²⁹

1.5 A enzima UDP-N-acetilglicosamina de *S. aureus* (SaMnaA)

Outra enzima que fornece uma perspectiva de alvo para a re-sensibilização de MRSA é a chamada MnaA, que é uma UDP-N-acetilglicosamina 2-epimerase. Esta enzima interconverte UDP-N-acetilglicosamina (UDP-GlcNac) em UDP-N-acetilmanosamina (UDP-ManNac), (Figura 12), realizando a modulação da geração de substratos para as enzimas TarA e TarO.

Figura 12 – Via de síntese dos ácidos teicóicos de parede. Em destaque, a proteína MnaA, a qual fornece substrato para as enzimas TarA e TarO.

Fonte: Adaptada de MANN.³¹

Evidências genéticas mostram que a enzima MnaA é essencial para a produção dos ácidos teicóicos de parede e para a resistência a beta-lactâmicos em MRSA. A presença de duas cópias funcionais redundantes desta enzima (MnaA e Cap5P) reforça a importância destas para a célula.³² A bactéria *S. aureus* também apresenta a capacidade de formação de biofilme, diminuindo assim seu metabolismo, tornando antibióticos que possuem como alvo vias metabólicas inativos. A formação de biofilme também impede a entrada de antibióticos nas células através da formação de uma matriz extracelular composta por polissacarídeos. Testes de perda de função (do inglês, *loss of function* ou LOF) em mutantes demonstram uma suscetibilidade a antibióticos beta-lactâmicos em modelos de infecção em ratos. Quando a enzima é modificada ou inibida, ocorre a depleção da WTA, impedindo também a formação de biofilme. A deleção de ambas Cap5 e MnaA em MRSA mostrou re-sensibilizar as células a beta-lactâmicos, diminuindo a concentração inibitória mínima (MIC) necessária para a ação antibiótica, como mostra a Tabela 4.³¹ A proteína MnaA é, portanto, importante para dois mecanismos de resistência de MRSA: formação de biofilme e coordenação da parede de peptidoglicanos.

Tabela 4 – Testes de re-sensibilização de bactérias cepa MRSA COL através da mutação e deleção dos genes que codificam para as proteínas Cap5 e MnaA.

Cepa	MIC Imipenemo ($\mu\text{g. mL}^{-1}$)	MIC Nafcillina ($\mu\text{g. mL}^{-1}$)	MIC Dicloxacilina ($\mu\text{g. mL}^{-1}$)
MRSA COL	32	> 64	> 64
Δcap5P	32	> 64	> 64
$\Delta\text{cap5P SaMnaA}^{\text{P12L}}$	1	8	4
$\Delta\text{cap5P SaMnaA}^{\text{Y194*}}$	1	8	2

Fonte: Adaptada de MANN.³¹

A estrutura do homodímero de MnaA do organismo *S. aureus* foi resolvida por cristalografia com resolução de 1,9 Å e depositada no PDB (do inglês, *protein data bank* ou banco de dados de proteínas), código 5ENZ³¹ (Figura 13). A estrutura resolvida apresenta uma molécula de UDP, tetraetileno glicol, TRIS e íons de sulfato no sítio ativo de cada monômero. A proteína também se liga à tunicamicina, conhecida por inibir a proteína TarO.³¹ Sua estrutura apresenta similaridades com outras 2-epimerases bacterianas, onde sua sequência de aminoácidos possui identidade de cerca de 61,33% com a proteína de *Bacillus subtilis* e de 49,59% com a de *Escherichia coli*. Além disso, as proteínas bacterianas apresentam interface de dimerização e estrutura quaternária similar entre si. O RMSD (do inglês, *root mean square deviation* ou raiz quadrática média do desvio posicional), calculado para os carbonos-alfa das proteínas de *S. aureus* e *E. coli* estão entre 1,40 e 2,90 Å, dependendo das cadeias que estão sendo comparadas. As principais diferenças estruturais entre estas se encontram na superfície ou na interface de dimerização das enzimas.

Figura 13 – Estrutura cristalográfica da cadeia A da proteína UDP-GlcNac 2-epimerase de *S. aureus* (PDB: 5ENZ), com uma molécula de UDP em sítio ativo, vista de frente (A) e rotacionada de 90° (B), como indicado na figura. Em vermelho, o domínio C-terminal e em azul, o domínio N-terminal.

Fonte: Elaborada pela autora.

Esta enzima faz parte do grupo de proteínas 2-epimerase, que realizam a conversão entre um açúcar e outro de forma não-hidrolisante. Estudos indicam que o mecanismo pelo qual a reação ocorre é iniciado pela eliminação *anti* do UDP para gerar um 2-acetoamidoglucal intermediário, seguido pela adição *syn* do UDP para gerar o produto UDP-ManNac, como mostra a Figura 14.³² Em *Staphylococcus aureus*, a conversão de UDP-ManNac em UDP-GlcNac se dá na proporção de 9:1 a favor do último.³¹

Figura 14 – Mecanismo de interconversão entre os compostos UDP-GlcNac e UDP-ManAc, para a enzima MnaA de *S. aureus*, que é não-hidrolisante.

Fonte: VELLOSO.³²

As enzimas 2-epimerase bacterianas realizam a regulação da sua função utilizando o substrato UDP-GlcNac, o qual age como ativador para a conversão de UDP-ManNac em UDP-GlcNac. Na ausência desta molécula, nenhum UDP-ManNac é convertido.³³⁻³⁵ Contudo, na presença de quantidades mínimas de UDP-GlcNac, a reação se encaminha para o equilíbrio. É possível que esta regulação seja feita por mudanças conformacionais que ocorrem na proteína quando o UDP-GlcNac entra no sítio ativo.³²

Outra característica observada nesta família de proteínas é o movimento de domínios, o qual causa a abertura e o fechamento do sítio ativo, observada também em proteínas da via das Tar, como por exemplo, a enzima TarM e a TarS.^{26, 28} A proteína UDP-GlcNac 2-epimerase de *Methanocaldococcus jannaschii* foi cristalizada sem ligantes em seu sítio ativo (PDB: 4NEQ) e com ambos UDP e UDP-GlcNac (PDB: 4NES). As estruturas apresentaram diferenças nas conformações entre seus domínios C-terminal e N-terminal, caracterizando os estados “aberto” e “fechado” (Figura 15).³⁶ Após a observação da estrutura de proteína de *S. aureus*, a qual possui apenas uma molécula de UDP ligada, foi possível caracterizar o estado “intermediariamente fechado” ou simplesmente “intermediário” (Figura 15).³¹

Figura 15 – Estruturas cristalográficas das enzimas UDP-GlcNac 2- epimerase de *S. aureus* (5ENZ, em verde)) e *M. jannaschii*, resolvida sem (4NEQ, em vermelho) e com ligantes (4NES, em azul). As estruturas foram alinhadas por seu domínio N-terminal, de modo a evidenciar as diferenças de ângulo de abertura entre os domínios. Os ligantes das estruturas foram removidos da figura para facilitar a visualização.

Fonte: Elaborada pela autora.

Na estrutura 5ENZ, o RMSD calculado para as posições de todos os carbonos alfa entre os monômeros difere de cerca de 1Å. Um se encontra em uma conformação mais próxima da considerada "aberta" e outro da considerada "fechada". Ressalta-se também a importância da mudança na conformação da alça His205-Gly211, que difere significativamente entre os monômeros, estando novamente um na conformação "aberta" e no outro mais próximo da forma "fechada". Estes resultados sugerem que a regulação da ação desta enzima ocorre com o equilíbrio dinâmico entre a forma "aberta" e "intermediário" na ausência de ligantes ou na presença de UDP. Este equilíbrio é deslocado na direção da forma fechada quando a proteína se encontra no complexo ternário com os ligantes UDP e UDP-GlcNac.

Esta proteína se mostra um interessante alvo para o desenvolvimento de novos fármacos. Os resíduos que coordenam a ligação com UDP-GlcNac das enzimas 2-epimerase bacterianas são altamente conservados, porém não apresentam identidade com os resíduos que realizam a mesma função em humanos.³² A inibição da enzima 2-epimerase resulta na depleção da WTA, a qual por sua vez afeta a susceptibilidade das bactérias a antibióticos beta-lactâmicos, tanto *in vitro* quanto em ratos.³¹

Inibidores para os primeiros passos não essenciais na biossíntese dos ácidos teicóicos de parede são particularmente interessantes como adjuvantes não bioativos ou agentes combinatórios, os quais, administrados em conjunto com beta-lactâmicos, caracterizam uma estratégia promissora para tratar infecções por MRSA.^{29, 31} Portanto, estudos sobre a estrutura destas proteínas e relacioná-la à sua função são passos importantes para o desenho de possíveis fármacos para o tratamento de doenças causadas por *S. aureus* resistentes.

2 OBJETIVOS

Visto que a busca por novos alvos moleculares e novos fármacos para o tratamento de infecções resistentes é algo urgente e de grande importância para a melhora na qualidade de vida e diminuição das altas taxas de internação e de mortalidade causadas pelas mesmas, os principais objetivos deste trabalho foram:

- Analisar a relação entre estrutura e função da proteína;
- Obter informações sobre a dinâmica da enzima em solução através de simulações de dinâmica molecular;
- Obter um modelo de estrutura estável com a presença do ligante UDP-N-acetilglicosamina;
- Buscar novos ligantes *in silico* e realizar testes para a avaliação da interação com a proteína alvo.

Dentre os objetivos específicos, se encontram, primeiramente, a avaliação do movimento entre domínios da proteína MnaA, explicitando os possíveis motivos para tal. Além disso, com a utilização e posterior comparação entre duas técnicas de dinâmica molecular acelerada, podemos obter e analisar informações mais detalhadas sobre a influência dos ligantes aos movimentos de abertura e fechamento da enzima. Estes resultados são importantes para a construção de um modelo estável da proteína e de seus ligantes para a realização da docagem molecular. Utilizando esta técnica computacional, temos como objetivo otimizar uma triagem de compostos, na busca de um ligante/inibidor para a enzima a partir de um banco de moléculas selecionado. Uma vez escolhido um número razoável de compostos, é interessante realizar testes *in vitro* para avaliar a interação entre as moléculas e a proteína. Para isso, precisamos obter a proteína enovelada, purificada e em quantidade suficiente para a realização dos testes, o que pode ser feito a partir da tecnologia do DNA recombinante. Uma vez obtida a SaMnaA, podemos seguir com os testes de interação para a busca de inibidores.

3 MATERIAIS E MÉTODOS

Devido ao grande interesse na busca de novos alvos moleculares para o tratamento de infecções causadas por *S. aureus* resistentes em humanos, o estudo de enzimas que participam da biossíntese dos ácidos teicóicos de parede se mostrou interessante para o desenvolvimento deste trabalho. Duas principais frentes foram escolhidas: a análise da dinâmica molecular da proteína *SaMnaA* e a clonagem, expressão e purificação de diversas enzimas envolvidas na via, dentre as quais muitas não possuem estrutura resolvida até o momento. Foi realizada também a clonagem de genes que codificam para enzima UDP-GlcNac 2-epimerase, para a produção da proteína de forma heteróloga, utilizada em testes de inibição *in vitro* e tentativas de cristalização nas conformações APO e HOLO fossem realizados. Diversos outros genes que codificam para proteínas envolvidas na síntese da WTA também foram clonados.

3.1 Metodologias computacionais

3.1.1 Parametrização da proteína e dos ligantes

A cadeia A da estrutura cristalográfica da proteína MnaA de *S. aureus* (PDB: 5ENZ) foi parametrizada utilizando o campo de forças AMBER FF14SB.³⁷ Foram adicionados átomos de hidrogênio e os estados de protonação das histidinas foram determinados com o auxílio do servidor H++³⁸ para o pH 7,0. A proteína foi imersa em uma caixa de água utilizando o modelo TIP3P.³⁹ Íons de sódio foram adicionados ao sistema de modo que a soma de cargas fosse zero.

As moléculas de UDP e UDP-GlcNac tiveram suas cargas calculadas utilizando o método de Hartree-Fock para o cálculo dos orbitais moleculares *ab initio* e a base 6-31G*,⁴⁰ disponível no programa Gaussian 09.⁴¹ As cargas atômicas obtidas foram então reajustadas utilizando o método RESP (do inglês, *Restrained Electrostatics Potential* ou Potencial Eletrostático Restrito),⁴² com o auxílio do programa ANTECHAMBER.⁴³ Os parâmetros de Lennard-Jones para estes ligantes foram definidos utilizando o campo de forças GAFF.⁴⁴⁻⁴⁵

A cadeia A da estrutura da proteína MnaA de *S. aureus*, sem ligantes em seu sítio ativo, foi utilizada como a estrutura inicial para as simulações de dinâmica molecular da proteína APO. A estrutura contendo os ligantes UDP e UDP-GlcNac em seu sítio ativo (HOLO) foi modelada utilizando a cadeia A da estrutura de *S. aureus* e a estrutura cristalográfica da MnaA de *Bacillus anthracis* (PDB: 3BEO) utilizando o programa Modeller.⁴⁶ A escolha da

estrutura da MnaA de *B. anthracis* foi devida ao fato de que sua estrutura foi resolvida com ambos os ligantes em seu sítio ativo. Deste modo, o propósito da modelagem das duas estruturas foi gerar uma conformação mais próxima daquela “fechada” para a proteína MnaA de *S. aureus*. Os sistemas finais continham 45.519 e 55.479 átomos para os sistemas APO e HOLO, respectivamente.

3.1.2 Dinâmica molecular em equilíbrio

A energia inicial do sistema APO foi minimizada em 1000 ciclos utilizando a técnica de minimização por gradiente conjugado. O sistema foi então aquecido a 300 K em 50 ps, a volume constante. A densidade do sistema foi equilibrada a 1 g/cm³ mantendo a pressão constante em 1 atm em 400 ps. Durante a minimização, aquecimento e equilíbrio da densidade, foram aplicadas restrições harmônicas a todos os resíduos, com peso de 2 kcal.mol⁻¹. Å⁻². Finalmente, o sistema foi equilibrado por 10 ns e a simulação produtiva foi realizada por 1 μs sem restrições. Todas as simulações foram realizadas utilizando o pacote de programas AMBER16.⁴⁷⁻⁴⁸

Uma simulação de equilíbrio similar também foi realizada para o sistema APO, porém utilizando o campo de forças SIRAH,⁴⁹⁻⁵⁰ o qual torna o sistema do tipo “*coarse grained*”, ou seja, assume-se que um conjunto de átomos se comportem como um único átomo no sistema (Figura 16). Simulações deste tipo, reduzem drasticamente o tamanho do sistema, tornando o custo computacional também mais baixo. A minimização ocorreu em duas etapas, ambas utilizando a técnica de *steepest descent* por 100 ciclos, mudando para minimização por gradiente conjugado até 5000 ciclos. Durante a primeira minimização, foi aplicada uma restrição aos átomos da cadeia principal de 2,4 kcal.mol⁻¹.Å⁻². Foram realizados também dois processos de equilíbrio, ambos a pressão e temperatura constantes. O primeiro foi conduzido por 5 ns com restrições de peso 2,4 kcal.mol⁻¹.Å⁻². Já o segundo, de 25 ns, aplicou restrições de peso 0,24 kcal.mol⁻¹.Å⁻² somente à cadeia principal da proteína. A simulação produtiva foi realizada logo após a segunda equilíbrio, por 23 μs.

Como mostrado na Equação 1, ΔV é definido com um potencial harmônico com força constante k , a qual depende da constante σ_0 (Equações 2 e 3).⁵¹ Nas Equações 2 e 3, $V_{\max}$ e $V_{\min}$ são o potencial máximo e mínimo do sistema, coletados antes que a GAMD se inicie. σ_v e V_{av} são o desvio padrão e a média da energia potencial, respectivamente. Como observado anteriormente, σ_0 deve ser escolhido de forma cuidadosa, de forma que se obtenha a maior aceleração possível, mas que permita um rebalanceamento preciso dos potenciais. Uma vantagem associada com simulações GAMD e dinâmicas moleculares aceleradas é que não é necessária a escolha de uma coordenada de reação ou uma variável coletiva (do inglês, *collective variable* ou CV). Além disso, como apontado por Miao *et al.*,⁵¹ a GAMD preserva a superfície de potencial original, permitindo o cálculo do PMF (*potential mean force* ou potencial de força média) menos ruidoso quando comparado com simulações aMD.⁵¹

Neste trabalho, foram aplicados potenciais de impulso aos potenciais total e de diedro, em uma abordagem *dual-boost*. Para ambos os potenciais, o σ_0 foi configurado para 3 kcal/mol, baseado em simulações anteriores, realizadas com diversos valores para σ_0 , onde 3 kcal/mol forneceu a maior aceleração.

O potencial de impulso resultante seguiu uma distribuição Gaussiana, a qual permitiu o rebalanceamento e o cálculo de um PMF. O rebalanceamento foi feito a partir da obtenção de uma expansão cumulante de segunda ordem para o potencial de impulso, como descrito previamente por Miao *et al.*⁵² Esse cálculo foi feito com base no ângulo de abertura observado, entre 10 e 80 graus, com incrementos de 4 graus.

3.1.4 Dinâmica molecular adaptativa

A dinâmica molecular adaptativa (Adaptively Biased Molecular Dynamics ou ABMD) é um método de amostragem umbrela, similar, em espírito, à metadinâmica. A força direcional aplicada ao sistema compensa o gradiente de energia livre, reproduzindo a superfície de energia anteriormente desconhecida para o sistema analisado na direção da variável coletiva (CV), escolhida pelo usuário.⁵³

Foi utilizado como CV o ângulo entre os carbonos alfa dos resíduos S128, G49 e P213, que descreve o movimento de abertura da estrutura de MnaA de *S. aureus*. A simulação foi conduzida com o método de amostragem *well-tempered umbrella*,⁵⁴ com uma pseudo-temperatura definida a 5000 K. O conjunto total de simulações utilizadas neste trabalho está descrito na Tabela 5.

Tabela 5 – Conjunto total de simulações de dinâmica molecular realizados no trabalho, divididos por método, sistema simulado e tempo de simulação.

Método	Sistema proteico	Tempo de simulação (ns)
Equilíbrio	APO	1000 + 4000
Equilíbrio	HOLO	1000
Equilíbrio	APO (SIRAH)	23000
GAMD	APO	3 x 500
GAMD	HOLO	3 x 500
ABMD	APO	3 x 250
ABMD	HOLO	3 x 250

Fonte: Elaborada pela autora.

3.1.5 Análise estrutural

A simulação para o sistema APO de 5 μ s foi analisado utilizando algoritmos disponíveis na biblioteca pyEMMA, que foi criada para a análise de trajetórias de dinâmica molecular utilizando modelos de estado de Markov (MSMs).⁵⁵⁻⁵⁶ Apenas as coordenadas dos carbonos-alfa da proteína foram utilizadas na análise, e o tempo de defasagem utilizado na análise de componentes independente do tempo (TICA) foi de 200 frames ou 10 ns. As forças eletrostáticas entre os domínios C-terminal e N-terminal da proteína foram realizados utilizando o servidor Delphi Force Web,⁵⁷ onde o primeiro domínio foi definido como a região da proteína que compreende os resíduos de 1 a 181. Três estados representativos das conformações “aberta”, “intermediária” e “fechada” da proteína foram utilizados. Nos cálculos que envolviam as estruturas contendo os ligantes no sítio ativo da proteína, os mesmos foram considerados como parte do domínio C-terminal. A análise geométrica das trajetórias das dinâmicas moleculares foi feita com o auxílio do programa CPPTRAJ⁵⁸ e as energias de interação foram computadas com o programa AmberEnergy++,⁵⁹ um programa escrito localmente para computar as energias potenciais de interação de acordo com o campo de forças AMBER.

3.1.6 Docagem de compostos

A docagem de compostos que eventualmente se liguem à proteína foi realizada utilizando a ferramenta de docagem baseada nas estruturas do receptor e dos ligantes,

LiBELa.⁶⁰⁻⁶¹ Como estrutura inicial para a docagem de compostos foram utilizados o UDP e a proteína MnaA de *S. aureus* do sistema HOLO, onde foram excluídas as águas, íons e UDP-GlcNac, o qual foi salvo em um arquivo MOL2 diferente e utilizada como referencial. Foram utilizados o método DELPHI para o cálculo do potencial eletrostático entre as interações polares do sistema, e um potencial de Lenard-Jones 6-12 para a modelagem de interações de van der Waals. A base de ligantes utilizada foi a ZDD (do inglês, ZINC drug database), a qual possui 2924 compostos disponíveis comercialmente e aprovados mundialmente.⁶²⁻⁶⁴ Os arquivos MOL2 utilizados foram preparados pela base de dados utilizando os métodos padrão do ZINC para a preparação dos ligantes. Foi utilizado também um potencial amortecido sem correção de solvatação, disponível no programa. Além disso, o modelo para os ligantes selecionado permitiu a amostragem de conformeros. Os demais parâmetros utilizados para a docagem de compostos estão listados no arquivo de entrada, disponível nos Anexos. Os resultados foram visualizados utilizando a ferramenta ViewDock, disponível no programa UCSF Chimera.⁶⁵ Os melhores candidatos foram selecionados a partir de uma inspeção visual das interações entre os ligantes e a enzima.

3.2 Metodologias experimentais

3.2.1 Clonagem e expressão das enzimas envolvidas na síntese da WTA de *S. aureus*

Para os procedimentos de clonagem e expressão, foi utilizado como molde para a produção de plasmídeos o DNA genômico do organismo *Staphylococcus aureus* cepa SA16, o qual foi gentilmente fornecido pela professora doutora Ilana L. B. C. Camargo, do IFSC/USP.⁶⁶ Foram clonados os genes para as proteínas TarA, TarB, TarS, TarL, TarO, TarI, TarJ e MnaA. Os primers foram desenhados contendo uma sequência para anelamento com o vetor pET/TRX-1a/LIC.⁶⁷ A sequências dos primers *forward* e *reverse* utilizados foram obtidas com o auxílio da ferramenta online HTP *oligo designer* e se encontram na Tabela 6.

Tabela 6 – *Primers* utilizados para a clonagem dos genes que codificam para as proteínas TarA, TarB, TarS, TarL, TarO, TarI, TarJ e MnaA de *S. aureus* SA16.

Proteína codificada	Primer <i>forward</i>	Primer <i>reverse</i>
TarA	CAGGGCGCCATGTTGCAAAT GGTTGAAAATATTAAC	GACCCGACGCGGTTACTATT TCGCTTTTTTTATTTTCTTT
TarB	CAGGGCGCCATGATTACGCC TCAAGTGATTG	GACCCGACGCGGTTATCATT TTGTCACCATCTTATCTATAT ATT
TarL	CAGGGCGCCATGTTGGTTAA AAGTAAGATATATAGATA AAATC	GACCCGACGCGGTTATTAGC TACCAAATAAATTTCTGACT AAA
TarO	CAGGGCGCCATGGTTACATT ATTACTAGTTGCAGTAACAA TG	GACCCGACGCGGTTACTAAT CTTCTTTGTGTGACGATTT
TarI	CAGGGCGCCATGAAATACGC TGGTATTCTAGCTG	GACCCGACGCGGTTATTAAT CATCGGCAATACCAC
TarJ	CAGGGCGCCATGATTAATCA AGTATATCAATTAGTTGCAC C	GACCCGACGCGGTTATTACA TAATCCATTTTAATACTGTTT TACC
MnaA	CAGGGCGCCATGATGAAAAA GATTATGACCATATTG	GACCCGACGCGGTTATCAGA AATCGCTTGGTTTT

Fonte: Elaborada pela autora.

A clonagem em células de *Escherichia coli* DH10B foi realizada de acordo com o protocolo estabelecido para o sistema LIC (do inglês *ligase independent cloning* ou clonagem independente de ligase).⁶⁷ Os plasmídeos contendo os genes de interesse foram então extraídos e transformados por choque térmico em células de *E. coli*, cepa BL21 (DE3) Rosetta. O cultivo das células foi realizado em meio Luria-Bertani (LB), na presença de 50 µg/mL do antibiótico canamicina.

A transformação e o crescimento de colônias positivas foram possíveis somente para os genes TarA, TarS e MnaA. As massas moleculares das proteínas codificadas pelos genes são, aproximadamente, 27 kDa (TarA), 66 kDa (TarS) e 42 kDa (MnaA). As construções possuem uma cauda de hexa-histidinas e uma proteína de carreadora tiorredoxina na região N-terminal da proteína de interesse, com massa molecular de aproximadamente 14 kDa (Figura 17).

Figura 17 - Construção para produção da proteína produzida, a qual possui uma cauda de hexa-histidinas em sua região N- terminal, seguida pela proteína carreadora tiorredoxina, um sítio para clivagem pela enzima TEV protease e a proteína de interesse.

Fonte: Elaborada pela autora.

Foram feitos testes de expressão com as proteínas clonadas anteriormente para a determinação das melhores condições de temperatura e concentração de indutor. O cultivo das células foi realizado em meio Luria-Bertani (LB), na presença de 50 µg/mL do antibiótico canamicina. A indução foi realizada por 16 horas a 18°C com a adição 0,1 mM de IPTG após a cultura ter atingido D.O. 600nm (densidade óptica a 600 nm) de 0,8. Os meios celulares foram centrifugados por 30 minutos a 7800 g e o pellet ressuspendido em tampão de lise (Tris-HCl 50 mM pH8,0, 500 mM NaCl, 1 mM PMSF, 1mg/mL lisozima, 5% glicerol). O tampão de lise utilizado para a MnaA foi diferente dos demais (50 mM NaH₂PO₄, pH 8,0, 300 mM NaCl, 10 mM imidazol, 1 mM PMSF, 1 mg/mL lisozima), de acordo com o protocolo utilizado para a caracterização da proteína UDP-GlcNac 2-epimerase de *Bacillus subtilis*.⁶⁸

3.2.2 Testes de solubilidade

Foram realizados testes de solubilidade das proteínas obtidas anteriormente, somente com o tampão lise e pela adição de detergentes (20 mM CHAPS, 0,1% Sarcosyl e 1% Triton X-100) ao mesmo. Os detergentes utilizados foram escolhidos por suas características físico-químicas.

As células contendo MnaA foram incubadas por 3 horas a 4 °C antes do processo de lise. Após, realizou-se a lise por ultra-sonicação com amplitude 20% da potência máxima, durante 15 minutos, em ciclos de 30 segundos ligado e 30 segundos desligado, em gelo. Após, o lisado foi centrifugado a 18000 g por 30 minutos a 4°C para separação da fração solúvel (sobrenadante) da fração insolúvel (*pellet*).

3.2.3 Purificação por afinidade

Somente as proteínas TarA e MnaA foram expressas em grandes quantidades e apresentaram solubilidade. A purificação da TarA foi realizada após a lise do material celular e solubilização da proteína com a adição de 40 mM do detergente CHAPS. Foi utilizada uma coluna HisTrap HP (GE Healthcare Life Sciences) a qual contém o metal níquel imobilizado em uma resina de sefarose. A coluna, de 1 mL de resina, foi inicialmente equilibrada com a passagem de 5 mL de tampão de ligação (50 mM Tris-HCl pH 8,0, 500 mM NaCl, 20 mM imidazol pH 8,0, 4 mM CHAPS, 5% glicerol). Após, foi aplicada a amostra e a coluna foi

lavada com o tampão de ligação até que o sinal de absorção a 280 nm (comprimento de onda em que há absorção de luz pelo triptofano) fosse estabilizado. A eluição foi realizada pela adição contínua de tampão de eluição (50 mM Tris-HCl pH 8,0, 500 mM NaCl, 500 mM imidazol pH 8,0, 4 mM CHAPS, 5% glicerol), com o auxílio do equipamento ÄKTA Purifier. (GE Healthcare Life Sciences)

Após o processo de lise celular e sem a adição de detergentes, a proteína MnaA foi purificada por afinidade utilizando uma coluna contendo 5 mL da resina Ni-NTA Agarose (McLab). A coluna foi equilibrada utilizando o tampão de lavagem (50 mM NaH₂PO₄, pH 8,0, 300 mM NaCl, 20 mM imidazol, 10% glicerol), sendo eluída em 500 mM de imidazol.

3.2.4 Clivagem pela enzima TEV protease

Após a purificação da TarA, foi realizada uma diálise para a retirada do imidazol, concentrando-se a amostra até a metade de seu volume inicial, utilizando o concentrador Vivaspín (GE Healthcare Life Sciences) que seleciona proteínas globulares com massa molecular superior a 10 kDa, e adição do tampão de diálise (50 mM Tris- HCl pH 8,0, 500 mM NaCl, 4 mM CHAPS, 5% glicerol). De acordo com as recomendações do fabricante, este procedimento foi realizado 3 vezes. A proteína foi então concentrada até 0,5 mg/mL e incubada com a adição de 0,01 mg da proteína TEV protease, a uma temperatura de 4°C, durante 16 horas.

O mesmo procedimento foi realizado com a proteína MnaA, onde a diálise e a concentração foram realizadas utilizando o concentrador Vivaspín. O tampão de diálise (50 mM NaH₂PO₄, pH 8,0, 300 mM NaCl, 10% glicerol) utilizado foi adicionado à amostra a ser dialisada. A proteína foi concentrada até cerca de 2 mg/mL e depois incubada com a TEV protease a 4 °C por 16 horas.

3.2.5 Fluorescência intrínseca

Foram realizados experimentos de fluorescência, onde foi monitorada a fluorescência intrínseca dos aminoácidos triptofano, tirosina e fenilalanina. Foram testados: o tampão de lise com a adição de 40 mM do detergente CHAPS, uma amostra contendo 0,5 mg/mL de proteína TarA em tampão de lise também com 40 mM de CHAPS e a mesma amostra anterior após aquecimento a 100°C por 10 minutos, com o auxílio do espectrofotômetro PC1 (ISS) com excitação no comprimento de onda de 280 nm.

3.2.6 Espectroscopia de dicroísmo circular (CD)

Foram realizados experimentos de espectroscopia de dicroísmo circular para a verificação da presença de estrutura secundária nas amostras. Para tal, as amostras foram preparadas em tampão fosfato de potássio, pH 7,0, contendo 150 mM de NaCl e 5% glicerol. Os espectros foram obtidos com a utilização de um espectropolarímetro JASCO J-815 (JASCO), equipado com um sistema de controle de temperatura do tipo PELTIER PTC 423S/15, localizado no Laboratório de Biofísica Molecular “Sérgio Mascarenhas”, IFSC-USP. Os espectros foram coletados na região do UV distante, no intervalo de comprimento de onda de 320 a 200 nm, com resolução de 0,1 nm e taxa de 100 nm min, com resposta de dois segundos. As medidas foram realizadas utilizando-se uma cubeta de quartzo de caminho óptico de 0,1 cm. A concentração de proteína utilizada foi de 0,2 mg/mL (aproximadamente 4,76 μ M), e a concentração de UDP e UDP-GlcNac foi de 4 mM e 0,2 mM, respectivamente.

3.2.7 Experimentos de Cristalização

Testes de cristalização para a proteína TarA a 0,5 mg/mL foram realizados com o auxílio do robô de cristalização HoneyBee 963 (Isogen Life Science) e kits comerciais Crystal Screen HT (Hampton Research), Index HT (Hampton Research) e MIDAS (Molecular Dimensions), utilizando a técnica de difusão de vapor em gota sentada em placas de 96 poços. Além das variações de tampão e de agente precipitante fornecidos pelos kits utilizados, foram feitas variações da proporção entre proteína e solução de cristalização de 1:2, 1:1 e 2:1. As três placas foram reservadas a 18°C. Após, foram variados pH e concentração de agente precipitante em condições próximas às aquelas em que foram obtidos cristais. Para este experimento, foi utilizada a técnica de difusão de vapor por gota pendurada. Cristais obtidos da proteína TarA foram testados em uma micro-fonte de cobre (Cu K α) de alta intensidade, tendo sua difração captada por um detector tipo CCD de alto desempenho, componentes de um difratômetro de raios-X Bruker.

Um procedimento similar foi realizado com a MnaA, onde foram utilizados os kits comerciais Crystal Screen HT (Hampton Research), MIDAS (Molecular Dimensions), Morpheus Crystallization Screen (Molecular Dimensions) e Structure Screen 1+2 HT-96 (Molecular Dimensions). Foram feitas também tentativas com a adição de UDP/UDP-GlcNac e ticarcilina às gotas de cristalização.

3.2.8 Fluorimetria de varredura diferencial (DSF) e Fluorimetria de varredura diferencial quantitativa (qDSF)

Os ligantes selecionados por meio da docagem de compostos, foram testados quanto à sua habilidade de se ligar à enzima MnaA de *S. aureus in vitro*. Para tal, a proteína foi incubada a uma concentração de 2 µM com cada ligante a uma concentração de 6 µM, por 30 minutos. Após, o complexo foi incubado com o corante SYPRO Orange *protein gel stain* (ThermoFischer) para o experimento de DSF, utilizando o equipamento CFX96 *real-time pcr detection system* (BioRad) em um intervalo de temperaturas entre 25 e 95 °C, com incrementos de 1 °C a cada 30 segundos. As temperaturas em que houve o desenovelamento da proteína (T_m) foram comparados com aquelas testadas para a interação com UDP e UDP-GlcNac, que são ligantes conhecidos da enzima. Foram selecionados aqueles compostos que mostraram T_m maiores ou semelhantes aos obtidos pela interação com os ligantes, indicando um processo de estabilização da estrutura da enzima.

Para as medidas de qDSF, foi realizado o mesmo procedimento, porém variando as concentrações do ligante. Os dados foram ajustados de acordo com o modelo proposto por Bai *et al.*, que relaciona a intensidade da fluorescência $Y(T)$ à temperatura T .⁶⁹

$$Y(T) = \left[\frac{1}{1 + K_U} (m_F T + b_F) \right] + \left[\frac{K_U}{1 + K_U} (m_U T + b_U) \right] \quad (4)$$

$$K_U = e^{\frac{-\Delta G}{RT}} \quad (5)$$

$$\Delta G = \Delta H \left(1 - \frac{T}{T_m} \right) - \Delta C_p \left(T_m - T + T \ln \left[\frac{T}{T_m} \right] \right) \quad (6)$$

O termo m_f/m_u na primeira equação é a relação das contribuições da fração de proteína enovelada (m_f) e desenovelada (m_u). Já o termo b_f/b_u leva em consideração as contribuições do corante quando a proteína se encontra nos estados enovelado/desenovelado, incorporando as linhas de base e a inclinação. K_u é a constante de equilíbrio entre os estados da proteína desenovelada e enovelada. Os valores de T_m foram extraídos dos dados de DSF. A constante ΔC_p foi ajustada para 3,21 kcal/(mol.K). Utilizando o modelo de Bai *et al.*, a fração de proteína desenovelada f_U é dada pela equação abaixo.⁶⁹

$$f_U = \frac{[U]}{[U] + [F] + [FL]} \quad (7)$$

Onde $[U]$ é a concentração de proteína desenovelada, $[L]$ é a concentração de ligante livre, $[F]$ e $[FL]$ são as concentrações da proteína enovelada sem ligantes e ligada, respectivamente. O modelo assume que apenas proteínas enoveladas se ligam às moléculas de ligante seguindo o seguinte equilíbrio:

Onde,

$$[P] = [F] + [U] + [FL] \quad (9)$$

$$[L]_T = [L] + [FL] \quad (10)$$

$$K_U = \frac{[U]}{[F]} \quad (11)$$

$$K_d = \frac{[F][L]}{[FL]} \quad (12)$$

Portanto a equação 7 pode ser escrita como:

$$f_U = \frac{1}{1 + [(1/K_U) \times (1 + [L]/K_d)]} \quad (13)$$

Depois de redefinir a concentração de ligante $[L]$ como uma função da concentração total de ligante $[L]_T$, f_U pode ser analisada em termos de apenas dois parâmetros, K_U e K_d , desde que $[L]_T$ e $[P]$ sejam conhecidas. Os dados obtidos foram analisados utilizando o *script*, escrito na linguagem python, chamado *DSF_fitting* e disponibilizado por Bai *et al.*⁶⁹

4 RESULTADOS E DISCUSSÃO

4.1 Análise das estruturas cristalográficas existentes de MnaA

A enzima MnaA de *S. aureus* possui apenas uma estrutura, ligada a moléculas de UDP, resolvida por cristalografia. Contudo, diversas outras estruturas de proteínas UDP-GlcNac não-hidrolisantes se encontram disponíveis no banco de dados PDB. As sequências de aminoácidos das proteínas listadas possuem identidade com a de *SaMnaA* entre 60 e 30%, onde a mais semelhante é a enzima de *Bacillus subtilis*, com 97% de cobertura e 61% de identidade.

As enzimas foram analisadas quanto à presença ou ausência de ligantes e ao ângulo de abertura (também utilizado neste trabalho como variável coletiva, ou CV), definido pelos carbonos-alfa dos resíduos S128, G349 e P213 em *SaMnaA* (Figura 18), ou por resíduos em posições equivalentes nas outras estruturas (Tabela 7). Estes resíduos foram escolhidos pois representam bem o movimento de abertura/fechamento observados para a *SaMnaA*.

A enzima de *Methanocaldococcus janaschii*, por exemplo, foi resolvida com (PDB: 4NES) e sem ligantes (PDB: 4NEQ) em seu sítio ativo.³⁶ A primeira estrutura está no estado “fechado”, com um ângulo CV de 28 graus. A estrutura sem ligantes se encontra no estado definido como “aberto”, com ângulo CV de 54 graus. A estrutura de *S. aureus*, que contém apenas uma molécula de UDP em seu sítio ativo, apresenta um ângulo CV de 31 graus para a cadeia A.

Figura 18 – Estrutura cristalográfica da cadeia A da proteína MnaA de *S. aureus*. Em vermelho, os aminoácidos S128, G349 e P213, que compõe o ângulo de abertura medido (CV).

Fonte: Elaborada pela autora.

Tabela 7 – Estado de abertura das estruturas cristalográficas conhecidas das enzimas UDP-GlcNac 2-epimerase.
*cadeia sem ligantes.

Código PDB	Organismo	Ligantes	Estado oligomérico	Ângulo CV cadeia A (graus)	Ângulo CV cadeia B (graus)
Enzimas não-ligadas					
4NEQ	<i>M. jannaschii</i>	-	Monômero	53	-
1V4V	<i>T. thermophilus</i>	-	Dímero	53	40
3OT5	<i>L. monocytogenes</i>	-	Dímero	47	33
1°6C	<i>B. subtilis</i>	-	Dímero	53	58
4HWG	<i>R. bellii</i>	-	Monômero	25	-
3DZC	<i>V. cholerae</i>	-	Dímero	48	46
Enzimas ligadas às moléculas de UDP e UDP-GlcNac					
1VGV	<i>E. coli</i>	UDP-GlcNac	Dímero	40*	25
3BEO	<i>E. coli</i>	UDP, UDP-GlcNac	Dímero	30	30
4FKZ	<i>B. subtilis</i>	UDP, UDP-GlcNac	Dímero	25	25
4NES	<i>M. jannaschii</i>	UDP, UDP-GlcNac	Monômero	29	-
5DLD	<i>B. vietnamiensis</i>	UDP, UDP-GlcNac	Monômero	24	-
Enzimas ligadas somente à UDP					
5ENZ	<i>S. aureus</i>	UDP	Dímero	31	26
1F6D	<i>E. coli</i>	UDP	Dímero	40	27

Fonte: Elaborada pela autora.

Pela análise dos resultados da Tabela 7, pode-se observar que os ângulos das estruturas cristalográficas da enzima sem ligantes (PDB: 1O6C, 1V4V, 3OT5, 4NEQ e 3DZC) variam entre 47 e 53°, indicando a tendência ao estado “aberto”, com exceção da estrutura cristalográfica de *Rickettsia bellii* (PDB: 4HWG). A estrutura cristalográfica da última pode apresentar essa diferença no ângulo de abertura devido à simetria cúbica observada. Nesta simetria, as interações intermoleculares podem estabilizar a conformação fechada da proteína, mesmo na ausência de ligantes.

Nas estruturas cristalográficas em que ambos UDP e UDP-GlcNac se encontram no sítio ativo (PDB: 1VGV, 3BEO, 4FKZ, 4NES e 5DLD), o ângulo de abertura é menor do que aquele para o estado não-ligado da enzima, variando entre 24 e 30°. Proteínas que possuem o ângulo CV próximo a este intervalo, são, portanto, caracterizadas como pertencentes ao estado “fechado”. As estruturas de *SaMnaA* (PDB: 5ENZ) e de *E. coli* (PDB: 1F6D), que foram resolvidas apenas com o UDP em seu sítio ativo, apresentam ângulos intermediários.

Cada monômero das duas enzimas possui ângulos de abertura diferentes entre si. Isto indica que a ligação das moléculas de UDP pode gerar uma variedade de subestados, variando entre os estados fechado e aberto, com ângulos desde 26-27° até 30-40°.

A análise destas estruturas indica a existência de uma mudança estrutural associada à ligação das moléculas de UDP e/ou UDP-GlcNac. Contudo, as estruturas cristalográficas representam conformações únicas de um conjunto de estruturas termodinamicamente acessíveis. Neste sentido, foram realizadas simulações de dinâmica molecular para a obtenção de uma perspectiva mais detalhada dos movimentos realizados pela proteína.

4.2 Simulações de dinâmica molecular

A dinâmica molecular é uma técnica muito utilizada para o estudo dos movimentos de macromoléculas biológicas com resolução atômica. Desde muito se sabe que as moléculas como proteínas, DNA e RNA, não se mantêm estáticas em solução, mas passam por flutuações entre diferentes estados e subestados.⁷⁰ Os movimentos das macromoléculas biológicas influenciam sua função e suas interações intermoleculares.⁷¹ A utilização de simulações de dinâmica molecular realizadas com a estrutura cristalográfica da proteína MnaA de *S. aureus*, fornece, portanto, informações sobre a influência das interações realizadas com seus ligantes em seu movimento de abertura/fechamento, bem como o efeito deste em sua função.

Para a análise dos movimentos entre os domínios da proteína, foram realizadas diversas simulações de dinâmica molecular. Ao todo, foram realizados 33,5 μ s de simulação. A análise das simulações, que foram realizadas utilizando métodos de dinâmica molecular diferentes, geraram um rico conjunto de dados a serem analisados. A partir das trajetórias e da análise energética da interação proteína-ligante, é possível explicar a natureza das forças que mantêm o ligante no sítio ativo e que também são responsáveis pelo movimento de abertura e fechamento da enzima.

4.2.1 Simulações de equilíbrio

Inicialmente, foram realizadas simulações em que se garantiu que o sistema estivesse em equilíbrio termodinâmico, por isso denominadas simulações de equilíbrio. O sistema APO da enzima SaMnaA foi simulado inicialmente por 1 μ s. Durante a simulação, observou-se um movimento entre os domínios C- e N-terminal, onde sua magnitude pôde ser medida pelas

mudanças no ângulo de abertura. As Figuras 19 A e 19 B mostram que a enzima acessa o estado “aberto”, com CV de aproximadamente 53° e o estado “intermediário”, com CV de aproximadamente 40°. O sistema partiu do estado “intermediário” da estrutura cristalográfica, transicionando para o estado “aberto” logo na etapa de equilíbrio do sistema. A proteína não acessa o estado fechado durante todo o tempo de simulação.

Esta foi então estendida até 5 μ s para o aumento da amostragem de estados termodinamicamente disponíveis. Mesmo com o aumento do tempo de simulação, o estado “fechado” nunca foi visitado (Figuras 19 E e 19 F). O estado “intermediário” foi o mais populado, gerando uma distribuição bimodal do ângulo CV, com picos em aproximadamente 40 e 53°, como mostra a Figura 19 F.

Figura 19 – Simulações de equilíbrio para a proteína *SaMnaA*. (A) Ângulo de abertura observado durante a simulação de 1 μ s do sistema APO. (B) Histograma dos ângulos de abertura observados durante a simulação de 1 μ s do sistema APO. (C) Ângulo de abertura observado durante a simulação do sistema HOLO. (D) Histograma dos ângulos de abertura observados durante a simulação do sistema HOLO. (E) Ângulo de abertura observado durante a simulação de 5 μ s do sistema APO. (F) Histograma dos ângulos de abertura observados durante a simulação de 5 μ s do sistema APO.

Fonte: Elaborada pela autora.

A trajetória de 1 μ s foi decomposta utilizando a análise de componentes principais (PCA) para a visualização dos movimentos principais realizados pela proteína durante a simulação. As duas primeiras componentes principais (PCs), indicam um movimento de abertura (primeira PC), associado a um movimento de torção entre os domínios (segunda PC), como indicado pelas setas na Figura 20. Os dados observados estão de acordo com as

observações realizadas da trajetória e no ângulo CV computado, que indica o movimento da proteína para o estado “aberto”. A enzima *SaMnaA* possui, portanto, um movimento de baixa frequência entre seus domínios, favorecendo o estado “aberto” quando na ausência dos ligantes.

Figura 20 – Análise das componentes principais da trajetória do sistema APO durante a simulação de equilíbrio. Projeções da primeira (A) e da segunda (B) componentes da análise.

Fonte: Elaborada pela autora.

A trajetória contendo 5 μ s de simulação do sistema APO foi analisada também utilizando o pacote PyEMMA.⁵⁶⁻⁵⁷ Como citado anteriormente, o pacote utiliza MSMs para a análise das

trajetórias das simulações de dinâmica molecular, que é altamente acurado, caso os estados MSM sejam capazes de tornar as coordenadas coletivas dos processos mais lentos discretas. Felizmente, proteínas podem ser caracterizadas utilizando apenas as componentes dos processos lentos em um dado tempo de atraso τ .⁵⁶ Uma análise de MSM consiste na decomposição de um conjunto de estados e na obtenção da probabilidade de interconversão entre estes. Há dois passos principais no processo de construção de um MSM. O primeiro é o agrupamento dos dados em algum conjunto pequeno de estados, e o segundo é a estimativa das probabilidades de transição entre os estados.⁷²

A escolha das variáveis coletivas para a realização da determinação dos diferentes estados da proteína foi realizada utilizando o processo chamado análise da componente independente baseada na estrutura temporal (ou com atraso de tempo) (tICA). Este é comumente utilizado para a diminuição da dimensionalidade dos dados coletados a partir das posições dos átomos na trajetória a ser analisada (Figura 21). São realizados testes em tempos de atraso diferentes, de modo que seja escolhido aquele em que o sistema possui uma dinâmica *markoviana*, isto é, em que há uma descorrelação temporal entre os estados.

A trajetória é decomposta em estados discretos, determinados pelas componentes independentes baseadas na estrutura temporal (tICs), as quais formam uma matriz de transição, que mantém informações sobre quantas vezes ocorreram transições de um estado de uma coluna para uma linha. Após, é calculada a média de transições observadas, onde sua transposição gera a matriz de probabilidade de transição reversível, de modo que cada transição *forward* aconteça um número de vezes igual às transições *reverse*. É realizada então uma análise das populações ajustadas, que são caracterizadas pelo primeiro autovetor da matriz de probabilidade de transição. Estados com maiores populações são mais termodinamicamente estáveis. A partir das análises realizadas é possível, portanto, obter os autofluxos dos autoprocessos dominantes do conjunto de dados, representando o movimento de um estado para o outro (Figura 22).

No caso da análise da SaMnaA, a trajetória de dinâmica molecular foi reduzida às coordenadas de seus carbonos-alfa, a qual foi decomposta em duas tICs. Foram identificados três macroestados, como mostra a Figura 23. As estruturas representativas de cada estado mostram que durante a simulação foi do estado “intermediária” para o “aberto”. A primeira TIC, que descreve o movimento mais lento do sistema está de acordo com o comportamento do ângulo CV escolhido, o que indica que esta caracteriza de forma satisfatória o movimento da proteína nestas condições. A análise dos conjuntos permitiu uma primeira estimativa do tempo de fechamento do sistema, desde o estado aberto até o intermediário, que ficou em

torno de 879 ns. O tempo para o movimento de abertura, indo do estado intermediário para o aberto, foi calculado em torno de 423 ns. Como não foi possível acessar o estado “fechado” durante a simulação do sistema APO da proteína, não foi possível estimar o tempo típico para o movimento de abertura/fechamento completo.

Figura 21 – Métodos para a redução de dimensionalidade. Evolução temporal das técnicas utilizadas para a decomposição das posições atômicas provenientes de uma simulação de dinâmica molecular em conjuntos de dados representativos menores e posterior análise por MSM.

Fonte: Adaptada de HUSIC.⁷³

Figura 22 – Processo de análise de dados de uma simulação por MSM. (a) Uma ou mais trajetórias de dinâmica molecular, onde cada frame é separado por um intervalo de tempo da simulação. (b) É realizada uma decomposição discreta dos estados encontrados na trajetória. (c) É construída uma matriz de contagem das transições observadas, a qual registra quantas vezes o sistema transiciona de uma posição na linha para uma posição na coluna. (d) A partir da matriz anterior, é calculada uma matriz de probabilidade de transições reversíveis, que é a transposição da anterior, de modo a garantir que as transições *forward* ocorram um número igual de vezes que as transições *reverse*. (e) São calculadas as populações ajustadas que são mais termodinamicamente estáveis. (f) Representação esquemática dos autoprocessos dominantes da trajetória, evidenciando os autofluxos entre os estados.

Fonte: Adaptada de HUSIC.⁷³

Os resultados obtidos da simulação de equilíbrio para a proteína APO sugerem que esta é encontrada primariamente no estado “aberto”, acessando também o estado “intermediário”. O estado fechado não foi observado mesmo durante a simulação de 5 μ s, o que concorda com as análises realizadas de outras estruturas cristalográficas disponíveis no PDB, onde ângulos de abertura que caracterizam o estado “fechado” só ocorrem com os ligantes presentes no sítio ativo.

Figura 23 – Análise de MSM da trajetória da simulação de 5 μ s para o sistema APO da proteína *SaMnaA*. (A) Comparação entre os ângulos CV medidos (gráfico superior) e a tIC 1 da análise tICA (gráfico inferior). (B) Mapa da tIC 1 versus tIC 2, onde os estados encontrados estão representados por três agrupamentos de pontos vermelhos, verdes e azuis. (C) Estruturas representativas dos três estados encontrados para o sistema. As esferas representam os resíduos utilizados para o cálculo de CV.

Fonte: Elaborada pela autora.

Para o sistema APO, ainda, foi realizada uma simulação em equilíbrio termodinâmico do tipo “*coarse grained*” (Figura 24). É possível observar que, mesmo com a utilização de um campo de forças diferente e uma tempo de simulação de 23 μ s, uma amostragem de conformações muito similar às anteriores foi obtida. Esta similaridade indica que, mesmo em uma simulação onde os parâmetros das moléculas e o campo de forças utilizados são definidos de forma diferenciada, uma força semelhante àquela encontrada na simulação atomística continua a agir na proteína, fazendo-a realizar os movimentos de abertura e fechamento entre os domínios.

Figura 24 – Ângulos acessados durante a simulação *coarse-grained*. O gráfico à esquerda mostra os ângulos CV durante os 25 μ s de simulação. À direita, a distribuição de ângulos. É possível notar que o pico de distribuição se encontra em aproximadamente 75°, configurando o estado “aberto”, esperado para a proteína APO.

Fonte: Elaborada pela autora.

A simulação de equilíbrio para o sistema HOLO mostrou comportamento diferente quando comparado com o sistema APO (Figuras 19 C e 19 D). A enzima permanece no estado “fechado”, com um ângulo CV típico de aproximadamente 28-30 graus. Novamente, o comportamento observado condiz com aquele encontrado nas estruturas cristalográficas analisadas, onde a presença das moléculas de UDP e UDP-GlcNac modifica o equilíbrio da

distribuição de populações para o favorecimento da conformação fechada, que é mantida durante toda a simulação.

A trajetória também foi decomposta utilizando PCA (Figura 25). As duas primeiras componentes representadas na estrutura da proteína não mostram um movimento claro entre os domínios, como o que foi observado na proteína *SaMnaA* não-ligada.

Figura 25 - Análise das componentes principais da trajetória do sistema HOLO durante a simulação de equilíbrio. Os ligantes estão representados em azul. Projeções da primeira (A) e da segunda (B) componentes da análise.

Fonte: Elaborada pela autora.

Figura 26 – Análise da interação entre UDP-GlcNac, UDP e *SaMnaA*. Mapa das interações entre UDP (A) e UDP-GlcNac (B) com a proteína *SaMnaA*. Contribuição da energia potencial total de interação entre UDP (C) e UDP-GlcNac (D) e os os principais aminoácidos interagentes da proteína *SaMnaA*.

Fonte: Elaborada pela autora.

Para o sistema HOLO, ainda, foi realizada uma análise das energias de interação entre a proteína-UDP e entre a proteína-UDP-GlcNac. A análise destas energias mostra uma rede de interações polares envolvendo os ligantes, a enzima e o solvente. A Figura 26 mostra um diagrama realizado utilizando o programa LIGPLOT,⁷⁴ que representa as interações envolvendo UDP e a enzima a partir de uma conformação selecionada da simulação de equilíbrio.

As interações entre UDP e a enzima envolvem principalmente dos resíduos Glu288, Arg206 e Arg11, além de diversas moléculas de solvente e da interação com a molécula de UDP-GlcNac. As contribuições individuais destes resíduos para a energia potencial de interação total estão mostradas na Figura 26, onde Arg11 e Arg206 se destacam como as maiores contribuições energéticas devido às interações com os grupos fosfato da molécula de UDP. Há também contribuições desfavoráveis, como a realizada pelo resíduo Glu288, devido à repulsão de cargas realizada também com os fosfatos.

A molécula de UDP-GlcNac interage com os resíduos Arg206, His40, Arg207, Lys239, His238 e Glu132. Além disso, interage também com Leu68, Gln66 e Met98 através de sua cadeia principal. Novamente, as maiores contribuições para a energia potencial total de interação são devidas aos resíduos carregados positivamente: Arg206, Arg 207, Lys239, His40 e His238. A análise energias de interação de ambos UDP e UDP-GlcNac com a proteína sugere, portanto, que os resíduos carregados positivamente possuem papel de destaque no posicionamento dos ligantes no sítio ativo.

As simulações realizadas no equilíbrio sugerem que os movimentos e abertura e fechamento da enzima ocorrem têm influência das interações com ambos ligante e cofator. A ligação do substrato representa uma grande contribuição energética favorável à interação com diversos aminoácidos localizados no sítio ativo. Mesmo em simulações mais longas, não foi possível amostrar todos os estados e subestados possíveis para a proteína. Foram feitas ainda simulações com a proteína ligada somente ao UDP, porém em todas as tentativas de montagem do sistema, o cofator saiu do sítio ativo ainda na termalização do sistema. Isto pode indicar que o UDP se liga fracamente à enzima, não havendo interações suficientes com o sítio ativo para mantê-lo ligado. Para que outros estados fossem acessados pelos sistemas APO e HOLO, foram utilizadas técnicas de dinâmica molecular aceleradas pela adição de potenciais externos. Estas técnicas de simulação fornecem também informações sobre o perfil energético da proteína, o que pode fornecer indícios sobre fatores importantes no posicionamento do UDP-GlcNac no sítio ativo.

4.2.2 Simulações de dinâmica molecular aceleradas

Para entender melhor o perfil energético associado às mudanças conformacionais devidas à ligação do substrato, foram realizadas também simulações aceleradas para os sistemas APO e HOLO. Estas forçam o sistema a amostrar um conjunto maior de estados da proteína pela adição de um potencial externo. Como o valor para o potencial é conhecido, pode-se também reconstruir o perfil de energia livre (potencial de força média ou PMF) da proteína na situação de equilíbrio inicial.

As simulações de GAMD representam uma metodologia interessante para a amostragem de estados, pois não requerem uma CV pré-definida para a aplicação do potencial de aceleração. O potencial harmônico é adicionado à energia total do sistema e/ou de diedro. Três simulações independentes do sistema APO foram realizadas, cada uma partindo de uma conformação inicial diferente. Foram escolhidas estruturas amostradas durante a simulação de equilíbrio nos tempos 0, 50 e 100 ns. Cada conformação foi simulada por 500 ns. Podemos observar que as duas simulações que foram iniciadas em ângulos próximos de 55°, foram para o estado “intermediário” com ângulos de abertura entre 40-43°. Na terceira simulação, a enzima se mostrou mais flexível, com ângulos de abertura entre 45 e 65°, como mostrado na Figura 27. Os histogramas das simulações aceleradas por gaussianas indicam uma clara preferência do sistema ao estado “intermediário”. Mesmo com o potencial de impulso, a enzima APO nunca visita o estado fechado, sugerindo uma alta barreira energética a ser vencida para que essa conformação seja alcançada.

Os perfis de PMF foram computados a partir das simulações de GAMD (Figura 27.). Em todos os casos, as energias aumentaram rapidamente para ângulos de abertura menores do que 35°. A simulação GAMD-APO-I (curva em preto nas Figuras 27 A e 27 C) mostra um perfil de energia favorável para as conformações que abrangem ângulos entre 35 e 55 graus. A simulação GAMD-APO-II (curvas em vermelho) apresenta baixas energias associadas a ângulos de abertura entre 35 e 45°. A terceira, GAMD-APO-III, (curvas em azul) mostra energias favoráveis associadas a ângulos entre 35 e 70°. Todas as simulações concordam que o estado intermediário é um estado de baixa energia para o sistema APO, e as simulações I e III também sugerem que o estado “aberto” (ângulo CV > 50°) pode ser acessado pela proteína não-ligada. Apesar de os resultados obtidos com GAMD serem muito ruidosos, como um todo estes concordam com os resultados obtidos durante a simulação de 5 μ s, sugerindo que, uma vez que a proteína se desliga de seus cofator e substrato, um equilíbrio putativo entre as

conformações “aberta” e “fechada”, com ângulos de abertura entre 40 e 45° (intermediário) e 50-60 graus (conformação aberta) são esperados em solução.

Similarmente, três simulações independentes para o sistema HOLO foram realizadas e analisadas. As simulações também mostraram que a proteína ligada tende a se manter na conformação fechada, com ângulo de abertura de aproximadamente 30 graus, como mostrado na Figura 28. Aqui, duas simulações permaneceram todo o período da simulação na conformação fechada (GAMD-HOLO-I e GAMD-HOLO-II, mostrados nas Figuras em vermelho e azul, respectivamente) enquanto a outra simulação (GAMD-HOLO-I, em preto) mostrou um ângulo de abertura que se iniciou em 26 graus, chegando até 48 graus. Curiosamente, nessa simulação, foi observada a saída do ligante UDP-GlcNac do sítio ativo da enzima, em um tempo de simulação de 175 ns, como indicado nas energias de ligação e RMSD computados a partir dessa trajetória (Figuras 28 D e 28 E).

Figura 27 – Simulações de GAMM para o sistema APO da estrutura SaMnaA. Os resultados das três simulações estão mostrados: GAMM-APO-I em preto, GAMM-APO-II em vermelho e GAMM-APO-III em azul. (A) Ângulos CV observados em função do tempo de simulação. (B) Distribuição dos ângulos CV durante a simulação. (C) PMF computada após o rebalanceamento.

Fonte: Elaborada pela autora.

Figura 28 – Dados da simulação GAMMD para o sistema HOLO. Os resultados das três simulações independentes estão mostrados: GAMMD-HOLO I em preto, GAMMD-HOLO II em vermelho GAMMD-HOLO III em azul. (A) Ângulos CV observados durante a simulação. (B) Histograma dos ângulos observados em função do tempo de simulação. (C) PMFs calculadas após o rebalanceamento. (D) RMSD calculado para as moléculas de UDP (preto) e UDP-GlcNac (vermelho). (E) Energias totais de interação computadas entre a proteína e as moléculas de UDP (preto) e UDP-GlcNac (vermelho).

Fonte: Elaborada pela autora.

As PMFs calculadas para as simulações de GAMD do sistema HOLO (Figura 28 C) indicam um mínimo na energia livre associada ao ângulo de abertura de aproximadamente 25°. Para as simulações GAMD-HOLO-I, onde o evento de desligamento foi observado, uma mudança no mínimo de energia em direção ao estado “intermediário” foi observado, consistente com o mínimo esperado para uma enzima não-ligada. O desligamento da molécula de UDP-GlcNac causa uma mudança na posição do UDP no sítio ativo, forçando-o a perder interações com a proteína. O histograma do ângulo de abertura para as três simulações de GAMD indicam uma preferência pelo estado “fechado” para o sistema ligado, o que é consistente com as simulações de equilíbrio realizadas. Além disso, o evento de desligamento observado devido à aceleração aplicada em nossa simulação fora do equilíbrio sugere que o evento de abertura está ligado à ausência do ligante.

A descrição ruidosa do perfil de energia da proteína *SaMnaA* pode ser associado com uma amostragem insuficiente da transição entre os estados, os quais são raramente observados em nossas simulações. Neste sentido, uma descrição energética mais acurada pode ser obtida com a utilização de outros métodos nos quais a CV é sistematicamente amostrada durante a simulação. Para tal, a dinâmica da proteína foi explorada utilizando o método de Dinâmica Molecular Adaptativa (ABMD). Três simulações de ABMD independentes foram realizadas para o sistema APO, com as mesmas coordenadas iniciais escolhidas para GAMD. Cada simulação explorou ângulos de abertura CV em um intervalo de 10-100 ° e resolução de 1° em um tempo de dinâmica de 250 ns. Os resultados obtidos, mostrados na Figura 29, indicam que os ângulos de CV os mínimos nos gráficos de PMF são obtidos no intervalo entre 40-60 graus, com barreiras energéticas de aproximadamente 10 kcal/mol (~ 16 KT) para atingir o estado “fechado”, definido como um ângulo de CV de cerca de 25 graus. A primeira simulação (ABMD-APO-I) mostrou um mínimo largo, cobrindo toda a região de ângulo definidos como os estados “aberto” e “intermediário”, enquanto as simulações ABMD-APO-II e ABMD-APO-III mostraram um mínimo mais definido em aproximadamente 55 °, consistente com o estado “aberto”. Conjuntamente, essas simulações mostraram um mínimo claro de energia cobrindo os estados “aberto” ou “intermediário” para a proteína não-ligada.

Figura 29 – Simulações de ABMD para o sistema APO. Gráficos representativos das simulações denominadas ABMD-APO-I (linha superior), ABMD-APO-II (linha média) e ABMD-APO-III (linha inferior). (A) Evolução temporal do ângulo CV. (B) Distribuição dos ângulos CV. (C) Potencial de força média. As linhas em cinza indicam a evolução temporal e convergência da PMF em 10 simulações de 25 ns cada.

Fonte: Elaborada pela autora.

É possível notar que os perfis de PMF observados nas simulações de GAMD e ABMD (Figuras 27 e 29) concordam como um todo, mas não são similares. As diferenças são esperadas, visto que as metodologias são baseadas em premissas distintas. Na GAMD, a simulação é acelerada pela introdução de potenciais de impulso que não consideram as “coordenadas de reação” do evento molecular a ser avaliado, enquanto na ABMD, o potencial de viés é direcionado a estas coordenadas. O efeito final é um perfil energético mais ruidoso, porém gerado sem viés para a GAMD e um perfil mais suave, porém enviesado, para a ABMD. Neste contexto, é interessante que ambas as técnicas sejam utilizadas em conjunto.

As simulações de ABMD para o sistema HOLO acessaram um intervalo maior de ângulos de CV, incluindo o estado aberto. Novamente, uma vez forçado no sentido de ângulos de CV maiores o substrato foi dissociado da estrutura da enzima, resultando em seu desligamento e abertura da enzima. A correlação entre esta perda de interação e a abertura se torna clara quando a Figura 30 é comparada com a Figura 31, que mostra as energias potenciais de interação entre a enzima e o substrato/cofator. Os dados de simulação indicam que a presença do ligante é crucial para a manutenção de uma conformação fechada estável. Uma vez que esta conformação é desestabilizada (por forças externas, neste caso), o ligante é espontaneamente dissociado da enzima.

Após as análises das simulações em equilíbrio e aceleradas dos sistemas APO e HOLO, ainda resta o esclarecimento do mecanismo envolvido na interação entre o ligante e enzima, bem como a influência desta interação no movimento dos domínios, ou mesmo qual a natureza da força que realiza o movimento entre os domínios na presença de UDP ou UDP-GlcNac. Uma análise visual da estrutura da enzima revela uma rede de aminoácidos carregados positivamente nas vizinhanças do sítio ativo. Estes resíduos estabilizam a interação com os grupamentos fosfato do UDP-GlcNac, tornando as contribuições eletrostáticas mais relevantes para a energia total de interação. Por outro lado, na ausência de um substrato, ou produto, a interação destes resíduos carregados positivamente é muito desfavorável e pode ser aliviada se os resíduos se afastam pelo movimento entre os domínios. Portanto, nossa hipótese é de que a principal força para esse movimento de grande escala possa ser a energia eletrostática associada com a interação de resíduos carregados positivamente.

Figura 30 - Simulações de ABMD para o sistema HOLO. Gráficos representativos das simulações denominadas ABMD-HOLO-I (linha superior), ABMD-HOLO-II (linha média) e ABMD-HOLO-III (linha inferior). (A) Evolução temporal do ângulo CV. (B) Distribuição dos ângulos CV. (C) Potencial de força média. As linhas em cinza indicam a evolução temporal e convergência da PMF em 10 simulações de 25 ns cada.

Fonte: Elaborada pela autora.

Figura 31 – Energia total de interação potencial computada entre UDP (preto) e UDP-GlcNac (vermelho) com a proteína *SaMnaA* para as três simulações ABMD.

Fonte: Elaborada pela autora.

Para que esta hipótese seja avaliada, as estruturas APO e HOLO foram analisadas utilizando o servidor web Delphi Force,⁷⁵ para acessar a magnitude das forças eletrostáticas entre os domínios N-terminal e C-terminal da *SaMnaA*. Para a enzima HOLO, os ligantes foram definidos como parte do domínio C-terminal. As forças líquidas computadas estão representadas como flechas na estrutura da proteína na Figura 32. Aqui observamos que, na ausência do substrato/cofator, um vetor de força líquida é observado apontando para a direção do movimento entre os domínios (seta laranja). Na presença do substrato, a força líquida é modificada para o movimento de torção entre os domínios (também observado na simulação de equilíbrio para o sistema HOLO). Portanto, a análise das forças eletrostáticas fornece um indício de que a força que realiza os movimentos observados para a proteína *SaMnaA* tenha uma natureza, ao menos parcialmente, eletrostática.

Figura 32 – Forças eletrostáticas nas estruturas de *SaMnaA*. A soma das forças eletrostáticas de interação entre os domínios C-terminal e N-terminal para os estados aberto (laranja), intermediário (cinza) e fechado (roxo), com os ligantes presentes no sítio ativo (ciano).

Fonte: Elaborada pela autora.

Figura 33 – Distribuição de ângulos CV acessados pelos mutantes construídos para o sistema APO, no tempo de 350 ns de simulação para cada mutante. No gráfico à esquerda se encontram em destaque as distribuições para a proteína sem mutações (curva em preto), para o mutante M62E-Q66E-L68D (curva em verde) e para o mutante K39E-M62E-Q66E-L68D-M43E (curva em vermelho). É possível observar que as mutações resultaram em um deslocamento das distribuições de ângulos, indicando uma estabilização da proteína em seus estados “intermediário” e “fechado”.

Fonte: Elaborada pela autora.

Foram realizadas também mutações em diversos aminoácidos presentes no sítio ativo, de modo a reduzir o número de resíduos positivamente carregados, ou adicionando resíduos negativos em posições próximas aos aminoácidos de interesse de modo a gerar uma força de atração entre os mesmos. Dos mutantes analisados (Figura 33), pelo menos dois (M62E-Q66E-L68D e K39E-M62E-Q66E-L68D-M43E) se mostraram efetivos na manutenção do

estado “fechado”, mesmo sem ligantes no sítio ativo da proteína, reforçando a hipótese de que a força que atua nos domínios para a realização dos movimentos de abertura/fechamento possui uma componente eletrostática importante.

Além disso, as mudanças observadas no perfil energético da proteína, causadas pela interação entre substrato e proteína, concordam com o conceito de um perfil energético dinâmico,⁷⁶ que é uma característica importante para a função da enzima, pois a macromolécula pode ajustar sua população de acordo com as mudanças no ambiente, tais como a presença do substrato. São propostos dois principais modelos para este perfil de energia dinâmico: tipo I e tipo II. O tipo I, é caracterizado por uma mudança no ambiente, que causa um deslocamento no mínimo do perfil energético, modificando também a conformação preferencial da proteína. O tipo II, por outro lado, descreve a situação em que a mudança no ambiente causa uma mudança na rugosidade da superfície, onde as conformações que se encontravam presas em um mínimo energético agora possuem uma barreira menor para superar e assumir uma nova conformação. Dentre as duas definições, o perfil dinâmico de energia observado para a *SaMnaA* parece ser mais próximo do tipo I, que é normalmente observado em enzimas que são reguladas alostericamente ou em complexos multimoleculares.

Como descrito anteriormente, o movimento de domínios pode ser um requisito estrutural essencial para a catálise. O estado fechado, observado na presença do substrato, faz com que o sítio ativo exclua moléculas de solvente, resultando em um meio que favorece a ação das forças eletrostáticas, permitindo que a catálise ocorra. Após a ocorrência da reação, o movimento de abertura permite a difusão dos produtos e a entrada de novas moléculas de substrato. Além disso, a simulação *coarse grained* realizada também evidencia o papel das forças eletrostáticas no movimento realizado pela *SaMnaA*, pois mesmo com um campo de forças e definições de massas atômicas distintas, foi observado comportamento semelhante àquele das simulações utilizando o campo de forças AMBER.

Os resultados apresentados sobre as simulações de dinâmica molecular da proteína em seus estados APO e HOLO, bem como as conclusões a que chegamos com este trabalho foram publicadas em um artigo denominado “*Energy landscape of the domain movement in Staphylococcus aureus UDP-N- acetylglucosamine 2-epimerase*”, submetido à revista “*Journal of Structural Biology*” e aceito no ano de 2019. O mesmo se encontra disponível integralmente na seção de Apêndices.

4.2.3 Conclusões parciais

A partir das simulações realizadas, é possível concluir que a proteína tem seus movimentos de abertura e fechamento influenciados pela força eletrostática repulsiva gerada pelos aminoácidos carregados positivamente, localizados no sítio ativo da enzima. A entrada dos ligantes anula essas forças de repulsão pela interação com os grupamentos fosfato presentes em ambos UDP e UDP-GlcNac. O movimento de fechamento da enzima, bem como a interação gerada pela entrada do substrato são importantes para a epimerização do composto.

Outra informação importante obtida pela análise das trajetórias de dinâmica molecular, em especial as aceleradas, é o perfil de energia livre da proteína ligada e não-ligada. As PMFs obtidas mostram que este é um perfil energético dinâmico, onde a entrada do substrato modifica o mínimo energético da proteína, privilegiando o estado “fechado” dentre as populações de estados conformacionais em solução. O oposto ocorre quando não há ligantes. Neste caso, o mínimo de energia é deslocado de forma que é encontrada com maior frequência o estado “aberto”. Esta relação entre a entrada do substrato e a mudança no perfil energético da proteína podem começar a elucidar os mecanismos de entrada de UDP-GlcNac e saída de UDP-ManNac, através da difusão do produto e consequente abertura da proteína para a entrada de um novo substrato.

4.3 Docagem e seleção de compostos

Os resultados das simulações de dinâmica molecular mostraram que a enzima tem uma mobilidade intrínseca e que esta mobilidade é dirigida pela presença de um ligante no sítio ativo. Por este modelo, o complexo proteína-ligante em posição de catálise é formado quando da interação e quando a conformação enzimática se encontra no estado fechado. Portanto, este estado deve ser o estado mais apropriado para a análise prospectiva de novos compostos capazes de interagir com o sítio ativo da enzima, na busca de potenciais inibidores que possam ser adjuvantes na antibioticoterapia contra infecções por *S. aureus*.

Neste contexto, propusemos uma campanha de docagem de compostos usando o modelo obtido nos estudos de dinâmica molecular para o estado fechado como o modelo inicial, já devidamente parametrizado. Para a realização dos experimentos de docagem de compostos, utilizamos a base *Zinc Drug Database* (ZDD), um catálogo de aproximadamente 3.000 compostos disponíveis comercialmente, os quais foram aprovados para o uso em humanos em

diversos países. O propósito para a escolha desta base de dados envolve a identificação de ligantes, mesmo fracos, que são conhecidamente não-tóxicos (ou pouco tóxicos) em seres humanos e que possuem um perfil farmacocinético tolerável.

A realização da docagem de compostos é uma estratégia que tem aumentado a taxa de sucesso dos testes para a descoberta candidatos a fármacos. Quando comparada com a técnica de amostragem *high-throughput*, a docagem de compostos mostra um ganho no número de moléculas interagentes de mais de 1000 vezes (0,021% contra 34,8%).⁷⁷ A realização de um ensaio inicial *in silico* descarta diversos candidatos que não possuem as características necessárias para se tornarem possíveis ligantes. Portanto, a docagem torna processo da busca por inibidores mais eficiente, tanto em termos de tempo quanto de orçamento, visto que não é necessário adquirir ou testar um catálogo inteiro de moléculas.

A biblioteca foi avaliada com o auxílio do programa LiBELA. Em seguida, os compostos docados foram analisados e selecionados de acordo com suas energias de interação calculadas, pelas interações semelhantes àquelas encontradas na ligação do substrato e pelo preenchimento do sítio ativo, com o auxílio do programa UCSF Chimera (Figura 34).⁶⁵ Os seguintes compostos foram selecionados para a avaliação experimental: metocarbamol, pantoprazol sódico, 5-azacitidina, pyridoxal 5'-fosfato, sulfametizol, suprofen, sitaxentano sódico, citarabina, olsalazina, ticarcilina, nilutamida, adenosina, timidina, tolbutamida, L-histidina, sulfadiazina e sulfametoxazol (Tabela 8).

Figura 34 – Interação prevista pelo programa de docking LiBELA para a molécula de ticarcilina e a proteína MnaA. (A) Pose obtida para o composto ticarcilina (em amarelo) e UDP (em azul) no sítio ativo; (B) Interações de hidrogênio (em azul) calculadas pelo programa de visualização Chimera entre a ticarcilina e a proteína.

Fonte: Elaborada pela autora.

Os compostos selecionados foram também analisados com o auxílio do servidor web PharmaGist, ⁷⁸ para a detecção de farmacóforos, que descrevem o arranjo espacial das características que são essenciais para a interação entre as moléculas de interesse e um alvo específico no receptor. A estrutura do UDP-GlcNac foi utilizada como molécula chave para a detecção. A partir dos resultados obtidos pelo servidor, é possível observar que há características comuns entre a maior parte das moléculas. O alinhamento de 16 moléculas mostrou três principais características espaciais, que é a presença de 1 doador e 2 aceptores de ligações de hidrogênio. Outro alinhamento, com 15 das moléculas fornecidas, possuem 1 grupamento aromático e 2 aceptores. Estas características, comuns à maior parte dos compostos selecionados e ao UDP-GlcNac, são importantes para as interações realizadas no sítio ativo.

A escolha de compostos que possam interagir com a proteína-alvo é apenas o primeiro passo na busca por inibidores. Após, são necessários testes *in vitro* e a análise mais detalhada da interação proteína-ligante. Neste trabalho, realizamos ensaios de DSF e qDSF de modo a obter informações qualitativas e quantitativas sobre a interação (ou não) das moléculas. Para isso, foi necessária também a produção da proteína MnaA recombinante.

Tabela 8 – Ligantes selecionados durante o processo de inspeção visual dos compostos após o processo de docagem.

Ligante	Estrutura	Função conhecida
Metocarbamol		Inibidor da enzima anidrase carbônica I de <i>H. sapiens</i> , comumente utilizada como relaxante muscular. ⁷⁹
Pantoprazol sódico		Inibidor da enzima ATPase transportadora de potássio de <i>H. sapiens</i> , comumente utilizada no tratamento de refluxo gastroesofágico, esofagite e síndrome de Zollinger-Ellison. ⁷⁹

(Continua)

(Continuação)

Ligante	Estrutura	Função conhecida
5-Azacitidina		Inibidor de DNA citosina-5 metiltransferase 1 e 3A, DNA e RNA, indicada pra o tratamento de diversas doenças, incluindo neoplasmas, leucemia e carcinoma ovariano. ⁷⁹
Piridoxal 5'-fosfato		Forma ativa da vitamina B6, que é cofactor de uma variedade de processos enzimáticos. ⁸⁰
Sulfametizol		Inibidor da enzima di-hidropteroato sintase bacteriana, utilizada no tratamento de osteomielite. ⁷⁹

(Continua)

(Continuação)

Ligante	Estrutura	Função conhecida
Suprofeno		Inibidor da enzima ciclo-oxigenase humana, comumente utilizada para o tratamento de doenças reumáticas. ⁷⁹
Sitaxentano sódico		Antagonista competitivo da molécula endotelina-1, que causa vasoconstrição pulmonar. ⁷⁹
Citarabina		Inibidor da enzima DNA polimerase, DNA e RNA, indicada para o tratamento de diversos linfomas, leucemia e sarcoma. ⁷⁹

(Continua)

(Continuação)

Ligante	Estrutura	Função conhecida
Olsalazine		Agonista da enzima receptor gama proliferadora ativada peroxissoma e inibidor das proteínas ciclo-oxigenase e lipo-oxigenase araquidonato, de <i>H. sapiens</i> , indicada para espondilose. ⁷⁹
CTicarcilina		Inibidor da proteína ligadora de penicilina bacteriana, indicada para o tratamento de infecções, sepse e osteomielite. ⁷⁹
Nilutamida		Antagonista do receptor de andrógeno, utilizado no tratamento de neoplasmas e neoplasma prostático. ⁷⁹

(Continua)

(Continuação)

Ligante	Estrutura	Função conhecida
Timidina		Nucleosídeo de DNA, descrita como medicamento antiviral. ⁸¹
Adenosina		Agonista da proteína receptora de adenosina, indicado para o tratamento de doenças coronária e síndrome de Wolff-Parkinson-White. ⁷⁹
Timidina		Nucleosídeo de DNA, descrita como medicamento antiviral. ⁸¹

(Continua)

(Continuação)

Ligante	Estrutura	Função conhecida
Tolbutamida		Bloqueador do receptor sulfoniluréia 1, utilizado no tratamento de neoplasmas, diabetes melitus, hepatite crônica, entre outros. ⁷⁹
L-histidina		Aminoácido essencial, precursor da histamina, que é um neurotransmissor. Pode ter ação imunomodulatória e atividade antioxidante. Pode ser indicado para o tratamento de artrite reumatoide. ⁸²
Sulfadiazina		Inibidor da proteína di-hidropteroato sintase bacteriana, indicado para o tratamento de infecções, toxoplasmose e osteomielite. ⁷⁹

(Continua)

(Continuação)

Ligante	Estrutura	Função conhecida
Sulfametoxazol		Inibidor da proteína di-hidropteroato sintase bacteriana, indicada para o tratamento de dermatite atópica, osteomielite, diabetes melitus, entre outros. ⁷⁹

Fonte: Elaborada pela autora.

4.4 Clonagem, expressão e purificação da enzima MnaA

A enzima *SaMnaA*, apesar de já possuir estrutura conhecida, apresenta diversas possibilidades para estudo com a utilização de técnicas experimentais, como sua cristalização sem ligantes ou com UDP-GlcNac em seu sítio ativo, ou mesmo para testes de atividade e inibição.

O gene que codifica para a proteína MnaA de *S. aureus* SA16 foi clonado em plasmídeo pET/TRX-1A/LIC. Após, foi transformado em células para a propagação do DNA (*E. coli* DH10B) e para a expressão de proteínas (*E. coli* BL21 (DE3) Rosetta). A presença do plasmídeo contendo o gene de interesse foi confirmada utilizando a técnica de reação de polimerase em cadeia (PCR). Foi realizado também o sequenciamento do DNA clonado. Após, foram realizados testes de expressão em duas temperaturas, 18 °C e 37 °C, e com concentrações do indutor IPTG de 0,1, 1 e 10 mM. As melhores condições para expressão destas proteínas foram aquelas em que se adicionou 0,1 mM de IPTG ao meio de cultivo celular, a uma temperatura para indução de 18 °C.

A proteína foi purificada por afinidade em uma coluna Ni-NTA, com concentração de imidazol para eluição de 500 mM. A MnaA possui 375 aminoácidos, com peso molecular calculado pelo programa ProtParam de aproximadamente 42 kDa e pI 6,09 (Figura 35).

Figura 35 - Expressão e purificação da proteína MnaA. Gel de poliacrilamida 15%, onde 1) Marcador de massa molecular; 2) Amostra antes da indução por 0,1 mM de IPTG; 3) Amostra após a indução por 0,1 mM IPTG; 4) Sobrenadante após lise por ultra-sonicação; 5) Pellet após lise por ultra-sonicação; 6) Amostra não ligada à coluna; 7 e 8) Lavagem da coluna de níquel com tampão de ligação; 9) Eluição com tampão adicionado de 500 mM de imidazol; 10) Proteína após a clivagem. Fonte: Elaborada pela autora.

A enzima não apresenta triptofanos em sua estrutura primária. Portanto, para a verificação da presença de estrutura secundária na amostra, indicativo de que a proteína se encontra enovelada em solução, foram realizadas medidas de espectroscopia de dicroísmo circular (Apêndices). O espectro mostra um vale em aproximadamente 220 nm. Esse perfil indica a presença de estruturas secundárias.⁸³

Uma análise o espectro de absorção utilizando o servidor BeStSel,⁸⁴ é possível decompor os dados nos tipos de estrutura secundária presentes na proteína medida. O programa “*single spectrum analysis and fold recognition*” realiza a decomposição da curva experimental pela combinação linear das curvas típicas de uma base de componentes fixada. O programa reconhece 2 tipos de hélices-alfa, *helix1* (alfa-hélice normal) e *helix2* (alfa-hélice com extremidade distorcidas), 3 tipos de folhas-beta, *anti1* (torcido à esquerda), *anti2* (relaxada), *anti3* (torcida à direita) e *para* (paralela), voltas (*turn*) e uma categoria denominada *others*, que inclui hélices 3_{10} e regiões irregulares. Uma análise da estrutura secundária da proteína cristalizada (5ENZ) também foi realizada utilizando o mesmo servidor, para comparação.

As quantidades de estruturas secundárias calculadas a partir do espectro de dicroísmo circular (experimental) e a partir do arquivo PDB da enzima *SaMnaA* mostram diferenças entre porcentagens de hélices-alfa e de folhas-beta, o que pode decorrer de uma predição exagerada da última, uma vez que o programa foi desenvolvido especialmente para a predição mais acurada deste tipo de estrutura secundária. Há uma diminuição de cerca de 11,2% na quantidade de hélices-alfa com relação à análise da estrutura cristalográfica, ao passo que a

quantidade de folhas-beta aumentou cerca de 17,9%. Os gráficos de distribuição se encontram na seção de Apêndices.

Apesar das diferenças encontradas, é possível concluir que a proteína possui estruturas secundárias em solução, após o processo de purificação. Estes indícios de enovelamento da enzima são importantes, pois estas amostras serão utilizadas durante os testes de triagem dos candidatos selecionados pela docagem. Não foi possível, contudo, no decorrer deste trabalho, realizar ou padronizar testes de cinética enzimática entre a proteína e seu substrato ou com os possíveis ligantes.

Foram realizados ainda testes de cristalização. As amostras utilizadas para os testes continham aproximadamente 35,2 mg/mL de proteína. Foram avaliados também 3 kits comerciais, onde na primeira gota foi adicionada somente a proteína, na segunda gota, proteína e UDP-GlcNac na proporção 1:2, e na terceira gota, proteína e ticarcilina, também na proporção 1:2. Uma análise visual das gotas testadas mostra que não houve cristalização.

4.5 Triagem experimental dos candidatos

A triagem experimental dos candidatos selecionados por meio da docagem molecular foi realizada utilizando ensaios de DSF. Inicialmente, a curva de desenovelamento foi caracterizada no estado não-ligado, bem como ligada ao UDP, ao UDP-GlcNc e à combinação de UDP+UDP-GlcNac. Como indicado na Figura 36, as temperaturas de desenovelamento para a MnaA não-ligada e para a enzima ligada somente à UDP apresentam diferenças da ordem de 3 °C, sendo o T_m medido de 51 e 54 °C, respectivamente. Além disso, as diferenças entre a temperatura para a proteína ligada somente ao substrato e ligada ao substrato e cofator variam somente 1 °C. A pequena variação causada pela adição de UDP à solução sugere que a enzima se liga fracamente ao cofator. Outra possibilidade seria a ligação de UDP bacteriano à proteína durante o processo de produção da enzima.

Figura 36 – Temperatura de desenovelamento para a MnaA ligada às moléculas selecionadas durante a docagem.

Fonte: Elaborada pela autora.

Os ensaios realizados indicam que mesmo a ligação do próprio substrato da enzima resulta em uma diferença pequena na temperatura de desnaturação, de cerca de 5 °C, quando comparado com a proteína ligada somente ao UDP. Como os próprios ligantes da proteína não geram uma grande diferença no T_m medido, buscamos moléculas que causem um aumento no T_m da proteína de, pelo menos, 3-4 °C, que é a variação encontrada para o UDP, que se liga fracamente à *SaMnaA*. Foram selecionados, portanto, os compostos metocarbamol, pantoprazol, sulfametizol, suprofeno, olsalazina ((E)-5,5'-(diazene-1,2-Diyl)bi), ticarcilina, sulfacetamida e sitaxetan, os quais apresentaram aumentos no T_m entre 3 e 4 °C.

A possível interação entre os ligantes e a enzima foi melhor investigada utilizando a técnica de qDSF. O propósito destes experimentos foi obter a constante de dissociação (K_d) da proteína para cada molécula selecionada. Esta constante fornece informações sobre a força da interação proteína-inibidor, que é um bom parâmetro para uma primeira análise da proteína, visto que esta não é uma enzima que realiza a interconversão entre UDP-GlcNac e UDP-ManNac de forma rápida (k_{cat} de aproximadamente $0,1 \text{ s}^{-1}$). Os compostos foram então testados nas concentrações descritas Tabela 9. As concentrações foram selecionadas de acordo com a solubilidade das moléculas a serem testadas. Apenas a ticarcilina apresentou comportamento em que o aumento da concentração da molécula resultou no aumento da estabilidade da proteína. Os demais compostos apresentaram comportamento que indica a ausência de ligação e/ou a desestabilização da proteína. Todas as medidas foram realizadas em duplicata.

Tabela 9 – Compostos e concentrações utilizados para a realização dos experimentos de DSF quantitativo.

Composto	Concentrações (μM)											
UDP	0	1	2	4	8	16	32	64	128	256	512	2000
UDP-GlcNac	0	1	2	4	8	16	32	64	128	256	512	2000
Metocarbamol	0	1	5	25	125	625	3125	15625	25000	-	-	-
Pantoprazol	0	1	5	25	125	625	3125	15625	25000	-	-	-
Sulfametizol	0	1	3	9	27	81	243	729	2187	6561	12500	-
Suprofeno	0	1	3	9	27	81	243	729	2187	6561	9600	-
Diazeno (Olsalazina)	0	1	2	4	8	16	32	64	128	256	512	1500
Ticarcilina	0	1	2	4	8	16	32	64	128	256	512	1300
Sulfacetamida	0	1	3	9	27	81	243	729	2187	6561	1050	-
Sitaxetano	0	1	5	25	125	625	3125	15625	25000	-	-	-

Fonte: Elaborada pela autora.

A análise das curvas de desenovelamento mostrou um aumento de T_m para o UDP-GlcNac, o qual acompanha o aumento da concentração de ligante na solução, como mostrado na Figura 37. De acordo com o ajuste dos dados, medidos em duplicata, quando $T_m=52\text{ }^{\circ}\text{C}$, um valor de K_u igual a 0,39 foi obtido, com um K_d de 31 μM . A uma temperatura de 54 $^{\circ}\text{C}$, foi observado um K_u de 1,10 e K_d de 32 μM . K_u é uma constante de equilíbrio entre as concentrações de proteína enovelada e desenovelada. Portanto se K_u é 1, há uma concentração igual entre as duas espécies em solução. O K_M para o substrato UDP-GlcNac, medido por Mann *et al.*,³¹ foi de $411 \pm 57\text{ }\mu\text{M}$.

Estas diferenças entre as constantes que medem a afinidade do substrato à enzima se devem ao fato de que os cálculos realizados para o método de qDSF assumem que há um equilíbrio entre as concentrações em solução, diferentemente de um modelo cinético, onde o substrato é consumido, diminuindo sua concentração e aumentando a concentração de produto. Neste caso, portanto, para o cálculo de K_u , assume-se que não há conversão do substrato.

Figura 37 – Dados de DSF quantitativo para a enzima UDP-GlcNac 2-epimerase ligada ao seu substrato. (A) Variação da temperatura em função do aumento da concentração de ligante. (B) Fluorescência medida para cada concentração de ligante. (C) Curvas isotérmicas no intervalo de temperaturas de 50 a 64 °C, onde a fração de proteína desnovelada está descrita em função da concentração de ligante, em μM . (D) Sobreposição das curvas isotérmicas para todo o intervalo de temperaturas definido anteriormente.

Fonte: Elaborada pela autora.

O gráfico de $\ln K$ em função de $1/T$ (Figura 38) não mostra uma dependência linear típica esperada para a relação de van't Hoff, indicando que a entalpia pode não ser independente da temperatura. Isto pode ser devido ao fato de que a atividade da enzima é modificada no intervalo de temperatura amostrado, resultando nesta relação não-linear.

Figura 38 – Relação de van't Hoff com relação ao K_d calculado para os ligantes UDP-GlcNac e ticarcilina.

Fonte: Elaborada pela autora.

Após a realização dos testes em que se variou a concentração dos ligantes, apenas em que foi medida a ligação entre a proteína e a ticarcilina apresentou aumento na temperatura de desenovelamento com o aumento da concentração do ligante (Figura 39). Os demais compostos mostrar uma diminuição na temperatura, que pode ser devida a algum efeito de desestabilização causado pela adição de grandes quantidades do ligante Além disso, diversos compostos precisaram ser dissolvidos em etanol (sulfametoxazol, metocarbamol, nilutamida), em DMSO (tolbutamida, 5'-azacitadina, suprofeno, adenosina), em HCl 1 M (piridoxal 5'-fosfato) ou em NaOH 1M (sulfadiazina), que em altas concentrações possuem efeito desnaturante.

Figura 39 – Variação de T_m em função do aumento da concentração para os compostos testados.

Fonte: Elaborada pela autora.

Durante os ensaios realizados com a adição da ticarcilina, foi observado também um aumento na temperatura de desenovelamento da proteína. Os dados analisados são mais ruidosos do que os anteriores, provavelmente devido ao fato de este ser um ligante fraco e não haver uma mudança tão expressiva de T_m com o aumento da concentração quanto com a adição do substrato. Quando $T_m=52$ °C, um K_u de 0,41 e um K_d de 564 μ M foram observados (Figura 40). Quando a temperatura medida foi de 54 °C, o K_u observado foi de 1,10 e o K_d , 452 μ M. Para a ticarcilina, ao contrário do que foi observado para UDP-GlcNac, uma relação linear foi observada para $\ln K$ versus $1/T$, o que pode ser devido ao pequeno número de pontos calculados (Figura 38).

Os dados obtidos a partir das medidas de qDSF indicam que a ticarcilina se liga fracamente à enzima, com um K_d no intervalo de 300-600 μ M, em nossas condições experimentais, o que é cerca dez vezes maior do que o K_d medido para o substrato, que varia de 10 a 30 μ M. O método de qDSF não é capaz de medir a constante de afinidade em uma situação de equilíbrio, porém dados anteriores mostram que as afinidades de ligação são próximas daquelas determinadas por outros experimentos, como a calorimetria de titulação isotérmica (ITC).⁶⁹

É esperado que a ticarcilina alcance concentrações séricas de até 124 μ g/mL após injeção intravenosa de 1 g,⁸⁵ o que significa uma concentração sérica acima de 300 μ M. Portanto, a essas concentrações, é esperado que a *SaMnaA* se encontre preferencialmente ligada ao composto. Considerando-se os valores de K_d calculados e as concentrações séricas do fármaco, parece viável que esta proteína esteja relacionada com parte do efeito bacteriostático do beta-lactâmico. Ainda, a interação entre a enzima e a ticarcilina pode fornecer explicações para a boa atividade encontrada em bactérias Gram-negativas, como a *Pseudomonas aeruginosa*. Utilizando a ferramenta de busca e alinhamento de sequências BLAST,⁸⁶ é possível verificar que a proteína de *S. aureus* e de *P. aeruginosa* compartilham cerca de 51% de identidade em sua estrutura primária. O mesmo vale para as espécies do gênero *Acinetobacter*. Uma hipótese para o efeito pronunciado deste fármaco em bactérias Gram-negativas é que o antibiótico se ligue à proteína UDP-GlcNac 2-epimerase, possivelmente inibindo-a. Outro ponto importante é que a atividade em bactérias como *P. aeruginosa* é somente alcançada em altas concentrações da molécula,⁸⁵ o que é coerente com os dados observados, pois sugere uma baixa afinidade com o alvo.

Figura 40 - Dados de DSF quantitativos para a enzima UDP-GlcNac 2-epimerase ligada à ticarcilina. (A) Variação da temperatura em função do aumento da concentração de ligante. (B) Fluorescência medida para cada concentração de ligante. (C) Curvas isotérmicas no intervalo de temperaturas de 50 a 64 °C, onde a fração de proteína desenovelada está descrita em função da concentração de ligante, em μM . (D) Sobreposição das curvas isotérmicas para todo o intervalo de temperaturas definido anteriormente.

Fonte: Elaborada pela autora.

4.5.1 Conclusões parciais

Os ensaios de DSF e DSF quantitativo realizados com a proteína *SaMnaA* heteróloga produzida indicam uma interação fraca com o antibiótico beta-lactâmico ticarcilina. Este antibiótico possui ação em um amplo espectro de bactérias Gram-Positivas e Gram-negativas. Sua interação com a enzima estudada e com enzimas similares pode indicar um mecanismo de ação antibiótica ainda não conhecida, gerando possibilidades para novos usos do fármaco. Para testar sua influência não somente em *S. aureus*, mas em outros organismos que possuam composição da parede celular diferente desta, está sendo iniciada uma parceria com o Instituto de Ciências Biológicas, ICB-USP. O objetivo desta parceria é testar o efeito da ticarcilina no crescimento celular de *Mycoplasma sp.*, que é uma bactéria que não possui parede celular, e que não é suscetível ao efeito dos antibióticos beta-lactâmicos, mas que possui uma enzima com cerca de 90% de cobertura e 60% de identidade com a MnaA de *S. aureus*.

4.6 Outras enzimas da via de biossíntese de WTA

Foram selecionadas diversas proteínas envolvidas na biossíntese dos ácidos teicóicos de parede para expressão heteróloga e posterior estudo de suas características biofísicas e estruturais. A WTA e as enzimas que participam de sua síntese representam alvos interessantes para o desenvolvimento de novos fármacos, pois a depleção da parede mostrou re-sensibilizar *S. aureus* resistentes a antibióticos beta-lactâmicos.

Os genes que codificam para as proteínas TarA, TarB, TarS, TarL, TarO, TarI e TarJ de *S. aureus* SA16 foram clonadas em plasmídeo pET/TRX-1A/LIC e transformados em células para a propagação do DNA (*E. coli* DH10B) e para a expressão de proteínas (*E. coli* BL21 (DE3) Rosetta). Porém, apenas os plasmídeos contendo os genes para as enzimas TarA e TarS foram clonados com sucesso. A presença dos plasmídeos contendo os genes de interesse foi confirmada utilizando a técnica de reação de polimerase em cadeia (PCR). Foi realizado também o sequenciamento do DNA clonado. Após, foram realizados testes de expressão em duas temperaturas, 18 °C e 37 °C, e com concentrações do indutor IPTG de 0,1, 1 e 10 mM.

As melhores condições para expressão destas proteínas foram aquelas em que se adicionou 0,1 mM de IPTG ao meio de cultivo celular, a uma temperatura para indução de 18 °C.

Foram realizados testes de solubilidade sem aditivos e com a adição dos detergentes CHAPS, Sarkosyl e Triton X-100, comumente utilizados na solubilização de proteínas (Figura 41).⁸⁷ O detergente zwitteriônico 3-[(3-colamidopropil)dimetilam-monio]-1-propanesulfonato (Chaps) é um detergente popular para a solubilização de proteínas de membrana por diversos motivos, dentre os quais estão suas propriedades não-desnaturantes, sua solubilidade em grande intervalo de pH e o tamanho das micelas que forma, o que torna possível a diálise para a retirada do composto da solução. Além disso, este inibe interações proteína-proteína.⁸⁸ O detergente sarkosyl (N-laurilsacosina) possui propriedades semelhantes, interagindo com regiões aniônicas e hidrofóbicas, prevenindo assim a coagregação de proteínas em solução.⁸⁹ O detergente não-iônico Triton X-100 também tem sido utilizado em pequenas concentrações para a purificação de proteínas de membrana, que podem ser insolúveis na ausência de aditivos, sem a necessidade de procedimentos de renovelamento.⁹⁰

Figura 41 - Teste de solubilidade da proteína TarA. Gel de poliácridamida 15%, no qual 1) Marcador de massa molecular; 2) Amostra antes da indução por 0,1 mM de IPTG; 3) Amostra após a indução por 0,1 mM de IPTG; 4) Pellet após lise sem detergentes; 5) Sobrenadante após lise sem detergentes; 6) Pellet após lise utilizando o detergente CHAPS 20 mM; 7) Sobrenadante após lise utilizando o detergente CHAPS 20 mM; 8) Pellet após lise utilizando o detergente Sarkosyl 0,1%; 9) Sobrenadante após lise utilizando o detergente Sarkosyl 0,1%; 10) Marcador de massa molecular; 11) Pellet após lise utilizando o detergente Triton X-100 1%; 12) Sobrenadante após lise utilizando o detergente Triton X-100 1%. A seta indica a proteína de interesse.

Fonte: Elaborada pela autora.

Apenas a proteína TarA foi solubilizada após a adição de 20 mM do detergente CHAPS. Esta foi então purificada por afinidade, com concentrações de imidazol para eluição de 500 mM (Figura 42). Não foi possível realizar a clivagem da cauda de histidinas da proteína TarA. Esta ausência de clivagem pode ser explicada pela presença do detergente, o qual, em alguns casos, pode impedir a ação da protease ou diminuir sua eficiência.⁸⁷ A enzima TarA de *S.*

aureus SA16 possui 231 aminoácidos, tendo seu peso molecular calculado pelo programa ProtParam⁹¹ de aproximadamente 26 kDa e pI teórico de 9,55 (Figura 43).

Figura 42 – Purificação por cromatografia de afinidade da proteína TarA. Dados obtidos com o auxílio do equipamento ÄKTA Purifier (GE Healthcare Life Sciences), o qual mostra a absorção em 280 nm da amostra em função do volume de eluição.

Fonte: Elaborada pela autora.

Figura 43 - Purificação da proteína TarA por cromatografia de afinidade. Gel de poliacrilamida 15%, onde 1) Proteína TarA purificada; 2) Amostra após a indução por 0,1 mM de IPTG; 3) Sobrenadante após lise com tampão de lise e detergente CHAPS 40 mM; 4) Pellet após lise com tampão de lise e detergente CHAPS 40 mM; 5 e 6) Amostra não ligada à coluna; 7) Lavagem da coluna de níquel com tampão de ligação; 8) a 13) Primeiro pico de absorção observado durante o gradiente de imidazol; 14) a 22) Segundo pico de absorção observado durante o gradiente de imidazol.

Fonte: Elaborada pela autora

Devido à necessidade de adição do detergente para a purificação da proteína TarA, foi necessária a busca por indicativos da integridade estrutural da proteína purificada. Para tal, foram realizados testes de fluorescência intrínseca da enzima (Apêndices). É possível observar que a amostra desnaturada por aquecimento possui uma mudança tanto na intensidade da fluorescência como no comprimento de onda em que ocorre. Isto se deve ao fato de que houve uma mudança no ambiente químico do aminoácido apolar e fluorescente triptofano, sugerindo que houve uma exposição do resíduo ao solvente polar. Este é um indicativo, porém não uma garantia, de que a proteína se encontra enovelada em solução.

Além disso, a fluorescência pode ser devida à presença da proteína de carreadora tiorredoxina, que possui 2 triptofanos em sua estrutura primária.

Como não foi possível obter a enzima TarA de *S. aureus* solúvel sem a adição de detergentes e visto que a partir dos experimentos realizados não é possível afirmar que a mesma se encontra enovelada em solução, não foram feitos mais testes com a mesma. Além disso, em 2019, Kattke *et al.*⁹² publicaram a resolução da estrutura cristalográfica da proteína TagA de *Thermoanaerobacter italicus*, a qual possui cerca de 37% de identidade com a enzima de *S. aureus*, incluindo o domínio WecB/TagA/CpsF. Kattke *et al.*⁹² produziram a enzima com uma deleção de 49 aminoácidos na região C-terminal. A autora indica que os resíduos deletados são responsáveis por ligar a proteína à membrana da célula. A construção para a produção da proteína completa gerou uma amostra pouco solúvel, sendo necessária a adição de grandes quantidades de sal e glicerol. Isto pode explicar as dificuldades encontradas na solubilização da proteína de TarA, que desempenha a mesma função em *S. aureus*.

Foram realizados também testes de cristalização da enzima TarA purificada e concentrada a aproximadamente 0,5 mg/mL. Não foi possível obter uma maior concentração da proteína, sendo observada agregação da amostra, o que concorda com a necessidade de detergentes para sua solubilização. A agregação deve ser proveniente de interações entre proteínas em solução, que são evitadas na presença de aditivos como o CHAPS. O fato de haver agregação entre as moléculas (zona de precipitação) foi motivador para a realização de testes de cristalização, mesmo em baixas concentrações, pois a variação entre as concentrações de proteína e agente precipitante podem resultar na nucleação e/ou formação de cristais (zonas de nucleação e metaestável) (Figura 44).

Figura 44 - Diagrama de fases para a cristalização de proteínas. Baseado na variação da concentração de proteína e de precipitante.

Fonte: CHAYEN.⁹³

Foram avaliados três kits comerciais, contendo 96 diferentes condições de tampões e agentes precipitantes, e três gotas de diferentes proporções entre amostra e tampão (1:2, 1:1 e 2:1) para cada condição, resultando em 864 diferentes gotas. Dentre as condições utilizadas, as que continham citrato de sódio 1,4 M e tampão HEPES 0,1 M pH 7,5 do kit Crystal Screen HT (Hampton Research) geraram pequenos cristais de forma arredondada (esferulitos), como mostra a Figura 45. Estas formações ocorrem normalmente em baixas concentrações de proteína, podendo ser utilizadas como “semente” para a formação de cristais maiores.⁹⁴ Os cristais obtidos na terceira gota (proporção amostra:tampão 2:1) foram submetidos à difração de raios-X, onde foi possível observar poucas reflexões em baixa resolução, insuficientes para a resolução de estrutura (Figura 46).

Figura 45 - Gota na condição de cristalização com tampão HEPES 0,1 M pH 7,5, citrato de sódio 1,4 M, proporção de amostra:tampão 2:1.
Fonte: Elaborada pela autora.

Figura 46 – (A) Cristal obtido na condição 1 M citrato de sódio e 0,1 M HEPES pH 7,8, proporção de amostra:tampão 2:1. (B) Difração de raios-X dos cristais obtidos na condição tampão HEPES 0,1 M pH 7,5, citrato de sódio 1,4 M, onde o círculo cinza marca a resolução de 7,7 Å e o círculo vermelho marca pontos de difração observados.

Fonte: Elaborada pela autora.

5 Conclusões

Este trabalho teve como objetivos principais investigar as características estruturais das proteínas envolvidas na biossíntese dos ácidos teicóicos de parede de *S. aureus*, em especial a enzima UDP-GlcNac 2-epimerase. A estrutura da proteína SaMnaA e o movimento realizado entre seus domínios são essenciais para a atividade de interconversão entre seu substrato (UDP-GlcNac) em produto (UDP-ManNac). O movimento de abertura/fechamento realizado pela proteína se mostrou essencial para a interação e manutenção do substrato em seu sítio ativo. Além disso, com base nas informações encontradas, foi realizada também uma busca por moléculas que interagem com a enzima, visto que a enzima apresenta grande importância para a resistência a antibióticos beta-lactâmicos nesta bactéria. Para esta finalidade, foram utilizados métodos computacionais e experimentais.

Como métodos computacionais, foram realizadas extensivas simulações de dinâmica molecular, em condições de equilíbrio termodinâmico ou aceleradas por um potencial externo. Estas simulações indicam que as cargas presentes no sítio ativo da proteína possuem papel fundamental em sua mudança conformacional do estado “aberto”, onde há repulsão de cargas causada por aminoácidos carregados positivamente, para o estado “fechado”, onde as cargas negativas do substrato neutralizam estas cargas positivas. A força eletrostática que age sobre os domínios N- e C-terminal da enzima SaMnaA devem, portanto, ser consideradas no processo de triagem de inibidores. A entrada do substrato no sítio ativo também modifica o perfil energético da proteína, indicando que há um equilíbrio entre populações de estados na proteína não-ligada, onde há uma preferência (mínimo energético) bem definida pelo estado “aberto”. A ligação do UDP-GlcNac, contudo, gera uma mudança do mínimo de energia dinâmico da proteína para o estado “fechado”. Estas mudanças no perfil energético geram hipóteses sobre como ocorre a ligação do substrato e a posterior liberação do produto, após a catálise.

Os dados coletados a partir das simulações de dinâmica molecular forneceram informações importantes para a busca de possíveis inibidores, realizada pelo processo de docagem molecular. Foram selecionadas 17 possíveis ligantes, dentre os quais 8 foram testados *in vitro*. Apenas a ticarcilina apresentou indícios de interação com a proteína, sendo necessária ainda a realização de novos experimentos para verificar se há inibição da atividade proteica. Além disso, a interação ticarcilina-SaMnaA gera novos questionamentos, como a atividade inesperada do beta-lactâmico em bactérias Gram-negativas, como *Pseudomonas aeruginosa*.

que possuem enzimas similares àquela de *S. aureus*. Testes de susceptibilidade destas bactérias a antibióticos beta-lactâmicos após o cultivo na presença de ticarcilina podem indicar a interação com a proteína UDP-GlcNac 2-epimerase de *P. aeruginosa*, trazendo também questionamentos sobre a função da enzima em bactérias que não produzem WTA.

REFERÊNCIAS

- 1 BROWN, E. D.; WRIGHT, G. D. Antibacterial drug discovery in the resistance era. **Nature**, v. 529, n. 7586, p. 336–343, 2016.
- 2 BLAIR, J. M. A. *et al.* Molecular mechanisms of antibiotic resistance. **Nature Reviews Microbiology**, v. 13, n. 1, p. 42–51, 2015.
- 3 BENNETT, P. M. Plasmid encoded antibiotic resistance: Acquisition and transfer of antibiotic resistance genes in bacteria. **British Journal of Pharmacology**, v. 153, n. SUPPL. 1, p. 347–357, 2008.
- 4 VAN HOEK, A. H. A. M. *et al.* Acquired antibiotic resistance genes: an overview. **Frontiers in Microbiology**, v. 2, p. 203, 2011. doi: 10.3389/fmicb.2011.00203.
- 5 BALABAN, N. Q. *et al.* Definitions and guidelines for research on antibiotic persistence. **Nature Reviews Microbiology**, v. 17, n. 7, p. 441–448, 2019.
- 6 MADIGAN, M. T. *et al.* **Microbiologia de Brock**. 12. ed. Porto Alegre: Artmed Editora S.A., 2010.
- 7 MCCLURE, N. S.; DAY, T. A theoretical examination of the relative importance of evolution management and drug development for managing resistance. **Proceedings of the Royal Society of London B: biological sciences**, v. 281, n. 1797, 2014. doi: 10.1098/rspb.2014.1861 .
- 8 CENTRE FOR DISEASE CONTROL AND PREVENTION. **Drug resistance**. Disponível em: <https://www.cdc.gov/drugresistance/pdf/5-2013-508.pdf>. Acesso em: 28 fev. 2020.
- 9 WORLD HEALTH ORGANIZATION. **Infographics: antibiotic resistance**. Disponível em: <https://www.who.int/mediacentre/events/2015/world-antibiotic-awareness-week/infographics/en/>. Acesso em: 28 fev. 2020.
- 10 BOUCHER, H. W. *et al.* Bad Bugs, No Drugs: No ESKAPE! an update from the Infectious Diseases Society of America. **Clinical Infectious Diseases**, v. 48, n. 1, p. 1–12, 2009.
- 11 CENTRE FOR DISEASE CONTROL AND PREVENTION. **Antibiotic development is dwindling**. Disponível em: <https://www.cdc.gov/antibiotic-use/community/materials-references/graphics.html>. Acesso em: 28 fev. 2020.
- 12 ETEBU, E.; ARIKEKPAR, I. Antibiotics: classification and mechanisms of action with emphasis on molecular perspectives. **International Journal of Applied Microbiology and Biotechnology Research**, v. 4, p. 90–101, 2016.
- 13 WANG, H. *et al.* Discovery of wall teichoic acid inhibitors as potential anti-MRSA beta-lactam combination agents. **Chemistry & Biology**, v. 20, n. 2, p. 272–284, 2013.

- 14 MCMILLAN, K. *et al.* Characterization of *Staphylococcus aureus* isolates from raw milk sources in Victoria, Australia. **BMC Microbiology**, v. 16, n. 1, p. 169, 2016.
- 15 PEACOCK, S. J.; PATERSON, G. K. Mechanisms of Methicillin resistance in *Staphylococcus aureus*. **Annual Review of Biochemistry**, v. 84, p. 577–601, 2015.
- 16 WEINER, L.M. *et al.* Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2011-2014. **Infection Control and Hospital Epidemiology**, v. 37, n. 11, p. 1288–1301, 2016.
- 17 WILLIAMS, J. D. F. β -Lactamases and β -lactamase inhibitors. **International Journal of Antimicrobial Agents**, v. 12, n. Supplement 1, p. S3–S7, 1999.
- 18 CENTRE FOR DISEASE CONTROL AND PREVENTION. **How antibiotic resistance happens**. Disponível em: <https://www.cdc.gov/drugresistance/about/how-resistance-happens.html>. Acesso em: 28 fev. 2020.
- 19 ROEMER, T.; SCHNEIDER, T.; PINHO, M. G. Auxiliary factors: a chink in the armor of MRSA resistance to β -lactam antibiotics. **Current Opinion in Microbiology**, v. 16, n. 5, p. 538–548, 2013.
- 20 DA COSTA, T. *et al.* PBP4: a new perspective on *Staphylococcus aureus* β -Lactam resistance. **Microorganisms**, v. 6, n. 3, p. 57, 2018.
- 21 DELEO, F. R.; DIEP, B. A.; OTTO, M. Host defense and pathogenesis in *Staphylococcus aureus* Infections. **Infectious Disease Clinics of North America**, v. 23, n. 1, p. 17–34, 2009.
- 22 CAVALCANTI, S. M. de M. *et al.* Estudo comparativo da prevalência de *Staphylococcus aureus* importado para as unidades de terapia intensiva de hospital universitário, Pernambuco, Brasil. **Revista Brasileira de Epidemiologia**, v. 9, n. 4, p. 436–446, 2006.
- 23 DAVIES, O. L.; BENNETT, S. **WHO publishes list of bacteria for which new antibiotics are urgently needed**. Disponível em: <https://www.who.int/news-room/detail/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed>. Acesso em: 28 fev. 2020.
- 24 DE KRAKER, M. E. A. *et al.* Mortality and hospital stay associated with resistant *Staphylococcus aureus* and *Escherichia coli* bacteremia: estimating the burden of antibiotic resistance in Europe. **PLoS medicine**, v. 8, n. 10, p. e1001104, 2011.
- 25 THEURETZBACHER, U. *et al.* The global preclinical antibacterial pipeline. **Nature Reviews Microbiology**, p. 1–11, 2019. doi: 10.1038/s41579-0288-0.1.
- 26 SOBHANIFAR, S. *et al.* Structure and mechanism of *Staphylococcus aureus* tars, the wall teichoic acid β -glycosyltransferase involved in methicillin resistance. **PLOS Pathogens**, v. 12, n. 12, p. e1006067, 2016.

- 27 FORMSTONE, A. *et al.* Localization and interactions of teichoic acid synthetic enzymes in *Bacillus subtilis*. **Journal of Bacteriology**, v. 190, n. 5, p. 1812–21, 2008.
- 28 SOBHANIFAR, S. *et al.* Structure and mechanism of *Staphylococcus aureus* TarM, the wall teichoic acid α -glycosyltransferase. **Proceedings of the National Academy of Sciences of the United States of America**, v.112,n.6,p. E576–E585, 2015.
- 29 SEWELL, E. W. C.; BROWN, E. D. Taking aim at wall teichoic acid synthesis: new biology and new leads for antibiotics. **Journal of Antibiotics**, v. 67, n. 1, p. 43–51, 2014.
- 30 D'ELIA, M. A. *et al.* The N-acetylmannosamine transferase catalyzes the first committed step of teichoic acid assembly in *Bacillus subtilis* and *Staphylococcus aureus*. **Journal of Bacteriology**, v. 191, n. 12, p. 4030–4034, 2009.
- 31 MANN, P. A. *et al.* Chemical genetic analysis and functional characterization of staphylococcal wall teichoic acid 2-epimerases reveals unconventional antibiotic drug targets. **PLOS Pathogens**, v. 12, n. 5, p. e1005585, 2016.
- 32 VELLOSO, L. M. *et al.* A structural basis for the allosteric regulation of non-hydrolysing UDP-GlcNAc 2-epimerases. **EMBO Reports**, v. 9, n. 2, p. 199–205, 2008.
- 33 KAWAMURA, T.; ISHIMOTO, N.; ITO, E. Enzymatic synthesis of uridine diphosphate glucose. **Journal of Biological Chemistry**, v. 254, n. 17, p. 8457–8465, 1979.
- 34 MORGAN, P. M. ;SALA, R. F.; TANNER*, M. E. Eliminations in the reactions catalyzed by udp-n-acetylglucosamine 2-epimerase.**Journal American Chemical Society**, v.119, n.43,p.10269, 1997.
- 35 SAMUEL, J.; TANNER, M. E. Active site mutants of the “non-hydrolyzing” UDP-N-acetylglucosamine 2-epimerase from *Escherichia coli*. **Biochimica et Biophysica Acta - proteins and proteomics**, v. 1700, n. 1, p. 85–91, 2004.
- 36 CHEN, S.C. *et al.* Crystal structures of the archaeal UDP-GlcNAc 2-epimerase from *Methanocaldococcus jannaschii* reveal a conformational change induced by UDP-GlcNAc. **Proteins**, v. 82, n. 7, p. 1519–1526, 2014.
- 37 MAIER, J. A. *et al.* ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. **Journal of Chemical Theory and Computation**, v. 11, n. 8, p. 3696–3713, 2015.
- 38 ANANDAKRISHNAN, R.; AGUILAR, B.; ONUFRIEV, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. **Nucleic Acids Research**, v. 40, n. W1, p. W537–W541, 2012.
- 39 PRICE, D. J.; BROOKS, C. L. A modified TIP3P water potential for simulation with Ewald summation. **Journal of Chemical Physics**, v. 121, n. 20, p. 10096–10103, 2004.
- 40 RASSOLOV, V. A. *et al.* 6-31G* basis set for third-row atoms. **Journal of Computational Chemistry**, v. 22, n. 9, p. 976–984, 2001.

- 41 FRISCH, M. J. *et al.* **Gaussian 09**. s.l.: Wallingford CT Gaussian, Inc., 2009.
- 42 COMELL, W. D. *et al.* Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. **Journal of American Chemistry Society**, v. 115, n.21, p. 9620–9631, 1993.
- 43 WANG, J. *et al.* Antechamber: an accessory software package for molecular mechanical calculations. **Journal of Chemical Information and Computer Sciences**, p. 1–41, 2001.
- 44 PONDER, J. W.; CASE, D. A. Force fields for protein simulations. **Advances in Protein Chemistry**, v. 66, p. 27–85, 2003. DOI:10.1016/s0065-3233(03)66002-x.
- 45 WANG, J. *et al.* Development and testing of a general amber force field. **Journal of Computational Chemistry**, v. 25, n. 9, p. 1157–1174, 2004.
- 46 WEBB, B.; SALI, A. Comparative protein structure modeling using MODELLER. **Current Protocols in Bioinformatics**, v. 54, n. 1, p. 5.6.1-5.6.37, 2016.
- 47 SALOMON-FERRER, R.; CASE, D. A.; WALKER, R. C. An overview of the Amber biomolecular simulation package. **WIREs Computational Molecular Science**, v. 3, n.2, p. 198–210, 2013.
- 48 CASE, D. A. *et al.* The Amber biomolecular simulation programs. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1668–1688, 2005.
- 49 MACHADO, M. R. *et al.* The SIRAH 2.0 force field: Altius, Fortius, Citius. **Journal of Chemical Theory and Computation**, v. 15, n. 4, p. 2719–2733, 2019.
- 50 MACHADO, M. R.; PANTANO, S. SIRAH tools: mapping, backmapping and visualization of coarse-grained models. **Bioinformatics**, v. 32, n. 10, p. 1568–1570, 2016.
- 51 MIAO, Y.; FEHER, V. A.; MCCAMMON, J. A. Gaussian accelerated molecular dynamics: unconstrained enhanced sampling and free energy calculation. **Journal of Chemical Theory and Computation**, v. 11, n. 8, p. 3584–3595, 2015.
- 52 MIAO, Y. *et al.* Improved reweighting of accelerated molecular dynamics simulations for free energy calculation. **Journal of Chemical Theory and Computation**, v. 10, p.n.7, 2677–2689, 2014.
- 53 BABIN, V. *et al.* Adaptively biased molecular dynamics: an umbrella sampling method with a time-dependent potential. **International Journal of Quantum Chemistry**, v. 109, n. 15, p. 3666–3678, 2009.
- 54 BARDUCCI, A.; BUSSI, G.; PARRINELLO, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. **Physical Review Letters**, v. 100, n. 2, p. 1–4, 2008.
- 55 SCHERER, M. K. *et al.* PyEMMA 2: a software package for estimation, validation, and analysis of Markov Models. **Journal of Chemical Theory and Computation**, v. 11, n.11, p. 5525–55242, 2015.
- 56 WEHMEYER, C. *et al.* Introduction to Markov state modeling with the PyEMMA

software [Article v1.0]. **Living Journal Computational Molecular Science**, v.1, n.1, p.5165,2018.

57 LI, L. *et al.* DelPhiForce web server: electrostatic forces and energy calculations and visualization. **Bioinformatics**, v. 33, n. 22, p. 3661–3663, 2017.

58 ROE, D. R.; CHEATHAM, T. E. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. **Journal of Chemical Theory and Computation**, v. 9, n. 7, p. 3084–3095, 2013.

59 NASCIMENTO, A. S. **AmberEnergy++**. Disponível em: <https://github.com/alessandronascimento/amberenergy>. Acesso em: 13 dez. 2018.

60 MUNIZ, H. S.; NASCIMENTO, A. S. Ligand- and receptor-based docking with LiBELa. **Journal of Computer-Aided Molecular Design**, v. 29, n. 8, p. 713–723, 2015.

61 MUNIZ, H. S.; NASCIMENTO, A. S. Towards a critical evaluation of an empirical and volume-based solvation function for ligand docking. **PloS one**, v. 12, n. 3, p. e0174336, 2017.

62 IRWIN, J. J.; SHOICHET, B. K. ZINC - a free database of commercially available compounds for virtual screening. **Journal of Chemical Information and Modeling**, v. 45, n. 1, p. 177–82, 2005.

63 IRWIN, J. J. *et al.* ZINC: a free tool to discover chemistry for biology. **Journal of Chemical Information and Modeling**, v. 52, n. 7, p. 1757–1768, 2012.

64 STERLING, T.; IRWIN, J. J. ZINC 15 – ligand discovery for everyone. **Journal of Chemical Information and Modeling**, v. 55, n. 11, p. 2324–2337, 2015.

65 PETTERSEN, E. F. *et al.* UCSF Chimera: a visualization system for exploratory research and analysis. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1605–1612, 2004.

66 DABUL, A. N. G. *et al.* Draft genome sequence of methicillin-resistant staphylococcus aureus strain sa16, representative of an endemic clone from a brazilian hospital. **Genome A**, v. 1, n. 5, p. e00754-13, 2013.

67 CAMILO, C. M.; POLIKARPOV, I. High-throughput cloning, expression and purification of glycoside hydrolases using Ligation-Independent Cloning (LIC). **Protein Expression and Purification**, v. 99, p. 35–42, 2014. doi: 10.1016/j.pep.2014.03.008..

68 SOLDO, B. *et al.* Characterization of a Bacillus subtilis thermosensitive teichoic acid-deficient mutant: Gene *mnaA* (*yvyH*) Encodes the UDP-N-Acetylglucosamine 2-Epimerase. **Journal of Bacteriology**, v. 184, n. 15, p. 4316–4320, 2002.

69 BAI, N. *et al.* Isothermal analysis of thermofluor data can readily provide quantitative binding affinities. **Scientific Reports**, v. 9, n. 1, p. 2650, 2019.

70 HILDEBRAND, P. W.; ROSE, A. S.; TIEMANN, J. K. S. Bringing molecular dynamics simulation data into view. **Trends in Biochemical Sciences**, v. 44, n. 11, p. 902–913, 2019.

- 71 HOLLINGSWORTH, S. A.; DROR, R. O. Molecular dynamics simulation for all. **Neuron**, v. 99, n. 6, p. 1129–1143, 2018.
- 72 SCHWANTES, C. R.; SHUKLA, D.; PANDE, V. S. Markov state models and tICA reveal a nonnative folding nucleus in simulations of NuG2. **Biophysical Journal**, v. 110, n. 8, p. 1716–1719, 2016.
- 73 HUSIC, B. E.; PANDE, V. S. Markov state models: from an art to a science. **Journal of American Chemical Society**, v.140, n.7, p.2386-2398,2018.
- 74 LASKOWSKI, R. A.; SWINDELLS, M. B. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. **Journal of Chemical Information and Modeling**, v. 51, n. 10, p. 2778–2786, 2011.
- 75 PANDAY, S. K. *et al.* Modeling electrostatics in molecular biology: a tutorial of DelPhi and associated resources [Article v1.0]. **Living Journal of Computational Molecular Science**, v. 1, n. 2, p. 1–24, 2019.
- 76 KUMAR, S. *et al.* Folding and binding cascades: dynamic landscapes and population shifts. **Protein Science**, v. 9, n. 1, p. 10–19, 2000.
- 77 SHOICHET, B. K. Virtual screening of chemical libraries. **Nature**, v. 432, n. 7019, p. 862–5, 16 dez. 2004.
- 78 SCHNEIDMAN-DUHOVNY, D. *et al.* PharmaGist: a webserver for ligand-based pharmacophore detection. **Nucleic Acids Research**, v. 36, n. Web Server issue, p. W223, 2008.
- 79 EMBL-EBI. **ChEMBL**. Disponível em: <https://www.ebi.ac.uk/chembl/>. Acesso em: 3 mar. 2020.
- 80 WHO Collaborating Centre for Drug Statistics Methodology. Disponível em: https://www.whocc.no/atc_ddd_index/?code=A11HA06&showdescription=yes. Acesso em: 3 mar. 2020.
- 81 PUBCHEM DATABASE. **Thymidine**. Disponível em: <https://pubchem.ncbi.nlm.nih.gov/compound/Thymidine#section=Drug-and-Medication-Information>>. Acesso em: 3 mar. 2020.
- 82 PUBCHEM DATABASE. **Histidine**. Disponível em: <https://pubchem.ncbi.nlm.nih.gov/compound/6274#section=Drug-and-Medication-Information>. Acesso em: 3 mar. 2020.
- 83 JOHONSON JR., W. C. J. Secondary structure of proteins through circular dichroism spectroscopy. **Annual Reviews of Biophysics and Biophysics Chemistry**, v. 17, p. 145–166, 1988.
- 84 MICSONAI, A. *et al.* BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. **Nucleic Acids Research**, v. 46, n. W1, p. W315–W322, 2018.

- 85 TAVARES, W. **Manual de antibióticos e quimioterápicos antiinfeciosos**. 3. ed. São Paulo: Editora Atheneu, 2001.
- 86 BLAST - Basic Local Alignment Search Tool. Disponível em: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>. Acesso em: 16 mar. 2020.
- 87 TAO, H. *et al.* Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS. **BioTechniques**, v. 48, n. 1, p. 61–64, 2010.
- 88 HANNAM, C.; LANGE, G. L.; MELLORS, A. Synthesis of a radiolabeled zwitterionic detergent and its use in protein purification. **Analytical Biochemistry**, v. 258, n. 2, p. 246–250, 1998.
- 89 FRANKEL, S.; SOHN, R.; LEINWAND, L. The use of sarkosyl in generating soluble protein after bacterial expression. **Proceedings of the National Academy of Sciences**, v. 88, p. 1192–1196, 1991. DOI: **10.1073/pnas.88.4.1192**.
- 90 SCHNAITMAN, C. A. Solubilization of the cytoplasmic membrane of *Escherichia coli* by Triton X-100. **Journal of Bacteriology**, v. 108, n. 1, p. 545–552, 1971.
- 91 GASTEIGER, E. *et al.* Protein identification and analysis tools on the ExPASy Server. In: WALKER, J. M. (ed.). **The proteomics protocols handbook**. New York: Humana Press, 2005. p. 571–607.
- 92 KATTKE, M. D. *et al.* Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria. **PLOS Pathogens**, v. 15, n. 4, p. e1007723, 2019.
- 93 CHAYEN, N. E. Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. **Acta Crystallographica D: biological crystallography**, v. 54, n. 1, p. 8–15, 1998.
- 94 COLEMAN, J. E.; ALLAN, B. J.; VAALLEE, B. L. Protein spherulites. **Science**, v. 131, n. 3397, p. 350–2, 1960.

APÊNDICE A – Artigo elaborado pela autora

Journal of Structural Biology 207 (2019) 158–168

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi Energy landscape of the domain movement in *Staphylococcus aureus* UDP-N-acetylglucosamine 2-epimerase

Erika Chang de Azevedo, Alessandro S. Nascimento*

Instituto de Física de São Carlos, Universidade de São Paulo, Av. Trabalhador São-Carlense, 400, Parque Arnold Schmidt, São Carlos, SP 13566-590, Brazil

ARTICLE INFO

Keywords:

Staphylococcus aureus
UDP-N-acetylglucosamine 2-epimerase
MnaA
Domain movement
Molecular dynamics
Energy landscape

ABSTRACT

Staphylococcus aureus is an important cause of resistant healthcare-associated infections. It has been shown that the wall teichoic acid (WTA) may be an important drug target acting on antibiotic-resistant cells. The UDP-N-acetylglucosamine 2-epimerase, MnaA, is one of the first enzymes on the pathway for the biosynthesis of the WTA. Here, detailed molecular dynamics simulations of *S. aureus* MnaA were used to characterize the conformational changes that occur in the presence of UDP and UDP-GlcNAc and also the energetic landscape associated with these changes. Using different simulation techniques, such as ABMD and GAMD, it was possible to assess the energetic profile for the protein with and without ligands in its active site. We found that there is a dynamic energy landscape that has its minimum changed by the presence of the ligands, with a closed structure of the enzyme being more frequently observed for the bound state while the unbound enzyme favors an opened conformation. Further structural analysis indicated that positively charged amino acids associated with UDP and UDP-GlcNAc interactions play a major role in the enzyme opening movement. Finally, the energy landscape profiled in this work provides important conclusions for the design of inhibitor candidates targeting *S. aureus* MnaA.

1. Introduction

Staphylococcus aureus is a naturally occurring Gram-positive bacteria that can be often found in the skin and mucosa of humans and other animals (McMillan et al., 2016). Although usually harmless, this organism can cause a wide range of infections in humans, such as folliculitis, furuncles, and carbuncles, impetigo, mastitis, wound infections, and staphylococcal scalded skin syndrome. More dangerous pathological conditions caused by *S. aureus* include bacteremia, pneumonia, endocarditis, bone, and joint infections, and toxic shock syndrome (Peacock and Paterson, 2015). *S. aureus* is also an important cause of hospital infections. According to data collected by the CDC between 2011 and 2014, this microorganism was responsible for approximately 10% of the overall healthcare-associated infections analyzed in the United States (Weiner et al., 2016).

The β -lactams, such as methicillin and penicillin, are among the antibiotics considered as the first choice for the treatment of this Gram-positive bacterial infection. These antibiotics act on the peptidoglycan layer of the cell, inhibiting transglycosylases and transpeptidases, collectively known as penicillin-binding proteins or PBPs, that are responsible for the synthesis and crosslink of specific polymers of the cell wall (Farha et al., 2013). However, the treatment of Staphylococcal

infections is becoming more challenging due to the emergence of drug-resistant strains (Lowy, 1998). These resistant strains were first reported only a few years after the introduction of penicillin and have increased in number of occurrences faster after the introduction of new antibiotics (Sewell and Brown, 2014). The CDC data, taken from 4,515 reporting hospitals and over 365,000 healthcare-associated infections, show that about 50% of the infections caused by *S. aureus* in healthcare units had isolated strains that were reported to be resistant to the β -lactam antibiotics like oxacillin, methicillin, and cofoxitin (Weiner et al., 2016).

S. aureus antibiotic resistance has a direct impact in the human lives but also accounts for an important burden in the health systems. Data taken in 2007 from 31 countries that participate in the European Antimicrobial Resistance Surveillance System (EARSS) showed that 27,711 bloodstream infections caused by MRSA resulted in 5503 excess deaths and 255,683 excess hospital days, thus generating in a burden of 44 million Euros (63.1 million dollars) (de Kraker et al., 2011).

The emergence of antibiotic resistance highlights the need for new drugs or even new ways to re-sensitize resistant bacterium. In this context, the cell wall remains as an interesting target for the discovery of new drugs and agents that fight β -lactam resistance (Farha et al., 2013). The cell wall of Gram-positive bacteria, such as *S. aureus*, is

* Corresponding author.

E-mail address: asnascimento@ifsc.usp.br (A.S. Nascimento).<https://doi.org/10.1016/j.jsb.2019.05.004>

Received 7 March 2019; Received in revised form 18 April 2019; Accepted 7 May 2019

Available online 11 May 2019

1047-8477/© 2019 Elsevier Inc. All rights reserved.

typically composed by peptidoglycan and wall teichoic acid (WTA). The WTA is located in the outer layer of the peptidoglycan and its role is associated the maintenance of the cell shape, virulence, coordination of the peptidoglycan synthesis, biofilm formation and, more recently, it has been linked to antibiotic resistance (Brown et al., 2012; Roemer et al., 2013; Sewell and Brown, 2014; Sobhanifar et al., 2016). In *S. aureus*, the WTA is typically made of a poly(ribitol-phosphate) polymer that is synthesized by the sequential action of several enzymes, called Tar (Teichoic acid ribitol) enzymes, starting with the glycosyl-transferases TarA and TarO. Another important enzyme involved in the biosynthesis of WTA and in the *S. aureus* resistance to β -lactams is the N-acetylglucosamine 2-epimerase or MnaA (Mann et al., 2016). This enzyme interconverts UDP-N-acetylglucosamine and UDP-N-acetylmannosamine, which are the substrates for TarO and TarA, respectively, thus regulating the balance of the starting points of the Tar biosynthesis pathway (Mann et al., 2016).

The *S. aureus* MnaA (SaMnaA) binds simultaneously to a UDP molecule and to the substrate UDP-N-acetylglucosamine (UDP-GlcNac) or UDP-N-acetylmannosamine (UDP-ManNac) (Mann et al., 2016; Velloso et al., 2008). Analysis of *S. aureus* MnaA (SaMnaA) crystal structure (Mann et al., 2016) and comparison with previously determined crystal structures revealed that the protein exhibits a domain movement upon binding the substrate, defining at least three observable conformations for the enzyme as opened, semi-closed and closed (Velloso et al., 2008). The opened state was observed in the absence of the substrate or product in the active site. The semi-closed state is observed in the presence of UDP. The closed state is typically observed when the protein has both UDP and UDP-GlcNac bound in its active site (Mann et al., 2016; Velloso et al., 2008). Interestingly, this movement between domains is also observed for several enzymes, including some enzymes of the WTA biosynthesis, such as TarM (Brown et al., 2012).

Here, we evaluated the domain movement of SaMnaA at the atomic level using different molecular dynamics simulations and assessing the free energy landscape associated to such movement. The simulations showed that the protein has a shift in the minimum of the free energy landscape profile when associated with the ligands UDP and UDP-GlcNac. Additionally, a major driving force that causes this shift seems to come from the electrostatic interactions between the enzyme and the substrate. These results can lead to a binding energy profile that indicates a dynamic energy landscape (Kumar et al., 2000) for the open/close movement of the protein, relevant for the enzymatic action.

2. Materials and methods

2.1. Protein and ligand parameterization

The chain A of the *S. aureus* MnaA structure (PDB ID: 5ENZ) was parameterized using AMBER FF14SB force-field (Maier et al., 2015). Hydrogen atoms were added and the histidine protonation states determined at the optimum pH of 7.0 using the H⁺ server (Anandakrishnan et al., 2012; Bashford and Karplus, 1990). The protein was immersed in a water box using TIP3P water model (Price and Brooks, 2004) and sodium ions were also included to keep the systems electrically neutral. UDP and UDP-GlcNac atomic charges were calculated using the Hartree-Fock method, with 6-31G⁺ basis set as available in Gaussian 09 (Rassolov et al., 2001). The obtained atomic charges were then refitted with RESP (Restrained Electrostatics Potential) (Cornell et al., 1993) procedure using ANTECHAMBER (Wang et al., 2006). The Lennard-Jones parameters for these ligands defined using GAFF force field (Wang et al., 2004). The crystal structure of SaMnaA (PDB ID 5ENZ (Mann et al., 2016)) was used as the starting structure for the APO MD simulations. The UDP and UDP-GlcNac bound structure (HOLO) was modeled using the chain A of the crystal structure of the SaMnaA (PDB ID: 5ENZ) and the crystal structure of *Bacillus anthracis* MnaA (BaMnaA) bound with UDP and UDP-GlcNac (PDB ID: 3BEO (Velloso et al., 2008)), using the program Modeller (Šali and Blundell,

1993; Webb and Sali, 2016). The purpose of this initial modeling was to generate a 'bound-state' (closed) conformation of SaMnaA with UDP and UDP-GlcNac coordinates from BaMnaA. The final systems comprised 45,519 and 55,479 atoms for the APO and HOLO system, respectively.

2.2. Equilibrium simulations

The initial energy of the APO system was minimized in 1000 cycles of conjugated gradient minimization. The system was then heated to 300 K in 50 ps, keeping the volume constant. The density of the system was afterward equilibrated to 1 g/cm³ by keeping constant pressure at 1 atm in 400 ps. During the minimization, heating and density equilibration, harmonic restraints were applied to all residues, with weight of 2 kcal.mol⁻¹.Å⁻². Finally, the system was equilibrated for 2 ns, without restraints, and the productive simulation was performed for 1 μ s in an NPT ensemble and, afterward, extended up to 5 μ s. Periodic boundary conditions were applied using a 12 Å cutoff radius for truncating short-range potentials. The simulation protocol for the HOLO system followed a similar procedure, with 4000 cycles of steepest descent energy minimization followed by 4000 steps of conjugated gradient minimization. The system was then heated to 300 K in 50 ps, with constant volume. The density was equilibrated until 1 g/cm³ in 400 ps keeping constant pressure at 1 atm. For minimization, heating and density equilibration, a harmonic positional restraint was applied to the entire protein and for the two ligands with a weight of 2 kcal.mol⁻¹.Å⁻². Finally, the system was equilibrated for 10 ns and the productive simulation was performed for 1 μ s without restraints. All the simulations were performed using AMBER16 (Case et al., 2005; Salomon-Ferrer et al., 2013).

2.3. Gaussian Accelerated Molecular Dynamics

In Gaussian Accelerated Molecular Dynamics (GAMD) (Miao et al., 2015), the trajectory is propagated in a non-negative Gaussian boost potential. This potential is applied to reduce the potential barriers and accelerate the state transitions, thus enhancing the conformational sampling. When the system potential V is lower than a threshold E , a harmonic boost potential ΔV is introduced to the total energy. The energy boost ΔV is chosen to result in the highest acceleration as possible but still allowing an accurate reweighting and free energy profile estimation. As shown in Equation (1), ΔV is defined as a harmonic potential with force constant k , which depends on the constant σ_0 (Eqs. (2) and (3)) (Miao et al., 2015). In Eqs. (2) and (3), $V_{\max}$ and $V_{\min}$ are the maximum and minimum system potential collected before the GAMD begins. σ_v and V_{av} are the standard deviation and average of the system potential energy, respectively. As previously observed, σ_0 must be chosen carefully in a compromise to get the best possible acceleration but still allowing an accurate reweighting of the accelerated potentials. One advantage associated with GAMD and with Accelerated MD (aMD) is the lack of a pre-chosen reaction coordinate or a collective variable (CV). Also, as pointed out by Miao and coworkers, GAMD preserves the original potential surface, allowing the computation of a PMF with less noise as compared to aMD (Miao et al., 2015). Here, we applied the boost potential to the dihedral and the total potential of the system, in a dual-boost approach. For both potentials, the σ_0 was set to 3 kcal/mol, based on preliminary simulations with varying values for σ_0 , where 3 kcal/mol provided the higher acceleration.

$$\Delta V(\vec{r}) = \frac{1}{2}k(E - V(\vec{r}))^2, \quad V(\vec{r}) < E \quad (1)$$

$$k = k_0 \left(\frac{1}{V_{\max} - V_{\min}} \right) \quad (2)$$

$$k_0 = \min(1.0, k'_0) = \min \left(1.0, \frac{\sigma_0}{\sigma_v} \frac{V_{\max} - V_{\min}}{V_{\max} - V_{av}} \right) \quad (3)$$

Table 1

Simulations used in this work. The simulation methods and times used are shown.

Method	Protein System	Time (ns)
Equilibrium	APO	1000 + 4000
Equilibrium	HOLO	1000
GAMD	APO	3 × 500
GAMD	HOLO	3 × 250
ABMD	APO	3 × 250
ABMD	HOLO	3 × 250

The resulting boost potential followed a Gaussian distribution that allowed a reweight and the calculation of a potential of mean force. The reweighting was made by obtaining the second-order of the cumulant expansion for the potential boost, as previously described (Miao et al., 2014). This calculation was made based on an opening angle computed between 10 and 80 degrees, with an angle increment of 4 degrees.

2.4. Adaptively Biased Molecular Dynamics

The Adaptively Biased Molecular Dynamics (ABMD) is an umbrella sampling method similar, in spirit, to metadynamics. The biasing force compensates for the free energy gradient, reproducing the previously unknown free energy surface for the system analyzed in the direction of the collective variable (CV), chosen by the user (Babin et al., 2009). Here, we used the angle between the α -carbon of the residues S128, G349, and P213, describing the opening angle of the SaMnaA structure. The simulation was conducted with a well-tempered umbrella sampling method (Barducci et al., 2008), with pseudo-temperature T defined at 5000 K. The total set of simulations used in this work are provided in Table 1.

2.5. Structural analysis

The APO system simulation of 5 μ s was analyzed using the algorithms available in the pyEMMA library (Scherer et al., 2015). Only the carbon alpha coordinates from the protein were used in the analysis, and the lag time used in the time-lagged independent component analysis (TICA) was set 200 frames or 10 ns. The electrostatic forces between the C-terminal and N-terminal domains of the protein were conducted using the Delphi Force Web server (Li et al., 2017), where the first domain was defined as the protein region encompassing residues 1 to 181. Three representative states of the protein were used in the calculations, open, intermediate and closed, with the ligands considered as part of the C-terminal domain. Geometric analysis of the MD

trajectories were carried out with CPPTRAJ (Roe and Cheatham, 2013) and interaction energies were computed with AmberEnergy++, a locally written program to compute interaction potential energies according to the AMBER force field (Nascimento, 2018).

3. Results

The states defined as open, semi-closed and closed can be better characterized by the opening angle observed for protein structures available on the PDB. The UDP-GlcNac 2-epimerase of *Methanocaldococcus jannaschii* structure was solved without ligands (PDB ID 4NEQ) (Chen et al., 2014) and with both UDP and UDP-GlcNac (PDB ID 4NES) (Chen et al., 2014). The first structure is in the opened state, with an opening angle (CV) of 54 degrees. Here, the opening angle is defined as the angle between the C α atoms of the residues S128, G349, and P213 in SaMnaA, or residues in similar positions for the other epimerases. The second structure is in a closed state, with a CV angle of 28 degrees. The SaMnaA was in the semi-closed state, with CV angle of 31 degrees, for chain A. A description of the opening angles, defined by our chosen CV, for all the UDP-GlcNac 2-epimerases with available structures to date is given in Table 2.

From the analysis of the results shown in Table 2, it can be observed that (i) for crystal structures of UDP-N-acetylglucosamine 2-epimerases in the absence of ligands (1O6C, 1V4V, 3OT5, 4NEQ and 3DZC) the opening angle varies between 47 and 53°, indicating the tendency to a opened state without substrate and UDP. The exception here is the crystal structure of *Rickettsia bellii* (4HWG). We hypothesize that the cubic symmetry observed in this crystal structure may stabilize the closed conformation in the absence of a ligand, due to an increased number of interchain contacts. We also observe that (ii) in all cases where both substrate (UDP-GlcNac) and cofactor (UDP) are found bound in the crystal structure (1VGV, 3BEO, 4FKZ, 4NES, and 5DLN) the observed opening angle is reduced to 24–30°, significantly reduced as compared to the unbound state for MnaA. Finally, (iii) SaMnaA (5ENZ) and 1F6D are bound to UDP only. Here, both proteins have different opening angles for each monomer, leaving opening questions for this scenario. A possible explanation is that UDP could promote a variety of intermediate states going from a closed state with an opening angle of 26–27° up to more opened states with angles of 30–40°.

Altogether, the analysis of the available structural data indicates the existence of a major structural change in the MnaA that seems to be associated with the binding of UDP/UDP-GlcNac. However, the crystal structures are single conformations of an ensemble of thermally accessible structures that could be sampled in solution. In order to get a more detailed perspective of this structural change, we used MD simulations to generate an ensemble of conformations in unbound (APO)

Table 2

Opening state of known crystal structures of UDP-GlcNac 2-epimerases. *chain without ligands.

PDB ID	Organism	Ligands	Oligomeric state	CV Angle Chain A (deg)	CV Angle Chain B (deg)
Enzymes in the unbound state					
4NEQ	<i>M. jannaschii</i>	–	Monomer	53	–
1V4V	<i>T. thermophilus</i>	–	Dimer	53	40
3OT5	<i>L. monocytogenes</i>	–	Dimer	47	33
1O6C	<i>B. subtilis</i>	–	Dimer	53	58
4HWG	<i>R. bellii</i>	–	Monomer	25	–
3DZC	<i>V. cholerae</i>	–	Dimer	48	46
Enzymes bound to UDP and UDP-GlcNac					
1VGV	<i>E. coli</i>	UDP-GlcNac	Dimer	40*	25
3BEO	<i>E. coli</i>	UDP, UDP-GlcNac	Dimer	30	30
4FKZ	<i>B. subtilis</i>	UDP, UDP-GlcNac	Dimer	25	25
4NES	<i>M. jannaschii</i>	UDP, UDP-GlcNac	Monomer	29	–
5DLN	<i>B. vietnamiensis</i>	UDP, UDP-GlcNac	Monomer	24	–
Enzyme bound to UDP only					
5ENZ	<i>S. aureus</i>	UDP	Dimer	31	26
1F6D	<i>E. coli</i>	UDP	Dimer	40	27

Fig. 1. Equilibrium simulations of SaMnaA. (A) Opening angle observed during the simulation time for the 1- μ s simulation of the APO SaMnaA structure. (B) Histogram of observed opening angle during the 1- μ s simulation of the APO SaMnaA. (C) Opening angle observed during the simulation of the HOLO SaMnaA structure. (D) Histogram of observed opening angle in HOLO SaMnaA simulation. (E) Opening angle observed during the 5- μ s simulation of the APO SaMnaA structure. (F) Histogram of the observed opening angle during the 5- μ s APO SaMnaA simulation.

and ligand-bound (HOLO) scenarios.

3.1. Equilibrium simulations

The APO SaMnaA enzyme was simulated for 1 μ s. During this simulation, a clear domain movement was observed in the MD trajectory and its magnitude could be measured by the change in the opening angle chosen as the collective variable (CV). As shown in Fig. 1A and B, the enzyme accessed the opened conformation with an opening angle of about 53 degrees and a more intermediate state with CV angle of about 40 degrees. The time evolution of the opening was very fast, with opening happening still during the equilibration. On the other hand, once the structure was opened, it remained opened (or in the intermediary state) during the entire simulation, never visiting the closed state over the 1 μ s simulation time.

In order to increase the sampling, the initial 1 μ s simulation was further extended up to 5 μ s. Even in this longer timescale, the closed conformation was never visited (Fig. 1E and F). Here, the intermediary state was mostly populated, resulting in a bimodal distribution of the CV angle with peaks at about 40 and 53 degrees, as shown in Fig. 1F.

The 1- μ s MD trajectory was analyzed using Principal Component Analysis (PCA) (Roe and Cheatham, 2013), and it was observed that the first two components described an opening movement of the enzyme domains (first PC) coupled to a torsion between them (second PC), as indicated by the arrows in Fig. 2. Confirming the observations made

based on the CV angle and the visual inspection of the trajectories, SaMnaA has a low-frequency domain movement towards an opened state in the absence of a ligand and cofactor.

The 5- μ s MD trajectory of the APO system was further analyzed using PyEMMA (Scherer et al., 2015). Within this package, the MD trajectory (reduced to the representation of C α atoms only) was clustered to define three major macrostates, as shown in Fig. 3B. The representative structures show that the structures are going from an intermediate state to the opened one (Fig. 3A and 3C). The first time-independent component (TIC), that describes the slowest movement of the system is in agreement with the behavior of the CV angle chosen, which indicates that this was a good choice of the CV indeed (Fig. 3A). The Markov-State Modeling of the clusters allowed us to estimate a typical time for the closing movement, i.e., transitioning from the opened to the intermediary state, as 879 ns, and for the opening, i.e., going from the intermediary state to the opened state, as 423 ns. Since the closed state was never sampled in the productive simulation, it was not possible to estimate the typical timescale for the full opening of the enzyme.

In summary, the results obtained from MD simulation for the APO state suggest that the unbound enzyme is found primarily in the opened state, but also sampling an intermediary state (very similar in structure to the opened state). The closed state was never observed, even in longer simulation times, consistently with the predominance of opened conformations observed in crystal structures available in the PDB.

Fig. 2. PCA Analysis of apo SaMnaA equilibrium simulation. Projections of the first (A) and second (B) components of the PCA analysis of the APO SaMnaA system MD trajectory.

Fig. 3. PyEmma Analysis of the 5 μ s simulation of apo SaMnaA. (A) Comparison between the CV angle measured (upper plot) and the TIC 1 from the TICA analysis (lower plot). (B) Heatmap of TIC1 versus TIC2 where the systems stated were coarse-grained, resulting in the separation of three macrostates, shown here as cluster defined by colors red, green and blue. (C) Representative structures of the three states found for the system, with atoms used for CV definition atoms in evidence.

Fig. 4. PCA analysis of holo SaMnaA. Projections of the first (A) and second (B) components of the PCA analysis for the HOLO SaMnaA system MD trajectory.

The equilibrium molecular dynamics for the HOLO system showed a markedly different behavior as compared with the APO one (Fig. 1C and D). Here, the enzyme shows a very well-defined preference for the closed state, with a typical CV angle of about 28–30 degrees. In agreement with crystal structures, the presence of the substrate and of the cofactor drives the equilibrium towards a closed conformation that is maintained during the entire simulation time. The trajectory was also analyzed using PCA (Fig. 4) and the two first components represented on the protein structure do not show a clear coordinated movement towards the domain movement for opening, as observed for the unbound SaMnaA.

For the HOLO SaMnaA, an analysis of the interaction energies involved in protein-UDP and protein-UDP-GlcNac interaction revealed a complex network of polar interactions involving the ligands, the enzyme and solvent atoms. Fig. 5A shows a LIGPLOT (Laskowski and Swindells, 2011) representation of the interactions involving UDP for a snapshot of the equilibrium simulation. The interactions between UDP and the enzyme mainly involve the residues Glu288, Arg206, Ser282 and Arg11, in addition to several solvent molecules and an interaction between UDP and UDP-GlcNac. The individual contributions of these residues to the total interaction potential energy are shown in Fig. 5C, where Arg11, and Arg206 stand out as major contributors due to their interactions with the phosphate groups of UDP. It is also interesting to note the positive (unfavorable) contribution from Glu288 for the UDP interaction (Fig. 5C), due to the charge-charge repulsion with the UDP phosphate groups.

For UDP-GlcNac (Fig. 5B), the set of interacting residues include Arg206, His40, Arg207, Lys239, His238 and Glu132 in addition to the interactions with Leu68, Gln66, and Met98 through their main chain atoms. Again, as expected, the major contributors to the total interaction potential are the positively charged residues Arg206, Arg207 and Lys239, followed by the histidine residues His40 and His238 (Fig. 5D). In conclusion, the interaction with UDP and UDP-GlcNac is maintained mainly through polar residues with the special contribution of positively charged residues in the vicinity of the active site.

3.2. Energy landscape of the opening for the unbound enzyme

In order to better understand the energy landscape associated with the conformational changes coupled to substrate binding, we also run non-equilibrium simulations of the APO and HOLO systems. Although GAMD simulations generate noisier free energy profiles (potential of mean forces) when compared to umbrella sampling techniques, it does not require a pre-chosen CV. Considering this advantage, we used GAMD to probe the energy landscape of the domain movement in SaMnaA.

Three independent simulations of the APO system were performed with three different starting conformations. These conformations are three different coordinates extracted from the equilibrium simulation (0 ns, 50 ns and 100 ns of the productive simulation). Each conformation was simulated for 500 ns. In the accelerated simulations, we observed that in two simulations where the starting conformations with opening angles were close to 55° were reduced to intermediary states with opening angles close to 40–43°. In the third simulation, the enzyme evolved more flexibly, with the opening angle continuously varying between 45 and 65°, as shown in Fig. 6A. The histogram of the accelerated simulations indicates a clear preference of the system for the intermediary state (Fig. 6B). Interestingly, even with the boost potential, the APO enzyme never visits the closed state, suggesting a high energy barrier to reach this conformation for the unbound enzyme.

Potential of mean force (PMF) profiles were computed from the GAMD simulations and are shown in Fig. 6C. Two slightly different profiles can be traced from the computed PMFs: (i) in all cases, the energies increase rapidly when the opening angle is below 35°. The simulation GAMD-APO-I (black curves in Fig. 6A and C) shows a low energy for the conformations encompassing opening angles between 35 and 55 degrees, GAMD-APO-II (red curves in Fig. 6A and C) shows low energies associated with opening angles between 35 and 45 degrees and GAMD-APO-III (blue curves) show low energies associated with opening angles between 35 and 70 degrees. All simulations agree that the intermediary state is a low energy state for the apo system, and

Fig. 5. Analysis of UDP-N-acetylglucosamine and UDP interaction with SaMnaA. LIGPLOT depiction of the interactions between UDP (A) and UDP-GlcNac (B) with SaMnaA. Total interaction (potential) energy between UDP (C) and UDP-GlcNac (D) and its interacting amino acids are also shown.

Fig. 6. GAMD simulation for APO SaMnaA structure. The results from three simulations are shown: GAMD-APO-I in black curve, GAMD-APO-II in red curve and GAMD-APO-III in blue curve. (A) Observed CV angle versus simulation time. (B) Distribution of observed CV angles during the simulation. (C) PMF computed after reweighting.

Fig. 7. GAMMD simulation data for the HOLO SaMnaA. The results from the three independent simulations are shown: GAMMD-HOLO-I in black curve, GAMMD-HOLO-II in red and GAMMD-HOLO-III in blue curve. (A) Observed CV angles during the simulation time. (B) Histogram of the observed angles during the simulation. (C) Computed PMFs from the energy reweighting. (D) RMSD calculated for the UDP (black) and UDP-GlcNac (red). (E) Total interaction energies computed for UDP (black) and UDP-GlcNac (red) with the protein. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

simulations I and III also suggest that the opened state (CV angle > 50°) can be thermally accessed by the unbound enzyme. Although the PMFs computed using GAMMD data are coarse, the overall results agree with the 5- μ s simulation data, suggesting that, once unbound, a putative equilibrium between conformations with opening angles between 40 and 45 degrees (intermediate) and 50–60 degrees (opened conformation) are expected to be found in solution.

Similarly, three independent GAMMD simulations for the HOLO SaMnaA were conducted and analyzed. The simulations also showed that the bound protein tends to stay in a closed conformation, with an opening angle of about 30 degrees, as shown in Fig. 7A. Here, two simulations remained the entire simulation time in the closed conformation (GAMMD-HOLO-II and GAMMD-HOLO-III, shown in Fig. 7A in red and blue curves, respectively) while the other simulation (GAMMD-

HOLO-I, shown in black curve in Fig. 7A) showed a continuously opening angle, starting with about 26 degrees at the beginning of the GAMMD simulation and reaching about 48 degrees at the end of the simulation. Interestingly, in this simulation, the ligand (UDP-GlcNac) was observed to unbind from the enzyme at a simulation time of 175 ns, as indicated in the binding energies and RMSD computed from this trajectory (Fig. 7D and E).

The PMFs calculated for the GAMMD simulations of the HOLO system (Fig. 7C) indicated a clear minimum in the free energy associated with an opening angle of about 25 degrees. For the simulations GAMMD-HOLO-I, where the unbinding event was observed, a shift in the minimum towards the intermediary state was observed, consistent with minimum expected for an APO enzyme. The unbinding of the UDP-GlcNac causes a major shift in the UDP position inside the active site,

Fig. 8. ABMD of the APO system for simulations ABMD-APO-I (upper row), ABMD-APO-II (middle row) and ABMD-APO-III (bottom row). (A) Time evolution of the CV angle; (B) CV angle distribution; (C) Potential of mean force. The gray lines indicate the time evolution (and convergence) of the PMF in 10 step ABMD simulations of 25 ns each.

forcing it to lose some of the interactions with the protein. Altogether, the histogram of the opening angle over the three GAMD simulations clearly indicates a preference for the closed state for the bound system, consistently with the equilibrium simulations. In addition, the unbinding event observed due to the acceleration applied in our non-equilibrium simulation clearly suggests that the opening event is coupled with the absence of the ligand.

The coarse description of the energy profile associated with the conformational state of SaMnaA may be associated with the poor sampling of the transition between the states, which is rarely observed in our simulations. In this way, a finer energetic description can be obtained for enhanced sampling techniques where the reaction coordinate or collective variable is systematically sampled over the simulation. Looking for this finer description of the opening/closing movement in SaMnaA, we explored the enzyme dynamics with Adaptively-Biased MD (ABMD).

Three independent ABMD simulations were conducted for the APO SaMnaA, with the same starting coordinates as chosen for GAMD. Each simulation explored the opening angle as a CV with a range of 10–100 degrees and resolution of 1 degree in 250 ns of ABMD simulations. The results obtained, shown in Fig. 8, indicate that the CV angles in the range of 40–60 degrees lie in the minimum of the free energy landscape, with energy barriers of about 10 kcal/mol (~ 16 kT) to reach a closed state, defined as a CV angle of about 25 degrees. The first simulation (ABMD-APO-I) showed an interesting broad minimum covering the entire region defined as the intermediary and opened state, while simulations ABMD-APO-II and ABMD-APO-III showed more defined minimum at about 55 degrees, consistent with an opened state. Together, these simulations indicate a clear energy minimum covering opened or intermediary-to-opened state for the unbound enzyme. We note that the PMF profiles observed in GAMD and ABMD simulations (Fig. 6 and Fig. 8) agree in general but are not similar. These differences are somewhat expected, since the methodologies rely on different assumptions. In GAMD, the simulation is accelerated by the introduction of boost potentials that do not consider the ‘reaction coordinate’ of the molecular event under evaluation, while in ABMD, the biasing potential is directed to this reaction coordinate. The final effect is a rougher, but

bias-free PMF for GAMD and a finer, though biased PMF for ABMD. In this context, having both experiments set may be a good strategy.

For the HOLO system, ABMD simulations accessed a wide range of CV angles, including the opened state. Again, once forced towards higher CV angles, the substrate was dissociated from the enzyme structure, resulting in the unbound substrate and the opened enzyme. The correlation between unbinding and opening becomes clear when Fig. 9 is compared with Fig. S1 (Supplementary material), that shows the interaction energy between the enzyme and substrate. The simulation data indicate that the presence of the ligand is crucial for holding the closed conformation stable and, once the conformation is destabilized (by external forces, in this case), the ligand is spontaneously dissociated from the enzyme.

An important question remains unanswered: what is the mechanism involved in ligand binding and enzyme domain movement? Or, what is the nature of the driving force for the domain movement observed in the presence of a substrate? A visual analysis of the enzyme structure reveals an extensive network of positively charged amino acids in the vicinity of the substrate pocket. These residues stabilize the interaction with the phosphate groups of the UDP-GlcNAc, thus making the electrostatic contribution of the binding energy as the largest contribution for binding. In the other hand, in the absence of a substrate, or product, the interaction of these positively charged residues is very unfavorable and could be alleviated if the residues are moved apart in a domain movement. Thus, we hypothesized a major driving force for this large-scale movement may be the electrostatic energy associated with the interaction of positively charged residues.

In order to evaluate this hypothesis, the APO and HOLO structures were analyzed using the Delphi Force web server (Li et al., 2017) to assess the magnitude of electrostatic forces between the N-terminal and C-terminal domains of SaMnaA. For the HOLO enzyme, the ligands were set as part of the C-terminal domain. The computed net forces are represented as arrows in protein structure in Fig. 10. Here we observe that, in the absence of the substrate/cofactor, a net force vector is observed as pointing towards the direction of the domain movement (orange arrow in Fig. 10). In the presence of the substrate, the net force is shifted to the twisting movement between the domains (also observed

Fig. 9. ABMD of the HOLO system for simulations ABMD-HOLO-I (upper row), ABMD-HOLO-II (middle row) and ABMD-HOLO-III (bottom row). (A) CV angle evolution during the simulation. (B) Histogram of CV angle. (C) Potential of mean force computed for the ABMD simulation data. The gray lines indicate the time evolution (and convergence) of the PMF in 10 step ABMD simulations of 25 ns each.

Fig. 10. Electrostatic forces on SaMnaA structures. The sum of the resulting interaction electrostatic forces between the C-terminal and N-terminal domains for the open (orange), intermediary (gray) and closed (purple) states of the protein, with ligands in the active site (cyan). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in the equilibrium simulation for the HOLO system). So, the analysis of the electrostatic forces provides a piece of evidence favoring the electrostatic term as a major driving force for the domain movement.

4. Discussion

In this work, we used over 10 μ s of simulation data to evaluate the mechanisms involved in the structural rearrangement of *S. aureus* UDP-N-acetylglucosamine 2-epimerase as a function of the binding of the substrate. (Table 1). Our equilibrium and non-equilibrium simulation data showed that the enzyme is remarkably stabilized in a closed conformation when bound to substrate and cofactor, while the unbinding event leads to a change in the enzyme equilibrium structure

towards a more opened structure. The analysis of the internal forces for the opened and closed conformation strongly suggested that the electrostatic term may be a major driving force for the enzyme opening, alleviating the unfavorable interactions between positively charged residues in the absence of the electron-rich substrate.

Considering the electrostatic forces as the main molecular driving forces acting to drive the ligand-dependent conformational changes, several additional experiments could be designed to further explore these molecular driving forces. We extended our analysis of this molecular mechanism using two additional experimental evidences. First, electrostatic forces should be captured by simpler molecular models, as such used in coarse-grained models. We used SIRAH force field (Darré et al., 2015; Machado et al., 2019) to generate a simpler (coarse-grained) model but still retaining most of the polar description of the protein. After 23 μ s of equilibrium simulation, we observed the APO enzyme to only sample the opened conformation (Supplementary Fig. S2) with a similar distribution of opening angle as observed in the atomistic simulation, suggesting that the molecular driving force that forces the apo enzyme to the opened conformation is kept even in this simpler model. As a second experimental evidence, we introduced set of mutations in enzyme surface to reduce the number of positively charged residues close to each other in the enzyme domains and/or to introduce ‘charged clamps’ by introducing negatively charged residues able to interact with positive residues in the opposite domain. As shown in Supplementary Fig. S3, from the mutants evaluated, at least two mutants were shown to be very effective in maintaining the closed conformation even in the absence of a ligand, thus reinforcing the electrostatic force as a major driving force for the domain movement in SaMnaA.

The shift in the energy landscape, caused by the change in the environment (substrate in the active site) agrees with the previously proposed concept of a dynamic energy landscape (Kumar et al., 2000), which is an important feature for the protein function since the macromolecule can adjust its population accordingly to changes in the environment, such as pH shifts or the presence of substrates. Kumar and coworkers (Kumar et al., 2000) proposed two main types of dynamic

energy landscapes. The Type I is characterized by an environment change that causes a shift in the energy minimum of the energy landscape, thus shifting the preferential conformation of the protein. Type II, on the other hand, describes the situation where the change in the environment causes a change of the roughness in the surface, where the conformations that were trapped in an energy minimum now have a lower barrier to overcome and change into another conformation. Within these definitions, the dynamic landscape observed for SaMnaA seems to be closer to type I, that is usually observed in enzymes that are allosterically regulated, or in multimolecular complexes.

As previously observed for other enzymes, the domain movement may also be an outstanding structural requirement for catalysis. The closed state, observed in the presence of the substrate, allows the active pocket to exclude solvent molecules, resulting in a low dielectric medium with higher magnitudes of electrostatic forces acting and finally allowing the catalytic reaction to proceed. Additionally, once the reaction is completed, the opening movement allows the diffusion of products and new substrate molecules (Kumar et al., 2000; Tsai et al., 1999). As far as these authors are aware, there is no experimental evidence of the mechanism of product scaping from the active site. By judging by the catalytic cleft structure, it seems plausible that the product could diffuse after a spontaneous opening of the enzyme, with UDP-ManNac scaping first and UDP remaining bound or scaping after UDP-ManNac. However, further investigations are still necessary to better elucidate this structural mechanism.

SaMnaA may be an important target for the development of drug candidates to fight resistant infections. The molecular simulations and previous data show that the ligands may enter the active site when the protein is in the open/intermediate conformation, being closed by the interactions made between the molecules and the protein. Also, positively charged residues are involved in the most energetically favorable interactions. This indicates that the open conformation should be targeted, and the electrostatic forces should be taken into account when screening for new drug candidates.

In conclusion, here we explored the dynamics of the *S. aureus* MnaA that plays an important role in the biosynthesis of the wall teichoic acid and in the cellular resistance to β -lactam antibiotics. The protein has an open/close movement, important for the maintenance of the substrate in the active site during the catalysis. This movement is mediated by the ligands, which interact with positive amino acids of the protein, shifting the energy profile for the system. The dynamic energy landscape for the systems APO and HOLO suggest that there is a population of preferential states that shifts with the change in the environment. Finally, a structure-based mechanism is proposed, providing a rationale for the domain movement observed for SaMnaA.

Declaration of Competing Interest

None.

Acknowledgements

We thank the fruitful contributions made by prof. Milton T. Sonoda (in memoriam) and also thank the anonymous reviewers for their valuable contributions. The authors thank the financial support provided by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) through grants 2017/18173-0, 2015/26722-8 and 2015/13684-0 and by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), through grant # 303165/2018-9. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jsb.2019.05.004>.

References

- Anandakrishnan, R., Aguilar, B., Onufriev, A.V., 2012. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucl. Acids Res.* 40, W537–W541. <https://doi.org/10.1093/nar/gks375>.
- Babin, V., Karpusenka, V., Moradi, M., Roland, C., Sagui, C., 2009. Adaptively biased molecular dynamics: an umbrella sampling method with a time-dependent potential. *Int. J. Quantum Chem.* 109, 3666–3678. <https://doi.org/10.1002/qua.22413>.
- Barducci, A., Bussi, G., Parrinello, M., 2008. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* 100, 1–4. <https://doi.org/10.1103/PhysRevLett.100.020603>.
- Bashford, D., Karplus, M., 1990. pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry* 29, 10219–10225. <https://doi.org/10.1021/bi00496a010>.
- Brown, S., Xia, G., Luhachack, L.G., Campbell, J., Meredith, T.C., Chen, C., Winstel, V., Gekeler, C., Irazoqui, J.E., Peschel, A., Walker, S., 2012. Methicillin resistance in *Staphylococcus aureus* requires glycosylated wall teichoic acids. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1–6. <https://doi.org/10.1073/pnas.1209126109>.
- Case, D.A., Cheatham III, T.E., Darden, T., Gohlke, H., Luo, R., Merz Jr., K.M., Onufriev, A., Simmerling, C., Wang, B., Woods, R.J., 2005. The amber biomolecular simulation programs. *J. Comput. Chem.* 26, 1668–1688.
- Chen, S.-C., Huang, C.-H., Shin Yang, C., Liu, J.-S., Kuan, S.-M., Chen, Y., 2014. Crystal structures of the archaeal UDP-GlcNAc 2-epimerase from *Methanocaldococcus jannaschii* reveal a conformational change induced by UDP-GlcNAc. *Proteins Struct. Funct. Bioinforma.* 82, 1519–1526. <https://doi.org/10.1002/prot.24516>.
- Cornell, W.D., Cieplak, P., Bayly, C.L., Kollman, P.A., 1993. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* 115, 9620–9631.
- Darré, L., Machado, M.R., Brandner, A.F., González, H.C., Ferreira, S., Pantano, S., 2015. SIRAH: a structurally unbiased coarse-grained force field for proteins with aqueous solvation and long-range electrostatics. *J. Chem. Theory Comput.* 11, 723–739. <https://doi.org/10.1021/ct5007746>.
- de Kraker, M.E.A., Davey, P.G., Grundmann, H., BURDEN study group, on behalf of the B. study, 2011. Mortality and hospital stay associated with resistant *Staphylococcus aureus* and *Escherichia coli* bacteremia: estimating the burden of antibiotic resistance in Europe. *PLoS Med.* 8, e1001104. <https://doi.org/10.1371/journal.pmed.1001104>.
- Farha, M.A., Leung, A., Sewell, E.W., D'Elia, M.A., Allison, S.E., Ejim, L., Pereira, P.M., Pinho, M.G., Wright, G.D., Brown, E.D., 2013. Inhibition of WTA synthesis blocks the cooperative action of pbps and sensitizes MRSA to β -lactams. *ACS Chem. Biol.* 8, 226–233. <https://doi.org/10.1021/cb300413m>.
- Kumar, S., Ma, B., Tsai, C., Sinha, N., Nussinov, R., 2000. Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci.* 9, 10–19.
- Laskowski, R.A., Swindells, M.B., 2011. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* 51, 2778–2786. <https://doi.org/10.1021/ci200227u>.
- Li, L., Jia, Z., Peng, Y., Chakravorty, A., Sun, L., Alexov, E., Valencia, A., 2017. DelPhiForce web server: electrostatic forces and energy calculations and visualization. *Bioinformatics* 33, 3661–3663. <https://doi.org/10.1093/bioinformatics/btx495>.
- Lowy, F.D., 1998. *Staphylococcus aureus* infections. *N. Engl. J. Med.* 339, 520–532. <https://doi.org/10.1056/NEJM199808203390806>.
- Machado, M.R., Barrera, E.E., Klein, F., Sónora, M., Silva, S., Pantano, S., 2019. The SIRAH 2.0 Force Field: Altius, Fortius, Citius. *J. Chem. Theory Comput.* 15, 2719–2733. <https://doi.org/10.1021/acs.jctc.9b00006>.
- Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., Simmerling, C., 2015. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 11, 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>.
- Mann, P.A., Müller, A., Wolff, K.A., Fischmann, T., Wang, H., Reed, P., Hou, Y., Li, W., Müller, C.E., Xiao, J., Murgolo, N., Sher, X., Mayhood, T., Sheth, P.R., Mirza, A., Labroli, M., Xiao, L., McCoy, M., Gill, C.J., Pinho, M.G., Schneider, T., Roemer, T., 2016. Chemical genetic analysis and functional characterization of staphylococcal wall teichoic acid 2-epimerases reveals unconventional antibiotic drug targets. *PLOS Pathog.* 12, e1005585. <https://doi.org/10.1371/journal.ppat.1005585>.
- McMillan, K., Moore, S.C., McAuley, C.M., Fegan, N., Fox, E.M., 2016. Characterization of *Staphylococcus aureus* isolates from raw milk sources in Victoria, Australia. *BMC Microbiol.* 16, 169. <https://doi.org/10.1186/s12866-016-0789-1>.
- Miao, Y., Feher, V.A., McCammon, J.A., 2015. Gaussian accelerated molecular dynamics: unconstrained enhanced sampling and free energy calculation. *J. Chem. Theory Comput.* 11, 3584–3595. <https://doi.org/10.1021/acs.jctc.5b00436>.
- Miao, Y., Sinko, W., Pierce, L., Bucher, D., Walker, R.C., McCammon, J.A., 2014. Improved reweighting of accelerated molecular dynamics simulations for free energy calculation. *J. Chem. Theory Comput.* (ACS Publications) 10, 2677–2689.
- Nascimento, A.S., 2018. AmberEnergy++ [WWW Document]. AmberEnergy. URL <https://github.com/alessandronascimento/amberenergy> (accessed 12.13.18).
- Peacock, S.J., Paterson, G.K., 2015. Mechanisms of methicillin resistance in *Staphylococcus aureus*. *Annu. Rev. Biochem.* 84, 577–601. <https://doi.org/10.1146/annurev-biochem-060614-034516>.
- Price, D.J., Brooks, C.L., 2004. A modified TIP3P water potential for simulation with Ewald summation. *J. Chem. Phys.* 121, 10096–10103. <https://doi.org/10.1063/1.1808117>.
- Rassolov, V.A., Ratner, M.A., Pople, J.A., Redfern, P.C., Curtiss, L.A., 2001. 6-31G* basis set for third-row atoms. *J. Comput. Chem.* 22, 976–984. <https://doi.org/10.1002/jcc.1058>.

- Roe, D.R., Cheatham, T.E., 2013. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* 9, 3084–3095. <https://doi.org/10.1021/ct400341p>.
- Roemer, T., Schneider, T., Pinho, M.G., 2013. Auxiliary factors: a chink in the armor of MRSA resistance to β -lactam antibiotics. *Curr. Opin. Microbiol.* 16, 538–548. <https://doi.org/10.1016/j.mib.2013.06.012>.
- Šali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815. <https://doi.org/10.1006/jmbi.1993.1626>.
- Salomon-Ferrer, R., Case, D.A., Walker, R.C., 2013. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* 3, 198–210. <https://doi.org/10.1002/wcms.1121>.
- Scherer, M.K., Trendelkamp-Schroer, B., Paul, F., Pérez-Hernández, G., Hoffmann, M., Plattner, N., Wehmeyer, C., Prinz, J.-H., Noé, F., 2015. PyEMMA 2: a software package for estimation, validation, and analysis of Markov models. *J. Chem. Theory Comput.* 11, 5525–5542. <https://doi.org/10.1021/acs.jctc.5b00743>.
- Sewell, E.W.C., Brown, E.D., 2014. Taking aim at wall teichoic acid synthesis: new biology and new leads for antibiotics. *J. Antibiot. (Tokyo)* 67, 43–51. <https://doi.org/10.1038/ja.2013.100>.
- Sobhanifar, S., Worrall, L.J., King, D.T., Wasney, G.A., Baumann, L., Gale, R.T., Nosella, M., Brown, E.D., Withers, S.G., Strynadka, N.C.J., 2016. Structure and mechanism of staphylococcus aureus tars, the wall teichoic acid β -glycosyltransferase involved in methicillin resistance. *PLOS Pathog.* 12, e1006067. <https://doi.org/10.1371/journal.ppat.1006067>.
- Tsai, C.J., Kumar, S., Ma, B., Nussinov, R., 1999. Folding funnels, binding funnels, and protein function. *Protein Sci.* 8, 1181–1190. <https://doi.org/10.1110/ps.8.6.1181>.
- Velloso, L.M., Bhaskaran, S.S., Schuch, R., Fischetti, V.A., Stebbins, C.E., 2008. A structural basis for the allosteric regulation of non-hydrolysing UDP-GlcNAc 2-epimerases. *EMBO Rep.* 9, 199–205. <https://doi.org/10.1038/sj.embor.7401154>.
- Wang, J., Wang, W., Kollman, P.A., Case, D.A., 2006. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* 25, 247–260. <https://doi.org/10.1016/j.jmglm.2005.12.005>.
- Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., Case, D.A., 2004. Development and testing of a general amber force field. *J. Comput. Chem.* 25, 1157–1174. <https://doi.org/10.1002/jcc.20035>.
- Webb, B., Sali, A., 2016. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* doi: 10.1002/cpbi.3.
- Weiner, L.M., Webb, A.K., Limbago, B., Dudeck, M.A., Patel, J., Kallen, A.J., Edwards, J.R., Sievert, D.M., 2016. Antimicrobial-resistant pathogens associated with health-care-associated infections: summary of data reported to the national healthcare safety network at the centers for disease control and prevention, 2011–2014. *Infect. Control Hosp. Epidemiol.* 37, 1288–1301. <https://doi.org/10.1017/ice.2016.174>.

APÊNDICE B – Gráfico referente ao experimento de fluorescência intrínseca

Figura 47 - Imagem da intensidade de emissão de radiação pela amostra, diminuída da intensidade de emissão do tampão. Em vermelho, emissão da amostra antes da desnaturação por calor e em preto, emissão da amostra após a desnaturação. Linhas verticais demarcam os máximos de intensidade de emissão de cada uma das amostras. A amostra foi excitada em 280 nm.

Fonte: Elaborada pela autora.

APÊNDICE C – Gráficos referentes aos experimentos de espectroscopia de dicroísmo circular

Figura 48 – Espectroscopia de dicroísmo circular para a enzima *SaMnaA*. (A) Gráfico dos resultados obtidos durante os experimentos. As setas indicam regiões de estrutura secundária. (B) Gráfico de absorção de diversas estruturas secundárias conhecidas quando submetidas à luz circularmente polarizada, onde o espectro para alfa-hélice está representado pela linha sólida, folhas-beta por pontos e traços, voltas-beta pela linha pontilhada e as regiões sem estrutura (*random coil*) pelos traços.

Fonte: Elaborada pela autora ; JOHNSON JR.⁸³

Figura 49 – Espectroscopia de dicroísmo circular para a enzima *SaMnaA*. (A) Gráfico dos resultados obtidos durante os experimentos. As setas indicam regiões de estrutura secundária. (B) Gráfico de absorção de diversas estruturas secundárias conhecidas quando submetidas à luz circularmente polarizada, onde o espectro para alfa-hélice está representado pela linha sólida, folhas-beta por pontos e traços, voltas-beta pela linha pontilhada e as regiões sem estrutura (*random coil*) pelos traços.

Fonte: Elaborada pela autora.