

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE FÍSICA DE SÃO CARLOS

Lívia Oliveira Dantas Clementino

***Enterococcus faecalis:***

*Cluster epa*

São Carlos

2023



**Lívia Oliveira Dantas Clementino**

***Enterococcus faecalis:***

*Cluster epa*

Tese de dissertação 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 Nascimento

**Versão Corrigida**

(Versão original disponível na Unidade que aloja o Programa)

**São Carlos**

**2023**

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.

Clementino, Livia Oliveira Dantas  
Enterococcus faecalis: Cluster epa / Livia Oliveira  
Dantas Clementino; orientador Alesssandro Nascimento -  
versão corrigida -- São Carlos, 2023.  
138 p.

Tese (Doutorado - Programa de Pós-Graduação em Física  
Biomolecular) -- Instituto de Física de São Carlos,  
Universidade de São Paulo, 2023.

1. Resistência bacteriana a antibióticos. 2.  
Enterococcus faecalis. 3. Cluster epa. 4. Biologia  
estrutural. I. Nascimento, Alesssandro , orient. II.  
Título.

Como ninguém chega a lugar nenhum sozinho, aqui está minha lista de agradecimento às pessoas e agências de fomento à pesquisa acadêmica, sem as quais eu definitivamente não teria escrito esta tese. Agradecer é só o que me resta.

Agradeço à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo auxílio financeiro, processo nº 2017/24901-8, assim como ao CNPq e CAPES. Agradeço também aos funcionários da Universidade de São Paulo de São Carlos pela gestão e manutenção do local onde desenvolvi minha pesquisa.

Agradeço ao professor Leonardo de Boni por sempre estar aberto a me ajudar na área de fotônica, além de ser uma pessoa muito agradável e simpática para um roqueiro, ao Prof. Dr. Francisco Guimarães pela ajuda com medidas confocais, à técnica Dra Andressa Patricia Alves Pinto por sua solicitude, os experimentos envolvendo o azul de molibdênio não ocorreriam sem ela me fornecer 3g do molibdato de amônio presente no Grupo de Biofísica. Agradeço ao Dr. Vinícius Ramos, por, dentre as discussões experimentais, sugerir a técnica de detecção de fosfato a partir do azul de molibdênio, à Dr. Luana Mourão, por sua disposição em ajudar todos os alunos do professor Alessandro durante a transição entre o Grupo de Biologia Molecular para o Grupo de Cristalografia. Muito obrigada.

Dentro do Grupo de Biologia Molecular, agradeço aos técnicos João Possato, Josimar Luiz Sartori, Livia Manzine e Maria Auxiliadora Morim Santos, por me ajudarem com um micro-ondas pifado, uma solda para um circuito Arduino, um procedimento novo ou só por uma conversa mesmo. Agradeço aos alunos do professor Igor Polikarpov, em especial, Gabriela, Vanessa, Paula, Anelyse, Caio, Malu, Milena e Natalia. São pessoas que pausam seus afazeres para te ajudar, são tão generosas quanto absurdamente competentes. Gabriela, você é uma pessoa maravilhosa e muito gentil, só não é muito perspicaz nas relações interpessoais, sua implicante! Vanessa, quem diria, a gente é tão diferente e, ainda assim, deu tão certo! Conte comigo para qualquer limpeza do laboratório no sábado! Paula, quem manja muito de Akta e de corujas, aprendi muito contigo! Anelyse, Caio, Milena, Malu e Natália, foi revigorante compartilhar o laboratório com a nova geração!

Agradeço ao professor Igor Polikarpov por disponibilizar toda infraestrutura do seu laboratório para os alunos do professor Alessandro, e pelas caipirinhas de final de ano que eram feitas para os alunos, exceto para Vanessa, ela tomava suco de limão.

Agradeço muito à Raissa por me auxiliar desde o início do doutorado, sempre aplicando uma ótima didática, e também à Angélica e ao Erick pelo apoio e companhia durante uma purificação ou outra de proteína. Agradeço aos ex-alunos do professor

Alessandro, Érika, Sthéffany e Victor, por me receberem tão bem ao grupo, pelas partidas de truco, pelos karaokês improvisados, pelas dicas de Ubuntu e por todas as confraternizações das quais participamos. Sthéffanny, obrigada pela ajuda com AmberTools, eu salvei tudo, espero conseguir ser tão hospitaleira quanto você, o seu altruísmo me comove. Érika, nossa Hannah Montana crossfiteira, conseguiu o melhor dos dois mundos entre o experimental e o teórico, queria ter feio o mesmo, mas não deu, vida que segue! Victor, tá aí alguém com quem eu adorava conversar, voa, cara, voa! Carol, obrigada por todas as idas em fast-foods, comer besteira nunca foi tão prazeroso. Lucas, muito obrigada pela paciência em me ajudar com o Arch Linux, pelas dicas culinárias e por me desestimular a jogar lol, você é um ótimo amigo. Manu, obrigada por ser tão sensata e ótima narradora de histórias, espero ler um livro seu um dia. Yuri, foi muito bom me reconectar com você, queria que você fosse meu irmão. Jenny, adoro chorar de rir contigo, e de me emocionar também, sua companhia foi importante para mim, muito obrigada. Menção (des)honrosa ao grupo de subs. Obrigada, professor Alessandro, por me ensinar coisas de computador, você é muito inteligente.

Obrigada, especialmente, para minha mãe, a melhor pessoa que eu conheço, eu te amo. À este período maravilhoso e absolutamente horrível: É só isso, não tem mais jeito, acabou, boa sorte.

*“Ando tão a flor da pele*

*Que qualquer beijo de novela me faz chorar.”*

*Gal Costa e Zeca Baleiro*





## RESUMO

CLEMENTINO, L. O. D. *Enterococcus faecalis: Cluster epa*. 2023. 138p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2023.

A resistência bacteriana a antibióticos é um problema de saúde pública, com tendência a se agravar ao longo do tempo. A consequência disto é refletida na quantidade de mortes provocadas por infecções de tais microrganismos, estimada em 7.7 milhões em 2019, número com tendência a aumentar. Neste contexto, iniciou-se o estudo de quatro enzimas presentes no *cluster* gênico *epa*. As proteínas *Epa* sintetizam e exportam carboidratos constituintes da parede celular de *Enterococcus*, envolvidos também na composição de biofilmes. Realizaram-se clonagem, expressão heteróloga e purificação com as enzimas *EpaI*, *EpaOX*, *EpaB* e *EpaE* de *Enterococcus faecalis* DSM20478. Apesar de diferentes estratégias aplicadas para otimizar a expressão e purificação dessas proteínas, a *EpaE* foi a única a apresentar rendimento e estabilidade razoáveis para ser submetida a uma triagem envolvendo múltiplas condições de cristalização. Um conjunto de dados foi coletado na linha Manacá do anel síncrotron brasileiro, o Sirius. O grupo espacial do cristal é o  $P2_1$ , apresentando fração de *twinning* pseudomeroedra. A estrutura refinada da *EpaE* foi determinada a 2.85 Å. Ela tem conformação tetramérica na unidade assimétrica, com um enovelamento do tipo Rossman. Análises realizadas pelo servidor PDBePISA indicaram que as interfaces formadas pelas cadeias A e B, e C e D são mais estáveis do que a B e C, e A e D, sugerindo se tratar de um dímero de dímeros. O sítio catalítico da *EpaE* foi identificado a partir da sobreposição estrutural com outra *RmlA* de maior resolução. Durante o refinamento da *EpaE*, foi percebida a presença de uma densidade eletrônica correspondente a molécula de timidina, presente no provável sítio alostérico da *EpaE*. O protocolo estabelecido para medir a atividade de *RmlAs* foi aplicado, ele detecta fosfato em solução a partir do reagente verde de malaquita. Entretanto, foi observado que o reagente verde de malaquita não é uma boa solução colorimétrica por não possuir intensidade das curvas de absorvância proporcional à concentração de fosfato no meio. Para sanar este problema, a técnica de detecção de fosfato com azul de molibdênio foi testada na medida da atividade específica da *EpaE*. Por fim, o programa *AlphaFold2* realizou predição das demais enzimas presentes no *cluster epa*, possibilitando elucidação de algumas funções dessas enzimas na síntese e transporte do antígeno polissacarídeo enterocócico de *E. faecalis*. Neste trabalho, determinamos a primeira estrutura cristalográfica de uma enzima do *cluster epa* e avançamos na identificação de um potencial inibidor da enzima *EpaE*.

**Palavras-chave:** Resistência bacteriana a antibióticos. *Enterococcus faecalis*. *Cluster epa*. Biologia estrutural.



## ABSTRACT

CLEMENTINO, L. O. D. *Enterococcus faecalis: Cluster epa*. 2023. 138p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2023.

Bacterial resistance to antibiotics is a public health problem, with a tendency to worsen over time. The consequence of this is reflected in the amount of deaths caused by such microorganisms, estimated at 7.7 million in 2019, a number with tendency to increase. In this context, the study of four enzymes present in the epa gene cluster began. Epa proteins synthesize and export constituent carbohydrates of the cell wall of *Enterococcus*, also involved in the composition of biofilms. The enzymes EpaI, EpaOX, EpaB and EpaE of *Enterococcus faecalis* DSM20478 were cloned, expressed and purified. Despite different strategies applied to improve the expression and purification of these proteins, EpaE was the only one to present reasonable yield and stability to be submitted to a screening involving multiple crystallization conditions. A dataset was collected on the Manacá line of the Brazilian synchrotron ring, Sirius. The space group of the crystal is  $P2_1$ , presenting twinning pseudomerohedral fraction. The refined EpaE structure was determined at resolution of 2.85 Å. It has a tetrameric conformation in the asymmetric unit, with a Rossmann-type fold. Analyses carried out by the PDBePISA server revealed that the interfaces formed by the chains A and B, and C and D are more stable than those formed by chains B and C, and A and D, suggesting that it is a dimer of dimers. The catalytic site of EpaE was identified from the structural alignment with another RmlA of higher resolution. During the refinement of EpaE, was noticed the presence of an electron density corresponding to the thymidine molecule, present in the probable allosteric site of EpaE. The established protocol for measuring RmlAs activity was applied, it detects phosphate in solution using the malachite green reagent. However, it was observed that the malachite green reagent is not a good colorimetric solution because its absorbance curves intensity aren't proportional to the phosphate concentration in the medium. To solve this problem, the molybdenum blue phosphate detection technique was tested to measure the specific activity of EpaE. Finally, the AlphaFold2 program predicted the structures of the other enzymes present in the cluster epa, enabling the elucidation of some of these enzymes functions in the synthesis and transport of the enterococcal polysaccharide antigen of *E. faecalis*. In this thesis, we determined the crystal structure of the EpaE enzyme, the first crystal structure of a member of the epa cluster, and also made advances in the identification of a potential inhibitor of this enzyme.

**Keywords:** Bacterial resistance. *Enterococcus faecalis*. Cluster epa. Structural biology.



Figura 1 – Mecanismos bacterianos de resistência a antibióticos. . . . .	27
Figura 2 – Intervalo de tempo entre o lançamento de antibióticos no mercado, em parênteses, e o aparecimento das primeiras cepas bacterianas resistentes. . . . .	28
Figura 3 – Infográfico disponibilizado pela Organização Mundial da Saúde, destacando as principais causas da resistência a antibióticos. . . . .	29
Figura 4 – Diminuição do número de antibióticos desenvolvidos e aprovados até 2014. . . . .	32
Figura 5 – <b>(a)</b> Imagem por microscopia de varredura eletrônica da organização celular de <i>E. faecalis</i> em um biofilme formado no tubo de celulose. As setas indicam possíveis canais de água. Escala de 1 $\mu\text{m}$ . <b>(b)</b> Fotomicrografia de um biofilme corado com DAPI que se desenvolveu em uma superfície metálica. . . . .	34
Figura 6 – Representação da parede celular de <i>Enterococcus</i> . Ancorados ao peptidoglicano, estão os ácidos teicoicos, os polímeros de ramnose, relacionados ao locus <i>epa</i> , e proteínas associadas a parede celular. . . . .	35
Figura 7 – Ilustração da região conservada do <i>cluster epa</i> em <i>E. faecalis</i> OG1RF. . . . .	36
Figura 8 – Esquematização da via de síntese de dTDP-L-ramnose, composta pelas enzimas RmlABCD. . . . .	37
Figura 9 – Interação entre o verde de malaquita e o grupo fosfato. Após a reação da RmlA, o PPI é hidrolisado pela pirofosfatase, gerando radical fosfato inorgânico, que é detectado pelo reagente verde de malaquita. . . . .	48
Figura 10 – Intensidade do reagente verde de malaquita proporcional à concentração de fosfato na solução. Concentrações crescentes de $\text{KH}_2\text{PO}_4$ da esquerda para direita. . . . .	48
Figura 11 – Formação do ácido fosfomolibdico. . . . .	49
Figura 12 – Modelo estrutural da EpaA, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a suposta região catalítica da EpaA. . . . .	52
Figura 13 – Modelo estrutural da <b>(a)</b> EpaA, em dourado, apresentando os potenciais resíduos de interação com a DPAGT1 de <i>Homo sapiens</i> (PDB ID: 6fwz), a partir do <b>(b)</b> alinhamento do modelo com a estrutura da enzima DPAGT1. Estão destacados com um $\bullet$ os resíduos diferentes entre a predição estrutural da EpaA e a DPAGT1. <b>(c)</b> Suposta cavidade catalítica da estrutura predita da EpaA. . . . .	53

Figura 14 – Modelo estrutural da EpaB, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a provável região catalítica da EpaB. . . . .	54
Figura 15 – Alinhamento estrutural entre o modelo predito da EpaB com a ppGalNAc-T2 (PDB ID: 5ndf). . . . .	55
Figura 16 – Modelo estrutural da EpaC, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a provável região catalítica da EpaC. . . . .	56
Figura 17 – Identificação dos potenciais resíduos catalíticos da EpaC a partir da sobreposição estrutural do modelo desta com a GlfT2 (PDB ID: 4fix). . . . .	56
Figura 18 – Modelo estrutural da EpaD, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a suposta região catalítica da EpaD. . . . .	57
Figura 19 – Identificação dos potenciais resíduos de interação da EpaD com a UDP-GlcUA, a partir da sobreposição entre o modelo predito da EpaD com a condroitina polimerase (PDB ID: 2z86). . . . .	58
Figura 20 – Modelo estrutural da EpaF, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a provável região catalítica da EpaF. . . . .	59
Figura 21 – <b>(a)</b> Modelo estrutural da EpaF, em dourado, sobreposto com a RmlC de <i>C. jejuni</i> (PDB ID: 7m14), em azul, ressaltando alguns dos potenciais resíduos catalítico da EpaF com um dos produtos da reação enzimática, o GDP, em roxo. <b>(b)</b> Resíduos da estrutura da RmlC (PDB ID: 7m14) que interagem com GDP. Estão destacados com um ● os resíduos diferentes entre a predição estrutural da EpaF e a RmlC de <i>C. jejuni</i> . . . . .	60
Figura 22 – Modelo estrutural da EpaG, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a provável região catalítica da EpaG. . . . .	61
Figura 23 – Identificação da região catalítica da EpaG através da sobreposição estrutural entre o modelo predito da EpaG com a RmlB de <i>S. suis</i> (PDB ID: 1ker). . . . .	62
Figura 24 – Modelo estrutural da EpaH, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. . . . .	63
Figura 25 – Modelo estrutural da <b>(a)</b> EpaH, em dourado, na provável região catalítica de interação com o NADH, em azul escuro, identificado a partir da sobreposição estrutural com a <b>(b)</b> RmlD de <i>E. coli</i> (PDB ID: 1vlo) NAD, em azul claro. . . . .	63

Figura 26 – Modelo estrutural da EpaI, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a suposta região catalítica da EpaI. . . . .	64
Figura 27 – Modelo estrutural da <b>(a)</b> EpaI, em dourado, sobreposto com a <b>(b)</b> dolicol fosfato manose sintase de <i>P. furiosus</i> (PDB ID: 5mm0), em azul, co-cristalizada com GDP-manose, em rosa, e íon $Mn^{+2}$ , em roxo. Modelo estrutural da <b>(c)</b> EpaI sobreposto com a <b>(d)</b> dolicol fosfato manose sintase (PDB ID: 5mm1), em azul, co-cristalizada com dolicol fosfato manose, em verde. Estão destacados com um ● os resíduos diferentes entre a predição estrutural da EpaI e as estruturas da dolicol fosfato manose sintase de <i>P. furiosus</i> . . . . .	64
Figura 28 – Modelo estrutural da EpaI, em dourado, e da GacI, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. . . . .	65
Figura 29 – <b>(a)</b> Modelo estrutural da EpaI, em dourado, sobreposta com os modelos da EpaB, em azul, EpaC, em roxo, e EpaD, em verde. <b>(b)</b> Foco no alinhamento entre as estruturas preditas da EpaI e EpaD. . . . .	66
Figura 30 – Modelo estrutural da EpaJ, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. . . . .	67
Figura 31 – <b>(a)</b> Modelo estrutural da GacJ, com coloração indicando do menor (em vermelho) para o maior (em azul) valor de pLDDT. <b>(b)</b> Sobreposição das estruturas preditas EpaJ, em dourado, e GacJ, em azul. . . . .	67
Figura 32 – Modelo estrutural da EpaL, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. . . . .	68
Figura 33 – Modelo estrutural da EpaL, em dourado, sobreposto com a estrutura do transportador ABC de <i>A. aeolicus</i> (PDB ID: 7k2t), em azul. . . . .	68
Figura 34 – <b>(a)</b> Modelo estrutural da GacD, com coloração indicando do menor (em vermelho) para o maior (em azul) valor de pLDDT. <b>(b)</b> Sobreposição das estruturas preditas EpaL, em dourado, e GacD, em azul. . . . .	69
Figura 35 – Modelo estrutural da EpaM, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. . . . .	70
Figura 36 – <b>(a)</b> Modelo estrutural da EpaM, em dourado, sobreposto com a estrutura do transportador ABC de <i>A. aeolicus</i> (PDB ID: 6m96), em azul. <b>(b)</b> Sobreposição do modelo estrutural da EpaL, em roxo, e da EpaM com o transportador ABC anterior (PDB ID: 6m96), indicando que EpaL e EpaM são duas proteínas de membrana acopladas, constituintes do transportador ABC. . . . .	70

Figura 37 – (a) Modelo estrutural da GacE, com coloração indicando do menor (em vermelho) para o maior (em azul) valor de pLDDT. (b) Sobreposição das estruturas preditas EpaM, em dourado, e GacE, em azul. . . . .	71
Figura 38 – Modelo estrutural da EpaN, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. . .	71
Figura 39 – (a) Modelo estrutural da EpaN, em dourado, sobreposto com um domínio da estrutura da glicosiltransferase de <i>L. borgpetersenii</i> (PDB ID: 6p61), em azul. (b) Modelos estruturais da EpaN, em azul, EpaB, em azul, EpaC, em roxo, e EpaD, em verde. . . . .	72
Figura 40 – Modelo estrutural da EpaO, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. . .	73
Figura 41 – (a) Predição estrutural da EpaO, em dourado, sobreposta com a estrutura da condroitina polimerase de <i>E. coli</i> (PDB ID: 2z86), em vermelho. (b) Sobreposição do modelo estrutural da EpaO, em dourado, com o da EpaD, em azul. . . . .	73
Figura 42 – Alinhamento dos domínios sanduíche $\alpha$ - $\beta$ - $\alpha$ da EpaO, em dourado, e da EpaN, em azul. . . . .	74
Figura 43 – Modelo estrutural da EpaR, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. . .	76
Figura 44 – Modelo estrutural da EpaR, em dourado, sobreposto com a estrutura da PglC (PDB ID: 5w7l), em azul. . . . .	77
Figura 45 – Esquematização da via de síntese proposta da parte inicial do EPA em <i>E. faecalis</i> ATCC19433 . . . . .	78
Figura 46 – Gel SDS-PAGE 15% corado com azul de <i>Coomassie</i> das etapas de purificação da proteína EpaI, expressa a partir dos vetores pETTrx-1a e pET21_NESG, e EpaOX, expressa a partir de pETTrx-1a. . . . .	82
Figura 47 – Gel SDS-PAGE 15% corado com azul de <i>Coomassie</i> das etapas de purificação das construções proteicas EpaI <sup>130</sup> e EpaOX <sup>220</sup> , expressas em vetor pETTrx-1a. . . . .	84
Figura 48 – Gel SDS-PAGE 15% corado com azul de <i>Coomassie</i> dos testes de expressão e solubilidade em <i>E. coli</i> LEMO21 (DE3). . . . .	85
Figura 49 – Gel SDS-PAGE 15% corado com azul de <i>Coomassie</i> das etapas da cromatografia por afinidade em resina de níquel, após o teste de reenvolvimento com 0.2% SDS das enzimas EpaI e EpaOX. . . . .	87
Figura 50 – Gel SDS-PAGE 15% corado com azul de <i>Coomassie</i> (a) da cromatografia por afinidade em resina de níquel e (b) da por exclusão de massa molecular da proteína EpaE (34 kDa), expressa em vetor pETM11 e <i>E. coli</i> Rosetta (DE3). . . . .	88



Figura 51 – (a), (b), (c) Fotos de cristais da enzima EpaE. (d) Padrão de difração obtido no anodo rotatório Rigaku Ultra X-18 e (e) na linha de luz Manacá, no Sirius. . . . .	89
Figura 52 – (a) Estrutura do tetrâmero da enzima EpaE refinada. (b) Visão lateral do monômero da cadeia A. . . . .	91
Figura 53 – Estrutura da EpaE, em dourado, sobreposta com a da RmlA de <i>M. tuberculosis</i> (PDB ID: 6b5e), em azul, co-cristalizada com o (a) produto dTDP-Glicose, em rosa, na cadeia F da RmlA (PDB ID: 6b5e), e (b) com o dTDP, em verde, na cadeia B da RmlA (PDB ID: 6b5e). . . . .	92
Figura 54 – Cadeia A da estrutura da EpaE, em dourado, sobreposta com a da RmlA de <i>M. tuberculosis</i> (PDB ID: 6b5e), em azul, co-cristalizada com o produto dTDP, em verde. (a) Identificação dos resíduos presentes no sítio alostérico da RmlA PDB (ID: 6b5e) para comparação com os da (b) EpaE. Estão destacados com um ● os resíduos diferentes entre a RmlA (PDB ID: 6b5e) e a EpaE. . . . .	92
Figura 55 – (a) Densidade eletrônica correspondente a da timidina, próxima às GLY144 das cadeias A e D da EpaE. (b) Potenciais resíduos que interagem com a timidina, em ciano. (c) Visão completa da região de interação da timidina com as cadeias A e D da EpaE. . . . .	93
Figura 56 – Sobreposição estrutural entre a EpaE e a RmlA de <i>E. coli</i> (PDB ID: 1h5r), com finalidade de comparação da localidade da timidina nas duas enzimas. . . . .	94
Figura 57 – Identificação dos resíduos do sítio catalítico da EpaE a partir da sobreposição estrutural da cadeia D da EpaE com a F da RmlA de <i>M. tuberculosis</i> (ID: 6b5e). . . . .	95
Figura 58 – Mecanismo da reação catalisada pela RmlA. O oxigênio O1P da glicose-1-fosfato promove um ataque nucleofílico no fosfato $\alpha$ do dTTP. A RmlA orienta essas duas moléculas de modo a facilitar essa substituição nucleofílica bimolecular. . . . .	97
Figura 59 – Cavidades identificadas na EpaE a partir da ferramenta <i>DoGSiteScorer</i> do <i>ProteinsPlus</i> , enumeradas em ordem decrescente do <i>DrugScore</i> de cada uma. A I se apresenta na interface entre as cadeias D e C da EpaE, a II na cadeia B, a III na C e a IV na cadeia D. . . . .	98
Figura 60 – Estrutura química dos compostos (a) Benztiázida, (b) Doripenem mono hidratado, e (c) Bicalutamida. . . . .	99
Figura 61 – Resultados das medidas de Fluorimetria Diferencial de Temperatura, realizadas para identificação de um ligante no sítio alostérico da EpaE. . . . .	100
Figura 62 – Teste da quantificação de pirofosfato formado a partir da reação catalisada pela EpaE, através do método reagente de malaquita. . . . .	101

Figura 64 – (a) Comportamento da curva de absorção do reagente colorimétrico do método de análise de fosfato, na presença de várias concentrações de fosfato. (b) Ajuste da intensidade de absorbância do azul de molibdênio em 830nm, na presença de fosfato, $R^2 = 0.99$ . . . . .	103
Figura 65 – Curvas de calibração da intensidade absorvida em 830 nm por concentração de fosfato em solução. . . . .	103
Figura 66 – Gráficos da produção de pirofosfato por tempo de reação da EpaE nas concentrações (a) $(41 \pm 6)$ nM (b) $(56 \pm 4)$ nM (c) $(79 \pm 2)$ nM. . . . .	104
Figura 67 – Teste da estabilidade do azul de molibdênio após diminuição da temperatura de incubação. . . . .	105
Figura 68 – Gráfico de produção de pirofosfato por tempo de reação de $(266 \pm 1)$ nM de EpaE. . . . .	106
Figura 69 – Comparação da região variável no locus epa das cepas OG1RF e V583 de <i>E. faecalis</i> . . . . .	122
Figura 70 – Mapa do vetor pETM11. . . . .	122
Figura 71 – Mapa do vetor pETTrx-1a visualizado a partir do programa SnapGene, disponível em <a href="http://snapgene.com">snapgene.com</a> . . . . .	123
Figura 72 – Mapa do vetor pET21_NESG visualizado a partir do programa SnapGene, disponível em <a href="http://snapgene.com">snapgene.com</a> . . . . .	123
Figura 73 – Mapa do vetor pLIC-EGFP visualizado a partir do programa SnapGene, disponível em <a href="http://snapgene.com">snapgene.com</a> . . . . .	124
Figura 74 – Modelo estrutural da EpaK, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. . . . .	127
Figura 75 – Modelo estrutural da EpaP, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. . . . .	128
Figura 76 – Modelo estrutural da EpaP, em dourado, alinhado com a estrutura da arabino-furanose transferase de <i>M. abscessus</i> (PDB ID: 6wby), em azul. . . . .	128
Figura 77 – (a) Modelo estrutural da GacK, à esquerda, e da GacL, à direita. (b) Modelo estrutural da WpsG, à esquerda, e da WpsG, à direita. A coloração indica do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Sobreposição da predição estrutural da EpaP, em dourado, com (c) as da GacK, em azul e da GacL, em roxo, e (d) WpsG, em laranja, e WpsJ, em ciano. . . . .	129
Figura 78 – (a) Modelo estrutural da EpaQ, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. (b) Modelo estrutural da EpaQ, em dourado, sobreposto com a estrutura da proteína de membrana polissacarídeo polimerase/ligase (PDB ID: 7tpg), em azul. . . . .	130

Figura 79 – Resultados das medidas de Fluorimetria Diferencial de Temperatura iniciais, realizadas para identificação de interações entre a EpaE e os potenciais ligantes. . . . .	134
Figura 80 – Resultados das medidas de Fluorimetria Diferencial de Temperatura, realizadas para identificação de um ligante no sítio alostérico da EpaE, padrão de cada medida. . . . .	135
Figura 81 – Espectro de fluorescência das células pós expressão no vetor pLIC-EGFP com IPTG, em <i>E. coli</i> Rosetta (DE3) e ArcticExpress. Medição realizada no espectrofotômetro <i>infinite</i> M200 PRO (TECAN). . . . .	136
Figura 82 – Microscopia confocal de fluorescência a partir da excitação em 488nm da eGFP fusionada com as proteínas <b>(a)</b> EpaB e <b>(b)</b> EpaE. . . . .	137



Tabela 1 – Resíduos presentes no provável sítio catalítico das glicosiltransferases EpaBCDION codificadas na região constante do <i>cluster</i> epa de <i>Enterococcus faecalis</i> . . . . .	75
Tabela 2 – Sumário das estatísticas da coleta de dados e do refinamento do cristal formado pela EpaE. . . . .	90
Tabela 3 – Genes epa catalogados nas cepas OG1RF e V583. . . . .	125
Tabela 4 – <i>Primers</i> utilizados para a clonagem dos genes codificantes das enzimas EpaI, EpaOX, EpaB e EpaE. . . . .	126



---

## LISTA DE ABREVIATURAS E SIGLAS

---

CDC	<i>Centers for Diseases Control and Prevention</i>
OMS	Organização Mundial de Saúde
ANVISA	Agência Nacional de Vigilância Sanitária
MRSA	<i>Staphylococcus aureus</i> resistente à meticilina
SARS-CoV-2	<i>Severe Acute Respiratory Syndrome Coronavirus 2</i>
VRE	<i>Enterococcus</i> resistente a vancomicina
EPA	<i>Enterococcal Polysaccharide Antigen</i>
GT2	Glicosiltransferases da família 2
RGP	L-Ramnose-Glicose polissacarídeo
LPS	Lipopolissacarídeo
NCBI	<i>National Center for Biotechnology Information</i>
LIC	<i>Ligation-Independent Cloning</i>
T <sub>m</sub>	<i>Melting Temperature</i>
D-Glc-1-P	D-Glicose-1-Fosfato
GBM	Grupo de Biotecnologia Molecular
PPi	Pirofosfato
pLDDT	<i>predicted Local-Distance Difference Test</i>
LB	<i>Luria Bertani</i>
IPTG	Isopropil $\beta$ -D-1-tiogalactopiranosídeo
DPAGT1	UDP-GlcNAc:dolicol-fosfato GlcNAc-1-fosfato transferase
PDB	<i>Protein Data Bank</i>
GlcNAc-1-P	N-acetil-glicosamina-1-fosforil
GlcNAc	N-Acetil-glicosamina

GalNAc	N-Acetil-galactosamina
Dol-P	Dolicol-Fosfato
Und-P	Undecaprenil-fosfato
ppGalNAc-T2	Polipeptídeo N-Acetil- $\alpha$ -galactose família 2 aminotransferase
GlfT1	D-Galacto-furanose transferase família 1
GlfT2	D-Galacto-furanose transferase família 2
UDP-Galf	UDP- $\alpha$ -D-galactofuranose
GalNAc-T2	Polipeptídeo GalNAc-transferase 2
CP	Condroitina Polimerase
GlcUA	Ácido Glicurônico
Glc	D-Glicose
Gal	D-Galactose
Rha	L-Ramnose



---

<b>1</b>	<b>INTRODUÇÃO</b>	<b>27</b>
<b>1.1</b>	<b>Resistência a antibióticos</b>	<b>27</b>
1.1.1	Cepas resistentes a antibióticos	29
1.1.2	COVID-19	30
1.1.3	Impacto Social e Econômico	30
<b>1.2</b>	<b><i>Enterococcus faecalis</i></b>	<b>32</b>
1.2.1	<i>Cluster epa</i>	34
<b>1.3</b>	<b>RmlA e a Síntese de Ramnose</b>	<b>36</b>
<b>2</b>	<b>OBJETIVOS</b>	<b>39</b>
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>41</b>
<b>3.1</b>	<b>Predição Estrutural das Enzimas do <i>Cluster epa</i></b>	<b>41</b>
<b>3.2</b>	<b>Clonagem</b>	<b>41</b>
3.2.1	Amplificação dos insertos e vetores	42
3.2.2	Anelamento e Transformação bacteriana	43
<b>3.3</b>	<b>Expressão de Proteínas</b>	<b>43</b>
<b>3.4</b>	<b>Testes de Solubilidade e Purificação</b>	<b>45</b>
<b>3.5</b>	<b>Teste de Cristalização e Coleta de Dados de Difração</b>	<b>46</b>
<b>3.6</b>	<b>Determinação Estrutural</b>	<b>46</b>
<b>3.7</b>	<b>Triagem de Potenciais Ligantes</b>	<b>46</b>
<b>3.8</b>	<b>Teste de Atividade Enzimática</b>	<b>47</b>
3.8.1	Reagente Verde de Malaquita	47
3.8.2	Método de Análise de Fosfato	49
<b>4</b>	<b>RESULTADOS DE MODELAGEM</b>	<b>51</b>
<b>4.1</b>	<b>Predição Estrutural das Enzimas do <i>Cluster epa</i></b>	<b>51</b>
<b>4.2</b>	<b>EpaA</b>	<b>51</b>
<b>4.3</b>	<b>EpaB</b>	<b>53</b>
<b>4.4</b>	<b>EpaC</b>	<b>55</b>
<b>4.5</b>	<b>EpaD</b>	<b>57</b>
<b>4.6</b>	<b>EpaF</b>	<b>59</b>
<b>4.7</b>	<b>EpaG</b>	<b>60</b>
<b>4.8</b>	<b>EpaH</b>	<b>62</b>
<b>4.9</b>	<b>EpaI</b>	<b>63</b>

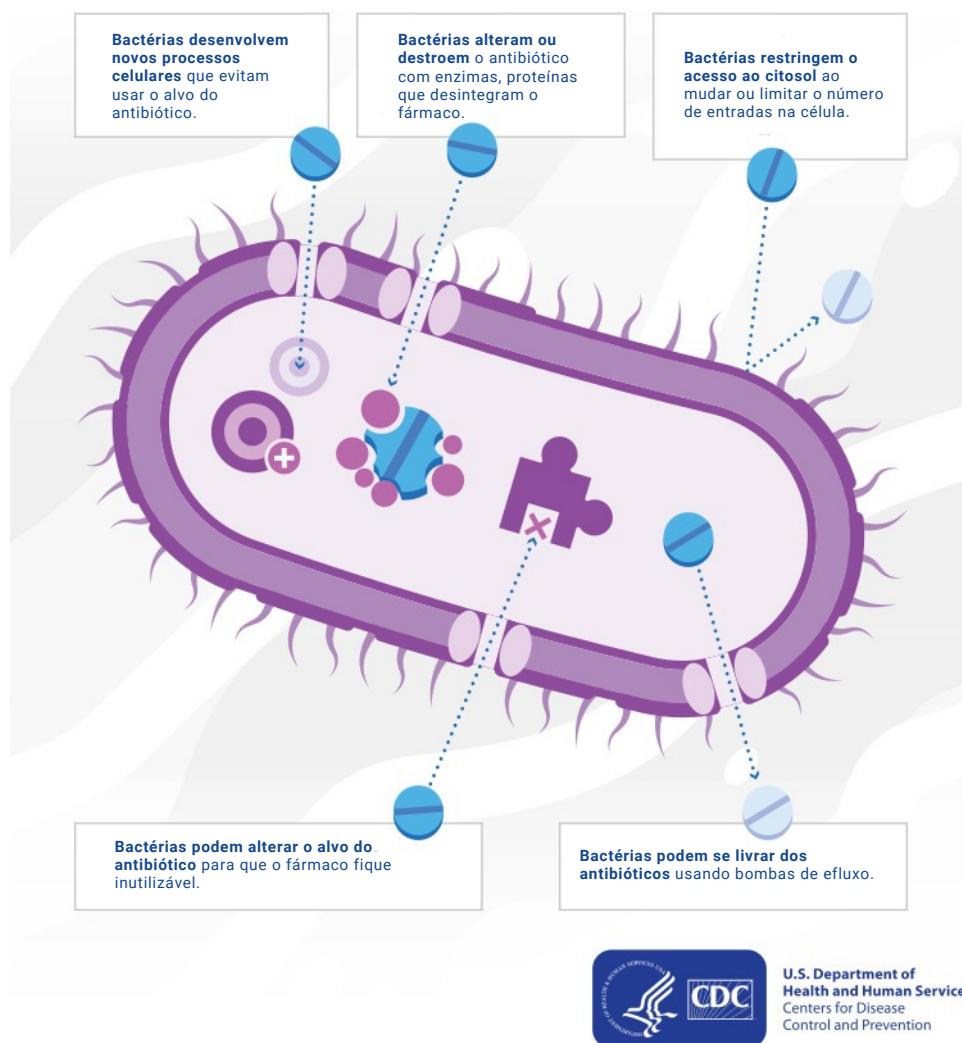
4.10	EpaJ . . . . .	66
4.11	EpaL . . . . .	67
4.12	EpaM . . . . .	69
4.13	EpaN . . . . .	71
4.14	EpaO . . . . .	72
4.15	Potenciais Resíduos Catalíticos das Glicosiltransferases . . . . .	73
4.16	EpaR . . . . .	76
4.17	Esboço da síntese da cadeia de ramnose . . . . .	77
<b>5</b>	<b>RESULTADOS EXPERIMENTAIS . . . . .</b>	<b>81</b>
5.1	Expressão das Proteínas EpaI e EpaOX e suas Construções . . . . .	81
5.2	Expressão das Proteínas EpaI e EpaOX em Outras Cepas de <i>E. coli</i> . . . . .	85
5.3	Solubilização com SDS e Renovelamento das Enzimas EpaI e EpaOX . . . . .	86
5.4	Purificação da Enzima EpaE . . . . .	87
5.5	Cristalização e Determinação Estrutural da enzima EpaE . . . . .	88
5.5.1	Sítio Alostérico . . . . .	91
5.5.1.1	<i>Timidina</i> . . . . .	93
5.5.2	Sítio Catalítico . . . . .	94
5.5.2.1	<i>dTDP-Glicose</i> . . . . .	94
5.5.2.2	<i>dTTP</i> . . . . .	95
5.5.3	Cavidades . . . . .	97
5.6	Triagem de Potenciais Ligantes . . . . .	97
5.7	Medidas de Atividade . . . . .	101
<b>6</b>	<b>CONSIDERAÇÕES FINAIS . . . . .</b>	<b>107</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>110</b>
	<b>APÊNDICE A FIGURAS E TABELAS . . . . .</b>	<b>122</b>
A.1	Região variável do <i>cluster epa</i> . . . . .	122
A.2	Vetor pETM11 . . . . .	122
A.3	Vetor pETTrx-1a . . . . .	123
A.4	Vetor pET21_NESG . . . . .	123
A.5	Vetor pLIC-EGFP . . . . .	124
A.6	Anotação dos genes do <i>cluster epa</i> . . . . .	125
A.7	<i>Primers</i> . . . . .	126
	<b>APÊNDICE B MODELAGEM - CONTINUAÇÃO . . . . .</b>	<b>127</b>
B.1	EpaK . . . . .	127
B.2	EpaP . . . . .	127

B.3	EpaQ . . . . .	130
APÊNDICE C	GRÁFICOS INDIVIDUAIS DAS MEDIDAS DE FLU- ORIMETRIA DIFERENCIAL DE TEMPERATURA . .	132
APÊNDICE D	MICROSCOPIA CONFOCAL . . . . .	136
D.1	EpaB e EpaE . . . . .	136



## 1.1 Resistência a antibióticos

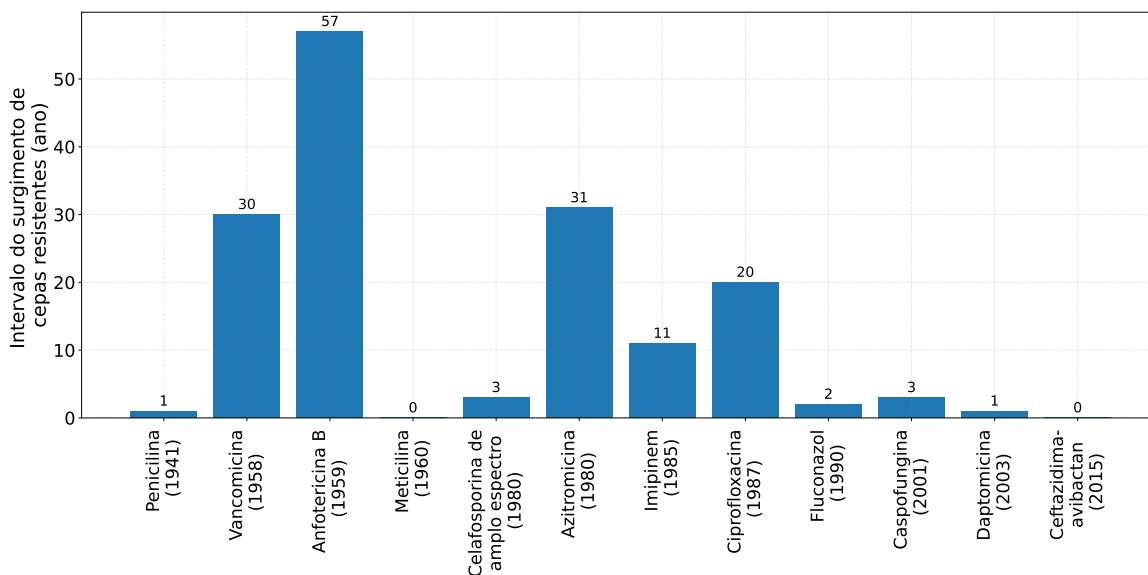
A resistência a antimicrobianos é a habilidade que bactérias, fungos, vírus e parasitas têm de anular o efeito de pequenas moléculas sintéticas ou naturais, as quais são empregadas na inibição do crescimento ou na morte desses patógenos.<sup>1</sup> Os genes que conferem resistência a antimicrobianos estiveram presentes ao longo da história,<sup>2</sup> e, mesmo em sistemas isolados, foram identificados genes bacterianos os quais conferem às células fatores de resistência (Figura 1),<sup>3-4</sup> demonstrando que a resistência a antibióticos é um mecanismo intrínseco de evolução em microrganismos.<sup>5-6</sup>



**Figura 1** – Mecanismos bacterianos de resistência a antibióticos.

Fonte: Adaptada a partir de *Centers for Disease Control and Prevention (CDC)*.<sup>7</sup>

Dentre os mecanismos de resistência a antibiótico, estão: evitar a aplicação do alvo do antibiótico em processos celulares, alteração do alvo do antibiótico, modificação ou destruição do antibiótico, restrição ao acesso do antibiótico ao meio celular e remoção do antibiótico do citosol.<sup>7</sup> Particularmente no caso de cepas bacterianas, a aquisição desses fatores de resistência a antibióticos tem ocorrido em intervalos de tempo cada vez menores, como mostram os dados levantados pelo *Centers for Disease Control and Prevention* (CDC), apresentados na Figura 2. Este gráfico informa, nas barras verticais, o intervalo de



**Figura 2** – Intervalo de tempo entre o lançamento de antibióticos no mercado, em parênteses, e o aparecimento das primeiras cepas bacterianas resistentes.

Fonte: Elaborada pela autora a partir de *Centers for Disease Control and Prevention* (CDC),<sup>8</sup> com Matplotlib.<sup>9</sup>

tempo entre o lançamento na sociedade de um antimicrobiano e a detecção do primeiro patógeno a ele resistente. Esse intervalo é pequeno para o caso de antibióticos da classe  $\beta$ -lactâmicos (Penicilina, Meticilina, Cefalosporina de amplo espectro e Cefazidima-avibactan), com exceção do Imipinem, em comparação com os antibióticos das classes macrolídeo (Azitromicina) e fluoroquinolona (ciprofloxacina). Os antifúngicos Fluconazol e Caspofungina também apresentam pequeno intervalo de tempo entre a administração desses em humanos e a aparição das primeiras cepas que são a eles resistentes, enquanto que o antifúngico Anfotericina B apresenta o maior intervalo de tempo desses dois acontecimentos entre todos os antimicrobianos analisados. Dentre os glicopeptídeos, esse intervalo é grande no caso da Vancomicina e pequeno no da Daptomicina. O padrão médio do gráfico da Figura 2 é que, para antimicrobianos lançados recentemente, a despeito da classe destes, há uma identificação mais rápida de cepas resistentes a eles.

Este cenário pode ser explicado, em parte, pelo uso excessivamente disseminado de antimicrobianos, o que acarreta no processo de seleção de cepas resistentes a essas

moléculas.<sup>10</sup> Além disso, segundo a Organização Mundial de Saúde (OMS), dentre os fomentadores da resistência a antimicrobianos, estão também a falta de acesso a saneamento básico, precária prevenção e controle de doenças, assim como insuficiência de vacinas e de testes clínicos, tanto em humanos quanto em animais<sup>1</sup> (Figura 3).



**Figura 3** – Infográfico disponibilizado pela Organização Mundial da Saúde, destacando as principais causas da resistência a antibióticos.

Fonte: Adaptada de OMS<sup>11</sup>

### 1.1.1 Cepas resistentes a antibióticos

No intervalo entre 2000 e 2015, o consumo mundial de antibióticos aumentou em 65% ao analisar as doses diárias.<sup>12</sup> Dentro do continente americano, o Brasil é o país com maior consumo de antibióticos, com 22.77 doses diárias por dia por 1000 habitantes.<sup>13</sup> No trabalho de Batolleti e colaboradores, 2016,<sup>14</sup> após promover tipagem de sequência multilocus e eletroforese em gel de campo pulsado em dez instalações médicas em São Paulo, foi observado o aumento percentual da frequência de isolados de *Klebsiella pneumoniae* resistentes a carbapenema (28.7%) e polimixina B (27.1%) no intervalo de 2011 a 2015, com total de 3085 isolados analisados.

O trabalho de Torres e colaboradores, 2011,<sup>15</sup> monitorou, durante 17 anos, o perfil genômico de *Streptococcus* do grupo A em pré-escolas de Curitiba. Neste trabalho, acompanhou-se o crescimento de isolados de *Streptococcus* do grupo A com resistência à eritromicina, de 1.9% a 4.0% durante o estudo, juntamente com níveis elevados de concentração inibitória mínima a este antibiótico, em um total de 1112 isolados examinados.

A Agência Nacional de Vigilância Sanitária (ANVISA),<sup>16</sup> entre 2004 e 2006, coletou 2710 amostras de carcaças de frangos a serem comercializadas em 14 estados brasileiros. Desta análise, constatou-se a prevalência de *Enterococcus* spp em 98.7% das amostras, das quais 61.4% eram *Enterococcus faecalis*. Do perfil de resistência das 2135 colônias de *Enterococcus* spp avaliadas, verificou-se elevada frequência de isolados não suscetíveis à estreptomicina (29.7%), tetraciclina (79.7%), eritromicina (58.3%), ciprofloxacina (27.5%), cloranfenicol (27.5%) e quinupristina-dalfopristina (95.2%).<sup>16</sup> Ainda dentro do gênero *Enterococcus*, entre 2003 e 2008, ocorreu o aumento de 10.5% de isolados Enterococci resistentes a vancomicina em centros clínicos no Brasil,<sup>17</sup> com total de 12324 isolados avaliados. Estes dados, levantados no país, mostram o nível crescente de resistência a antibióticos observada em diversos patógenos importantes.<sup>18</sup>

### 1.1.2 COVID-19

Em casos de infecções respiratórias virais, é comum o desenvolvimento de infecções bacterianas concomitantes à entrada do paciente em uma unidade de saúde, definida como co-infecção,<sup>19</sup> ou no decorrer da estadia do estabelecimento médico, a infecção secundária.<sup>19–20</sup> Isto é consequência da maior taxa de adesão bacteriana às células epiteliais do hospedeiro, resultante da infecção viral.<sup>21–23</sup> Nas pandemias gripe espanhola de 1918 e gripe H1N1 de 2009, foi observado que a mortalidade entre pacientes infectados simultaneamente com o vírus e o patógeno bacteriano era expressiva,<sup>19,21</sup> em especial, quando envolviam bactérias causadoras de pneumonia.<sup>24</sup> Dos casos de óbito em 2009, entre 30% e 50% envolveram infecções bacteriana.<sup>20</sup> Este conhecimento influenciou as práticas adotadas por algumas instalações médicas de ministrarem antibióticos em pacientes admitidos em unidades médicas com infecção do SARS-CoV-2 (*Severe Acute Respiratory Syndrome Coronavirus 2*), como medida profilática, sem de fato ter comprovação da infecção bacteriana, aumentando o consumo de antibióticos e favorecendo a seleção de cepas resistentes.<sup>25–26</sup>

### 1.1.3 Impacto Social e Econômico

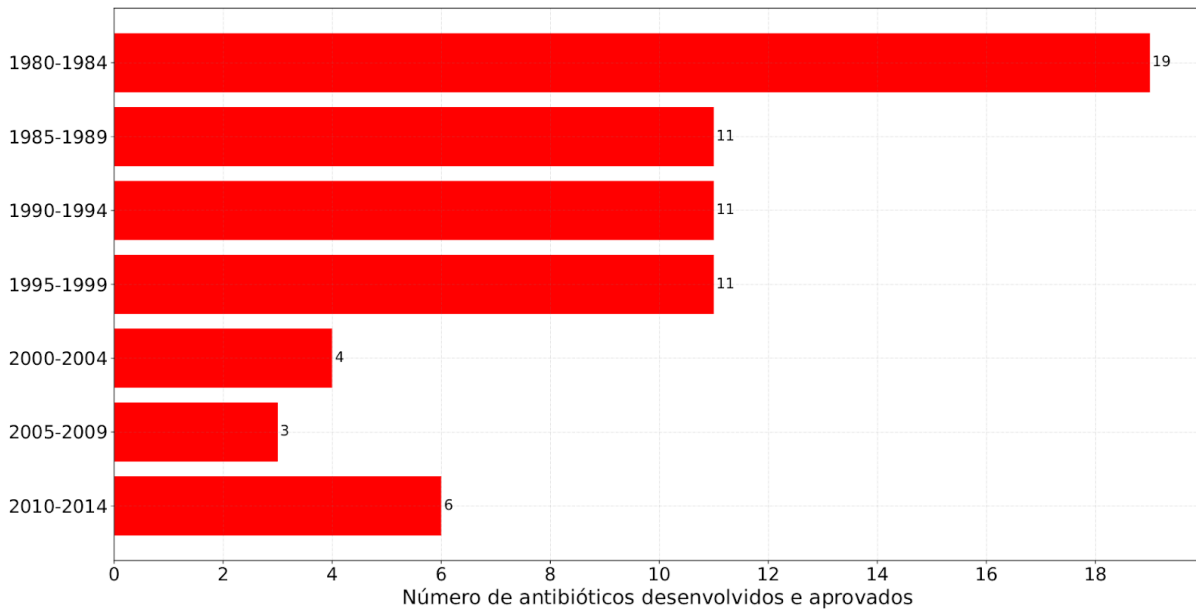
Ao adquirir infecção bacteriana em ambiente hospitalar, é necessário tratar essa infecção antes de submeter o paciente a procedimentos cirúrgicos para que a infecção não evolua para sepse, e nem dar alta ao paciente para não propagar essa infecção pela sociedade. O custo relacionado ao tratamento das infecções é resultante dos insumos aplicados e do trabalho do corpo médico até finalização do tratamento, assim como os dias que são deixados de trabalhar por parte do paciente.<sup>27</sup> A implicação econômica



gerada pelas infecções por patógenos insensíveis a antimicrobianos é abordada em alguns trabalhos. No estudo de Thorpe *et al.*, 2018,<sup>27</sup> nos Estados Unidos, chegou-se à estimativa de aproximadamente 1383 dólares adicionais no tratamento de cada paciente infectado por bactéria resistente a antibiótico, somando o total 2.2 bilhões de dólares anuais.<sup>27</sup> No trabalho de Bush *et al.*, 2011,<sup>28</sup> é citado o gasto absoluto de 55.8 bilhões de dólares por ano. Dados apontam tendência de aumento do número de óbitos decorrentes da resistência a antimicrobianos. Visto que a taxa de mortalidade é atrelada à economia, é importante controlar o problema da resistência a antibióticos atualmente para evitar grande número de mortes e crises econômicas no futuro.

A taxa de sucesso no tratamento de pacientes infectados com cepas resistentes pode ser duas vezes menor do que a com cepas suscetíveis a antibióticos.<sup>29</sup> Dados reportados na Europa mostraram que, de 27711 casos de infecções da corrente sanguínea causado por *Staphylococcus aureus* resistente à meticilina (MRSA), 5503 resultaram em óbitos,<sup>30</sup> além de 255683 dias adicionais na estadia hospitalar por parte dos pacientes. Dados disponibilizados pelo CDC, nos Estados Unidos, em 2019, indicam que 2.8 milhões de pessoas adquiriram infecção relacionada a resistência a antimicrobianos, e dessas, cerca de 35000 morreram.<sup>31</sup> Globalmente, estima-se que 7.7 milhões de mortes estão atreladas a uma bactéria resistente a antibióticos, configurando 13.6% das mortes ocorridas em 2019.<sup>32</sup> Em nível de comparação, desde janeiro de 2020 até junho de 2023, foram reportadas 6.94 milhões de mortes por SARS-CoV-2.<sup>33</sup> Ao analisar as espécies bacterianas resistentes mais letais, estão nas primeiras posições *Staphylococcus aureus* e *Escherichia coli*,<sup>32</sup> duas espécies que configuram como tema de diversas pesquisas. E, na nona posição, está a *Enterococcus faecalis*,<sup>32</sup> uma espécie com menos abordada em trabalhos científicos quando comparada com aquelas duas. Isto nos motivou a tê-la como objeto de estudo.

A resistência a antimicrobianos é intensificada pela administração indiscriminada dos antimicrobianos a pessoas e animais, não vacinação, condições precárias de saneamento e atendimento clínico insuficiente. Como este fenômeno é inerente aos microrganismos,<sup>2-6</sup> o desenvolvimento de novos agentes terapêuticos deveria ser uma constante no contexto das indústrias farmacêuticas. Contudo, a tendência de desenvolver novos antibióticos é decrescente, como mostrado na figura Figura 4. Isto decorre da maior rigidez regulatória para aprovação do antibiótico hoje em dia, além da baixa rentabilidade desses fármacos em comparação com outros lançados pela indústria farmacêutica.<sup>34</sup> Os dados apresentados indicam que a aquisição de resistência a antimicrobianos, em particular em patógenos bacterianos, tem um impacto direto na vida das pessoas. As implicações na economia são reflexos da principal consequência deste problema sanitário, a taxa de mortalidade de pessoas infectadas, com tendência a aumentar. Isto confere uma dimensão da relevância da resistência a antibióticos no aspecto da saúde pública.



**Figura 4** – Diminuição do número de antibióticos desenvolvidos e aprovados até 2014.

Fonte: Elaborada pela autora a partir de *The Antibiotic Resistance Crisis*,<sup>34</sup> com Matplotlib.<sup>9</sup>

## 1.2 *Enterococcus faecalis*

*Enterococcus faecalis* é uma bactéria Gram-positiva presente na microbiota intestinal dos mamíferos<sup>35</sup> e de alguns pássaros.<sup>36</sup> Esta espécie se destaca como um patógeno oportunista, promotor de infecções nosocomiais.<sup>36–37</sup> As linhagens de *E. faecalis* são intrinsecamente resistentes a vários antibióticos,<sup>38</sup> o que, em conjunto com a alta propensão de adquirirem outros fatores de resistência,<sup>38–39</sup> tanto dificulta o tratamento de infecções, quanto o controle de surtos hospitalares provocados por essa espécie bacteriana.<sup>38,40</sup>

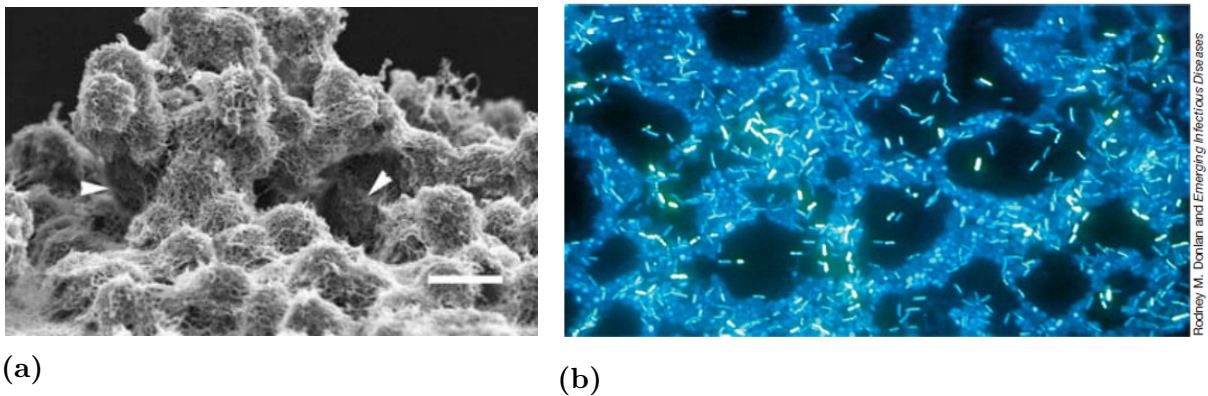
O trabalho de Weiner-Lastinger e colaboradores, 2020,<sup>41</sup> realizado entre 2015 e 2017, em recintos hospitalares, reporta a frequência dos patógenos nas seguintes infecções: sepse, infecção do trato urinário associada a cateter vesical, pneumonia associada a ventilação mecânica e infecção de sítio cirúrgico. Destas infecções bacterianas, *Enterococcus faecalis* está entre os 4 patógenos mais frequentes de cada categoria, exceto no caso de pneumonia associada a ventilação mecânica. A frequência absoluta de *E. faecalis* aumentou quando comparada com o trabalho anterior da mesma autora e colaboradores, entre 2011 e 2014.<sup>42</sup> Entretanto, a bactéria *E. faecalis* figura em posições de menor frequência entre os patógenos quando considerada a resistência a antibióticos, neste caso, a vancomicina.<sup>41–42</sup> Já para a *Enterococcus faecium*, o oposto acontece. *E. faecium* não é tão frequente nas infecções presentes nas unidades médicas, mas a maioria das cepas desta espécie é resistente à vancomicina.<sup>41–42</sup> Nos dois estudos citados, se analisada a frequência dos patógenos no nível de gênero, o de *Enterococcus* se destacaria. Apenas a contabilização das espécies *Enterococcus faecalis* e *Enterococcus faecium* faria com que esse gênero ficasse entre a 1ª e

a 3<sup>o</sup> posição como linhagens patogênicas mais frequentes em todas as categorias analisadas de ambos os trabalhos, exceto no caso de pneumonia associada a ventilação mecânica.

Visando identificar genes determinantes de resistência a antibióticos, em especial aqueles associados à formação de biofilme em *Enterococcus faecalis*, Dale e colaboradores, 2015,<sup>39</sup> criaram uma biblioteca mutante de transposons para triagem genética da linhagem OG1RF. Dentre os genes que se destacaram na análise dos mutantes transposons, estavam o *epaI* (NCBI ID: OG1RF\_11730) e o *epaOX* (NCBI ID: OG1RF\_11715). Notaram-se os seguintes impactos a partir da supressão desses genes em relação com a cepa parental: **(i)** redução do número de colônias em ambiente com ácidos biliares e detergente SDS, com objetivo de mimetizar o meio gástrico, revelando baixa integridade do envelope celular, **(ii)** aumento de susceptibilidade ao antibiótico daptomicina e **(iii)** diminuição da biomassa dos biofilmes produzidos .

Biofilmes são constituídos pela aglomeração de inúmeras células bacterianas, de mesma espécie ou não, aderidas umas às outras e a uma superfície em comum, a partir de uma matriz.<sup>43</sup> Os principais componentes desta matriz são polissacarídeos secretados, podendo apresentar também proteínas, lipídeos e ácidos nucleicos.<sup>43</sup> O biofilme se apresenta como um agente bacteriano de dessensibilização aos antibióticos,<sup>44–45</sup> visto que atuam como uma barreira física, evitando a interação do antibiótico com as células.<sup>46</sup> O biofilme também viabiliza a transferência horizontal de genes por reter qualquer material celular exportado e restos do lisado celular,<sup>43,46</sup> potencializando a transferência de plasmídeos contendo genes codificantes de fatores de resistência a antimicrobianos. Os biofilmes são facilmente encontrados em aparatos médicos, como cateteres venosos e urinários, e válvulas cardíacas danificadas,<sup>47–48</sup> facilitando a contaminação de pacientes em ambiente hospitalar. Durante o pico da pandemia de COVID-19, pacientes com piores quadros respiratórios exigiam uso de ventiladores e cateteres por mais tempo e com maior frequência, dificultando a apropriada limpeza desses equipamentos, e facilitando infecções de patógenos durante a estadia hospitalar.<sup>49</sup>

No trabalho seguinte de Dale *et al.*, 2017,<sup>52</sup> foi analisada a estrutura dos biofilmes produzidos pela cepa parental *Enterococcus faecalis* OG1RF e os mutantes,  $\Delta$ *epaI* e  $\Delta$ *epaI*. A autora e colaboradores descrevem que esses mutantes se comportaram igualmente em situação com ou sem os antibióticos daptomicina e gentamicina, adotando uma disposição celular descontínua e tridimensional. Na ausência dos antibióticos, os biofilmes das células parentais apresentaram aspecto bidimensional e contínuo. Já na presença dos antibióticos, os biofilmes produzidos pela cepa parental se assemelhavam aos gerados pelos mutantes  $\Delta$ *epaI* e  $\Delta$ *epaI* ao transicionar do estado bidimensional para o tridimensional, ainda apresentando maior união e continuidade em relação aos mutantes. Foi proposto que a desconexão entre a superfície e o biofilme facilitaria a disseminação celular durante a infecção, porém, não foi possível mensurar a desvantagem na exposição das células



**Figura 5** – (a) Imagem por microscopia de varredura eletrônica da organização celular de *E. faecalis* em um biofilme formado no tubo de celulose. As setas indicam possíveis canais de água. Escala de 1  $\mu$ m. (b) Fotomicrografia de um biofilme corado com DAPI que se desenvolveu em uma superfície metálica.

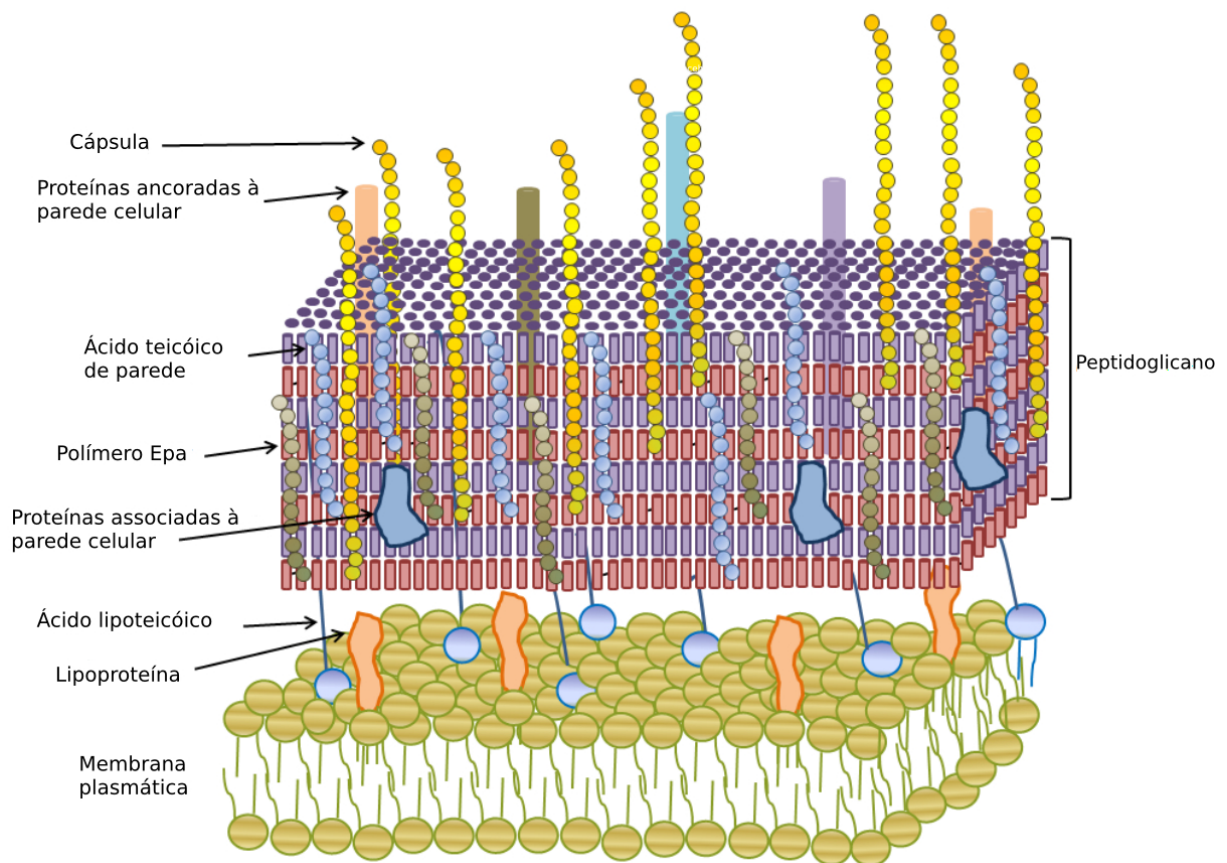
Fonte: ERLANDSEN; KRISTICH; DUNNY<sup>50</sup> e MADIGAN *et al.*<sup>51</sup>

desprendidas a antibióticos e respostas imunes. Esses dados sugerem que os genes *epaI* e *epaOX* estão diretamente relacionados na arquitetura de biofilmes em *E. faecalis*, os quais adquirem menor estabilidade quando dispostos na forma tridimensional.

### 1.2.1 *Cluster epa*

Os genes *epaI* e *epaOX* estão presentes no *cluster epa* (*enterococcal polysaccharide antigen*),<sup>37,53</sup> um conjunto de genes responsáveis por codificarem proteínas sintetizadoras e exportadoras de polissacarídeos, de modo a serem parcialmente responsáveis pela decoração da parede celular em *Enterococcus*<sup>54</sup> (Figura 6). Os genes do *cluster epa* são referidos, nesta tese, por letra minúscula inicial "epa", as proteínas codificadas por eles são apresentadas com letra inicial maiúscula "Epa", e o polímero sintetizado pelas enzimas Epa será mencionado com todas letras maiúsculas, "EPA".

O *cluster epa* é constituído de uma região conservada de 18 genes em *E. faecalis*, do gene *epaA* ao *epaR* (Figura 7), seguida por uma região variável própria de cada linhagem<sup>39</sup> (seção A.1). A região conservada do *cluster epa* codifica proteínas sintetizadoras e exportadoras do polímero de ramnose, que é ancorado ao peptidoglicano. Este polímero de ramnose está presente em todas cepas do tipo *Enterococcus faecalis*.<sup>54</sup> A região variável do *cluster epa* codifica enzimas responsáveis pela decoração de carboidratos do polímero de ramnose, esses carboidratos estão expostos na superfície da parede celular, e a composição deles varia entre cepas.<sup>55</sup> Dos 18 genes conservados, 15 homólogos foram identificados em algumas cepas de *E. faecium*,<sup>56</sup> em ordem distinta no locus *epa*, confirmando a presença do *cluster epa* também nesta espécie. A similaridade entre as sequências de resíduos de aminoácidos dos homólogos *epa* das duas espécies variam entre 31% e 92% de identidade.<sup>54</sup> A *EpaI*, por exemplo, apresenta até 45% de identidade com 21% de cobertura sequencial



**Figura 6** – Representação da parede celular de *Enterococcus*. Ancorados ao peptidoglicano, estão os ácidos teicoicos, os polímeros de ramnose, relacionados ao locus epa, e proteínas associadas a parede celular.

Fonte: Adaptada de GILMORE *et al.*<sup>54</sup>

com as enzimas de *E. faecium*, enquanto que a EpaOX apresenta até 49% de identidade numa cobertura de 98% com as sequências alinhadas de proteínas de *E. faecium*.

Os trabalhos publicados sobre o *cluster* epa se iniciaram no ano 2000 por Xu *et al.*,<sup>37</sup> seguido da observação de indícios conservativos entre linhagens de *E. faecalis*, assim como a participação de proteínas Epa na defesa celular, por Teng *et al.*, 2002.<sup>53</sup> A anotação genética (Figura 7) em conjunto com a descoberta da composição do polissacarídeo Epa foram realizadas somente em 2009 por Teng e colaboradores.<sup>57</sup> Posteriormente, alguns estudos ainda revelaram mais informações sobre a virulência atribuída à bactéria pelas enzimas Epa<sup>58–59</sup> e os efeitos deletérios na ausência de genes epa no crescimento celular.<sup>39,52,60</sup> No entanto, as funções biológicas da maioria das proteínas codificadas dentro do *cluster* epa não foram elucidadas,<sup>61</sup> há apenas anotação da atividade biológica delas, baseada no alinhamento das sequências proteicas,<sup>57</sup> como visto na seção seção A.6. Até o momento de escrita deste trabalho, nenhuma enzima participante do *cluster* epa teve sua estrutura cristalográfica resolvida.

As enzimas EpaI e EpaOX são anotadas como glicosiltransferases da família 2



**Figura 7** – Ilustração da região conservada do *cluster epa* em *E. faecalis* OG1RF.

Fonte: Elaborada de GILMORE *et al.*<sup>54</sup>

(GT2).<sup>37</sup> Desta categoria, dentre as sequências dos genes reportados, cerca de 0.081% das enzimas codificadas por eles foram caracterizadas bioquimicamente e 0.008% tiveram as estruturas das enzimas resolvidas, como indicado na base de dados CaZY.<sup>62</sup>

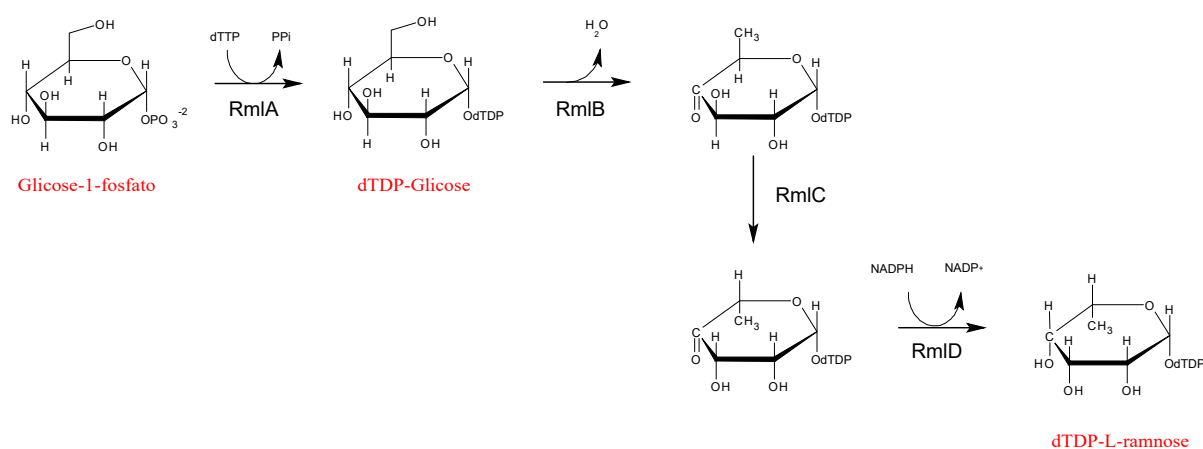
Teng e colaboradores<sup>53</sup> também observaram que a deleção de outros genes do *cluster epa*, incluindo o *epaB* e o *epaE*, diminuía a taxa de sobrevivência celular de *E. faecalis* OG1RF em 87% na presença de leucócitos polimorfonucleares, e triplicava o processo de fagocitose pelos glóbulos brancos. Xu *et al.*, 2000<sup>37</sup> notaram que camundongos os quais recebiam injeções peritoniais com células mutantes  $\Delta epaB$  e  $\Delta epaE$  apresentavam taxa de sobrevivência maior em comparação com os injetados com as células parentais. Já o trabalho de Singh *et al.*, 2009,<sup>58</sup> avaliou que a proliferação de colônias de *E. faecalis* OG1RF no trato urinário de um camundongo era maior do que no mutante  $\Delta epaB$  por um fator de  $10^3$  unidades formadoras de colônias. Estes resultados demonstram o papel de importância biológica que estas duas enzimas do *cluster epa* desempenham, auxiliando na determinação da virulência bacteriana.

A *EpaB* (NCBI ID: EF2197) também é uma GT2,<sup>37</sup> possivelmente mediadora do transporte de ramnose para os polissacarídeos ancorados na parede celular.<sup>57</sup> Já a *EpaE* (NCBI ID: EF2194) é definida como uma glicose-1-fosfato timidiltransferase,<sup>57</sup> identificada como uma *RmlA*<sup>63</sup> ou *RfbA*.<sup>64</sup> A *EpaB* apresenta até 53% de identidade e cobertura de 100% entre as glicosiltransferases de *E. faecium*, e a *EpaE* apresenta 92% de identidade e 100% de cobertura com as glicose-1-fosfato timidiltransferases de *E. faecium*.

### 1.3 RmlA e a Síntese de Ramnose

A enzima *RmlA*, *EpaE* em *Enterococcus faecalis*, integra a via de síntese de dTDP-L-ramnose (Figura 8), juntamente com as enzimas *RmlB*, *RmlC* e *RmlD*. A *RmlA* catalisa a formação de dTDP-glicose a partir de trifosfato de desoxitimidina (dTTP) e glicose-1-fosfato (D-Glc-1-P), a qual é desidratada pela *RmlB* e sofre uma epimerização pela *RmlC*, o produto dessas reações é então reduzido pela *RmlD*, resultando na dTDP-L-ramnose.<sup>61,63</sup> A via da dTDP-L-ramnose é conservada entre bactérias Gram-positivas e Gram-negativas<sup>63,65</sup> para a obtenção de L-ramnose. A L-ramnose, sacarídeo que será ancorado na parede celular, pode constituir o ramnose-glicose polissacarídeo (RGP) de

Gram-positivas<sup>61,66</sup> ou o O-antígeno do lipopolissacarídeo (LPS) de Gram-negativas,<sup>67</sup> por exemplo. Em eucariotos, a biossíntese de L-ramnose é presente em plantas.<sup>68</sup> Em humanos, essa via não foi identificada,<sup>63</sup> tornando as enzimas responsáveis pela síntese de L-ramnose alvos potenciais para o desenvolvimento de candidatos a antibióticos. Embora a L-ramnose se sobressaia como um constituinte celular importante, também foi proposta a possibilidade de que o acúmulo de deoxi-carboidratos por si só impacte negativamente a célula.<sup>66</sup> Foi observado que a redução da massa de biofilme produzido pela bactéria, acarretada pela deleção do gene *rmlA*,<sup>69</sup> codificante da enzima RmlA. Até o momento, há estruturas reportadas de RmlA de 10 organismos.



**Figura 8** – Esquematização da via de síntese de dTDP-L-ramnose, composta pelas enzimas RmlABCD.

Fonte: Adaptada de BROWN *et al.*<sup>70</sup> no ChemSketch™.

Pesquisas envolvendo RmlAs ocorrem desde 1997, e, mesmo assim, novas descobertas ainda são reportadas sobre estas enzimas. Achados recentes indicam a presença do gene *rmlA* na maioria dos isolados produtores de biofilme em *Stenotrophomonas maltophilia*,<sup>71–73</sup> e a RmlA, em conjunto com a RmlC e RmlD, configuram entre as nove proteínas mais significativas no fator de virulência de *Klebsiella pneumoniae* ESBL (*Extended-spectrum  $\beta$ -lactamases*).<sup>74</sup> Dados como estes motivam avanços nos estudos estruturais e funcionais sobre a via de síntese de L-ramnose. Apenas recentemente, o produto final da síntese de ramnose, a dTDP-ramnose, foi identificada como inibidor da RmlA,<sup>75</sup> assim como a presença da inibição não competitiva<sup>76</sup> da RmlA, cujo mecanismo ainda não foi elucidado. Isto, em adição a ausência de síntese e metabolismo da ramnose em humanos, tornam as enzimas RmlABCD atrativos alvos de estudo.





---

## Capítulo 2

### OBJETIVOS

---

O objetivo central deste trabalho é caracterizar estrutural e bioquimicamente enzimas codificadas por genes do *cluster epa* de *E. faecalis*, contribuindo para a ampliação do conhecimento sobre elas. O projeto envolve: **(i)** a determinação das estruturas de enzimas relevantes do *cluster epa* para a resistência a antibióticos, **(ii)** a caracterização funcional/bioquímica e **(iii)** a identificação de potenciais inibidores. Os experimentos desenhados abrangem, mais especificamente:

- Expressão e purificação das enzimas EpaI, EpaOX, EpaB e EpaE de *E. faecalis*, selecionadas pelos efeitos de virulência das mesmas, identificados nos trabalhos da Dale e colaboradores,<sup>39,52</sup> Teng e colaboradores<sup>53,57</sup> e Xu e colaboradores.<sup>37</sup>
- Ensaio de cristalização e co-cristalização com substratos, e estudos estruturais da EpaE, única proteína suficientemente solúvel para estudos estruturais.
- Triagem de potenciais ligantes por métodos computacionais e avaliação experimental.
- Padronização de ensaios de atividade enzimática.
- Aplicação do *AlphaFold2* para predição estrutural das demais enzimas participantes do *cluster epa*, as quais não tiveram estruturas determinadas.



### 3.1 Predição Estrutural das Enzimas do *Cluster* epa

Foi utilizado o *AlphaFold2*<sup>77</sup> para predição estrutural das enzimas do *cluster* epa. As sequências proteicas foram obtidas a partir do *National Center for Biotechnology Information* (NCBI), da cepa ATCC 19433 de *E. faecalis*. O *AlphaFold2* incorpora novas arquiteturas de rede neural e processos de treinamento baseadas nas restrições evolutivas, físicas e geométricas de estruturas proteicas. A predição é analisada a partir do *predicted local-distance difference test* (pLDDT), uma avaliação de quão bem as distâncias interatômicas de uma estrutura proteica são reproduzidas no modelo gerado. O pLDDT também examina a plausibilidade estereoquímica dos resíduos de aminoácidos.<sup>78</sup> Para facilitar a exposição dos resultados, foi realizada a média de pLDDT de cada modelo para expor a acurácia de cada caso. Por fim, a partir do servidor DALI,<sup>79</sup> foram selecionadas estruturas similares às dos modelos gerados pelo *AlphaFold2*, para auxiliar na inspeção de prováveis reações catalisadas pelas proteínas codificadas pelo *cluster* epa, por meio de sobreposições estruturais realizadas no UCSF Chimera.<sup>80</sup>

### 3.2 Clonagem

Para clonagem molecular, foi aplicada a técnica *Ligation-Independent Cloning* (LIC), como descrita em Camilo *et al.*, 2014.<sup>81</sup> As clonagens foram realizadas em vetor pETM11 (seção A.2), pETTrx-1a (seção A.3) e pLIC-EGFP (seção A.5), com os genes epaI (região 980 844-981 569 do NCBI ID: NZ\_KB944590), epaOX (região 1 794 302-1 795 279 do NCBI ID: NC\_017316), epaB (região 22 414-23 196 do NCBI ID: ASDA01000008) e epaE (região 984 201-985 067 do NCBI ID: NZ\_KB944590) de *E. faecalis*, amplificados a partir da biblioteca genômica DSMZ20478, ou ATCC 19433, disponível no Grupo de Biotecnologia Molecular (GBM) do IFSC/USP. o vetor pETTrx-1a co-expressa uma sequência de seis histidinas, visando a técnica de cromatográfica por afinidade em coluna de níquel, assim como a tioredoxina, cuja função é auxiliar na solubilização da proteína de interesse, cujo gene foi clonado neste vetor. O pETM11 é um vetor igual ao pETTrx-1a, os *primers* de amplificação desses dois vetores são os mesmos, só que no vetor pETM11 não há o gene que codifica a tioredoxina. O vetor pLIC-EGFP co-expressa a eGFP (*Enhanced green fluorescent protein*), que ao ser excitada em 488 nm, emite na região do verde. Este último vetor foi utilizado para experimentos com microscopia (seção D.1) envolvendo a localização das enzimas EpaB e EpaB na célula. O gene epaOX, ausente na cepa ATCC 19433, foi sintetizado e adquirido através da empresa Biomatik, já otimizado para a

expressão em *Escherichia coli*. Os *primers* empregados para a amplificação gênica estão mostrados na Tabela 4 da seção A.7, e foram desenhados com auxílio da ferramenta online HTP *oligo designer*.<sup>82</sup> Além das proteínas EpaI e EpaOX, também foram expressas as versões truncadas delas, em que a região C-terminal foi reduzida para resultar nas construções EpaI<sup>130</sup>, EpaI<sup>210</sup>, EpaOX<sup>220</sup> e EpaOX<sup>277</sup>, que possuem os 130, 210, 220 e 277 primeiros aminoácidos das proteínas originais, respectivamente. Isto foi realizado como tentativa de melhorar a solubilização dessas proteínas.

### 3.2.1 Amplificação dos insertos e vetores

Para cada reação de PCR, em conjunto com 40ng de DNA molde e 25 pmol dos *primers*, foram utilizados 1x Tampão *Phusion* DNA polimerase (New England BioLabs), 0.2 μmol dNTPs (Cellco), 0.5 μmol de cada *primer* (*forward* e *reverse*) e 1 unidade de enzima *Phusion* DNA polimerase (New England BioLabs), H<sub>2</sub>O milliQ foi adicionada para volume final de 50 μL. No termociclador (Bio-Rad), o processo de amplificação seguiu as seguintes etapas : 98 °C durante 30s; por 35 repetições, 98 °C durante 10s, durante 30s, a reação foi submetida a uma temperatura de anelamento equivalente à temperatura de *melting* de cada *primer* diminuída de 5 °C, seguido por 72 °C durante 30s/kb; a amplificação acabou depois de 10 min a 72 °C.

Os vetores pETM11 e pETTRx foram linearizados a partir de PCR. Os *primers* utilizados foram: *Forward* 5' TGGCGCCCTGAAAATAAAG 3' e *Reverse* 5' CCGCGTCGGGT-CAC 3'. Para um volume total de reação de 50 μL, adicionou-se 1x de Tampão *Phusion* DNA polimerase (New England BioLabs), 2.5 ng de DNA molde do vetor, 0.5 pmol de cada *primer*, 0.5 μmol do mix de dNTP (Cellco) e 1 unidade da enzima *Phusion* polimerase (New England BioLabs), o restante do volume foi completado com H<sub>2</sub>O milliQ. A reação de PCR ocorreu da seguinte forma: 98 °C durante 30s; por 35 ciclos, 98 °C durante 10s, 65 °C durante 30s para anelamento e 72 °C por 105s; por fim, 72 °C durante 10 min. Todos produtos de PCR foram tratados com 20 unidades da enzima DpnI (New England BioLabs) durante 16 h a 37 °C para remoção do do DNA molde. Executou-se eletroforese em gel agarose 0.8% (m/v) para verificação dos tamanhos de DNA, seguida da purificação dos mesmos, usando o kit *PCR CleanUp System* (Promega). Após a amplificação, os produtos de PCR dos insertos foram tratados com 6 unidades de T4 DNA polimerase (Promega), 1x Tampão T4 DNA Polimerase, 2.5 μmol dATP (Cellco), 4 μmol DTT para cerca de 200 ng de DNA e H<sub>2</sub>O milliQ, adicionada para volume final de 20 μL. O vetor pETM11/pETTRx-1a linear também foi tratado com T4 DNA polimerase (Promega) pela mesma reação, com diferença que foi utilizado 2.5 μmol dTTP (Cellco) e não dATP, para 500 ng de vetor.

### 3.2.2 Anelamento e Transformação bacteriana

Para realização do anelamento entre os insertos e os vetores, os produtos de PCR dos insertos e do vetor, anteriormente tratados com T4 DNA polimerase e seus respectivos desoxinucleotídeos, foram incubados por 30 min a temperatura ambiente, para realizar transformação bacteriana por choque térmico, na relação 1:2.5 ng/ $\mu$ L de inserto para vetor segundo o manual da Promega.<sup>83</sup> As células *E. coli* DH5 $\alpha$  passaram pelo tratamento padrão com CaCl<sub>2</sub> para gerar células competentes.<sup>84</sup> As células *Escherichia coli* DH5 $\alpha$  foram transformadas pelo método de choque térmico.<sup>85</sup> Neste processo, 100  $\mu$ L de células foram utilizados para 100 ng de DNA plasmidial, e após 10 minutos desse conjunto no gelo, durante 1 minuto e 30s, foi feito choque térmico em banho-maria a 42 °C, seguido por incubação no gelo por 5 minutos. Adicionou-se *Luria Bertani* (LB) e as células ficaram a 37 °C a 180 rpm durante 1 h. Por fim, as células foram plaqueadas em meio LB contendo o antibiótico apropriado à resistência do vetor, canamicina neste caso.

Posteriormente à clonagem molecular, realizou-se PCR de colônia das células que cresceram após a transformação bacteriana com DH5 $\alpha$ , para certificação de que o plasmídeos gerados pela fusão dos insertos gênicos com os vetores estavam presente nelas. Em um microtubo de 1 mL, adicionou-se 1 mL de meio LB e 50  $\mu$ g/mL de canamicina. Com uma ponteira estéril, retiraram-se de cada placa colônias individuais, as ressuspendendo em 20  $\mu$ L de meio LB contendo canamicina. A solução de reação foi composta por 0.5  $\mu$ L 10x Tampão Taq DNA polimerase (Cellco), 50 nmol de dNTPs (Cellco), 6.25 pmol dos *primers* utilizados na amplificação de cada gene, 20% (v/v) DMSO, 1  $\mu$ L da colônia que foi ressuspendida em meio LB com canamicina e 1 unidade de Taq DNA polimerase (Cellco). Água milliQ autoclavada foi utilizada para completar o volume de reação em 25  $\mu$ L. O protocolo utilizado para amplificação foi: 95 °C durante 5 min; 30 repetições de 95 °C durante 30 s, 56 °C durante 45 s e 72 °C durante 2 min; finalizando com 72 °C por 10 min. Executou-se eletroforese em gel agarose 0.8% (m/v) para verificação da presença de bandas e dos tamanhos de DNA. Para as colônias positivas, foi efetuada remoção plasmidial pelos kits comerciais *Wizard® Plus SV Minipreps DNA Purification System* (Promega) e *Fast-n-Easy Plasmid Mini-Prep* (Cellco). Os genes inseridos nos plasmídeos foram então sequenciados.

### 3.3 Expressão de Proteínas

Os plasmídeos foram utilizados para transformação em *E. coli* Rosetta (DE3) competentes pelo prévio tratamento com CaCl<sub>2</sub>.<sup>84</sup> A transformação bacteriana foi feita pelo modo supracitado (subseção 3.2.2), adicionando somente o uso do antibiótico cloranfenicol, uma vez que essa cepa possui marcador de resistência para este antibiótico. Com uma colônia de cada placa foi feito pré-inóculo. O pré-inóculo foi utilizado para gerar uma solução estoque de bactéria com glicerol 20% (v/v), armazenada em -80 °C. O restante

do pré-inóculo foi inoculado em 25 mL de meio LB com finalidade de renovação celular. Posteriormente, esses 25mL foram transferidos para 1 L de meio LB, que permaneceu no *shaker* a 37 °C e 180 rpm. Quando a  $DO_{600nm}$  celular atingiu entre 0.5 e 0.8, a expressão foi induzida ao adicionar 0.5 mM de IPTG (Isopropil  $\beta$ -D-1-tiogalactopiranosídeo) ao meio bacteriano. A indução ocorreu durante 18h, a 18 °C e 180 rpm. As células foram então centrifugadas a  $7.9k \times g$  por 30 min a 4 °C, e o *pellet* congelado a -80 °C. Para constatação da expressão, uma alíquota de 1 mL meio LB era coletada antes da adição de IPTG, centrifugada, seu sobrenadante era descartado e o *pellet* resultante armazenado a 4 °C. Após 18h de adição de IPTG ao meio de cultura, outra alíquota era coletada, seguindo o procedimento descrito anteriormente. Cada *pellet* foi ressuspensionado com água, em volume tal que a concentração de células se mantivessem equivalentes a partir da medida da  $DO_{600nm}$ . O *pellet* originado do meio de cultura, ressuspensionado em água, antes da indução é definido aqui como fração não-induzida, e o do meio de cultura após indução, ressuspensionado em água, é denominado fração induzida. As amostras desses dois *pellets* foram analisados por eletroforese de proteínas em gel SDS 15% corado com azul de *Coomassie*.

O plasmídeo pET21\_NESG (seção A.4) foi adquirido através da empresa DNASU, ele contém o gene para a enzima EpaI associada com a sequência de seis resíduos de histidina na posição C-terminal. Neste caso, o protocolo descrito por Rush *et al.*, 2017,<sup>86</sup> foi empregado para expressão e purificação.

Para averiguar o grau de solubilidade das proteínas EpaI e EpaOX, expressas em vetor pETTrx-1a, e da EpaI em vetor pET21\_NESG também, foram selecionadas as cepas de expressão LEMO21 (DE3), C43 (DE3) pLysS, BL21(DE3) pT-GroE, ArcticExpress e Rosetta-gami (DE3) de *E. coli* para produção das proteínas, empregando os protocolos adequados de cada cepa. A célula LEMO21 (DE3) é uma variação da BL21(DE3), a função dela é evitar o vazamento de expressão de proteínas tóxicas ou de membranas, que, em expressão elevada, chegam a ser danosa para a célula ou a formar corpos de inclusão, respectivamente.<sup>87</sup> Essa modulação de expressão é feita por meio da adição de L-ramnose. A célula C43 (DE3) pLysS também é indicada para expressão de proteínas heterólogas que apresentam citotoxicidade. Essa célula impede a expressão basal das proteínas heterólogas antes da indução proteica ao expressar uma pequena quantidade de T7 lisozima a partir de um plasmídeo da própria célula.<sup>88</sup>

A cepa BL21(DE3) pT-GroE contém o plasmídeo pT-GroE que codifica a expressão das chaperonas GroEL (57kDa) e GroES (10kDa)<sup>89</sup> para auxiliar no enovelamento das proteínas recombinantes e evitar corpos de inclusão. A célula ArcticExpress também co-expressa chaperonas, a Cpn10 (10kDa) e a Cpn60 (60kDa) de *Oleispira antarctica*. Essas duas chaperonas possuem ótima atividade de reenovelamento na faixa de temperatura 4 °C - 12 °C,<sup>90</sup> e, como no caso da BL21(DE3) pT-GroE, são aplicadas para casos de proteínas

heterólogas com dificuldade de enovelamento.

A cepa Rosetta-gami (DE3) é uma variação da Rosetta(DE3), além de carregar um plasmídeo que codifica códons raros, ela possui mutação nos genes *trxB* e *gor*, codificantes das proteínas tiorredoxina redutase e glutatona redutase, as quais auxiliam na correta formação de pontes de dissulfeto da proteína heteróloga no citoplasma.

Os testes de expressão foram realizados na plataforma automatizada Freedom EVO 200 (Tecan), segundo o método estabelecido por Camilo *et al.*,<sup>81</sup> com o auxílio da Dra. Livia Manzine e da Pós-Doutoranda do grupo na época, Dra. Luana Morão. Testes subsequentes, com cepas individuais também foram efetuados para a verificação dos resultados.

### 3.4 Testes de Solubilidade e Purificação

A lise celular foi realizada a partir dos *pellets* ressuspensos em tampão 50 mM Tris pH 8.5 (EpaI e EpaE) ou pH 7.5 (EpaOX), 300 mM NaCl e com 10% de glicerol. Essa solução era então submetida à sonicação, com amplitude 10% e um total de 10 ciclos compostos por 20s *ON* e 40s *OFF*. Logo em seguida, as células eram centrifugadas a  $14k \times g$  e 4 °C. Posteriormente à filtração do sobrenadante, o processo de cromatografia por afinidade em coluna de níquel foi realizado. Na bancada, essa primeira etapa de purificação foi realizada com 4 mL de resina de níquel (GE Healthcare). O procedimento consistia em uma passagem inicial de tampão de ligação, contendo a solução de lise complementada com 5 mM de imidazol, seguido pela aplicação da fração solúvel na coluna contendo a proteína a ser purificada. A incubação da proteína com a resina durava 1h sob agitação, e eram realizadas consecutivas lavagens com a solução de lise com adição de concentrações crescentes de imidazol, em um intervalo de 8 mM a 50 mM. A eluição da proteína ocorria em 300 mM e 400 mM de imidazol presentes na solução de lise. Todo o processo, desde expressão proteica à primeira etapa de purificação, foi analisado por eletroforese de proteínas em gel SDS 15% corado com azul de *Coomassie*.

Para as proteínas mais insolúveis, EpaI e EpaOX, a técnica de *refolding* aplicando detergente descrita por He *et al.*, 2017,<sup>91</sup> foi testada. Outros detergentes testados para a purificação das enzimas foram 20 mM CHAPS e 1% Triton X-100. Também foi utilizado tampão fosfato, em pH 8.0 no caso da EpaI, e em pH 7.5 no da EpaOX. Os pHs iniciais para as soluções das proteínas foram determinados a partir de valores maiores do que o pI das mesmas, predito pelo servidor *ProtParam tool*.<sup>92</sup> O pH de maior estabilização para a EpaE foi determinado posteriormente, pela técnica de Fluorimetria Diferencial de Temperatura. A proteína EpaE foi concentrada até volume de solução 5 mL, para então ser injetada na coluna *HiLoad 26/60 Superdex 200*, previamente equilibrada com 2 volumes de coluna de tampão 50 mM pH 8.5, 150 mM NaCl. A partir do gráfico de absorção em 288 nm por volume (mL) eluído da proteína, as frações de eluição foram coletadas e analisadas por eletroforese de proteínas em gel SDS 15% corado com azul de *Coomassie*. Para esta

finalidade, usou-se o GE AKTA *Purifier* e o programa UNICORN.

### 3.5 Teste de Cristalização e Coleta de Dados de Difração

A enzima EpaE foi concentrada até 10 mg/mL e levada a testes de cristalização. Isto foi auxiliado pelo robô de cristalização *HoneyBee* 963 (DigiLab), operado pelo técnico João Possato. Os kits comerciais aplicados foram: *Morpheus Screen*, *Structure Screen 1 + 2* HT-96 - MD1-30, *PACT Premier* HT-96 - MD1-36, *Morpheus Additive Screen* e *Morpheus III* - MD116-117, todos da empresa Molecular Dimensions. Para a cristalização, utilizou-se a técnica de difusão de vapor em gota sentada em placas de 96 poços, e as variações da proporção entre proteína e solução de cristalização (solução do poço) de 1:2, 1:1 e 2:1. Todas as placas foram mantidas a 18 °C, monitoradas pelo *Rock Maker*<sup>web</sup>. Experimentos de difração foram realizados com o Prof. Dr. Alessandro Nascimento no difratômetro Rigaku Ultra X-18 e, posteriormente, na linha Manacá do novo anel Síncrotron Brasileiro, o Sírius. Um conjunto completo foi coletado, com comprimento de onda 1.324 nm, contendo 3600 imagens, com uma oscilação de 0.1 grau por imagem, totalizando 360 graus coletados.

### 3.6 Determinação Estrutural

O programa XDS-GUI<sup>93</sup> foi aplicado para indexação e integração dos dados. Em seguida, para o escalonamento, usou-se o AIMLESS.<sup>94</sup> A análise da simetria sugeria inicialmente o grupo espacial  $P22_12_1$ . A enzima EpaE apresenta até 70% de identidade com algumas estruturas depositadas no PDB, então adotou-se o faseamento por substituição molecular através do programa PHASER-MR.<sup>95</sup> A estrutura modelo para esse procedimento foi gerada a partir da modelagem das estruturas PDB ID: 4ho0, PDB ID: 1iim e PDB ID: 3pkp do PDB no MODELLER.<sup>96</sup> O modelo foi então reconstruído no PHENIX AUTOBUILD,<sup>97</sup> e o refinamento no espaço recíproco nos programas PHENIX.REFINE<sup>98</sup> e REFMAC.<sup>99</sup> A inspeção visual foi realizada no Coot,<sup>100</sup> e o modelo final foi validado no *MolProbability*.<sup>101</sup>

### 3.7 Triagem de Potenciais Ligantes

A seleção das moléculas foi realizada a partir da similaridade estrutural com a molécula dTTP, pelo LiBELa,<sup>102–103</sup> utilizando as seguintes bases de dados: ChEBI, ChEMBL21, DBALL, HMDBDrug, HMDBMicrobe, NPACT, NuBBE, PDB, TargetMolNP e UEFS. Todas as moléculas selecionadas possuíam uma similaridade de ao menos 70% com o dTTP, calculada pelo índice de *Hodgkin*.<sup>104</sup> As moléculas foram então submetidas a inspeção visual. No final, oito moléculas foram selecionadas para compra: Hidrobenztiazida, Benzotiazida, Doripenem mono hidratado, Toltrazuril, Rimsulfuron, Bicalutamida, 4-amino-N-[4-(aminosulfonil)] e Piridoxal fosfato. Dessas, apenas três foram aplicadas experimentalmente, devido a solubilidade em água em quantidade suficiente para testar a



interação entre o ligante e EpaE na relação 3:1 de massa. As três são: Benztiázida (Merck ID: B7149), Bicalutamida (Merck ID: BP1115) e Doripenem mono hidratado (Merck ID: 32138).

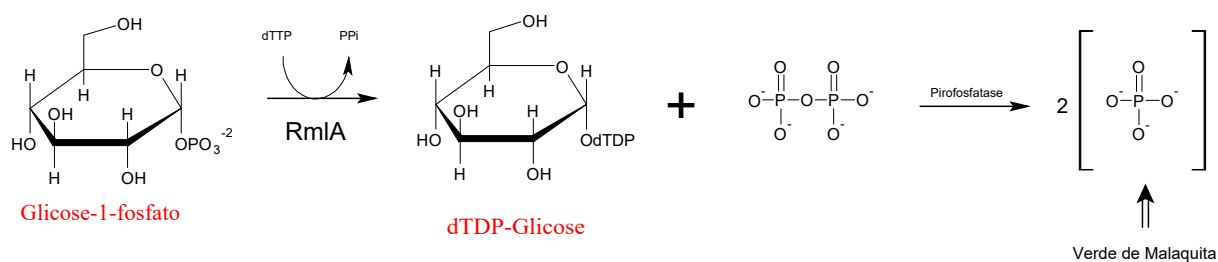
Para verificação rápida da interação da EpaE por parte dos compostos selecionados, foi aplicada a técnica Fluorimetria Diferencial de Temperatura,<sup>105</sup> que analisa a temperatura de desenovelamento proteico em diferentes condições na qual a proteína foi exposta. Compostos que geram Temperatura de *Melting* ( $T_m$ ) maior do que a da proteína em tampão indicam algum tipo de interação estabilizante da estrutura da proteína EpaE. A preparação da amostra consiste em 11 nM da EpaE, 4.2x da sonda JBS *ThermoFluor Dye* (Cellco) e tampão contendo os ligantes, para volume final de solução de 20  $\mu$ L. O Doripenem foi diluído em água, os testes com ele ocorreram nas concentrações 3.2  $\mu$ M. Já a Benztiázida e a Bicalutamida precisaram ser diluídas em DMSO, seguido por diluição em água. Os testes com a Benztiázida ocorreram nas concentrações 0.5% DMSO 24 nM, 2.5% DMSO 0.2  $\mu$ M e 0.5% DMSO 29 nM, e, com a Bicalutamida, nas concentrações 0.5% DMSO 81 nM, 2.5% DMSO 8.1  $\mu$ M e 0.5% DMSO 69 nM. As medidas foram realizadas no CFX *Connect Real-Time PCR System* (Bio-Rad), feitas de 25 °C a 95 °C, com as medidas realizadas após 30 s em cada temperatura deste intervalo. Também foi determinado o  $T_m$  da EpaE na presença de 5 mM  $MgCl_2$ , 0.2 mM dTTP e 1 mM D-Glc-1-P.

### 3.8 Teste de Atividade Enzimática

#### 3.8.1 Reagente Verde de Malaquita

Esta metodologia foi adaptada a partir dos trabalhos de Sha *et al.*, 2012,<sup>106</sup> Sherwood *et al.*, 2013,<sup>107</sup> Fisher *et al.*, 1994,<sup>108</sup> e Baykov *et al.*, 1988.<sup>109</sup> O método é baseado na formação de um complexo entre o reagente de malaquita e o grupo fosfato ( $PO_4^{-3}$ ) livre em solução. A quantidade desses complexos em solução é diretamente proporcional a  $DO_{630nm}$  medida.<sup>106</sup> Sendo assim, quando um dos produtos da reação entre a EpaE com D-Glc-1-P (D-Glicose-1-Fosfato) e dTTP for formado, o pirofosfato (PPi), a adição de uma pirofosfatase promove hidrólise do PPi, e a atividade da enzima pode ser indiretamente medida através da concentração de  $PO_4^{-3}$  no meio (Figura 9).

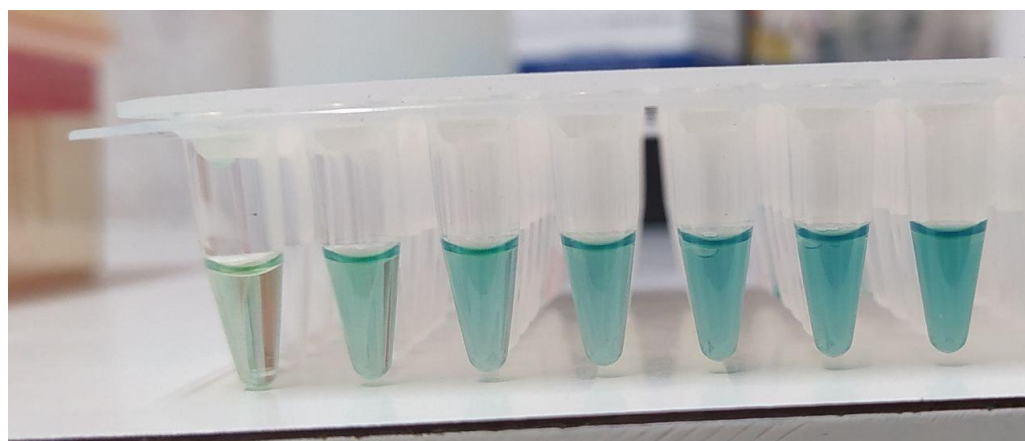
O reagente verde de malaquita utilizado para análise cinética da enzima EpaE era composto de 0.03% m/v verde de malaquita, 1N HCl e 0.05% Triton X-100 dissolvidos em água milli-Q. O tampão de reação continha 50 mM Tris pH 8.5, 1 mM DTT, 10% Glicerol, 5 mM  $MgCl_2$ , 0.2 mM dTTP, 1 mM D-Glc-1-P e 0.1 unidades de pirofosfatase (Thermo Fisher Scientific), para volume final de 50  $\mu$ L contando com a enzima. A quantidade de pirofosfatase foi tal que estivesse em excesso e não fosse um fator limitante da reação.<sup>106</sup> O tampão e a enzima foram incubados a 37 °C durante (min): 0, 1, 2, 3, 4, 5 e 6 minutos no termociclador (Bio-Rad). Em seguida, eram adicionados 50  $\mu$ L do reagente verde de malaquita à solução de reação, e o conjunto era incubado a 37 °C durante 5 min no



**Figura 9** – Interação entre o verde de malaquita e o grupo fosfato. Após a reação da RmlA, o PPi é hidrolisado pela pirofosfatase, gerando radical fosfato inorgânico, que é detectado pelo reagente verde de malaquita.

Fonte: Adaptada de SHA *et al.*<sup>106</sup> no ChemSketch™.

termociclador (Bio-Rad), seguido pela transferência para uma placa de leitura transparente de fundo plano de 96 poços e leitura no espectrofotômetro Multiskam Spectrum (Thermo Fisher Scientific) em 630 nm.



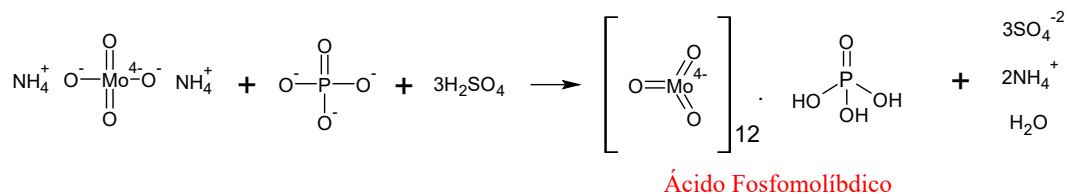
**Figura 10** – Intensidade do reagente verde de malaquita proporcional à concentração de fosfato na solução. Concentrações crescentes de  $\text{KH}_2\text{PO}_4$  da esquerda para direita.

Fonte: Elaborada pela autora.

Uma curva de calibração foi previamente criada a partir da diluição de uma solução 1 mM de  $\text{KH}_2\text{PO}_4$  em água milli-Q, no tampão de reação, nas concentrações (mM): 0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06 e 0.07. Após 0.07 mM  $\text{KH}_2\text{PO}_4$ , a reta perdia sua linearidade. Para cada 50  $\mu\text{L}$  de solução  $\text{KH}_2\text{PO}_4$  em tampão de reação, foram adicionados 50  $\mu\text{L}$  do reagente verde de malaquita para incubação a 37°C durante 5 minutos no termociclador (Bio-Rad). Cada medida foi repetida 3 vezes durante o mesmo experimento. Nas medidas relacionadas à cinética da EpaE, foram feitas triplicadas para cada tempo de incubação, e o valor do branco (apenas tampão de reação, excluindo a pirofosfatase) foi descontado do valor medido para cada caso. Para a realização do gráfico da concentração de PPi ( $\mu\text{M}$ ) por tempo da reação (min.), é considerada a relação 1:2 entre o PPi e o grupo fosfato.

### 3.8.2 Método de Análise de Fosfato

No Grupo de Biotecnologia Molecular, são realizados testes enzimáticos com fitases, utilizando o método de análise de fosfato. Esta metodologia também é baseada na quantificação de radicais fosfatos liberados após reação enzimática, a absorvância da solução colorimétrica utilizada é proporcional à quantidade de fosfato no meio. Este protocolo é baseado no trabalho de Chen *et al.*, 1956,<sup>110</sup> e foi disponibilizado pelo Pós-doutorando do grupo na época, o Dr. Vinícius Ramos. A preparação da amostra seguiu a mesma descrita na subseção 3.8.1, tanto para reação enzimática quanto para calibração, exceto no intervalo de concentração do  $\text{KH}_2\text{PO}_4$  dissolvido em apenas água milliQ, e no pH da solução, que, neste caso, foi pH 7.2. A solução colorimétrica do método análise de fosfato era composta pela relação 1:1:3 de 10% ácido ascórbico, 2.5% molibdato de amônio e 1M  $\text{H}_2\text{SO}_4$ .<sup>111</sup> Antes de adicionar a solução colorimétrica, a de reação era aquecida a 80 °C durante 20 s, para desnaturar a EpaE e cessar a atividade desta. Esse valor foi estipulado a partir dos experimentos de Fluorimetria Diferencial de Temperatura, envolvendo a estabilidade da EpaE na presença dos substratos e co-fator (Apêndice C). Após adição da solução colorimétrica na de reação, o conjunto era incubado a 50 °C durante 10 min, a medida era realizada em 830 nm no espectrofotômetro Multiskam Spectrum (Thermo Fisher Scientific). O molibdato de amônio, em meio ácido, reage com fosfato para formar ácido fosfomolibdico (Figura 11),<sup>112</sup> um íon do tipo Keggin.<sup>113</sup> Quando o ácido fosfomolibdico é reduzido, aqui, por ácido ascórbico, uma solução com coloração azul é gerada, chamada azul de molibdênio.<sup>112–113</sup> A incubação a 50 °C ajuda a estabilizar o azul de molibdênio.



**Figura 11** – Formação do ácido fosfomolibdico.

Fonte: Elaborada pela autora no ChemSketch <sup>TM</sup>.



#### 4.1 Predição Estrutural das Enzimas do *Cluster epa*

O *cluster epa* ainda não possui qualquer tipo de caracterização estrutural, pouco se sabe a despeito da sua importância na manutenção da homeostase da bactéria, da sua atribuição de virulência bacteriana e da sua capacidade de desenvolver resistência a alguns antibióticos.

Durante o período de desenvolvimento desta tese, a ferramenta *AlphaFold2* foi desenvolvida, validada e tornada pública. Os dados mostrados revelaram que a ferramenta é o melhor preditor de proteínas, capaz de lidar com a falta do contexto físico e produzir modelos precisos, como ficou evidente no CASP14.<sup>77</sup>

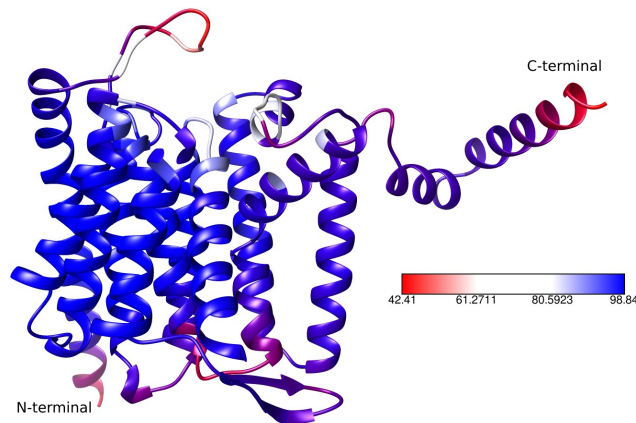
Neste contexto, a predição estrutural de todo o *cluster epa* foi realizado através da ferramenta *AlphaFold2*. Os modelos gerados foram submetidos ao banco de dados DALI<sup>79</sup> para buscar homólogos estruturais, a fim de auxiliar na caracterização das enzimas Epa. Nas seções seguintes, serão feitas breves apresentações da relação estrutura-função obtida a partir da análise destes modelos.

Neste capítulo, para fins de comparação, as imagens das predição estruturais estarão sempre ao lado das imagens das proteínas depositadas no PDB mais similares a elas, ressaltando os resíduos do sítio catalítico dessas para identificação dos potenciais resíduos da cavidade catalítica dos modelos estruturais. Quando necessário, um ● marcará resíduos diferentes entre a predição estrutural e a estrutura presente no PDB. Os resíduos das predições estão sempre em dourado, e os das estruturas depositadas no PDB, em azul, em roxo ou em verde.

#### 4.2 EpaA

O modelo estrutural da EpaA (NCBI ID: EOL33622) tem pLDDT médio de 87, medida de 0 a 100, com 100 indicando maior confiabilidade no modelo gerado. As regiões de menor confiabilidade encontram-se nas extremidades da proteína, como mostrado na Figura 12. O modelo da EpaA apresenta feixes de hélices consecutivas, seu homólogo mais próximo é anotado como uma proteína de membrana. A partir da sequência da EpaA, o programa DeepTMHMM<sup>114</sup> prediz que esta também é uma proteína de membrana.

O modelo da EpaA é similar à estrutura da UDP-GlcNAc:dolicol-fosfato GlcNAc-1-fosfato transferase (DPAGT1) (EC 2.7.8.15) de *Homo sapiens*, PDB ID: 6fwz, apresentando

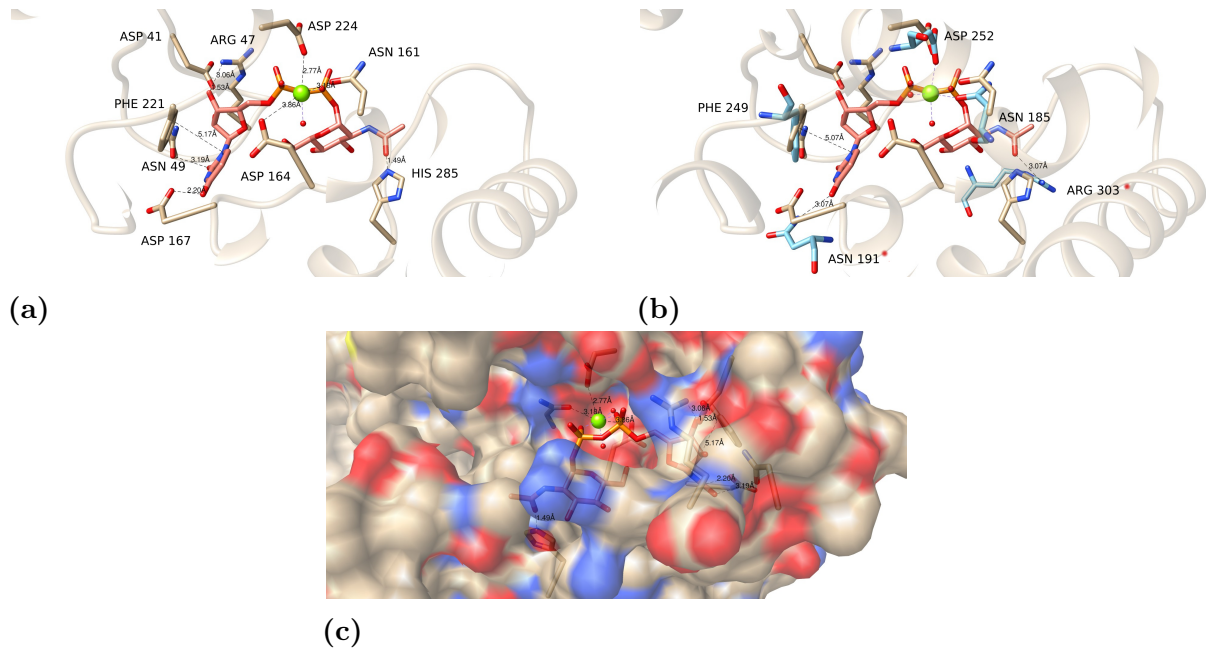


**Figura 12** – Modelo estrutural da EpaA, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a suposta região catalítica da EpaA.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

Z-score 32.2, RMSD 2.5 e 19% de identidade. DPAGT1 é uma proteína de membrana que transfere uma unidade de N-acetil-D-glicosamina-1-fosforil (GlcNAc-1-P) do UDP-N-acetilglicosamina (UDP-GlcNAc) para o dolicol fosfato (Dol-P). Esta reação configura a primeira etapa da N-glicosilação de resíduos de asparagina.<sup>115</sup> O modelo da EpaA e a DPAGT1 tem bom alinhamento na região do sítio catalítico desta última, onde está a UDP-GlcNAc. Considerando que o modelo gerado está na forma não relaxada, as distâncias podem divergir em maior ou menor grau das reais. Há 9 potenciais resíduos que estabilizam a UDP-GlcNAc (Figura 13). Com base nesse alinhamento dos resíduos da EpaA com a PDB ID: 6fwz (Figura 13b), os resíduos catalíticos em potencial da EpaA são: o PHE221, em empilhamento pi com o anel uracila da UDP-GlcNAc, e o grupo hidroxila da ASP167 em ligação hidrogênio com o oxigênio desse mesmo anel; ASP41 com o oxigênio do anel uracila; o nitrogênio do grupo guanidina da ARG47 e a hidroxila do ASP41 em ligação hidrogênio com uma hidroxila da ribose do UDP; o grupo amina da HIS285 com o oxigênio do grupo amida da N-acetilgalactose, e ASN161, ASP224 e ASP164 interagindo com a molécula de magnésio. A ASP164, aqui, substituindo uma molécula de água, como visto na Figura 13b. A cavidade catalítica da EpaA está representada na Figura 13c.

Gueradel e colaboradores<sup>55</sup> realizaram análises bioquímicas do conteúdo presente no polissacarídeo da parede celular de *E. faecalis* V583, assim como compilação dos dados disponibilizados até então das enzimas do *cluster* epa desta cepa, esquematizando a provável via de síntese do EPA. Tal qual a EF2198 da cepa V583, a EpaA também é homóloga da enzima TagO (NCBI ID: NP\_391433), a iniciadora de formação do polímero de ramnose.<sup>55,116</sup> As sequências da EpaA e TagO apresentam 39.06% de identidade e E-value  $1.e^{-84}$ , em 99% de cobertura. A TagO adiciona GlcNAc-1-P ao transportador lipídico undecaprenil-fosfato (Und-P), a partir do UDP-GlcNAc.<sup>55</sup> O resultado é a formação de GlcNAc-P-P-Und. Os dados obtidos a partir da modelagem e das informações disponíveis



**Figura 13** – Modelo estrutural da (a) EpaA, em dourado, apresentando os potenciais resíduos de interação com a DPAGT1 de *Homo sapiens* (PDB ID: 6fwz), a partir do (b) alinhamento do modelo com a estrutura da enzima DPAGT1. Estão destacados com um ● os resíduos diferentes entre a predição estrutural da EpaA e a DPAGT1. (c) Suposta cavidade catalítica da estrutura predita da EpaA.

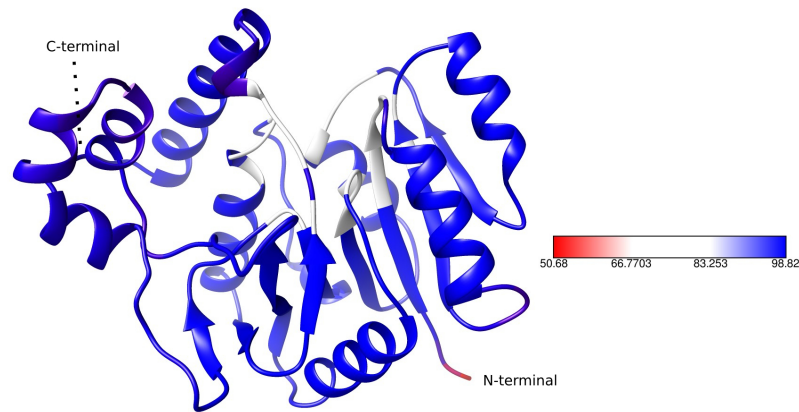
Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

da EpaA convergem, sendo possível inferir a ela uma atividade de transferência de um fosfo-açúcar a um carreador lipídico.

### 4.3 EpaB

O modelo estrutural da EpaB (NCBI ID: EOL33621) tem pLDDT médio de 95. A estrutura predita apresenta altos b-fatores na maioria dos resíduos, possuindo classe do tipo complexo  $\alpha$ - $\beta$  (Figura 14).

O modelo EpaB é mais similar à estrutura da polipeptídeo N-acetil- $\alpha$ -galactose aminotransferase (ppGalNAc-T2) (EC 2.4.1.41) de *H. sapiens*, PDB ID: 5ndf, uma das iniciadoras da glicosilação O-GalNAc em resíduos de treonina e serina. O alinhamento entre a EpaB e ppGalNAc-T2 apresenta *Z-score* 22.7, RMSD 2.5 e 16% de identidade. A sobreposição das estruturas ocorre na região catalítica da ppGalNAc-T2. Esta enzima transfere o grupo N-Acetil-galactosamina do UDP-N-acetil- $\alpha$ -D-galactosamina para um resíduo de serina ou treonina, usando  $Mn^{+2}$  como cofator.<sup>117</sup> Já que o modelo da EpaB foi sobreposto com a estrutura 5ndf (Figura 15), co-cristalizado com um dos produtos da reação, o UDP, foi possível identificar potenciais resíduos que interagem com essa molécula, são eles: ASN41 com o grupo amina do UDP, a TYR12 em empilhamento pi com

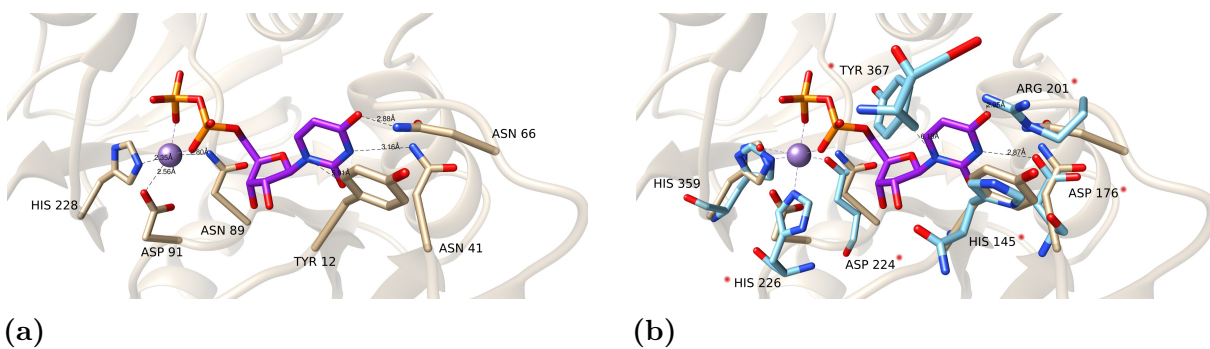


**Figura 14** – Modelo estrutural da EpaB, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a provável região catalítica da EpaB.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

o anel uracila, ASN66 em ligação hidrogênio com o oxigênio da uracila e, possivelmente a ASN89, o ASP91 e a HIS228 do modelo da EpaB façam o papel da HIS226, ASP224 e HIS359 da ppGalNAc-T2, respectivamente, na interação como o  $Mn^{+2}$ , estabilizador do grupo pirofosfato do UDP (Figuras 15a e 15b). Não há resíduo equivalente ao TYR367 da ppGalNAc-T2 na predição estrutural da EpaB.

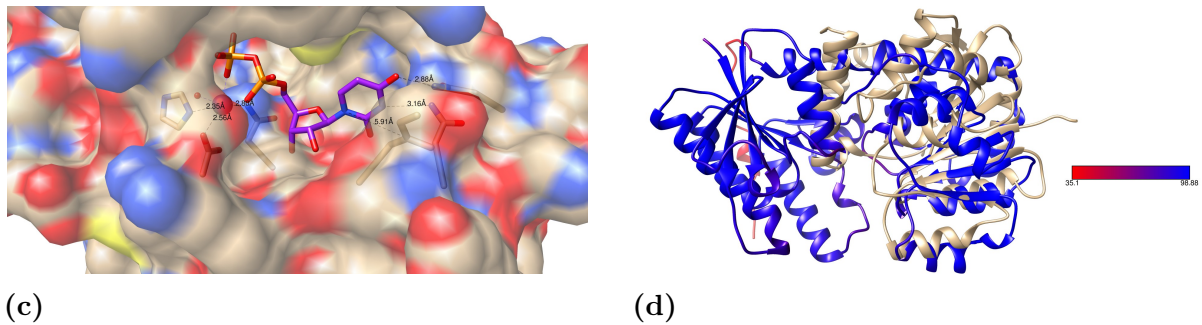
Em bactérias que apresentam cadeias de ramnose ancoradas, é comum a presença de um loci cromossomal denominado *rgp*, próximo do loci *rml*.<sup>116</sup> Os genes *rgp* codificam algumas ramnosil e glicosiltransferases. O trabalho de Guerardel *et al.*, 2020,<sup>55</sup> propõe que a EpaB em *Enterococcus faecalis* V583 adicione o primeiro resíduo de ramnose ao GlcNAc-P-P-Und, a partir da dTDP-ramnose sintetizada pelas enzimas EpaEFGH. Em bactérias *Lactococcus lactis*, cujo conhecimento sobre a biossíntese de ramnose é mais estabelecido, essa função é exercida por uma ramnosiltransferase, a RgpA (NCBI ID: BDH82874), a qual não possui identidade sequencial relevante com nenhuma enzima codificada pelo *cluster epa*, mas cuja predição estrutural se alinha com um sanduíche  $\alpha - \beta$  da predição EpaB (Figura 15d). Experimentos realizados na identificação os açúcares pre-



(continua)



(continuação)



**Figura 15** – Modelo estrutural da EpaB (a), em dourado, apresentando os potenciais resíduos de interação com o UDP a partir do (b) alinhamento do modelo com a estrutura da enzima ppGalNAc-T2 (PDB ID: 5ndf), em azul. Estão destacados com um ● os resíduos diferentes entre a predição estrutural da EpaB e a ppGalNAc-T2. (c) Cavidade catalítica do modelo EpaB. (d) Sobreposição das estruturas preditas da EpaB, em dourado, e RgpA, de *Lactococcus lactis*, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

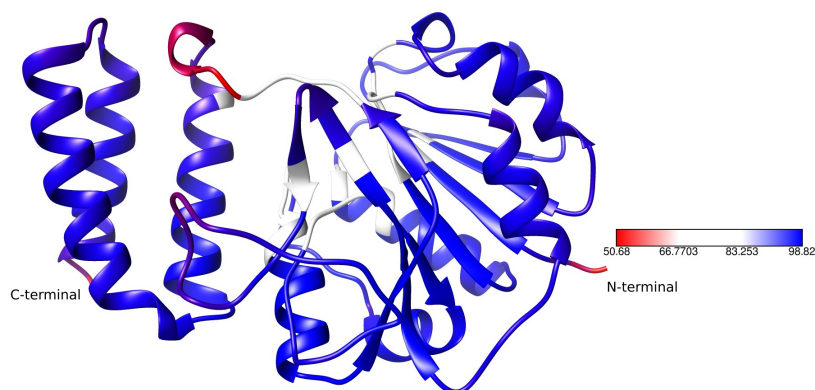
Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

sentos no polissacarídeo antígeno de *Enterococcus* afirmam que o polímero ancorado no peptidoglicano de *Enterococcus faecalis* é constituído apenas de ramnose, glicose e N-acetilglicose.<sup>55</sup> Sendo assim, há uma divergência entre os dados disponibilizados até então com as informações obtidas aqui, fundamentada na similaridade estrutural do modelo predito da EpaB e estruturas depositadas no PDB.

#### 4.4 EpaC

O modelo estrutural da EpaC (NCBI ID: EOL33620) possui pLDDT médio de 93, exceto por uma volta terminal da hélice  $\alpha$  e alguns *loops*, todos os resíduos possuem pLDDT alto (Figura 16). A EpaC apresenta feixes de hélice e sanduíche  $\alpha - \beta$ .

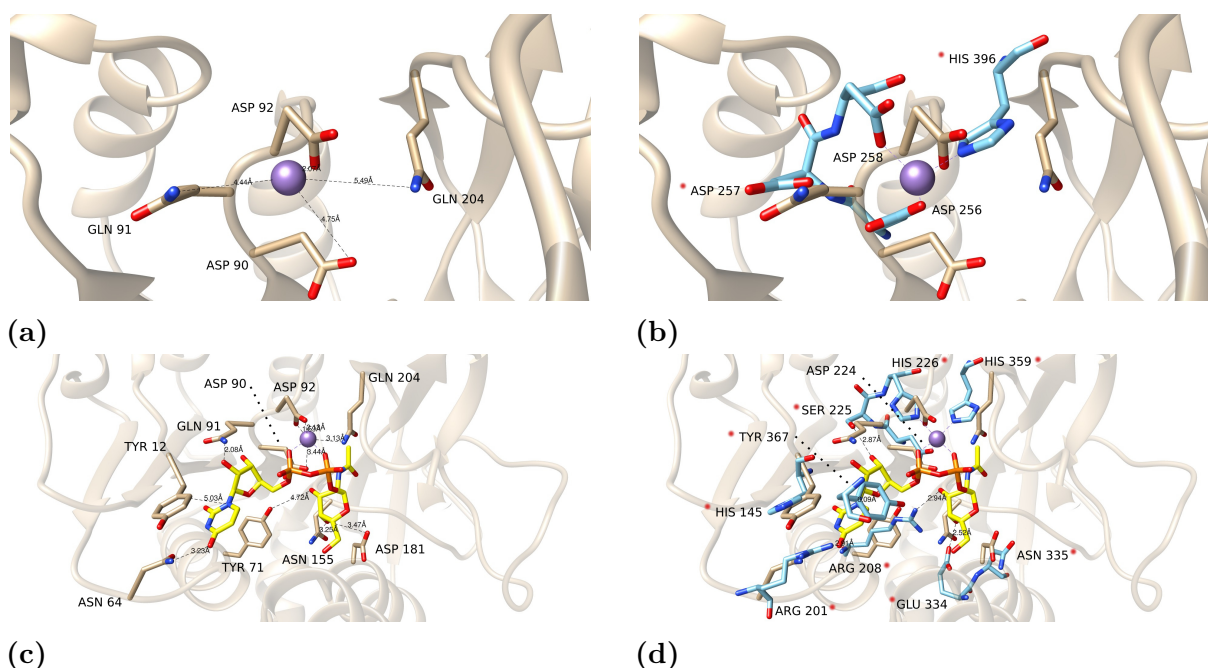
A estrutura predita da EpaC é mais similar à da galacto-furanose transferase (GlfT2) (EC 2.4.1.288) de *Mycobacterium tuberculosis*, PDB ID: 4fix, com *Z-score* 19.8, RMSD 3.0 e 10% de identidade. A sobreposição das estruturas ocorre na região catalítica da GlfT2. Esta enzima, juntamente com a GlfT1, adiciona a porção galactano ao micolarabinogalactano, componente da parede celular do gênero *Mycobacterium*.<sup>118</sup> A partir do UDP- $\alpha$ -D-galactofuranose (UDP-Galf), a GlfT1 inicia o processo acrescentando duas moléculas Galf a um oligossacarídeo ligado a poliprenol, e a GlfT2 produz ligações  $\beta(1 \rightarrow 5)$  e  $\beta(1 \rightarrow 6)$  entre os resíduos de Galf, anexando os 30 monossacarídeos restantes.<sup>118</sup> Visto que a GlfT2 (PDB ID: 4fix) formou cristal apenas com  $Mn^{+2}$ , foi possível apenas identificação dos potenciais resíduos que interagem com este cátion, que são: ASP90, GLN91, ASP92 e GLN204 (Figura 17a), alinhados, respectivamente, com ASP256, ASP257, ASP258 e



**Figura 16** – Modelo estrutural da EpaC, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a provável região catalítica da EpaC.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

HIS395 da GltT2.



**Figura 17** – Modelo estrutural da (a) EpaC, em dourado, sobreposto com a GltT2 (PDB ID: 4fix), em azul, apresentando os potenciais resíduos de interação com o  $Mn^{+2}$ . (b) Resíduos da GltT2 que interagem com o  $Mn^{+2}$ . Potenciais resíduos catalíticos da (c) EpaC, em dourado, identificados pela sobreposição destes com os (d) resíduos da GalNAc-T2 (PDB ID: 4d0t), em azul. Estão destacados com um • os resíduos diferentes entre a predição estrutural da EpaB com a GltT2 e a GalNAc-T2.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

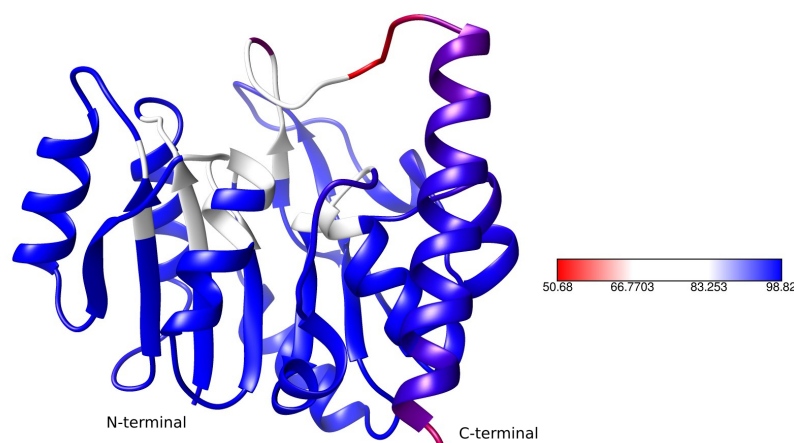
O modelo estrutural predito da EpaC também é similar à polipeptídeo GalNAc-transferase 2 (GalNAc-T2) (EC 2.4.1.41) de *H. sapiens*, PDB ID: 4d0t, co-cristalizada com UDP-GalNAc. A GalNAc-T2 apresenta *Z-score* 17.4, RMSD 3.1 e 13% de identidade

com a estrutura predita da EpaC, a sobreposição das estruturas ocorre na região catalítica da GalNAc-T2. Esta enzima transfere um resíduo de GalNAc de uma UDP-GalNAc para cadeias laterais Ser/Thr e, assim, inicia a O-glicosilação de algumas proteínas.<sup>119</sup> Neste caso, a sobreposição do modelo da EpaC com a GalNAc-T2 (PDB ID: 4d0t) é vista na região catalítica desta. A partir desse alinhamento estrutural, os potenciais resíduos catalíticos da EpaC (Figura 17c) são: GLN204, ASP90, ASP92 podem estar envolvidos na estabilização do íon  $Mg^{+2}$ , enquanto que ASN64 e TYR 12 em interações com o anel pirimídico da UDP-GalNAc, em ligação hidrogênio e empilhamento pi, respectivamente. A GLN91 pode realizar uma ligação hidrogênio com um dos oxigênios do anel imidazólico da UDP-GalNAc, uma vez que o nitrogênio está a 2.08 Å do oxigênio do anel, entretanto, não é observado um resíduo similar à GLN91 na GalNAc-T2 (PDB ID: 4d0t), somente uma SER225 (Figura 17d). Os resíduos ASP181 e ASN155 estão próximos da hidroxila do C-4 da N-acetil-galactose.

Dentre as estruturas do PDB, várias eram similares às estruturas preditas da EpaB e da EpaC, sendo possível compartilharem a mesma função biológica na célula. O trabalho de Guerardel *et al.*, 2020<sup>55</sup> propõe que além da EpaB, as EpaCDNO participam do alongamento da cadeia de ramnose.

#### 4.5 EpaD

O modelo estrutural da EpaD (NCBI ID: EOL33619) apresenta pLDDT médio de 93. Essa predição apresenta um longo *loop* com pLDDT baixo (Figura 18). O enovelamento da EpaD apresenta um sanduíche  $\alpha$ - $\beta$ .

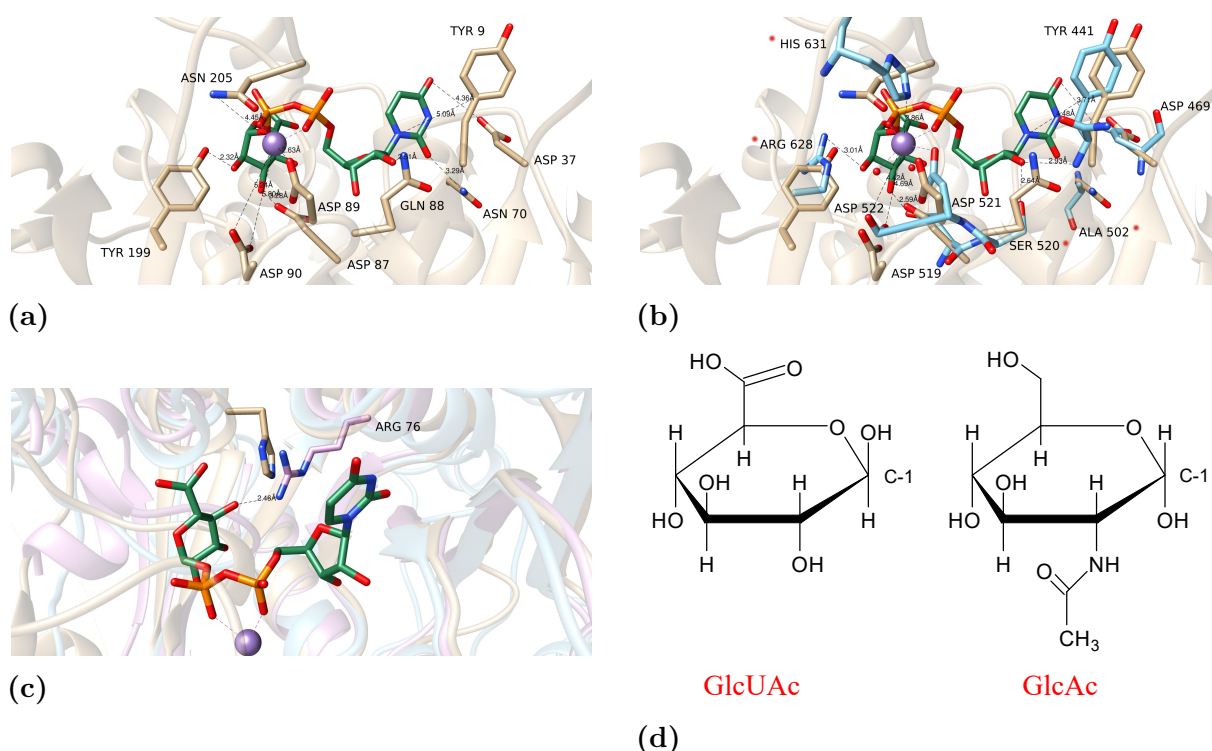


**Figura 18** – Modelo estrutural da EpaD, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a suposta região catalítica da EpaD.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

A estrutura predita da EpaD é mais similar à da condroitina polimerase (CP) (EC 2.4.1.288 e 2.4.1.226) de *E. coli*, PDB ID: 2z86, com *Z-score* 25.3, RMSD 2.6 e 23% de

identidade. A sobreposição das estruturas ocorre em uma das regiões catalítica da CP, a que transfere ácido glicurônico (GlcUA) de uma UDP-ácido glicurônico (UDP-GlcUA) a uma cadeia de condroitina.<sup>120</sup> Os resíduos do modelo EpaD (Figura 19a) que podem interagir com a UDP-GlcUA, baseada na sobreposição com a proteína CP (Figura 19b) são: ASP37, ASN70 e GLN88 estabelecendo interações hidrogênio com a uridina, e TYR9, numa interação empilhamento pi com o anel pirimídico. A TYR199 pode realizar ligação de hidrogênio com a hidroxila do carbono C-4 da GlcUA. As ASP87, ASP89, ASP90 e ASP205 estariam estabilizando o íon  $Mn^{+2}$ , função das ASP219, ASP521, ASP222, e HIS631 na estrutura da CP (PDB ID: 2z86).



**Figura 19** – Modelo estrutural da (a) EpaD, em dourado, sobreposto com (b) a estrutura da condroitina polimerase (PDB ID: 2z86), em azul, apresentando os potenciais resíduos de interação com a UDP-GlcUA, em verde, e íon  $Mn^{+2}$ , em roxo. Estão destacados com um ● os resíduos diferentes entre a predição estrutural da EpaD e a CP. (c) Sobreposição do modelo EpaD, CP (PDB ID: 2z86) e GalT1 (PDB ID: 5hea), em roxo. (d) Representação do Ácido Glicurônico (GlcUA) e N-Acetilglicosamina (GlcNAc). A ligação com um difosfato de nucleotídeo ocorre entre um fosfato e o oxigênio da hidroxila do carbono C-1.

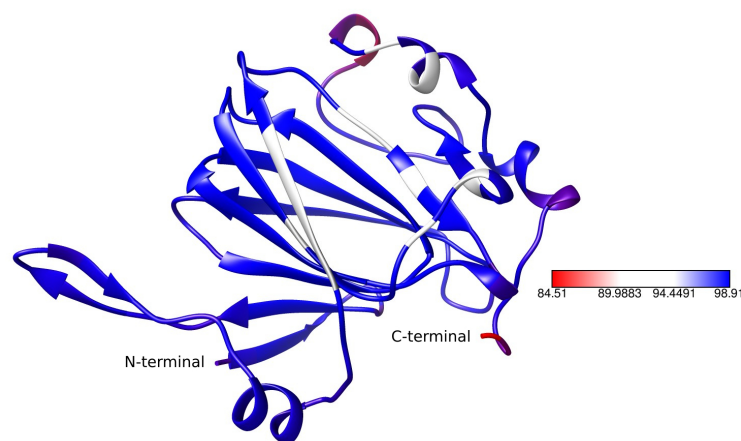
Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

O modelo da EpaD também é semelhante à estrutura da GalT1, citada na seção 4.4, apenas usando galactose como substrato. Uma das estruturas GalT1 é a PDB ID: 5hea, a qual possui *Z-score* 22.6, RMSD 2.7 e 23% de identidade quando estruturalmente comparada ao modelo predito da EpaD. Esta GalT1 não foi co-cristalizada com nenhuma molécula, mas utiliza como substrato UDP-GlcNAc ou UDP-Glc para transferir glicose a

glicoproteínas ricas em serina, ou *Serine-rich repeat glycoproteins* (SRRPs).<sup>121</sup> Na estrutura da GalT1 (PDB ID: 5hea), há a ARG76, capaz de formar uma interação hidrogênio com a hidroxila do carbono C-4 da molécula UDP-GlcUA, co-cristalizada com a proteína PDB ID: 2z86. Caso a EpaD fosse uma GalT1, seu substrato poderia ser GlcNAc. A GlcNAc é similar à GlcUA (Figura 19d), então caso a GlcNAc fosse co-cristalizada com a PDB ID: 5hea, essa mesma ARG76 poderia estar em uma interação de hidrogênio com a hidroxila do carbono C-4 da GlcNAc. O resíduo HIS175 do modelo da EpaD potencialmente agiria de forma semelhante à ARG76 da PDB ID: 5hea (Figura 19c). Dentre os açúcares encontrados no polissacarídeo Epa, estão glicose (Glc), galactose (Gal), ramnose (Rha), GlcNAc e GalNAc, o GlcUA não está presente.<sup>54</sup> Sendo assim, é mais viável que o substrato compartilhado entre EpaC e EpaD, na hipótese que são GlftT2 e GlftT1, respectivamente, seja um difosfato de nucleotídeo ligado a ou GlcNAc ou GalNAc. Como mencionado, o trabalho de Guerardel e colaboradores, 2020,<sup>55</sup> considera que a EpaD também tenha como substrato a dTDP-ramnose, a qual também é capaz de se ajustar no potencial sítio catalítico da EpaD. Outra possibilidade seria a EpaB, EpaC e EpaD desempenharem a mesma atividade catalítica, as predições estruturais delas possuem enovelamento parecido e se alinham bem na região de sanduíche  $\alpha - \beta$  dos três modelos estruturais.

#### 4.6 EpaF

O modelo estrutural da EpaF (NCBI ID: EOL33617) tem pLDDT médio de 97, também apresentando pLDDT menor nas extremidades da estrutura (Figura 20). O enovelamento da EpaF consiste por sanduíche  $\beta$ , formado por duas folhas  $\beta$  antiparalelas, e algumas hélices curtas.

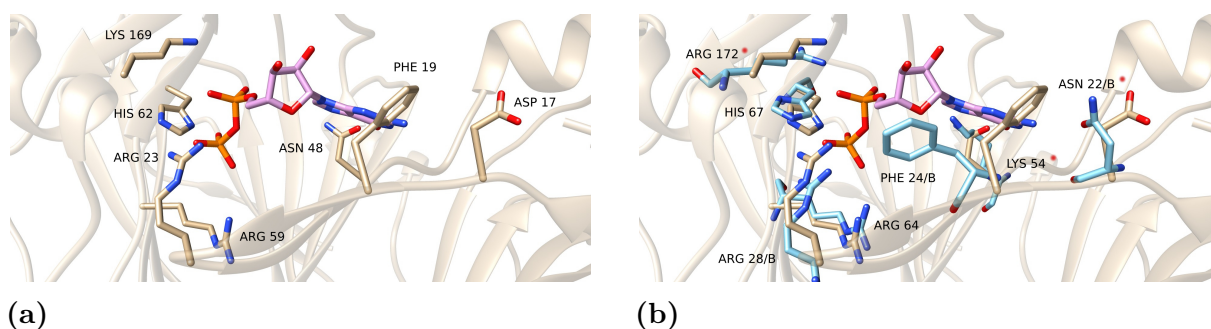


**Figura 20** – Modelo estrutural da EpaF, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a provável região catalítica da EpaF.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

A estrutura predita da EpaF é similar à da GDP-D-glicero-4-ceto-D-ly-heptose-

3,5-epimerase (RmlC) de *Campylobacter jejuni*, PDB ID: 7m14, co-cristalizada com uma molécula GDP. Esta enzima também é classificada como uma dTDP-4-dehidroramnose 3,5-epimerase (EC 5.1.3.13), mesma anotação da EpaF. O alinhamento entre elas apresenta *Z-score* 21.8, RMSD 1.5 e 29% de identidade com o modelo predito da EpaF. As EpaEFGH já estão bem estabelecidas como integrantes da via de síntese de dTDP-L-ramnose.<sup>70</sup> A EpaF é homóloga da RmlC, ela realiza duas epimerizações na dTDP-4-ceto-6-deoxi-D-glicose, nos carbonos C-3 e C-5 (Figura 8). A sobreposição estrutural da RmlC de *C. jejuni* (PDB ID: 7m14) pode ser vista na Figura 21. No sítio catalítico onde a GDP está presente, destacam-se os resíduos ASN48, ARG59, HIS62 e LYS169 do modelo predito da EpaF, análogos com os resíduos LYS54, ARG64, HIS67 e ARG172 respectivamente, da RmlC de *Campylobacter jejuni* (PDB ID: 7m14). O sítio catalítico desta última também é composto pela região de interface entre as cadeias A e B, com os resíduos ASN22, PHE24, ARG28, ao sobrepor a predição estrutural da EpaF com esta cadeia B da RmlC de *C. jejuni*, esses três últimos resíduos se alinham com a ASP17, PHE19 e ARG23 da EpaF.



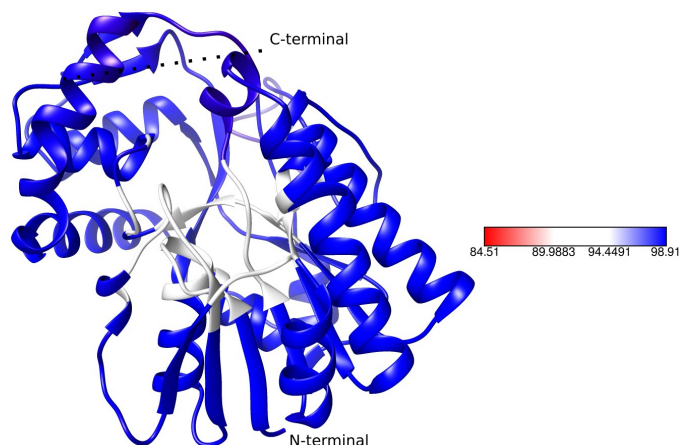
**Figura 21** – (a) Modelo estrutural da EpaF, em dourado, sobreposto com a RmlC de *C. jejuni* (PDB ID: 7m14), em azul, ressaltando alguns dos potenciais resíduos catalítico da EpaF com um dos produtos da reação enzimática, o GDP, em roxo. (b) Resíduos da estrutura da RmlC (PDB ID: 7m14) que interagem com GDP. Estão destacados com um ● os resíduos diferentes entre a predição estrutural da EpaF e a RmlC de *C. jejuni*.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

## 4.7 EpaG

O modelo estrutural da EpaG (NCBI ID: EOL33616) tem pLDDT médio de 95, o pLDDT de cada resíduo está apresentado na Figura 22. A estrutura predita da EpaG apresenta domínio sanduíches  $\alpha - \beta$ .

A segunda estrutura mais similar ao modelo predito da EpaG é a da dTDP-D-glicose de-hidratase (RmlB) de *Streptococcus suis* (EC 4.2.1.46), PDB ID: 1ker, com *Z-score* 55.7, RMSD 1.2 e 80% de identidade. A primeira é RmlB de *E. coli*, PDB ID: 1oc2, preterida em função da RmlB de *S. suis* (PDB ID: 1ker), a mais similar, por esta ter sido co-cristalizada com o co-fator e o substrato da EpaG, uma molécula nicotinamida adenina

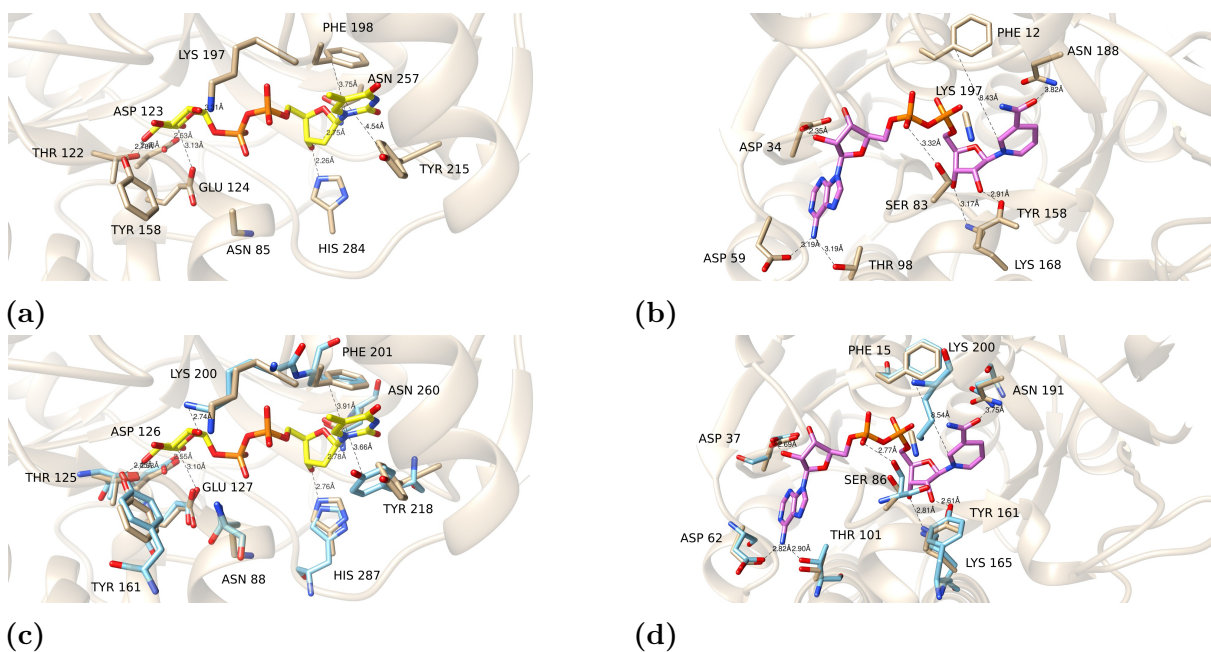


**Figura 22** – Modelo estrutural da EpaG, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a provável região catalítica da EpaG.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

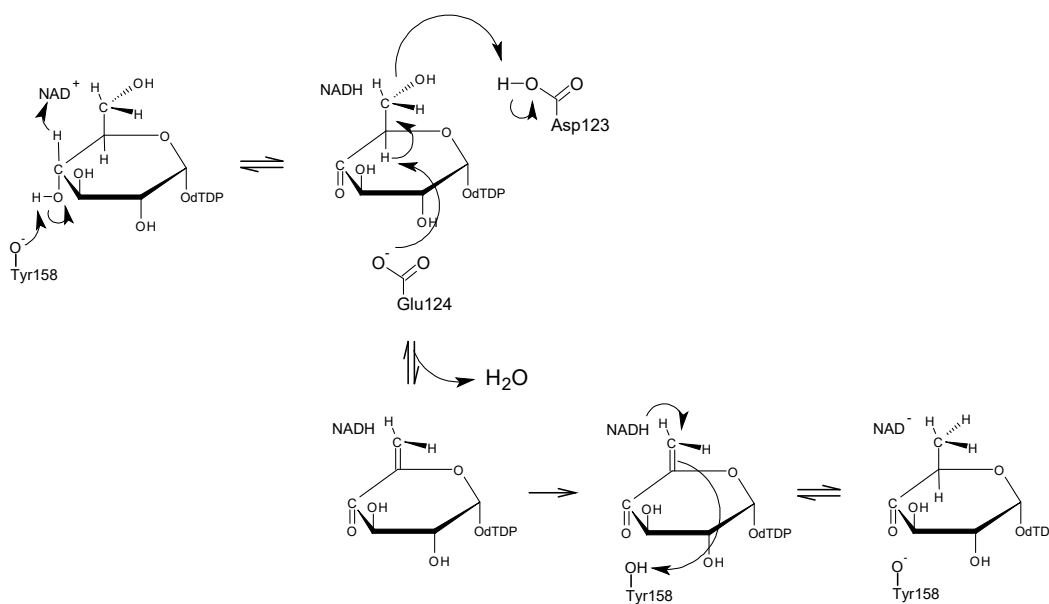
dinucleotídeo (NAD) e dTDP-D-glicose, respectivamente. A EpaG oxida e promove reação de desidratação na dTDP-glicose, formando a dTDP-4-ceto-6-deoxi-D-glicose<sup>65</sup> (Figura 8).

As seguintes análises dos resíduos catalíticos da EpaG (Figura 23a e Figura 23b) foram feitas baseadas na sobreposição da estrutura predita da EpaG e a RmlB de *S. suis* (PDB ID: 1ker) (Figura 23c e Figura 23d), e pelo mecanismo de reação proposto no trabalho de S. Allard e colaboradores.<sup>122</sup> Do ponto de vista do sítio catalítico da EpaG, na reação de oxidação, a molécula NAD é ajustada por interações envolvendo PHE12, ASP34, ASP59, THR98, TYR158, LYS168 e ASN188. A LYS197 e SER83 talvez estabilizem íons



(continua)

(continuação)



(e)

**Figura 23** – Modelo estrutural da (a) EpaG, em dourado, na suposta região catalítica de oxidação da dTDP-D-glicose, em amarelo, pelo (b) NAD, em roxo. Os potenciais resíduos catalíticos da EpaG foram identificados a partir da (b) sobreposição com os resíduos da RmlB de *S. suis* (PDB ID: 1ker), em azul, co-cristalizada com dTDP-glicose (c) e NAD (d). (e) Suposto mecanismo da enzima EpaG.

Fonte: Elaborada pela autora no UCSF-Chimera<sup>80</sup>; adaptada de S. ALLARD *et al.*<sup>122</sup> no ChemSketch™.

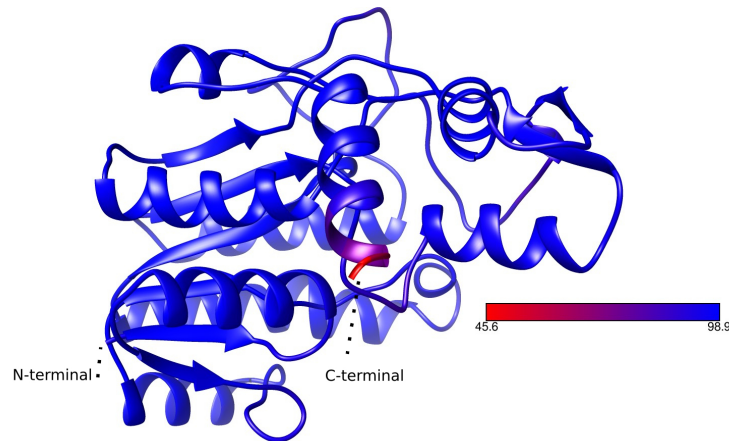
Mg<sup>+2</sup>/Mn<sup>+2</sup>, tanto da molécula NAD quanto a molécula de glicose. A TYR158, o GLU124 e ASP123 podem estar diretamente envolvidos nas reações de oxidação e desidratação da dTDP-glicose, a esquematização simplificada disto está representado na Figura 23e.

#### 4.8 EpaH

O modelo estrutural predito da EpaH (NCBI ID: EOL33615) tem pLDDT médio de 96, o pLDDT de cada resíduo está apresentado na figura Figura 24. O enovelamento da EpaH é constituído por sanduíches  $\alpha - \beta$ .

Uma das estruturas mais similares ao modelo predito da EpaH é a da dTDP-4-de-hidro-ramnose redutase (RmlD) de *E. coli* (EC 1.1.1.133), PDB ID: 1vlo, com *Z-score* 41.7, RMSD 1.1 e 43% de identidade. A RmlD promove a redução da dTDP-4-ceto-6- deoxi-D-glicose pós epimerização, resultando na dTDP-L-ramnose<sup>65</sup> (Figura 8). A RmlD (PDB ID: 1vlo) foi co-cristalizada com nicotinamida adenina dinucleotídeo reduzida (NADH), e alinhada com o modelo predito da EpaH. Baseando-se nisso, foram identificados alguns

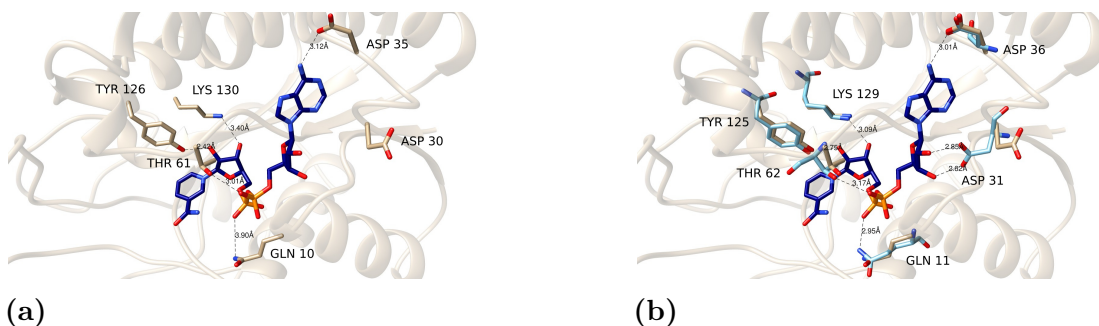




**Figura 24** – Modelo estrutural da EpaH, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

resíduos da EpaH que potencialmente interajam com a NADH (Figura 23a): GLN10, THR61, TYR126, LYS130, ASP35 e ASP30, se rotacionada em 180° (Figura 25b).



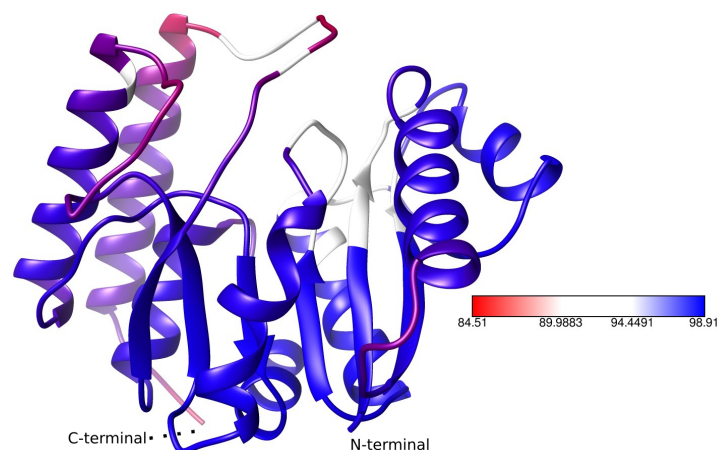
**Figura 25** – Modelo estrutural da (a) EpaH, em dourado, na provável região catalítica de interação com o NADH, em azul escuro, identificado a partir da sobreposição estrutural com a (b) RmlD de *E. coli* (PDB ID: 1vlo) NAD, em azul claro.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

## 4.9 EpaI

O modelo estrutural predito da EpaI (NCBI ID: EOL33614) tem pLDDT médio de 88, o padrão de pLDDT ao longo da estrutura predita está apresentado na Figura 26. O enovelamento da EpaI é constituído por sanduíche  $\alpha - \beta$ .

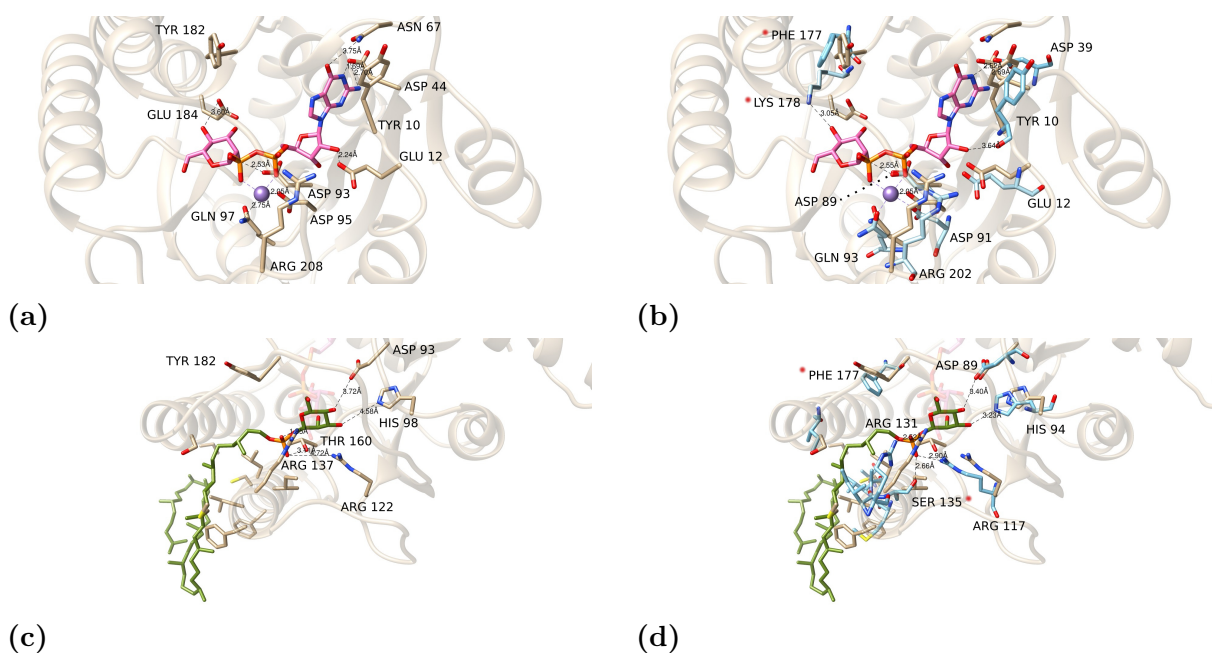
As duas estruturas mais semelhantes ao modelo predito da EpaI são da dolicol fosfato manose sintase de *Pyrococcus furiosus* (EC 2.4.1.83), co-cristalizada com GDP-manose (PDB ID: 5mm0) e com GDP e dolicol fosfato manose (PDB ID: 5mm1). Quando comparadas essas estruturas ao modelo estrutural EpaI, são apresentados *Z-score* 22.7, RMSD 2.4 e 27% de identidade. Para facilitar a análise estrutural do modelo gerado da EpaI, esta foi sobreposta com as duas estruturas de dolicol fosfato manose sintase de



**Figura 26** – Modelo estrutural da EpaI, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Em branco, está a suposta região catalítica da EpaI.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

*P. furiosus* (PDB ID: 5mm0 e 5mm1). Com base nisto e no trabalho envolvendo essas estruturas, de Gandini e colaboradores,<sup>123</sup> os potenciais resíduos catalíticos da EpaI foram identificados (Figura 27).

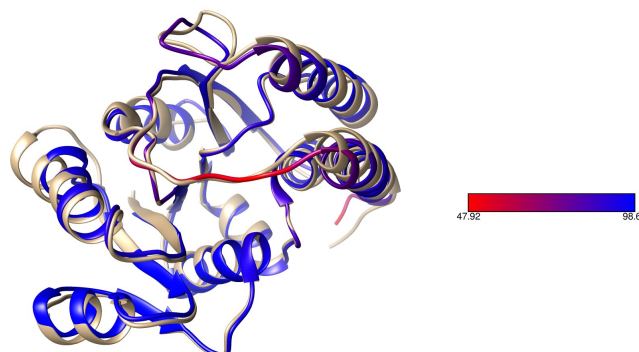


**Figura 27** – Modelo estrutural da (a) EpaI, em dourado, sobreposto com a (b) dolicol fosfato manose sintase de *P. furiosus* (PDB ID: 5mm0), em azul, co-cristalizada com GDP-manose, em rosa, e íon  $Mn^{+2}$ , em roxo. Modelo estrutural da (c) EpaI sobreposto com a (d) dolicol fosfato manose sintase (PDB ID: 5mm1), em azul, co-cristalizada com dolicol fosfato manose, em verde. Estão destacados com um • os resíduos diferentes entre a predição estrutural da EpaI e as estruturas da dolicol fosfato manose sintase de *P. furiosus*.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

A dolicol fosfato manose sintase transfere manose a partir da UDP-manose para o carreador dolicol fosfato. Os resíduos ASP95 e GLN97 da EpaI podem estar envolvidos na estabilização do cátion  $Mn^{+2}$ , enquanto que GLU12, ASP44, ASN67 e GLU184 estariam formando pontes de hidrogênio ao longo do açúcar nucleotídeo, enquanto que a TYR12 poderia realizar empilhamento pi com o anel guanina. A TYR182 da EpaI (Figura 27a) pode agir similarmente à PHE177 da dolicol fosfato manose sintase (Figura 27b), juntamente ao *loop* na região, no bloqueio da saída do açúcar nucleotídeo do sítio catalítico. Os resíduos ARG202 e GLU12 das estruturas da dolicol fosfato manose (PDB ID: 5mm0 e 5mm1) formam interação iônica com o açúcar nucleotídeo,<sup>123</sup> para também mantê-lo na cavidade catalítica. Na EpaI, isto pode ser realizado pelos resíduos ARG208 e GLU12. Em relação ao produto dolicol fosfato manose, os seguintes resíduos da EpaI se destacam por terem distância razoável desse para realizar com ele interações moleculares (Figura 27c): HIS98, ARG122, ARG137 e THR160. O ASP93 pode interagir com o grupo fosfato do trifosfato de nucleosídeo ou com o carreador lipídico. A interface exterior da hélice ao lado da cadeia alifática do dolicol fosfato manose, da dolicol fosfato manose sintase (PDB ID: 5mm0 e 5mm1), é formada majoritariamente por resíduos alifáticos (Figura 27d), assim como a interface exterior da hélice na estrutura predita da EpaI.

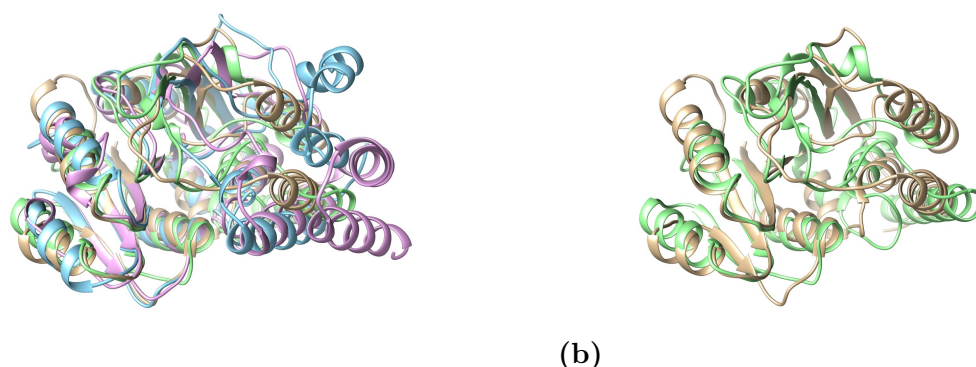
No trabalho de Guerardel e colaboradores, 2020,<sup>55</sup> é proposto que a enzima homóloga da EpaI na cepa V583 de *E. faecalis*, EF2190, seja uma undecaprenil-fosfato GlcNAc, sintetizadora de GlcNAc-P-Und a partir de UDP-GlcNAc e undecaprenil fosfato.<sup>116</sup> O GlcNAc-P-Und teria o papel de doador da porção GlcNAc para o polímero de ramnose, na região lateral deste. Isto é baseado na consideração que a EF2190 seja homóloga da GacI (AAZ51228), enzima de *Streptococcus pyogenes*. De fato, em uma cobertura de 98%, a EpaI e GacI possuem 48% de identidade, com *E-value*  $6e^{-70}$ , e o modelo estrutural das duas realizado pelo *AlphaFold2* se alinham bem entre si (Figura 28). Isto sugere que EpaI e GacI são proteínas homólogas.



**Figura 28** – Modelo estrutural da EpaI, em dourado, e da GacI, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

A EpaI é uma glicosiltransferase da família 2, assim como as EpaB, EpaC e EpaD. Ao alinhar as estruturas preditas dessas 4 enzimas, percebe-se que todas possuem estruturas terciárias muito parecidas entre si (Figura 29a). O enovelamento de glicosiltransferases é conservado, sendo normalmente uma variação de um sanduíche  $\alpha$ - $\beta$ - $\alpha$ .<sup>124</sup> Entretanto, no caso dessas GT2, elas se ajustam muito bem quando alinhadas, o enovelamento delas é quase idêntico. Em termo de similaridade estrutural, há proteínas que são comuns à quase todas as GT2 do *cluster epa* discutidas nesta tese, como a CP (PDB ID:2z86) ou a GalT1 (PDB ID: 5hea), por exemplo. Contudo, isto não é o caso para a EpaI. Nenhuma das 30 enzimas cuja estrutura mais se assemelham ao modelo da EpaI é estruturalmente semelhante às outras GT2 aqui citadas.



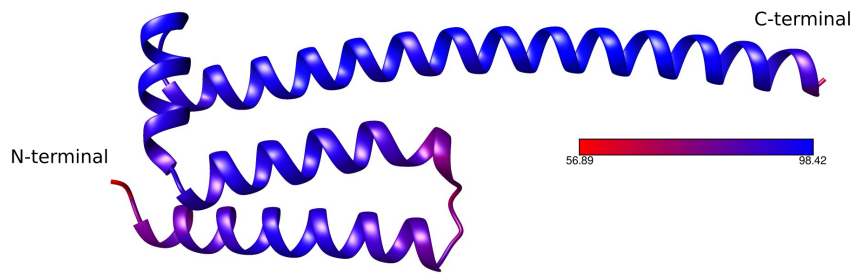
**Figura 29** – (a) Modelo estrutural da EpaI, em dourado, sobreposta com os modelos da EpaB, em azul, EpaC, em roxo, e EpaD, em verde. (b) Foco no alinhamento entre as estruturas preditas da EpaI e EpaD.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

#### 4.10 EpaJ

O modelo estrutural predito da EpaJ (NCBI ID: EOL33613) tem pLDDT médio de 90, o pLDDT de cada resíduo está apresentado na Figura 30. O enovelamento da estrutura predita da EpaJ é composto por feixes de hélice. O modelo estrutural da EpaJ possui similaridade com algumas estruturas de receptores de glicina depositadas no PDB, entretanto, não é significativo, a estrutura do receptor de glicina de *Danio rerio*, PDB ID: 6vm3, apresenta apenas *Z-score* 7.1, RMSD 2.6 e 11% de identidade com a predição estrutural da proteína EpaJ.

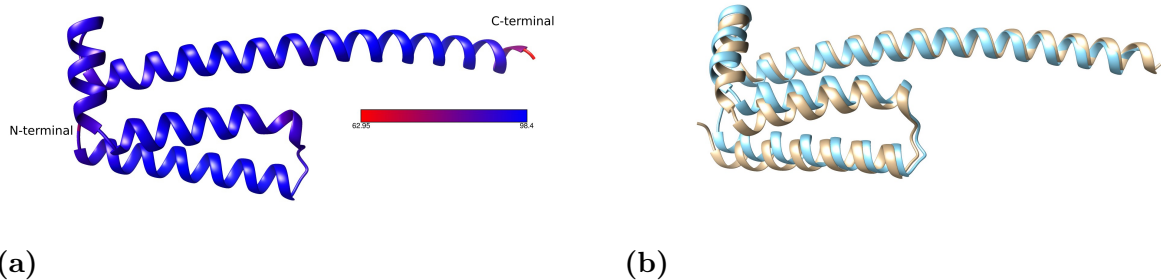
Em *Streptococcus pyogenes*, há uma pequena proteína de membrana, GacJ (NCBI ID: AAZ51229), a qual otimiza a reação catalisada por uma segunda enzima, GacI,<sup>86,116</sup> mencionada na seção anterior (seção 4.9). Quando alinhadas as sequências da proteínas EpaJ e GacJ, há apenas 24.24% de identidade entre elas, com cobertura 85% e *E-value*  $6.e^{-10}$ . A predição estrutural da GacJ pelo *AlphaFold2* possui pLDDT médio de 93 (Figura 31a), e bom alinhamento com a estrutura predita da EpaJ (Figura 31b), o que valida a hipótese de que EpaJ e GacJ são homólogas.<sup>86</sup>



**Figura 30** – Modelo estrutural da EpaJ, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

No trabalho de Guerardel e colaboradores, 2020,<sup>55</sup> é proposto que a enzima homóloga da EpaJ na cepa V583, EF2189, catalise a mesma reação da EpaI, como discutido na seção 4.9, só que sintetizando Glc-P-Und a partir de UDP-GlcNA e undecaprenil fosfato.<sup>116</sup> Baseado no mecanismo descrito Gandini e colaboradores<sup>123</sup> sobre a síntese de um dolicol fosfato açúcar, é improvável que a EpaJ exerça a atividade de uma Glc-P-Und, mesmo em um estado oligomérico. O mais plausível é que a EpaJ auxilie a reação catalítica da EpaI, de transferência de um açúcar para um carreador lipídico.<sup>86</sup>

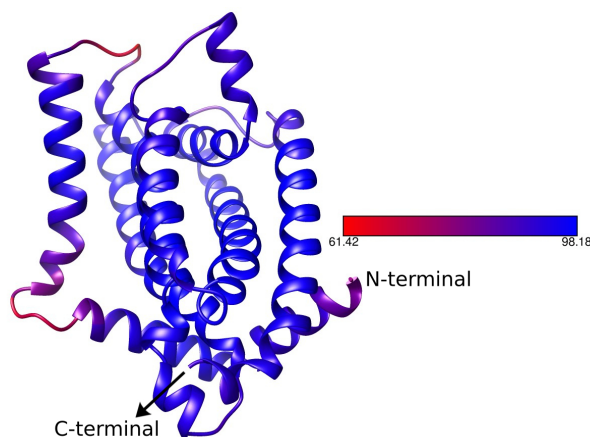


**Figura 31** – (a) Modelo estrutural da GacJ, com coloração indicando do menor (em vermelho) para o maior (em azul) valor de pLDDT. (b) Sobreposição das estruturas preditas EpaJ, em dourado, e GacJ, em azul.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

#### 4.11 EpaL

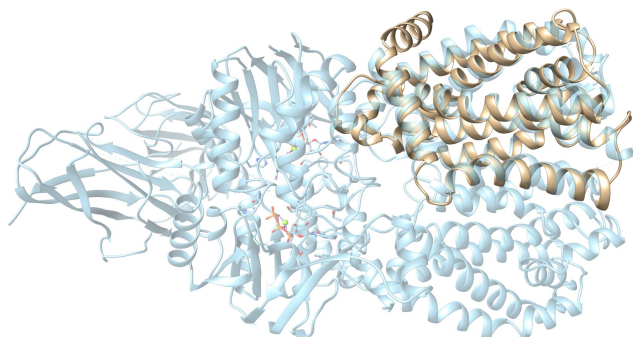
O modelo estrutural da EpaL (NCBI ID: EOL33611) tem pLDDT médio de 92. O padrão de pLDDT ao longo da estrutura predita está apresentado na Figura 32. O enovelamento da predição estrutural da EpaL contém feixes de hélice e alguns *loops* alongados. Este enovelamento é semelhante ao de vários transportadores ABC de membrana, segundo os dados obtidos do servidor Dali.<sup>79</sup>



**Figura 32** – Modelo estrutural da EpaL, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

Das estruturas depositadas no PDB, a do transportador ABC de *Aquifex aeolicus*, PDB ID: 7k2t, é uma das mais semelhantes à estrutura predita da EpaL, as duas estruturas possuem *Z-score* 28.0, RMSD 2.0 e 27% de identidade. O alinhamento ocorre em um dos domínios da do transportador ABC (Figura 33). Entretanto, o alinhamento não contempla o domínio sanduíche  $\beta$  do transportador ABC de *A. aeolicus* (PDB ID: 7k2t).

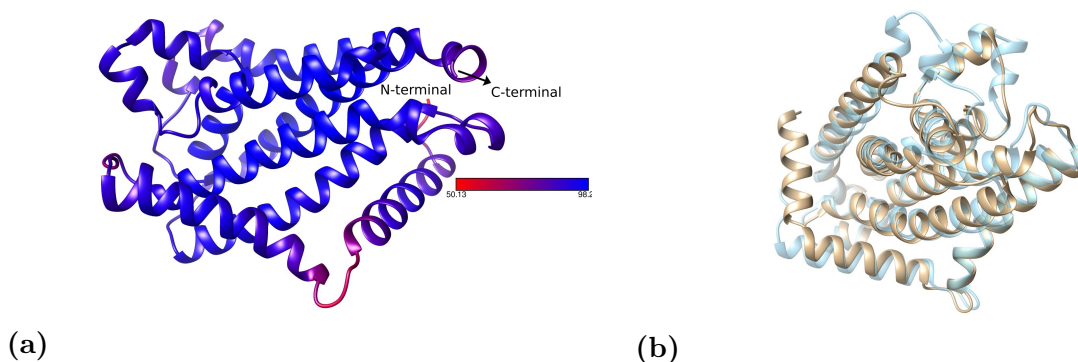


**Figura 33** – Modelo estrutural da EpaL, em dourado, sobreposto com a estrutura do transportador ABC de *A. aeolicus* (PDB ID: 7k2t), em azul.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

Guerardel e colaboradores<sup>55</sup> caracterizam os homólogos da EpaL e EpaM, proteínas EF2182 e EF2183, respectivamente, como formadoras do transportador ABC do polímero de ramnose para o exterior celular. As proteínas EF2182 e EF2183 são homólogos da EpaM e EpaL, respectivamente, da cepa *E. faecalis* V583. Ao analisar o provável papel de proteínas participantes de *clusters* análogos ao *cluster* epa, a EpaL é mais semelhante às sequências GacD (NCBI ID: AAZ51223) de *S. pyogenes* e RgpC (NCBI ID: BDH82876) de *Lactococcus lactis*. A sequência da EpaL apresenta 26.61% de identidade e *E-value*  $4e^{-22}$  com a da GacD, e 28.37% de identidade e *E-value*  $3e^{-14}$  com a da RgpD. A função dessas duas enzimas é relativamente bem caracterizada, são transporte permeases acopladas a uma

segunda proteína de ligação dNTP, o que compõe o transportador ABC responsável pela exportação da cadeia alongada do polissacarídeo a ser ancorado no peptidoglicano.<sup>116</sup> A predição estrutural da GacD pelo *AlphaFold2* apresenta pLDDT médio de 90 (Figura 34a), e ela se alinha bem ao modelo predito da EpaL (Figura 34b). Este alinhamento é semelhante ao dos modelos estruturais RgpD e EpaL, o qual foi aqui omitido. Foi escolhida a predição estrutural GacD para ser alinhada com a da EpaL, pois a confiabilidade do modelo da GacD é maior do que o da RgpD, quando analisado o pLDDT médio e o perfil do pLDDT ao longo desses dois modelos. Sendo assim, é factível que a EpaL seja uma transportadora envolvida no transporte do polímero de ramnose para o exterior da membrana plasmática.



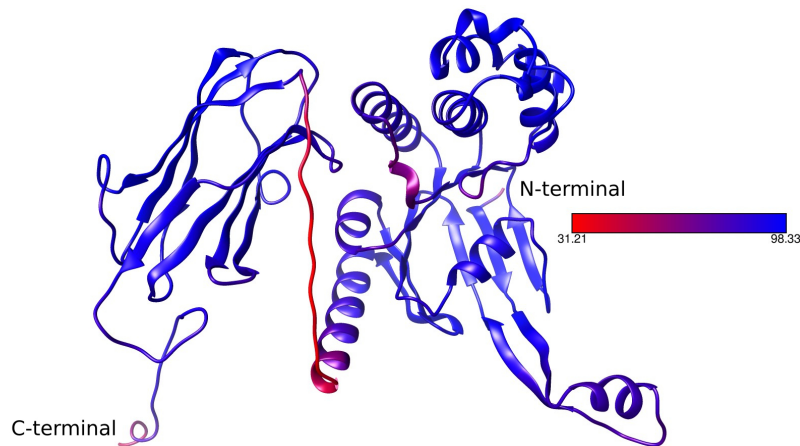
**Figura 34** – (a) Modelo estrutural da GacD, com coloração indicando do menor (em vermelho) para o maior (em azul) valor de pLDDT. (b) Sobreposição das estruturas preditas EpaL, em dourado, e GacD, em azul.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

## 4.12 EpaM

O modelo estrutural da EpaM (NCBI ID: EOL33610) tem pLDDT médio de 88. O padrão de pLDDT ao longo da estrutura predita está apresentado na figura Figura 35. O enovelamento da predição estrutural da EpaM contém sanduíche  $\beta$  e sanduíche  $\alpha$ - $\beta$ , assim como alguns *loops* alongados.

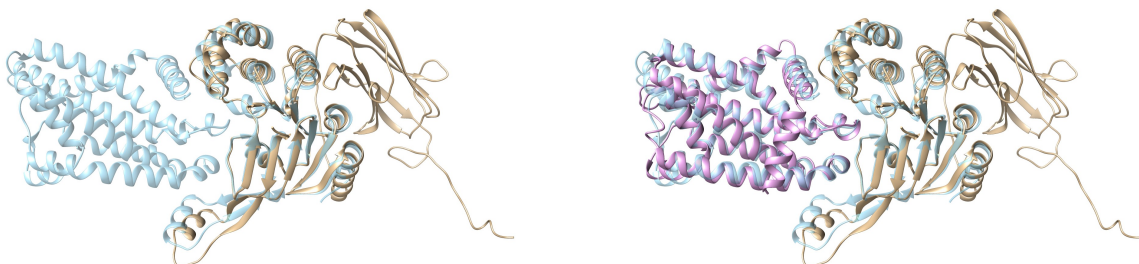
O modelo estrutural da EpaM é similar a vários transportadores ABC, assim como o modelo predito da EpaL. A estrutura do transportador ABC de *Aquifex aeolicus*, PDB ID: 6m96, é uma das mais semelhantes à estrutura predita da EpaM, as duas possuem *Z-score* 28.5, RMSD 1.9 e 46% de identidade. Essas duas estruturas possuem bom alinhamento em um dos domínios do transportador ABC de *A. aeolicus* (PDB ID: 6m96), como visto na Figura 36a. Na Figura 36b, é possível visualizar um bom alinhamento da predição estrutural da EpaL com o outro domínio do transportador ABC (PDB ID: 6m96), corroborando a conjectura da EpaL e EpaM constituírem juntas um transportador ABC, talvez envolvido na transferência do polímero de ramnose do citosol para o exterior da membrana plasmática.



**Figura 35** – Modelo estrutural da EpaM, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

A sequência da EpaM também apresenta similaridade com as sequências das enzimas GacE (NCBI ID: AAZ51224) de *S. pyogenes* e RgpD (NCBI ID: BDH82877) de *L. lactis*, duas proteínas dNTP-binding.<sup>116</sup> A sequência da EpaM apresenta 36.88% de identidade e E-value  $8e^{-72}$  com a da GacE, e 38.53% de identidade e E-value  $2e^{-72}$  com a



(a)

(b)

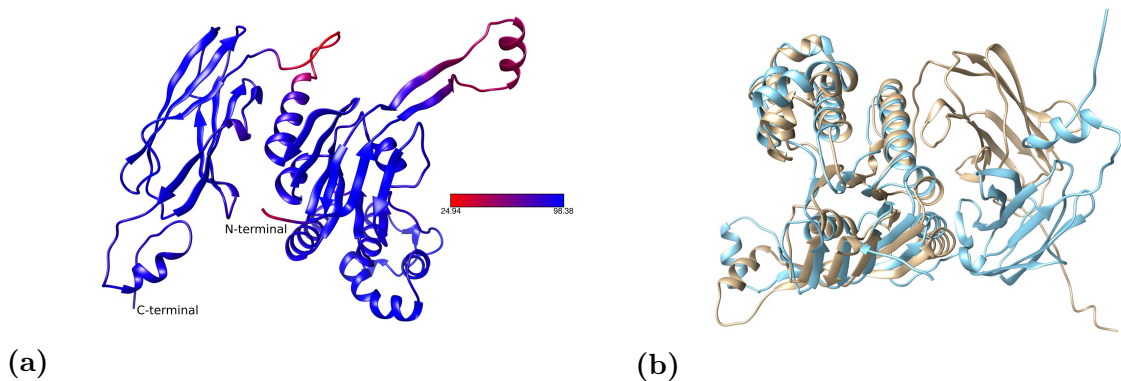
**Figura 36** – (a) Modelo estrutural da EpaM, em dourado, sobreposto com a estrutura do transportador ABC de *A. aeolicus* (PDB ID: 6m96), em azul. (b) Sobreposição do modelo estrutural da EpaL, em roxo, e da EpaM com o transportador ABC anterior (PDB ID: 6m96), indicando que EpaL e EpaM são duas proteínas de membrana acopladas, constituintes do transportador ABC.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

da RgpD. A predição estrutural da GacE apresenta pLDDT médio de 87 (Figura 37a), e ela se alinha a quase todos os domínios do modelo predito da EpaM (Figura 37b), exceto por um sanduíche  $\beta$ , que, mesmo presente nos dois modelos, está em conformação diferente em cada um deles. O modelo gerado da RgpD pelo *AlphaFold2* é semelhante ao da GacE, apresentando também pLDDT médio de 87. Visto que o alinhamento entre os modelos RgpD e EpaM também é similar ao das estruturas preditas GacE e EpaM, os resultados referentes à modelagem da RgpD foram aqui omitidos. Portanto, a EpaM pode



ser homóloga da GacF e da RgpD, as quais formam um transportador ABC com a GacD e a RgpC, respectivamente, citadas na seção anterior (seção 4.11).

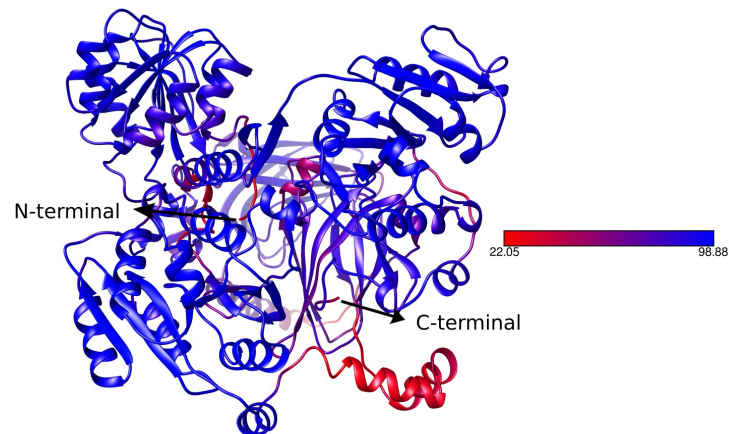


**Figura 37** – (a) Modelo estrutural da GacE, com coloração indicando do menor (em vermelho) para o maior (em azul) valor de pLDDT. (b) Sobreposição das estruturas preditas EpaM, em dourado, e GacE, em azul.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

#### 4.13 EpaN

O modelo estrutural da EpaN (NCBI ID: EOL33609) tem pLDDT médio de 80. O padrão de pLDDT ao longo da estrutura predita está apresentado na Figura 38. O enovelamento da predição estrutural da EpaN contém sanduíche  $\alpha$ - $\beta$ - $\alpha$  e sanduíche  $\beta$ .

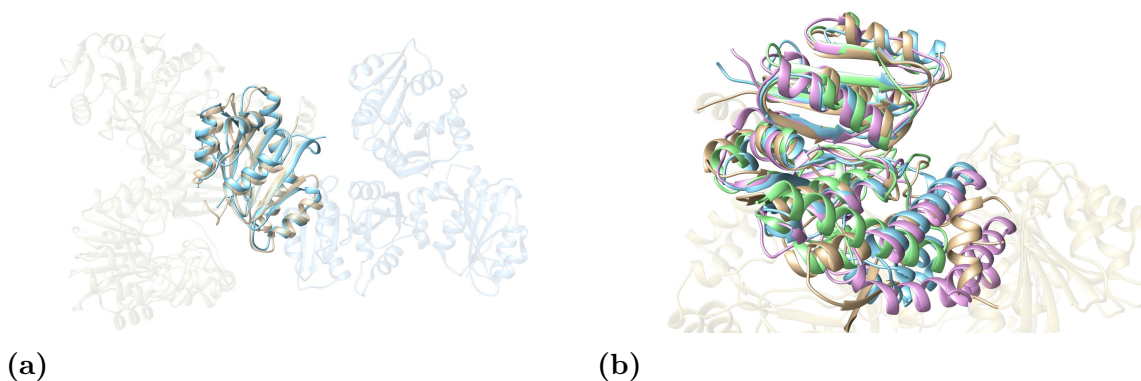


**Figura 38** – Modelo estrutural da EpaN, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

A estrutura mais similar ao modelo predito da EpaN é a da glicosiltransferase de *Leptospira borgpetersenii*, PDB ID: 6p61, apresentando *Z-score* 25.8, RMSD 1.8 e 24% de identidade com a estrutura predita da EpaN. O alinhamento dessas duas estruturas pode ser visualizado na Figura 39a, ele é realizado em um dos domínios de cada estrutura. Uma vez que essa glicosiltransferase de *L. borgpetersenii* não foi co-cristalizada nem com

substratos ou produtos da reação que catalisa, não é possível realizar a discussão sobre os prováveis resíduos catalíticos da EpaN. Guerardel e colaboradores<sup>55</sup> supõem que a enzima homóloga da EpaN, EF2181, também contribua para a polimerização da cadeia de ramnose. Dentre as estruturas mais semelhantes a predição estrutural da EpaN, algumas também são similares às estruturas preditas da EpaB, EpaC e EpaD. O alinhamento entre estas últimas estruturas e a do modelo estrutural da EpaN pode ser visualizado na Figura 39b. Enquanto que os modelos da EpaB, EpaC e EpaD se alinham por toda a estrutura predita de cada uma, a EpaN se alinha com essas três em apenas um domínio. Isso porque a EpaN é uma proteína formada por 1047 resíduos de aminoácidos, enquanto que as enzimas EpaB, EpaC e EpaD são constituídas por 260, 275 e 237 resíduos de aminoácidos, respectivamente. A região de alinhamento entre a EpaN e essas GT2 provavelmente abrange o sítio catalítico da EpaN, por ser uma estrutura terciária comum a várias glicosiltransferases,<sup>124</sup> como discutido na seção 4.9. A EpaN ainda possui dois mais outros domínios com o enovelamento sanduíche  $\alpha$ - $\beta$ - $\alpha$ , apresentando provavelmente mais duas cavidades catalíticas.



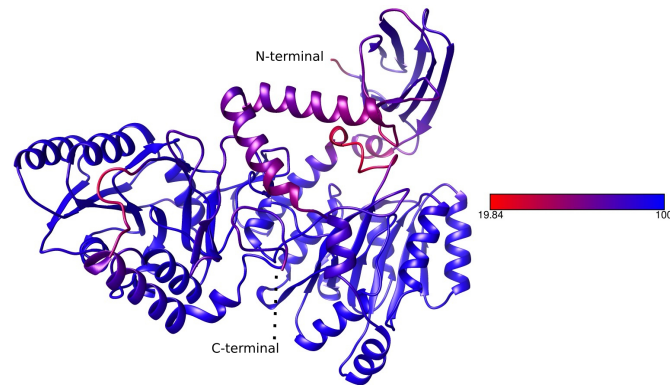
**Figura 39** – (a) Modelo estrutural da EpaN, em dourado, sobreposto com um domínio da estrutura da glicosiltransferase de *L. borgpetersenii* (PDB ID: 6p61), em azul. (b) Modelos estruturais da EpaN, em azul, EpaB, em azul, EpaC, em roxo, e EpaD, em verde.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

#### 4.14 EpaO

O modelo estrutural predito da EpaO (NCBI ID: EOL33608) tem pLDDT médio de 90, o padrão de pLDDT ao longo da estrutura predita está apresentado na Figura 40. O enovelamento da EpaO é constituído por sanduíches  $\alpha - \beta$  e folha  $\beta$ .

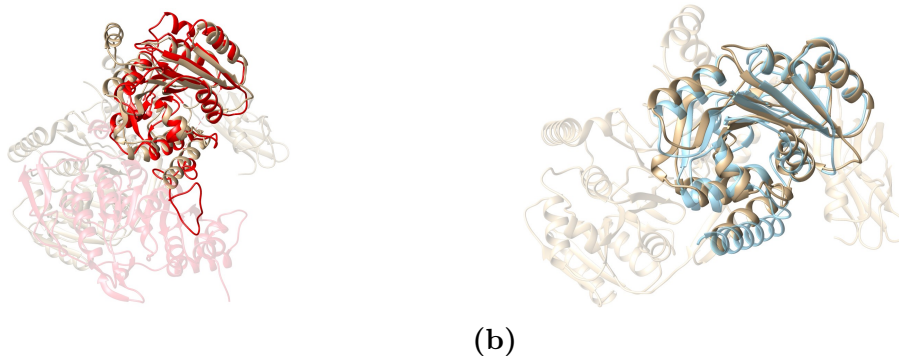
A estrutura predita da EpaO é mais similar à da condroitina polimerase (EC 2.4.1.288 e 2.4.1.226) de *E. coli*, PDB ID: 2z86, a qual apresenta *Z-score* 28.6, RMSD 2.4 e 25% de identidade com o modelo predito da EpaO. A estrutura da CP de *E. coli* (PDB ID: 2z86) é a mesma que foi apresentada na seção 4.5, sobre a EpaD. Como a região de alinhamento entre a condroitina polimerase (PDB ID: 2z86) e a predição estrutural da EpaO (Figura 41a) é a mesma entre os modelos estruturais EpaO e EpaD (Figura 41b), a



**Figura 40** – Modelo estrutural da EpaO, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

discussão sobre os potenciais resíduos catalíticos da EpaO seria redundante e foi dispensada. Isto também reforça proposta que a EpaO exerça a mesma função catalítica das EpaB, EpaC, EpaD e EpaN.<sup>55</sup> A EpaO pode comportar mais do que um sítio catalítico, assim como a EpaN. Os domínios sanduíche  $\alpha$ - $\beta$ - $\alpha$  da EpaO se alinham a dois dos domínios da EpaN (Figura 42).

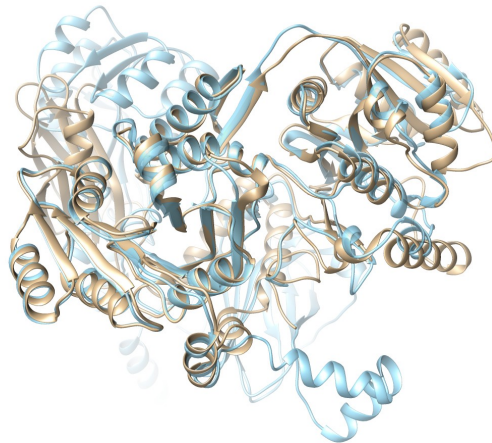


**Figura 41** – (a) Predição estrutural da EpaO, em dourado, sobreposta com a estrutura da condroitina polimerase de *E. coli* (PDB ID: 2z86), em vermelho. (b) Sobreposição do modelo estrutural da EpaO, em dourado, com o da EpaD, em azul.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

#### 4.15 Potenciais Resíduos Catalíticos das Glicosiltransferases

Analisando os prováveis sítios catalíticos das EpaBCDNO com o da EpaI, o melhor caracterizado até aqui, há a identificação de alguns resíduos de aminoácidos os quais são conservados nessas glicosiltransferases. O resumo desta análise é visto na Tabela 1, incluindo apenas observações dos domínios  $\alpha$ - $\beta$ - $\alpha$  da EpaN e da EpaO que melhor se alinharam com as estruturas preditas das demais enzimas, estes serão denominados domínios principais. Quando pertinente, as discussões sobre os outros domínios dessas duas será feita apenas



**Figura 42** – Alinhamento dos domínios sanduíche  $\alpha$ - $\beta$ - $\alpha$  da EpaO, em dourado, e da EpaN, em azul.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

nos parágrafos subsequentes. O domínio da EpaN contendo numeração mais baixa de aminoácido é aqui denominado secundário, e o com numeração alta, terciário.

A TYR10 da EpaI é conservada em todas as enzimas, exceto nos outros dois sítios catalíticos da EpaN, presentes nos outros domínios  $\alpha$ - $\beta$ - $\alpha$  desta. Já no caso do segundo domínio  $\alpha$ - $\beta$ - $\alpha$  da EpaO, há uma LYS447 alinhada a TYR10 da EpaI.

O GLU12 presente na EpaI não é observado nas demais estruturas, entretanto, há uma conservação de um resíduo de valina nesta posição, nos domínios principais da  $\alpha$ - $\beta$ - $\alpha$  da EpaN e da EpaO. Nos outros domínios  $\alpha$ - $\beta$ - $\alpha$  da EpaN, há uma LEU55 bem alinhada com a GLU12 da EpaI, e o ASP819 próximo dessa posição, assim como nessa mesma região há a ASN448 no segundo domínio da EpaO. Na posição da GLU12 da EpaI, há uma serina na EpaB e glicina na EpaD, na EpaC, há uma metionina.

O ASP44 da EpaI só não está presente na EpaB e EpaC, as quais possuem ambas ASN41. Na posição ASP44 da EpaI, está presente a LEU55 no domínio secundário da EpaN, enquanto que no segundo domínio da EpaO, há a ASN476. O terceiro domínio da EpaN não possui bom alinhamento com a EpaI nessa região.

A ASN67 é bem conservada nos *loops* das outras enzimas, exceto na EpaD, a qual não se alinha tão bem nessa região com as demais glicosiltransferases, e apresenta uma GLN65. No domínio  $\alpha$ - $\beta$ - $\alpha$  secundário da EpaN, há uma ASN74 próxima da ASN67 da EpaI, mesma situação da ASN508 do segundo domínio da EpaO.

O ASP93 da EpaI é bem conservado nas demais enzimas, salvo na EpaB, em que há uma ASN89 na mesma posição. Isto também acontece no outro domínio da EpaO e no terceiro da EpaN, com a ASN529 e a ASN882, respectivamente. O segundo domínio da EpaN apresenta um resíduo análogo, o qual não está bem alinhado com o ASP93 da EpaI, somente próximo desse, o ASP10.

O ASP95 da EpaI está presente em todas as cinco enzimas. O segundo domínio da EpaN não apresenta bom alinhamento com a EpaI nesse local, e, no terceiro, há a LEU885 em posição análoga ao do ASP95 da EpaI. Há um ASP531 no outro domínio da EpaO, próxima ao ASP95 da EpaI.

A GLN97 da EpaI não é conservada nas outras glicosiltransferases, mas o aspartato é em três delas, EpaDNO. A TYR182 da EpaI só é conservada na EpaB, na EpaC, nesta posição, está o ASP178, nas EpaD, EpaN e EpaO, a MET174, GLY726 e GLY352 estão, respectivamente, presentes na região GLN97 da EpaI. No domínio secundário da EpaN, há um GLU99 nessa posição, e, no terciário, há um GLU985. No outro domínio da EpaO, há um PHE632 alinhado com a TYR182 da EpaI.

O GLU184 da EpaI não é conservada entre as outras glicosiltransferases, e sim, um resíduo de aspartato nessa região de bom alinhamento entre as cinco enzimas. Essa conservação também é observada no outro domínio da EpaO, com a ASP634, mas o mesmo não é o caso com os outros domínios da EpaN, os quais não se alinham bem nessa região com as outras glicosiltransferases.

**Tabela 1** – Resíduos presentes no provável sítio catalítico das glicosiltransferases EpaBC-DION codificadas na região constante do *cluster epa* de *Enterococcus faecalis*.

<b>EpaI</b>	<b>EpaB</b>	<b>EpaC</b>	<b>EpaD</b>	<b>EpaN</b>	<b>EpaO</b>
Tyr10	Tyr12	Tyr12	Tyr9	Tyr563	Tyr189
Glu12	Ser14	Met14	Gly11	Val565	Val191
Asp44	Asn41	Asn41	Asp37	Asp592	Asp218
Asn67	Asn66	Asn64	Gln65	Asn621	Asn247
Asp93	Asn89	Asp90	Asp87	Asp644	Asp270
Asp95	Asp91	Asp92	Asp89	Asp646	Asp272
Gln97	Val92	Thr93	Asp90	Asp647	Asp273
Tyr182	Tyr201	Asp178	Met174	Gly726	Gly352
Glu184	Asp204	Asp181	Asp176	Asp729	Asp355
His98	Val92	Thr93	Asp90	Asp647	Asp273
Arg122	Val119	Val119	Leu117	Glu674	Glu300
Arg137	-	-	-	-	-
Thr160	Ser177	Ile154	Ile151	Thr704	Ser330

Fonte: Elaborada pela autora.

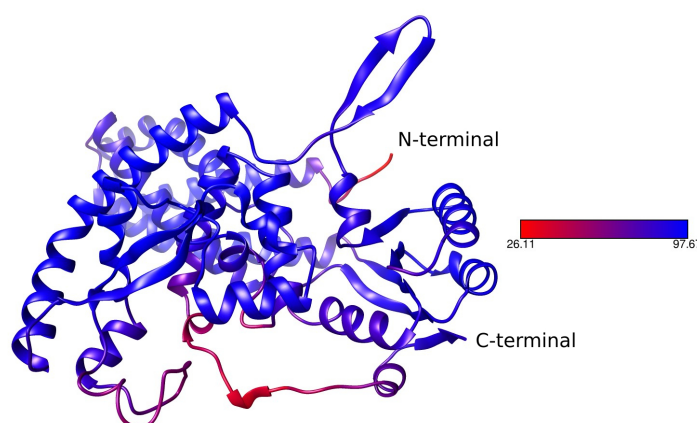
Na região da HIS98 da EpaI, o resíduo de aspartato é conservado entre as EpaD, EpaN e EpaO. No outro domínio da EpaO, a HIS658 está na mesma posição que a HIS98 da EpaI. A ARG122 está presente somente na EpaI, a qual se alinha com a valina da EpaB e EpaC, e com o glutamato da EpaN e EpaO nos domínios principais dessas. A ARG137 da EpaI não se alinha com nenhuma das outras glicosiltransferases, assim como a

hélice que esse aminoácido compõe. Entretanto, nessa região, as EpaBCDNO se alinham bem entre si. A THR160 da EpaI é conservada na EpaN, apesar do ajuste entre os *loops* nessa área não ser tão bom. A isoleucina é conservada na EpaC e EpaD, enquanto que a serina é conservada na EpaB e EpaO.

Há conservação ou dos resíduos de aminoácidos ou das propriedades dos resíduos no potencial sítio catalítico das glicosiltransferases EpaBCDION. Há maior similaridade dos resíduos entre as EpaBCDNO do que dessas com a EpaI. A conservação dos potenciais resíduos do sítio catalítico é menor nos outros domínios da EpaN e EpaO.

#### 4.16 EpaR

O modelo estrutural predito da EpaR (NCBI ID: EOL33605) tem pLDDT médio de 86, o padrão de pLDDT ao longo da estrutura predita está apresentado na figura Figura 78a. O enovelamento da EpaQ é formada por feixes de hélice, sanduíches  $\alpha - \beta$  e folhas  $\beta$ .

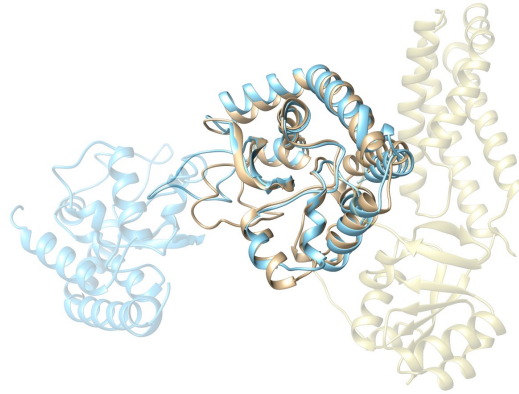


**Figura 43** – Modelo estrutural da EpaR, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

A predição estrutural da EpaR é mais similar à estrutura da N,N'-dimetil bacillus amina-1-fosfato transferase (PglC) (EC. 2.7.8.36) de *Campylobacter concisus*, PDB ID: 5w7l, apresentando *Z-score* 21.0, RMSD 2.5 e 42% de identidade com o modelo estrutural predito da EpaR. A PglC (PDB ID: 5w7l) é uma proteína de membrana, e, segundo a predição do DeepTMHMM,<sup>114</sup> a EpaR também é. O alinhamento estrutural do modelo EpaR e a estrutura da PglC (PDB ID: 5w7l) pode ser visto na Figura 44. Guerardel e colaboradores<sup>55</sup> sugerem que a segunda cadeia de polissacarídeos é iniciada pelo homólogo da EpaR, a proteína EF2177. Esta enzima seria uma undecaprenil-fosfato glicose fosfotransferase, responsável por adicionar Glc-1-P ao carreador lipídico undecaprenil fosfato, gerando Glc-P-P-Und. O conteúdo dessa segunda cadeia de polissacarídeos depende da região variável do *cluster* epa, das enzimas codificadas por ela. A cadeia de polissacarídeos

sinterizados nessa região é anexada ao polímero de ramnose, e permanece exposta ao meio extracelular. As outras estruturas depositadas no PDB similares ao modelo estrutural predito da EpaR apresentam variadas funções biológicas, apenas a PDB ID: 5w7l é uma fosfato fosfoglicosil transferase. Sendo assim, aqui não foi possível ratificar proposição de Guerardel e colaboradores.<sup>55</sup>



**Figura 44** – Modelo estrutural da EpaR, em dourado, sobreposto com a estrutura da PglC (PDB ID: 5w7l), em azul.

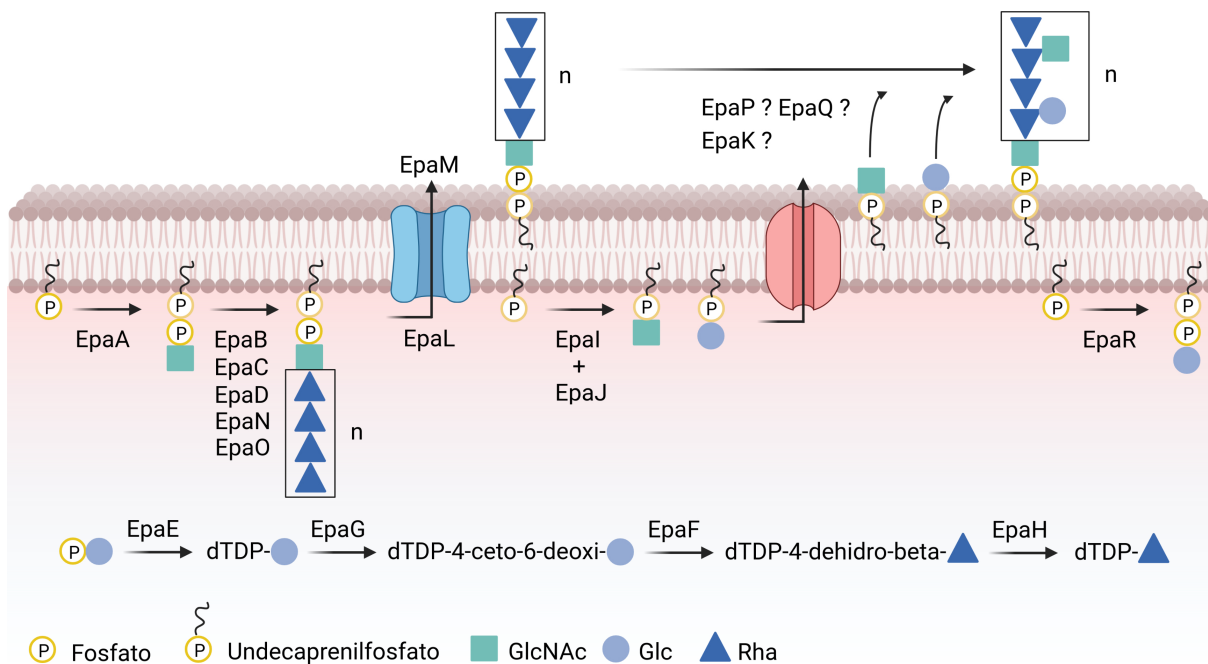
Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

#### 4.17 Esboço da síntese da cadeia de ramnose

A discussão nesse capítulo englobou a análise das funções enzimáticas das proteínas codificadas na região conservada do *cluster* epa de *E. faecalis*, com base na predição estrutural das enzimas Epa pelo *AlphaFold2*, na comparação com *clusters* homólogos ao epa e nos trabalhos de Guaredel *et al.*, 2020,<sup>55</sup> e Guérin *et al.*, 2022.<sup>116</sup>

A maioria dos modelos apresentou boa confiabilidade de predição, e, em algumas situações, foi possível identificar alguns potenciais resíduos catalíticos, mesmo com os modelos estruturais gerados no modo não relaxado. Destacam-se os seguintes casos: **(i)** EpaBCDNO, as quais apresentaram bom alinhamento estrutural entre si, apesar das sequências proteicas distintas, além da conservação dos potenciais resíduos dos sítios catalíticos dessas glicosiltransferases; **(ii)** corroboração da EpaJ ser um homólogo da GacJ, assim como a EpaI é da GacI, indicado por uma ótima sobreposição estrutural dos modelos preditos dos dois casos, e não do alinhamento da sequências da EpaJ e GacJ e EpaI e GacI, as quais não apresentam semelhanças significativa entre si; **(iii)** confirmação da EpaL e EpaM constituírem um transportador ABC da cadeia de ramnose; **(iv)** EpaN e EpaO podem apresentar mais do que um sítio catalítico. Para algumas proteínas Epa, não foi possível inferir os substratos das enzimas, ou nem atribuir nenhuma função biológica. Para estes últimos casos, a análise também foi realizada e consta nos anexos desta tese (Apêndice B).

O esquema da via de síntese da cadeia de ramnose com Glc e GlcNAc pode ser



**Figura 45** – Via de síntese proposta da parte inicial do EPA em *E. faecalis* ATCC19433. A EpaA sintetiza GlcNAc-P-P-Und, ao qual é adicionado L-ramnose pelas enzimas EpaBCDNO, a partir de dTDP-L-ramnose. A dTDP-L-ramnose é formada pelas enzimas EpaEGFH. A cadeia de ramnose contém  $n$  resíduos desse açúcar. O polímero de ramnose é então exportado para fora do citosol pela transportadora permease EpaL e *ATP-binding* EpaM, duas proteínas que geram o transportador ABC. A EpaI, possivelmente, auxiliada pela EpaJ, sintetiza Und-P-GlcNAc, e, talvez, Glc-P-Und. O transporte para a face externa da membrana plasmática dos GlcNAc-P-Und e Glc-P-Und, e a adição desses carboidratos talvez fique a cargo das proteínas EpaP, EpaQ e EpaK. Por fim, a EpaR catalisa a formação de Glc-P-P-Und, carreador do polissacarídeo sintetizado por proteínas codificadas na região variável do *cluster* epa. Este polissacarídeo será adicionado à cadeia de ramnose, e ficará exposto ao meio extracelular.

Fonte: Elaborada pela autora no BioRender.com.

visto na Figura 45, realizado desde a enzima EpaA à EpaR. A EpaA, uma provável proteína de membrana, gera GlcNAc-P-P-Und a partir de UDP-GlcNAc e do carreador lipídico undecaprenil fosfato. As enzimas EpaEGFH constituem a via de dTDP-L-ramnose, substrato das EpaBCDNO, as quais adicionam L-ramnose ao GlcNAc-P-P-Und. EpaB, EpaC e EpaD, EpaN e EpaO têm predições estruturais semelhantes a GalT1, uma enzima que inicia o processo de polimerização, e a EpaC a GlfT2, enzima responsável pelo alongamento da cadeia após ação da GlfT1. No mais, a sobreposição estrutural dos modelos preditos das EpaBCDNO possuem um bom alinhamento entre si, então há também possibilidade de todas catalisarem reações similares ou idênticas. A EpaN possui três domínios sanduíche  $\alpha$ - $\beta$ - $\alpha$ , enquanto que a EpaO apresenta dois desses domínios, os quais são conservados em glicosiltransferases e contêm a cavidade catalítica dessas



enzimas.<sup>124</sup> A potencial cavidade catalítica desses domínios não é tão conservada quanto na dos domínios principais.

Após completo, o polímero de ramnose é transportado, em conjunto, pelas proteínas EpaL e EpaM para fora do citosol. A EpaI sintetiza GlcNAc-P-Und a partir de UDP-GlcNAc e undecaprenil fosfato, a EpaJ, proteína ancorada à membrana plasmática, provavelmente otimiza a catálise da EpaI. Talvez a EpaI também seja responsável por sintetizar Glc-P-Und, ou a EpaR realize este papel, mas nada pode ser afirmado quanto a isto. Aqui, há um vácuo de conhecimento sobre quais proteínas efetivamente realizam o transporte de GlcNAc-P-Und e Glc-P-Und para o outro lado da membrana plasmática, assim como qual ou quais adicionam esses açúcares ao polímero de ramnose. Para estas funções, restam EpaK (seção B.1), uma pequena proteína constituída por feixes de hélice, EpaP (seção B.2), uma suposta glicosiltransferase cujo modelo estrutural apresenta enovelamento parecido a GacL, e EpaQ (seção B.3), única proteína que possui uma região não modelada pelo AlphaFold2. Estas três proteínas são preditas como proteína de membrana. Uma hipótese é que, assim como no caso das EpaI e EpaJ, a EpaK auxilie a atividade catalítica da EpaP de transferir Glc e GlcNAc ao polímero de ramnose, restando à EpaQ a função transportadora. Por fim, a EpaQ transfere Glc-1-P ao undecaprenil fosfato, resultando em Glc-P-P-Und, etapa inicial da síntese da parte decorativa do EPA, realizada por enzimas codificadas na região variável do *cluster* epa. Ao final de tudo, ainda resta transferir o EPA da membrana ao peptidoglicano. O deslocamento e a fixação do EPA ao peptidoglicano é realizada por enzimas codificadas fora do *cluster* epa.

Esta análise, realizada com o programa *AlphaFold2*, permitiu uma visão estrutural, assim como o mapeamento de possíveis funções baseadas na estrutura para a maioria das proteínas codificadas na região conservada do *cluster* epa. Embora esta designação funcional não seja possível para todas essas enzimas, este é o primeiro esforço neste sentido, apresentando o ineditismo ao trazer mais informações para a função de cada membro do *cluster* epa.



Dada a relevância dos membros do *cluster epa*, identificada pela análise de Dale e colaboradores,<sup>39,52</sup> selecionamos 4 genes para etapas experimentais: *epaI*, *epaOX*, *epaB* e *epaE*. Os experimentos foram iniciados com as enzimas EpaI e EpaOX. O gene *epaI* se encontra na região conservada do *cluster epa*, enquanto que a EpaOX é uma proteína codificada na região variável do *cluster epa*, única em algumas linhagens de *E. faecalis*. A EpaB (seção D.1) e EpaE se tornaram os objetos de estudo após o encerramento dos experimentos com a EpaI e EpaOX. A EpaE foi a única enzima a apresentar progresso na área de caracterização estrutural e bioquímica.

### 5.1 Expressão das Proteínas EpaI e EpaOX e suas Construções

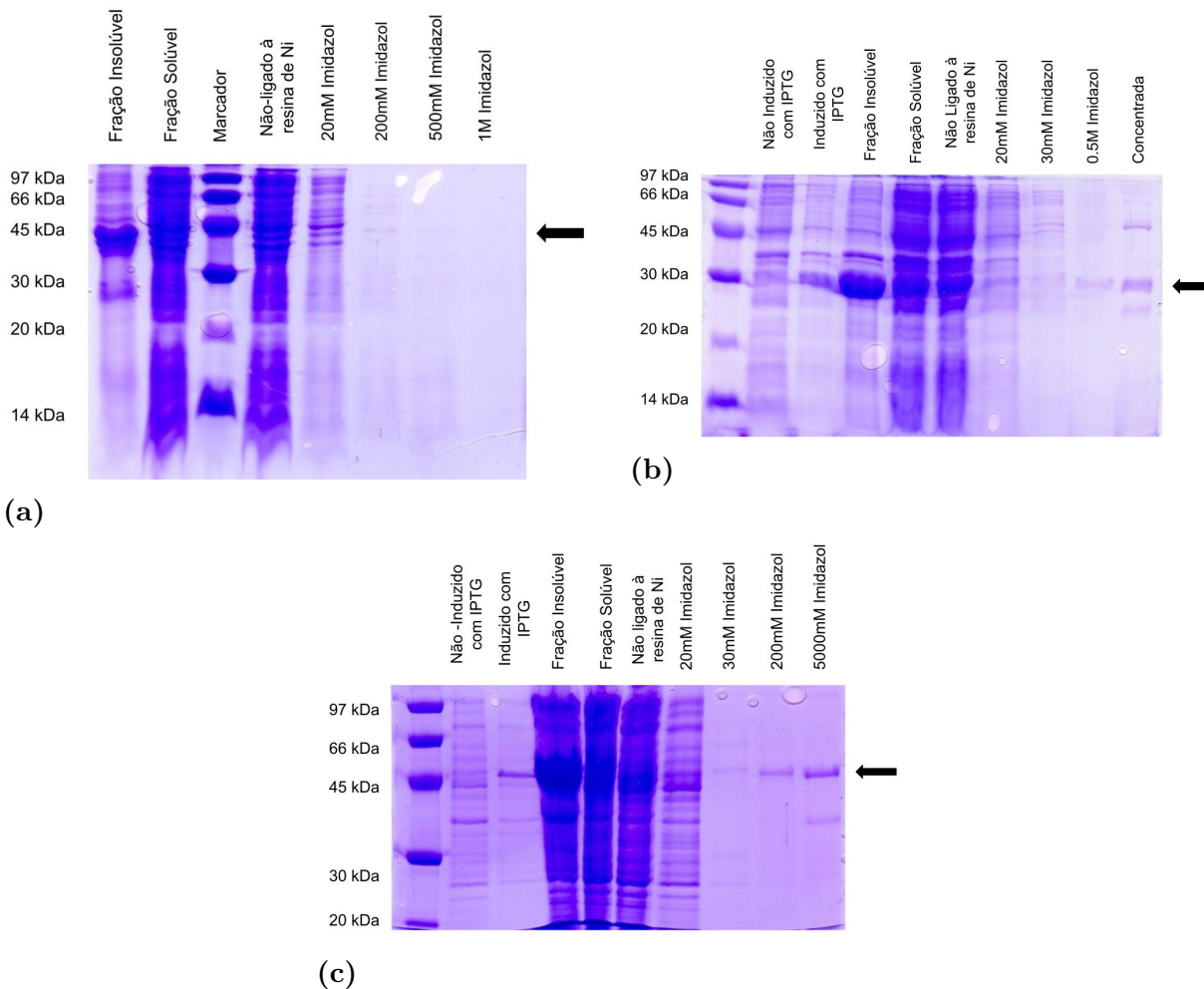
As enzimas EpaI e EpaOX apresentaram problemas de solubilidade quando expressadas no vetor pETTrx-1a e em *E. coli* Rosetta (DE3) (Figura 46). Ao final da cromatografia por afinidade em resina de níquel, a concentração final da proteína EpaI expressa no vetor pETTrx-1a (Figura 46a) não foi suficiente para os estudos estruturais, a concentração medida era da ordem do erro de leitura do equipamento NanoDrop (ThermoFischer), na faixa de  $10^{-2}$  mg.

Quando expressa no vetor pET21\_NESG, a EpaI resultante após cromatografia de afinidade apresentou uma concentração maior, em torno de  $10^{-1}$  mg por 2 L de meio de cultura, após remoção de imidazol. Ainda assim, a amostra estava impura, vide a presença de outras bandas no gel SDS-PAGE 15%, o que necessitaria de outras etapas de cromatografia e acarretaria na diminuição da quantidade de proteína resultante.

A proteína EpaOX, expressa no vetor pETTrx-1a (Figura 46c), também foi obtida impura e na mesma ordem de rendimento,  $10^{-1}$  mg por 2 L de meio de cultura após o primeiro passo de cromatografia. Outro problema identificado, tanto para a proteína EpaI quanto para a EpaOX, foi a precipitação das mesmas na presença de imidazol em concentrações mais altas. Condições distintas de expressão, como velocidade de rotação do meio, tempo de indução, temperatura de indução e concentração de IPTG foram testadas, porém nenhuma delas apresentou resultados significativos quanto à solubilidade.

Os géis SDS-PAGE 15% corados com azul de *Coomassie* seguem o seguinte padrão de colunas, da esquerda para a direita: massa molecular, meio LB antes da indução de expressão de proteína, meio LB após indução com IPTG, fração insolúvel do lisado celular, fração solúvel do lisado celular. Em seguida, estão mostradas as etapas de cromatografia

em coluna de níquel: solução não ligada à coluna de níquel e eluição da coluna com concentração crescente de imidazol.



**Figura 46** – Gel SDS-PAGE 15% corado com azul de *Coomassie* das etapas de purificação da proteína EpaI, expressa a partir (a) do vetor pETTrx-1a, com tamanho 41kDa e (b) do vetor pET21\_NESG, com tamanho 28 kDa. (c) Etapas de purificação da proteína EpaOX, expressa a partir do vetor pETTrx-1a, de tamanho 51 kDa. As expressões ocorreram em *E. coli* Rosetta (DE3).

Fonte: Elaborada pela autora.

A análise das sequências da EpaI e da EpaOX não revelou a presença de domínios transmembrana. Contudo, as predições estruturais dessas proteínas pelo PSIPRED<sup>125</sup> e XtalPred-RF<sup>126</sup> indicaram que elas possuíam consecutivas  $\alpha$ -hélices na porção C-terminal, o que também foi visualizado no modelo estrutural predito pelo *AlphaFold2*. Como mencionado no capítulo anterior (Capítulo 4), há descrições de uma proteína homóloga da EpaI, a GacI de *Streptococcus pyogenes*,<sup>86</sup> um proteína citosólica a qual forma um complexo com outra proteína ancorada à membrana plasmática, a GacJ. Segundo Rush *et al.*, 2017,<sup>86</sup> a formação desse complexo estimula a atividade da GacI. A predição estrutural das proteínas EpaJ e GacJ, e EpaI e GacI pelo *AlphaFold2* mostra que elas são muito parecidas entre si,

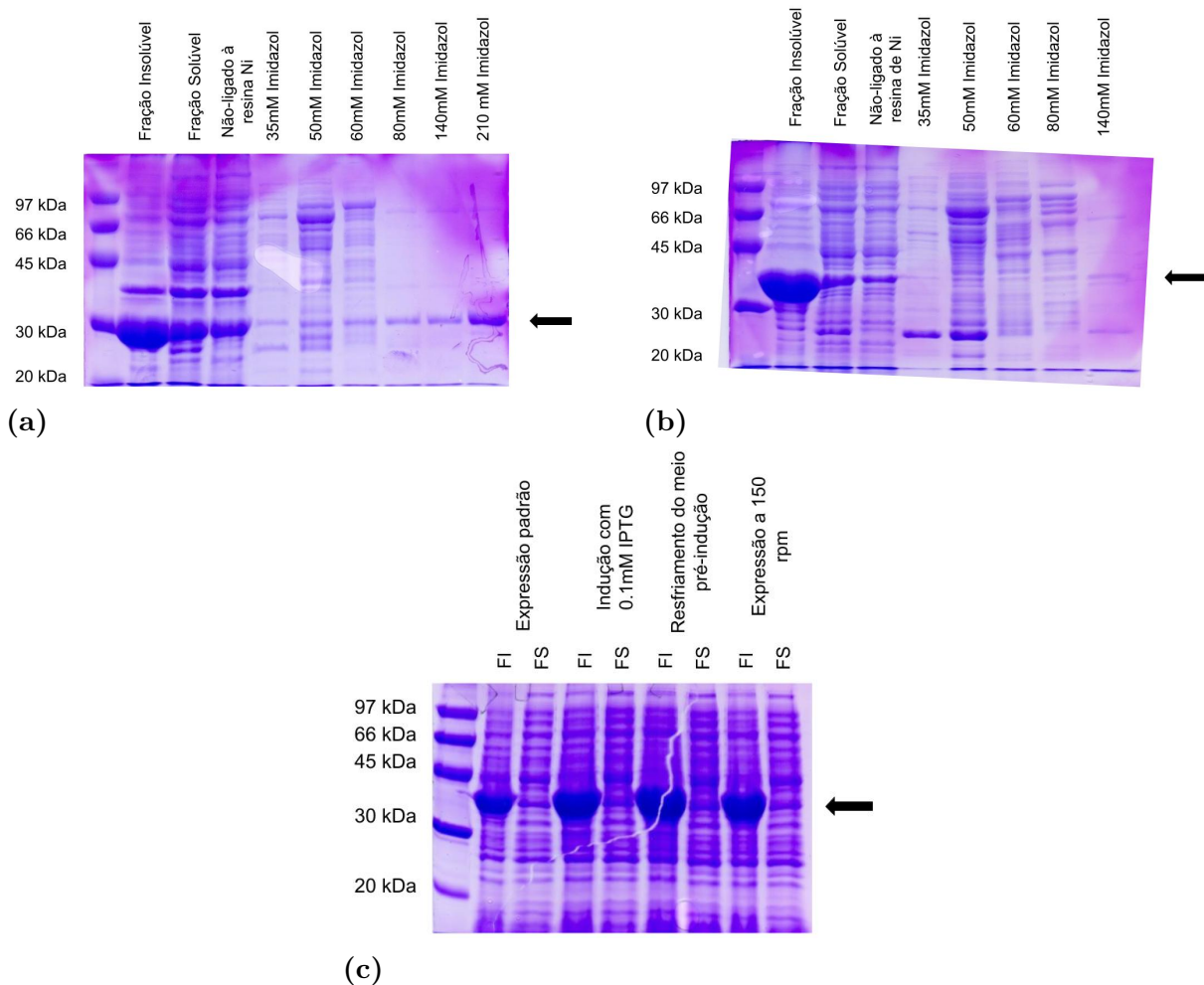
sugerindo que, assim como a GacI, a EpaI também seja associada à membrana plasmática de *E. faecalis*. No protocolo utilizado para purificação da GacI e EpaI,<sup>86</sup> há uma etapa adicional de centrifugação, com velocidade maior para separar a fração de membrana,  $30k \times g$ , velocidade não alcançada pelas centrífugas a disposição na época dos experimentos. Mesmo aplicando esta alta velocidade, não foi mencionado pelos autores o rendimento da purificação nem da GacI nem da EpaI, apenas que aplicaram entre 50  $\mu$ g a 100  $\mu$ g desta enzima nos testes bioquímicos.

Como tentativa de melhorar a solubilidade da EpaI e EpaOX, estas enzimas foram expressas após a remoção de algumas hélices terminais das mesmas. A decisão de eliminação de aminoácidos nas construções EpaI<sup>130</sup> e EpaI<sup>210</sup> foi pautada no alinhamento obtido através do BLAST<sup>127</sup> com todas as proteínas depositadas no *Protein Data Bank* (PDB).<sup>128</sup> As proteínas Dolicol fosfato manose sintase (PDB ID: 5mlz) e Poliisopreno-fosfato glicosiltransferase GtrB (PDB ID: 5ekp) apresentaram a maior identidade com a sequência de aminoácidos da EpaI, e foram alinhadas estruturalmente entre si pelo PyMOL.<sup>129</sup> Visto que os sítios catalíticos dessas glicosiltransferases se encontram na porção N-terminal, e que a similaridade estrutural entre elas terminava por volta do 130º aminoácido, a primeira construção da EpaI foi definida, a EpaI<sup>130</sup>. Para analisar o efeito de quatro  $\alpha$ -hélices na solubilidade após a posição 130, trinta aminoácidos foram removidos a partir da posição C-terminal da EpaI, gerando a construção EpaI<sup>210</sup>. As duas construções foram expressadas no vetor pETTrx-1a.

Ao contrário da proteína EpaI, as estruturas proteicas depositadas no PDB contendo melhor alinhamento com a proteína EpaOX não apresentavam similitude estrutural entre si, impossibilitando que esse método guiasse qualquer tipo de construção a partir da EpaOX. Assim, uma análise no servidor XtalPred-RF<sup>126</sup> foi feita, analisou-se a proteína EpaOX com proteínas similares, em termos de sequência de aminoácidos que favorecessem a formação de cristais de proteínas. As construções foram realizadas ao adequar o tamanho e o ponto isoelétrico (pI) tais que se enquadrassem nas médias estatísticas preditas de cristalização, ou seja, pI de 5 a 6 e número de resíduos entre 200 a 300 aminoácidos. Ao remover 48 aminoácidos do C-terminal da EpaOX, a construção EpaOX<sup>277</sup> foi criada, a qual apresentava pI 5.6 e tamanho de 277 aminoácidos. Uma segunda construção também foi sugerida, resultante da remoção de 105 aminoácidos do C-terminal da proteína EpaOX, denominada construção EpaOX<sup>220</sup>, também possuindo propriedades de cristalização alinhadas com aquelas que melhor proporcionaram a cristalização de proteínas.

Das duas construções da EpaI, a EpaI<sup>130</sup> foi eluída em maior quantidade (Figura 47a), a qual demonstrou ser mais solúvel do que a EpaI<sup>210</sup> (Figura 47b). A EpaI<sup>130</sup> ainda apresentou problemas de interação com a resina de níquel, visto pela eluição da mesma na presença de baixas concentrações de imidazol, assim como a EpaI<sup>210</sup>. Apesar do maior rendimento da EpaI<sup>130</sup>, a sensibilidade dela ao imidazol persistiu e ela precipitava

rapidamente na etapa de concentração da proteína, antes da etapa de injeção na coluna de cromatografia por exclusão de massa molecular. Isso impediu a remoção de imidazol por coluna *desalting*, a qual também requer prévia concentração da solução com a proteína para injeção de amostra em volume apropriado.



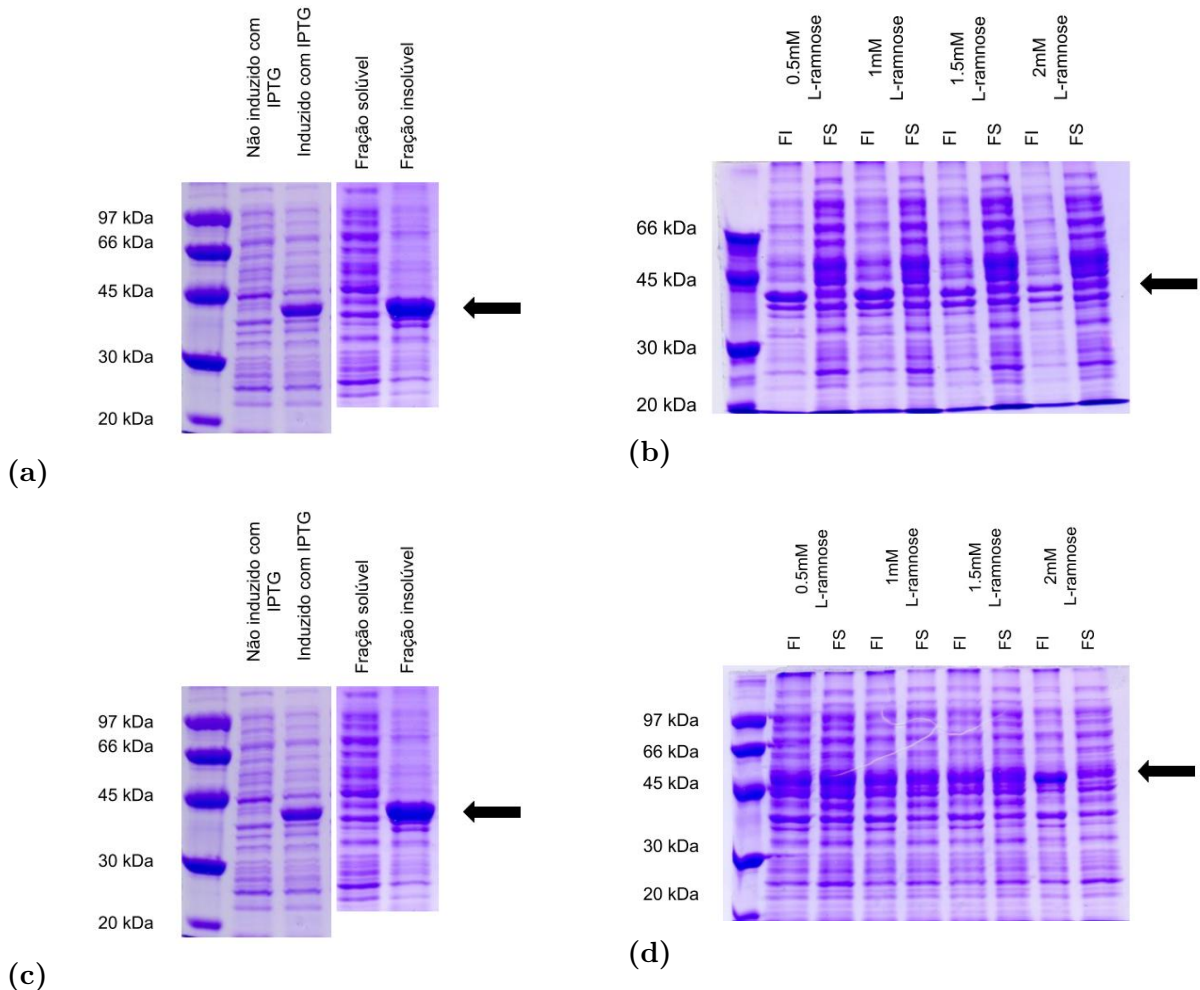
**Figura 47** – Gel SDS-PAGE 15% corado com azul de *Coomassie* das etapas de purificação (a) da construção proteica EpaI<sup>130</sup> (29 kDa), assim como (b) da EpaI<sup>210</sup> (38 kDa), e (c) dos testes de solubilidade da construção EpaOX<sup>220</sup> (39 kDa). Todas foram expressas a partir do vetor pETTrx-1a, em *E.coli* Rosetta (DE3). **FI** = Fração insolúvel, **FS** = Fração solúvel.

Fonte: Elaborada pela autora.

Não foi possível obter transformantes em *E.coli* Rosetta (DE3) a partir do vetor pETTrx-1a contendo o gene epaOX<sup>277</sup>. Em todas as tentativas de transformação bacteriana, nenhuma colônia se formou, embora a eletroforese em gel 0.8% de agarose indicasse boa integridade do DNA e o sequenciamento do plasmídeo confirmasse a correta clonagem molecular. Sendo assim, apenas a construção EpaOX<sup>220</sup> foi empregada em testes de expressão. Nos ensaios de solubilidade, essa última construção não indicou ser mais solúvel do que a EpaOX (Figura 47c).

## 5.2 Expressão das Proteínas EpaI e EpaOX em Outras Cepas de *E. coli*

Das cepas testadas, a expressão das enzimas EpaI (Figura 48a) e EpaOX (Figura 48c) só foi observada na LEMO21 (DE3) de *E. coli*. Para as linhagens restantes, BL21(DE3), pT-GroE, C43 (DE3) pLysS e ArcticExpress de *E. coli*, não foi visualizada diferença entre as bandas da amostras coletadas antes e depois da indução bacteriana no gel SDS-PAGE 15% corado com azul de *Coomassie*, apesar dos protocolos de indução terem sido aplicados.



**Figura 48** – Gel SDS-PAGE 15% corado com azul de *Coomassie* dos testes de expressão em *E. coli* LEMO21 (DE3), e solubilidade sem a presença de L-ramnose (a) da EpaI (41 kDa) e (c) da EpaOX (51kDa), e na presença de L-ramnose (b) da EpaI e (d) da EpaOX. Expressões realizadas a partir do vetor pETTrx-1a. **FI** = Fração insolúvel, **FS** = Fração solúvel.

Fonte: Elaborada pela autora.

A adição de L-ramnose anteriormente à indução bacteriana não alterou a solubilidade da EpaI (Figura 48b), diferentemente do que ocorreu com a EpaOX (Figura 48d), em que os melhores resultados foram atingidos com com 0.5 mM e 1 mM de L-ramnose. Contudo, mesmo otimizando a solubilidade das proteínas, o rendimento final delas após a cromatografia por afinidade em resina de níquel foi baixo. Neste caso, pela fraca interação

entre as caudas de histidinas co-expressas com as proteínas e a resina de níquel. Não foi observado diferença entre as proteínas EpaI e EpaOX e suas respectivas construções.

### 5.3 Solubilização com SDS e Renovelamento das Enzimas EpaI e EpaOX

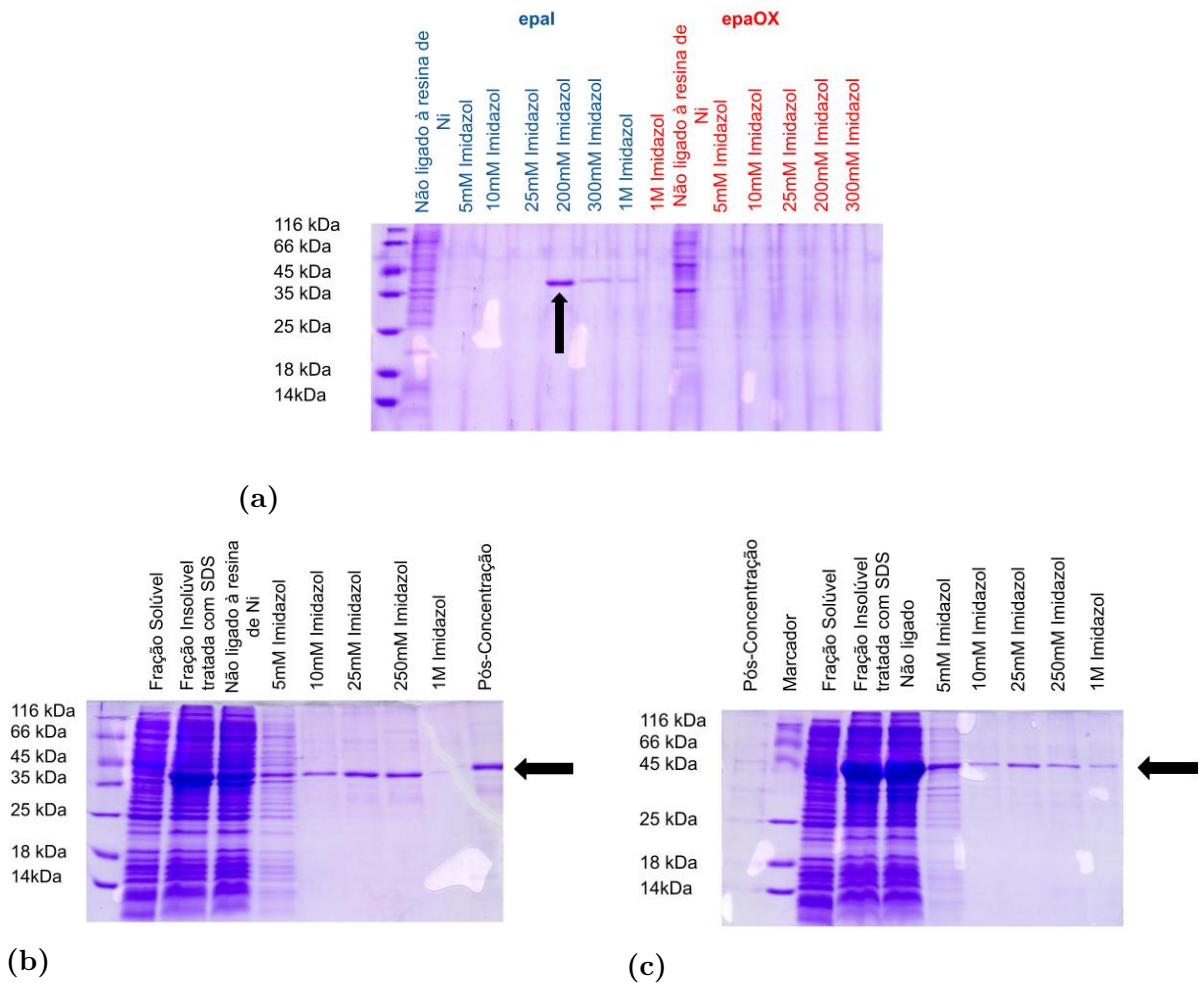
Inicialmente, ao testar a adição de detergentes para auxiliar na solubilização das enzimas, foram aplicados 20 mM CHAPS e 1% Triton X-100 no tampão de lise das proteínas EpaI e EpaOX. Entretanto, estas adições não alteraram a solubilidade dessas duas enzimas. Recorreu-se então ao protocolo de SDS para solubilização proteica. Esta metodologia é empregada para recuperar o enovelamento de proteínas que formam corpos de inclusão. Após a adição de SDS, este reagente é cuidadosamente removido de forma a manter o enovelamento das proteínas. Esta metodologia foi realizada para testar a recuperação das proteínas EpaI e EpaOX de corpos de inclusão. Inicialmente, o experimento com 0.2 % de SDS foi realizado com as enzimas EpaI e EpaOX, expressas a partir do vetor pETTrx-1a em *E. coli* Rosetta (DE3) como tentativa de melhor solubilizá-las (Figura 49a). Esta concentração de SDS não fez efeito para a EpaOX. Já para a EpaI, com 0.2% de SDS, uma maior quantidade de proteína conseguiu ser recuperada da fração insolúvel, e, em seguida, eluída em 200 mM. Ao executar a cromatografia de afinidade, a enzima EpaOX não se ligou à resina de níquel. O experimento foi repetido, aplicando agora 2% de SDS. Para a enzima EpaI (Figura 49b), não se nota grandes diferenças entre as quantidades de proteínas obtidas com 0.2% e 2% de SDS. A proteína EpaI, fusionada com sítio da TEV protease-Trx-6xHistag, foi incubada com 1 mM DTT e TEV protease na relação 1 mg de TEV protease para cada 50 mg de proteína a 10 °C, durante 12 h, entretanto, não foi observada proteína clivada. A proteína ainda precipitou após a tentativa de clivagem.

Com o tratamento de 2% de SDS com a EpaOX (Figura 49c), percebe-se que uma parte significativa da proteína foi recuperada da fração insolúvel, porém, o problema de interação com a resina de níquel ainda persistiu. O procedimento de renovelamento com 2% de SDS também foi aplicado à construção EpaOX<sup>220</sup>, expressa em vetor pETTrx-1a. Assim como no caso da proteína EpaOX, a EpaOX<sup>220</sup> apresentou problemas em se ligar e se manter ligada à resina de níquel, sendo eluída com baixas concentrações de imidazol, resultando também em rendimento insuficiente para prosseguimento experimental.

A técnica de *refolding* com SDS apresentou duas dificuldades. A primeira é a de remoção de SDS, que se precipitava a baixas temperaturas, apesar das lavagens consecutivas indicadas em He e Ohnishi, 2017.<sup>91</sup> A segunda é a necessidade de realizar experimentos enzimáticos para assegurar o enovelamento correto da proteína, o que é crítico para enzimas com ausência de protocolos sobre a medida das constantes cinéticas das mesmas, como no caso da EpaI e EpaOX.

Em resumo, as tentativas de expressão das enzimas EpaI e EpaOX envolveram aplicação vetores e cepas diferentes, além de suas construções, variando algumas condições





**Figura 49** – Gel SDS-PAGE 15% corado com azul de *Coomassie* das etapas da cromatografia por afinidade em resina de níquel, após o teste de renovelamento com (a) 0.2% SDS das enzimas EpaI (41 kDa) e EpaOX (51kDa) e os testes de 2% SDS (b) da EpaI e (c) da EpaOX, expressas a partir do vetor pETTrx-1a em em *E.coli* Rosetta (DE3).

Fonte: Elaborada pela autora.

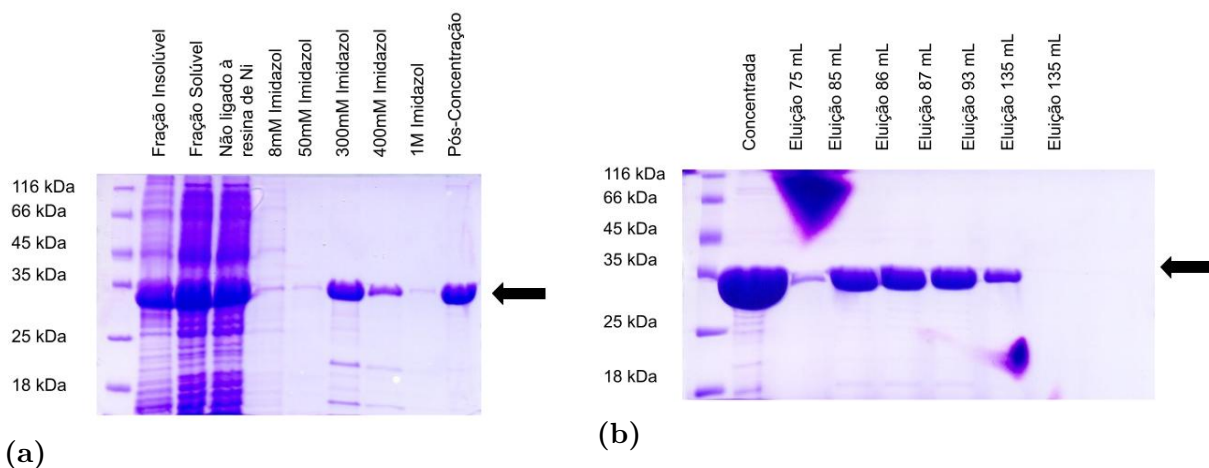
de expressão e protocolos de solubilização das mesmas. Em todos os resultados, não foi possível obter as proteínas em quantidade suficiente e estável para viabilizar ensaios de cristalização posteriores. Desta forma, os procedimentos experimentais com a EpaI e EpaOX foram interrompidos, assim como os envolvendo a EpaB (seção D.1), a qual também se apresentou pouco solúvel.

#### 5.4 Purificação da Enzima EpaE

Inicialmente, o gene *epaE* foi clonado em vetor pLIC-EGFP, o qual co-expressa, na porção C-terminal, a eGFP. Isto foi realizado para a identificação da localização da EpaE na célula de expressão, ou no citosol ou em região exterior ao citosol, a partir da fluorescência detectada em um microscópio óptico de fluorescência confocal. Este

experimento é descrito na seção D.1. Após a expressão da EpaE em vetor pLIC-EGFP, foram realizados experimentos com diferentes condições de clivagem entre a EpaE e a eGFP pela TEV protease. As condições envolveram incubação sob agitação ou não, 3h ou 18h de incubação com a TEV protease e temperatura de clivagem a 8 °C ou a 25 °C. Como não foi possível clivar a fusão proteica com a TEV, o gene *epaE* foi sub-clonado nos vetores pETM11 e pETTrx-1a.

Testes de expressão, solubilização e de cromatografia por afinidade na coluna de níquel foram realizados com os produtos da expressão de pETTM11::*epaE* e pETTrx-1a::*epaE* a partir de 100 mL de meio de cultura (Figura 50). Nos dois casos, o nível de proteína eluída na resina de níquel foi satisfatório. Pelo rendimento aparentemente maior no expresso pETTM11::*epaE*, e pela possibilidade de dispensar a etapa de clivagem com a TEV protease, uma expressão em maior escala, em 2 L de meio LB, foi efetuada, seguida pelas cromatografias de afinidade em resina de níquel e de exclusão molecular. Com a triagem de tampões pela técnica Fluorimetria Diferencial de Temperatura,<sup>105</sup> definiu-se que o tampão 50 mM Tris pH 8.5 200 mM NaCl é apropriado para a enzima EpaE.



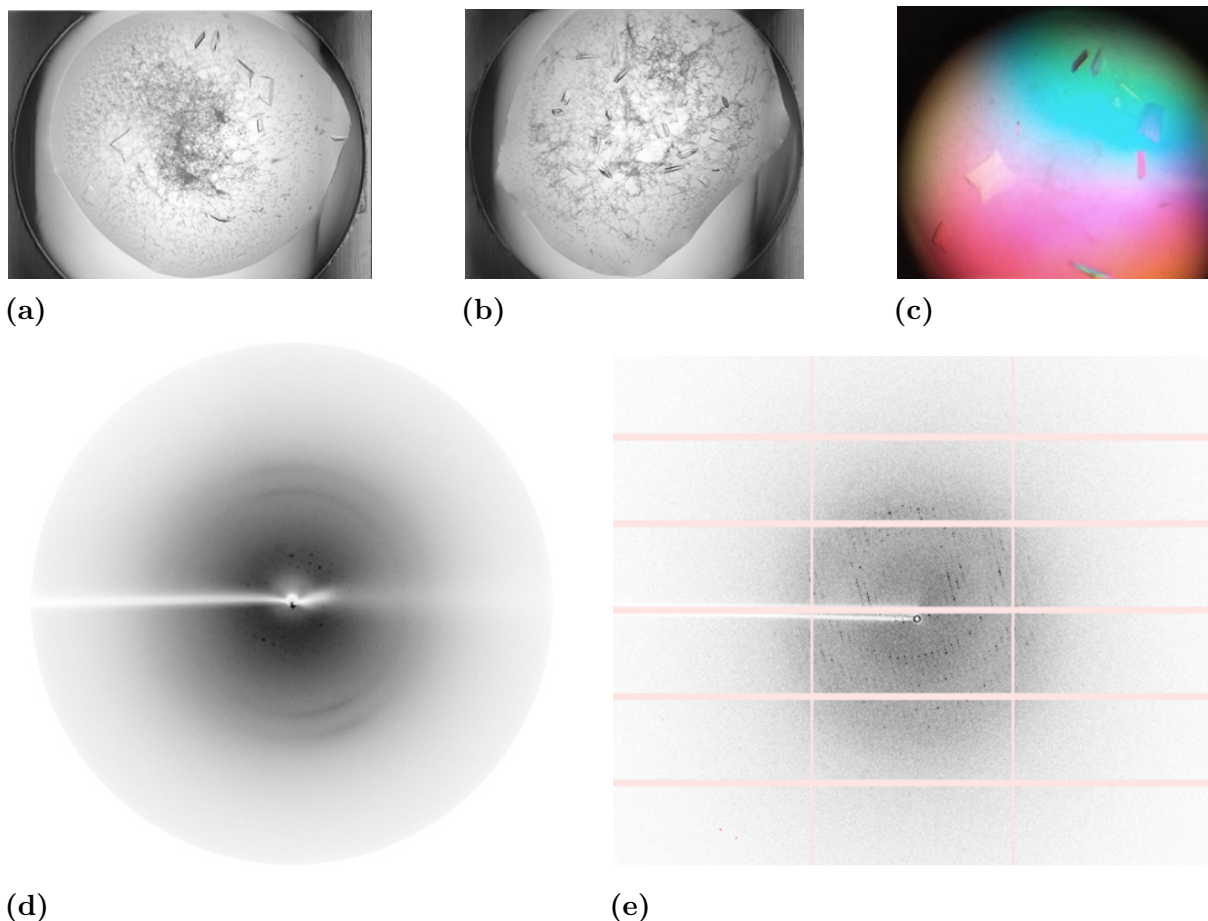
**Figura 50** – Gel SDS-PAGE 15% corado com azul de *Coomassie* (a) da cromatografia por afinidade em resina de níquel e (b) da por exclusão de massa molecular da proteína EpaE (34 kDa), expressa em vetor pETM11 e *E. coli* Rosetta (DE3).

Fonte: Elaborada pela autora.

## 5.5 Cristalização e Determinação Estrutural da enzima EpaE

A enzima EpaE purificada foi empregada em experimentos de cristalização. Na construção utilizada, pETM11::*epaE*, a enzima é expressa e purificada contendo uma sequência de seis histidinas na região N-terminal, o sítio de clivagem da enzima TEV protease e alguns aminoácidos adicionais, totalizando 23 aminoácidos co-expressos com a EpaE. Após a triagem de diversas condições de cristalização, a melhor foi a 0.1 M Tris pH 8.5, 0.2 M NaCl e 20% w/v PEG 4000, e aditivo BICINE, condição B3 do kit *Morpheus*

*Additive Screen* (Molecular Dimensions). Os cristais obtidos (Figura 51a, 51b, 51c) foram levados ao anodo rotatório do IFSC/USP, onde um padrão de difração foi obtido a baixa resolução, em torno de 8 Å. A linha Manacá, do Sirius foi disponibilizada no dia 04-12-2020, na qual um padrão de difração de maior resolução foi alcançado. Desta vez, uma resolução em torno de 2.85 Å foi obtida a partir do conjunto de dados (Figuras 51d e 51e).



**Figura 51** – (a), (b), (c) Fotos de cristais da enzima EpaE. (d) Padrão de difração obtido no anodo rotatório Rigaku Ultra X-18 e (e) na linha de luz Manacá, no Sirius.

Fonte: Elaborada pela autora.

O processamento dos dados no XDS-GUI<sup>93</sup> indicava um grupo pontual  $P222$  para o cristal. O escalonamento no AIMLESS<sup>94</sup> sugeria o grupo espacial  $P22_12_1$ . A resolução da estrutura por substituição molecular no PHASER-MR,<sup>95</sup> com um valor TFZ de 29 para 2 moléculas na unidade assimétrica, foi seguida por reconstrução do modelo pelo PHENIX.AUTOBUILD,<sup>97</sup> aplicando os parâmetros de célula do grupo  $P22_12_1$ . Neste caso, valores de *R-work* e de *R-free* não diminuíam durante o refinamento no PHENIX.REFINE<sup>98</sup> em conjunto com o refinamento manual no Coot. Análises com o REFMAC<sup>99</sup> sugeriram a possibilidade de *twinning*. O mesmo ocorreu nas análises com PHENIX.XTRIAGE.<sup>130</sup> Desta forma, optamos por reduzir a simetria para  $P2_1$  e realizar as etapas posteriores de

**Tabela 2** – Sumário das estatísticas da coleta de dados e do refinamento do cristal formado pela EpaE.

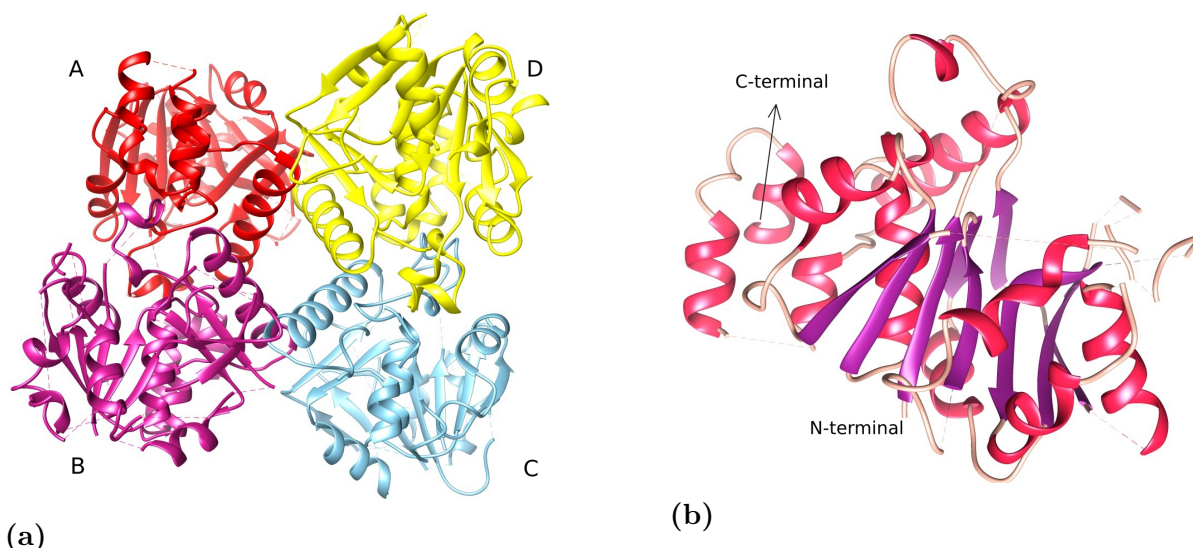
Dados da coleta	
Comprimento de onda (Å)	1.324
Limite de resolução (Å)	46.53 - 2.85 (2.952 - 2.85)
Grupo espacial	P 1 21 1
Célula unitária (Å, °)	53.585 134.32 64.523 90 90 90
Reflexões totais	42614 (4242)
Reflexões únicas	21340 (2123)
Multiplicidade	2.0 (2.0)
Completeza (%)	99.93 (99.91)
Média I/σ(I)	14.71 (1.12)
R-merge	0.03379 (0.7094)
R-means	0.04779 (1.003)
R-pim	0.03379 (0.7094)
CC1/2	0.999 (0.448)
CC*	1 (0.787)
Reflexões aplicadas no refinamento	21339 (2121)
Reflexões aplicadas R-free	1061 (88)
R-work	0.2429 (0.3372)
R-free	0.2900 (0.3465)
CC(work)	0.911 (0.483)
CC(free)	0.870 (0.467)
Números de átomos não-hidrogênio	6576
Macromoléculas	6559
Solvente	17
Resíduos da proteína	968
RMS(ligação)	0.003
RMS(ângulos)	0.70
Resíduos favoráveis no gráfico Ramachandran (%)	94.85
Resíduos permitidos no gráfico Ramachandran (%)	5.04
Ramachandran outliers (%)	0.11
Outliers de rotâmeros (%)	0.20
Clashscore	14.33
B-fator Médio (Å <sup>2</sup> )	94.66
Macromoléculas	94.87
Solvente	72.48
Número de grupos TLS	18

Fonte: Elaborada pela autora.

refinamento, considerando a regra de *twinning* pseudomeroedral “-h, -k, l”. Com a simetria para o grupo  $P2_1$ , os valores R-work/R-free diminuíam durante o refinamento do modelo. O modelo refinado apresentou valores 0.24 e 0.29, para R-work e R-free, respectivamente (Tabela 2). Tentou-se cristalizar a EpaE, nas mesmas condições descritas, adicionando

dTTP,  $Mg^{+2}$  e glicose-1-fosfato. Contudo, nenhum cristal foi formado.

Todas as enzimas RmlA com estrutura conhecida são tetraméricas ou formam um tetrâmero a partir das operações de simetria do cristal. A proteína EpaE também forma um tetrâmero com enovelamento do tipo Rossmann, constituído por uma alternância  $\alpha$ -hélice e fita  $\beta$  (Figura 52). Segundo PDBePISA,<sup>131</sup> cada monômero apresenta 243 aminoácidos, e a área de superfície de cada um deles varia entre 12199  $\text{\AA}^2$  e 12733  $\text{\AA}^2$ . Para o tetrâmero, a área de superfície total é de 41780.9  $\text{\AA}^2$ . O tetrâmetro possui área média de em de 1176.3  $\text{\AA}^2$  na interface entre as cadeias A e B, e entre as cadeias C e D, com energia livre de solvatação média ( $\Delta^iG$ ) do arranjo dessas interfaces de  $-22.2$  kcal/mol. Já para as interfaces entre cadeia B e C, e A e D, a área média é de 713.6  $\text{\AA}^2$ , e  $\Delta^iG$  médio =  $-4.5$  kcal/mol. A maior estabilidade entre as cadeias A e B, e C e D em comparação com as B e C, e A e D, indica que a EpaE também é um dímero de dímeros. Ao comparar cada monômero entre si, através do alinhamento estrutural, são observadas poucas dissimilarças entre eles, destacando diferentes regiões de descontinuidade da cadeia principal entre eles.



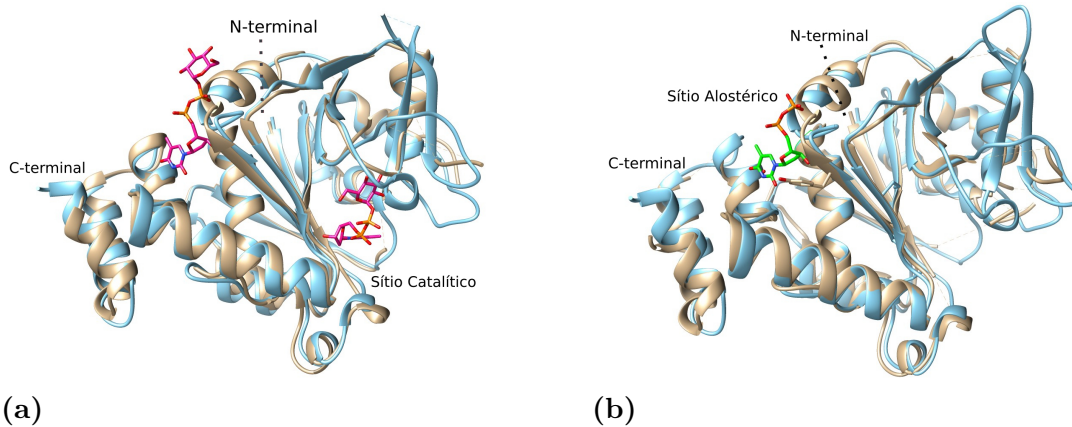
**Figura 52** – (a) Estrutura do tetrâmero da enzima EpaE refinada. (b) Visão lateral do monômero da cadeia A.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

Há reportado na literatura a presença de uma sítio alostérico nas RmlA,<sup>132–133</sup> o qual comporta algumas moléculas, como dTDP, timidina ou até o produto G-1-P.<sup>76</sup> Para identificá-lo na EpaE, a estrutura da EpaE foi sobreposta com a da RmlA de *M. tuberculosis*, PDB ID: 6b5e, assim como para identificação do sítios catalítico da mesma, indicados na Figura 53.

### 5.5.1 Sítio Alostérico

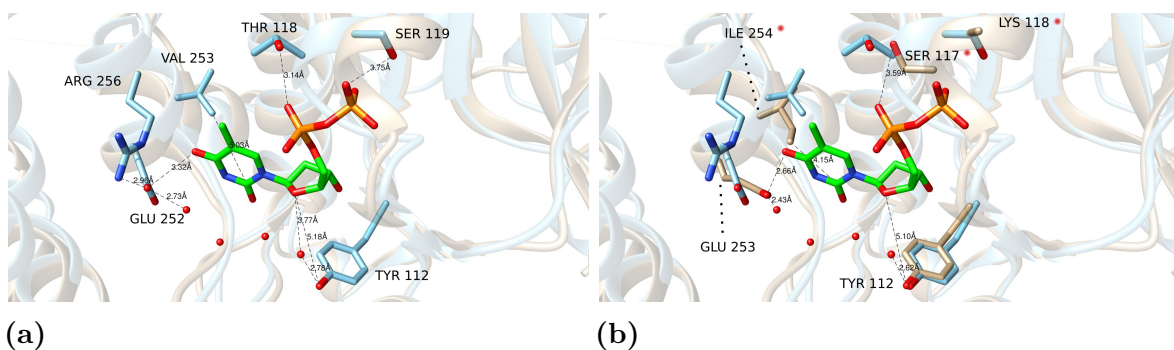
A RmlA de *M. tuberculosis* (PDB ID: 6b5e) é uma estrutura de boa resolução, com 1.85  $\text{\AA}$ . O alinhamento entre as estruturas da EpaE e RmlA de *M. tuberculosis* (PDB ID:



**Figura 53** – Estrutura da EpaE, em dourado, sobreposta com a da RmlA de *M. tuberculosis* (PDB ID: 6b5e), em azul, co-cristalizada com o (a) produto dTDP-Glicose, em rosa, na cadeia F da RmlA (PDB ID: 6b5e), e (b) com o dTDP, em verde, na cadeia B da RmlA (PDB ID: 6b5e).

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

6b5e) é melhor observado ao restringir uma das cadeias da EpaE para sobrepor a PDB ID: 6b5e, resultando em *Z-score* 30.1, RMSD 1.5 e 63% de identidade, de acordo com o servidor DALI.<sup>79</sup> No caso da análise da região do sítio alostérico da PDB ID: 6b5e, a cadeia B desta foi sobreposta com a cadeia A da EpaE. Exceto por dois, é possível observar os mesmo aminoácidos em posições parecidas nas duas estruturas. Os resíduos do sítio alostérico da PDB ID: 6b5e são TYR112, TYR118, SER119 e GLU252, os quais interagem com a molécula de dTDP (Figura 54a). A VAL253 também está presente nesta cavidade, único resíduo apolar desse sítio alostérico. Na EpaE, são vistos TYR112, SER117, GLU253 e ILE254 em posições parecidas aos resíduos do sítio alostérico da PDB ID: 6b5e.

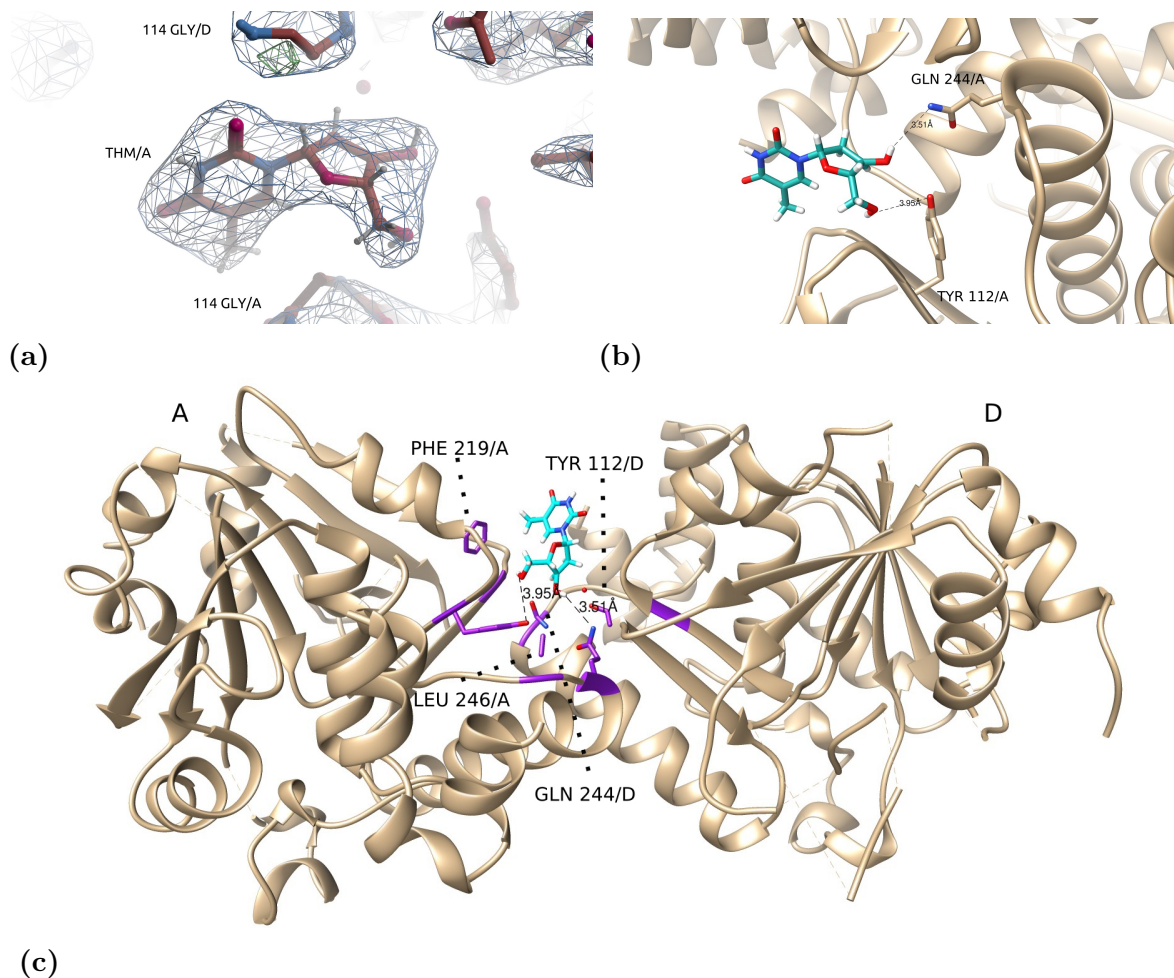


**Figura 54** – Cadeia A da estrutura da EpaE, em dourado, sobreposta com a da RmlA de *M. tuberculosis* (PDB ID: 6b5e), em azul, co-cristalizada com o produto dTDP, em verde. (a) Identificação dos resíduos presentes no sítio alostérico da RmlA PDB (ID: 6b5e) para comparação com os da (b) EpaE. Estão destacados com um ● os resíduos diferentes entre a RmlA (PDB ID: 6b5e) e a EpaE.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

### 5.5.1.1 Timidina

Durante o refinamento, foi identificada uma densidade eletrônica próxima das GLY144 das cadeias A e D da EpaE. Após testar o ajuste de algumas moléculas que interagem com outras RmlAs nessa região, foi observado que a densidade eletrônica era correspondente à molécula timidina (Figura 55). A região desta densidade eletrônica é reportada como sendo um sítio alostérico de RmlAs,<sup>76,132–133</sup> adjacente ao sítio descrito anteriormente da RmlA de *M. tuberculosis* (PDB ID: 6b5e).

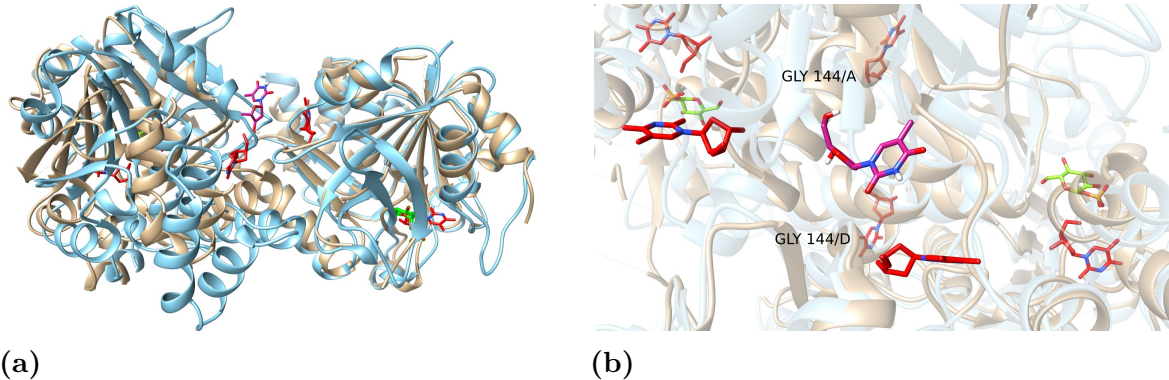


**Figura 55** – (a) Densidade eletrônica correspondente a da timidina, próxima às GLY144 das cadeias A e D da EpaE. (b) Potenciais resíduos que interagem com a timidina, em ciano. (c) Visão completa da região de interação da timidina com as cadeias A e D da EpaE.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

A timidina está em uma cavidade com volume de  $529.92\text{\AA}^3$  e superfície  $1026.42\text{\AA}^2$ . Há dois resíduos com capacidade de realizar ligação hidrogênio com a timidina, TYR112 e GLN244, sugerindo que este sítio é pouco específico, como o esperado de um sítio alostérico. Os seguintes resíduos também estão próximos da timidina: PHE219 e LEU246 da cadeia A, TYR112 e GLN244 da cadeia D, configurando uma cavidade polar. A timidina é também encontrada na interface entre dois monômeros de RmlAs, como por exemplo, no caso

da RmlA de *E. coli*, PDB ID: 1h5r (Figura 56). Este tetrâmero, assim como o da estrutura da RmlA de *M. tuberculosis* (PDB ID: 6b5e), não se alinha ao tetrâmero da EpaE, o alinhamento é visto em apenas um dos monômeros entre essas RmlAs. Para que os tetrâmeros se alinhassem entre si, seria necessária uma mudança conformacional entre um deles.



**Figura 56** – (a) Sobreposição entre a EpaE, em dourado, e a RmlA de *E. coli* (PDB ID: 1h5r), com destaque para a timidina ligada à EpaE, em magenta, timidinas, em vermelho, e glicose-1-fosfato, em verde, ligadas à RmlA PDB ID: 1h5r. (b) Ampliação do sítio de ligação da timidina com a EpaE. Alinhamento estrutural com *Z-score* 29.9, RMSD 1.4 e 72% de identidade.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

### 5.5.2 Sítio Catalítico

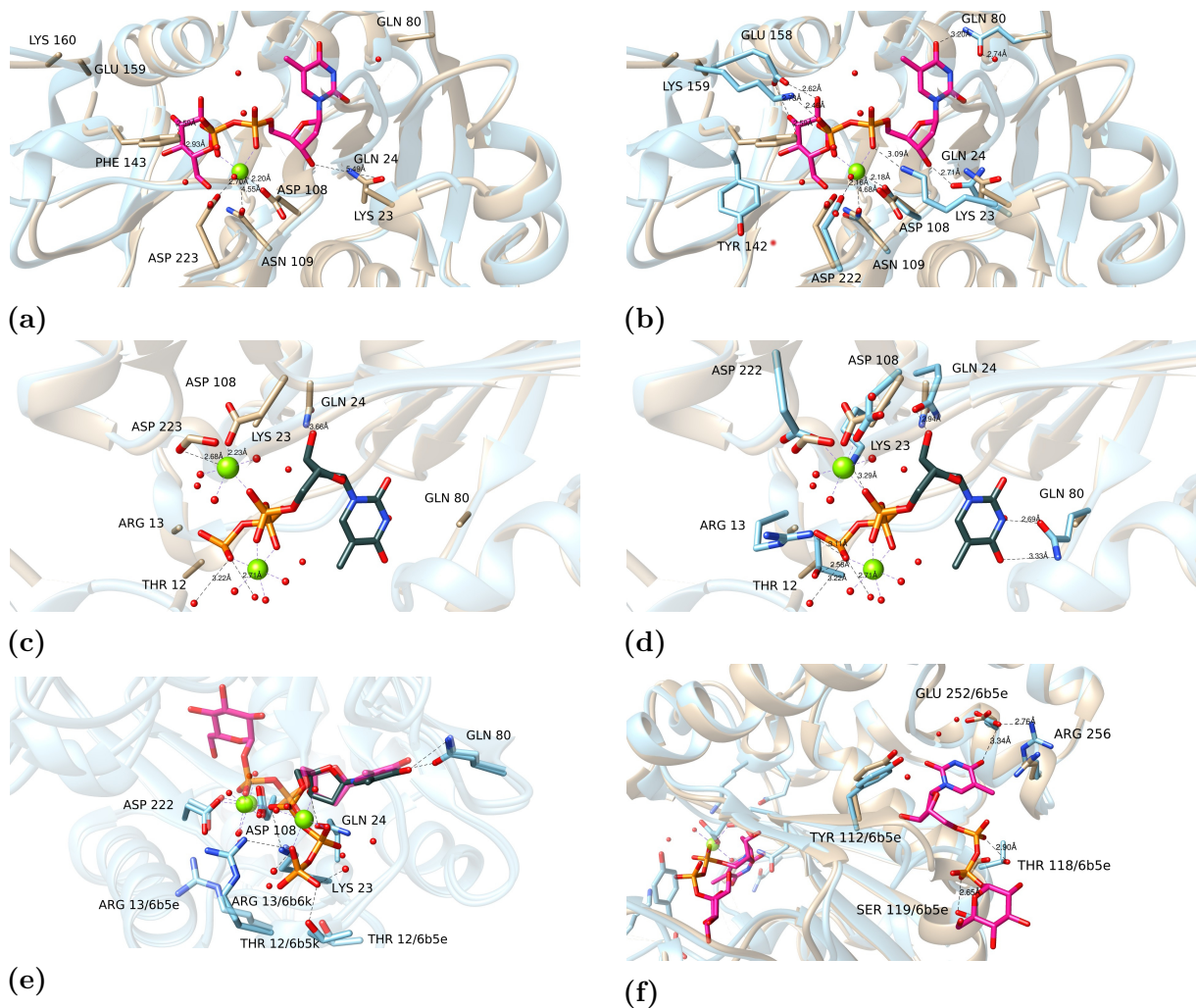
Os resíduos catalíticos da EpaE foram identificados (Figura 57a) a partir da análise dos da PDB ID: 6b5e e PDB ID: 6b5k (Figura 57b).<sup>70</sup> Os resíduos THR12, ARG13, LYS23, GLN80, GLU159 e LYS160 não possuem cadeia lateral reconstruída, mas, pelo alinhamento estrutural, é provável que estejam envolvidos na interação com o produto dTDP-glicose e produto dTTP.

#### 5.5.2.1 dTDP-Glicose

Nesta seção, são destacados os resíduos que interagem com o produto dTDP-glicose. Os resíduos envolvidos nas ligações de hidrogênio com esse produto são: a GLU159 com o oxigênio do carbono C-2 da glicose e uma molécula de água, e a LYS160 com o oxigênio O3P do fosfato  $\beta$ . A PHE143 pode realizar uma interação empilhamento pi com o anel de glicose e a GLN24 com o grupo hidroxila do carbono C-3. Os resíduos ASP108, ASP109 e ASP223 estariam envolvidos na estabilização do cátion  $Mg^{+2}$ , que interage com a glicose-1-fosfato. Aqui, a LYS23 está a uma distância de 3.09 Å o oxigênio O1P do fosfato  $\alpha$ , e o GLN80 a 3.20 Å do oxigênio do carbono C-4 da timina. Esses dois resíduos também estão próximos do dTTP, co-cristalizada com a mesma RmlA de *M. tuberculosis*, a PDB ID:



6b5k. Para entender melhor o papel desses dois resíduos, é feita a análise da interação entre a RmlA e o dTTP a seguir.



**Figura 57** – Cadeia D da estrutura da EpaE, em dourado, sobreposta com a cadeia F da RmlA PDB ID: 6b5e, em azul, co-cristalizada com o produto dTDP-Glicose, em rosa. (a) Identificação dos resíduos do sítio catalítico da EpaE a partir dos (b) resíduos do sítio catalítico da RmlA PDB ID: 6b5e que interagem com a dTDP-Glicose. (c) Identificação dos resíduos do sítio catalítico da EpaE a partir dos (d) resíduos do sítio catalítico da RmlA PDB ID: 6b5k que interagem com o dTTP, em cinza escuro. (e) Sobreposição das estruturas da RmlA de *M. tuberculosis*, PDB ID: 6b5e com 30% de transparência, e PDB ID: 6b5k, ressaltando os resíduos em comum entre elas. (f) Produto dTDP-Glicose ligado a uma região distinta do sítio catalítico da RmlA PDB ID: 6b5e. Estão destacados com um ● os resíduos diferentes entre a RmlA PDB (ID: 6b5e e 6b5k) e a EpaE.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

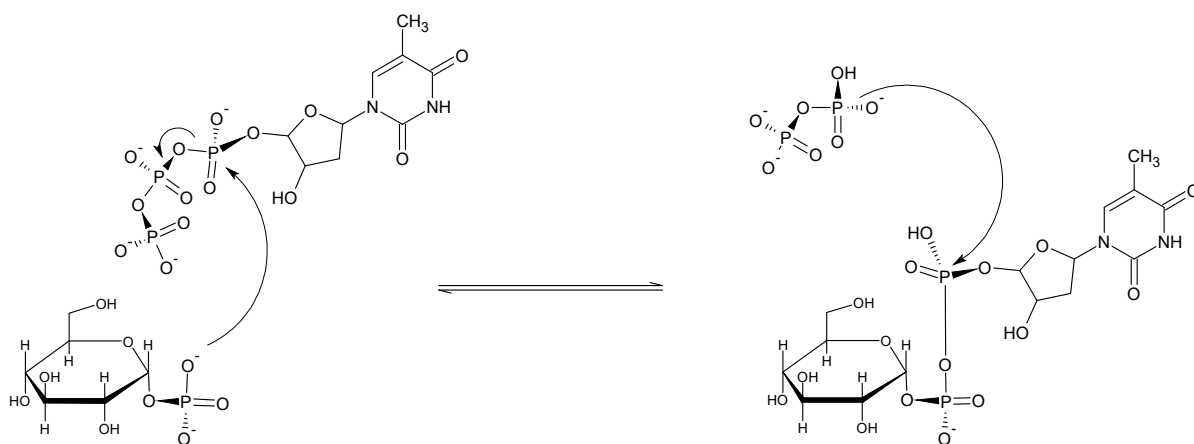
### 5.5.2.2 dTTP

Os resíduos da EpaE relacionados na orientação do dTTP para receber o ataque nucleofílico do sítio catalítico estão destacados na Figura 57c. Esta identificação é baseada

na sobreposição estrutural entre a EpaE e a RmlA de *M. tuberculosis* (PDB ID: 6b5k) (Figura 57d), especialmente nos casos dos resíduos nos quais as cadeias laterais não foram reconstruídas. Os resíduos que podem realizar ligação de hidrogênio com o dTTP são: THR12 com o oxigênio O3G do fosfato  $\alpha$ , ARG13 com o oxigênio O3G do fosfato  $\gamma$ , GLN24 com a hidroxila do anel de ribose, LYS23 com o oxigênio O1A do fosfato  $\alpha$ , GLN80 com o oxigênio do carbono C-4 e grupo amina do anel de timina. Os resíduos ASP108 e ASP223 estão envolvidos na estabilização de um dos íons  $Mg^{+2}$  presentes no sítio catalítico da PDB ID: 6b5k, responsável pela estabilização do grupo fosfato de tanto da dTDP-glicose quanto do dTTP (Figura 57e).

Ao alinhar as estruturas da RmlA de *M. tuberculosis* (PDB ID: 6b5e e 6b5k), através da sobreposição da dTDP-glicose e dTTP, é possível perceber os resíduos que interagem com os substratos dTTP e G-1-P (Figura 57e). Sendo assim, na EpaE, os resíduos THR12, ARG13, LYS23, GLN24 e GLN80 podem estar diretamente envolvidos no direcionamento do dTTP durante a catálise, e os PHE143, GLU159 e LYS160 no da glicose-1-fosfato.<sup>133</sup> Já os resíduos ASP108, ASN109 e ASP222 podem estar auxiliando a orientação do íon  $Mg^{+2}$  para estabilização dos grupos fosfatos de ambos G-1-P e dTTP. O mecanismo da reação catalisada pela EpaE consiste de uma substituição nucleofílica bimolecular (Figura 58). A ARG13 se dobra em torno do dTTP, enquanto que o lado oposto é ocupado pela G-1-P.<sup>134-133</sup> A conformação tensa do dTTP, promovida pela ARG13, ativa-o para receber o ataque nucleofílico no fosfato  $\alpha$  pelo oxigênio O1P da G-1-P levando à quebra de ligação entre o fosfato  $\alpha$  e o pirofosfato do dTTP.<sup>133-134</sup> É possível ver que a THR12 e a ARG13 adotam outra conformação quando o produto dTDP-glicose é formado (Figura 57e). Na literatura, é descrito que uma lisina na posição 162 é importante para o correto posicionamento e estabilização da carga negativa da G-1-P, através de uma interação eletrostática de aproximadamente 2.5 Å entre o oxigênio carregado O4P da G-1-P.<sup>132</sup> Na RmlA (PDB ID: 6B5e), há a LYS159 a uma distância de 2.46 Å do oxigênio O3P da dupla ligação com o átomo de fósforo da G-1-P, capazes de formarem uma ligação de hidrogênio. Na EpaE, há a LYS160.

A cadeia F da estrutura da RmlA de *M. tuberculosis* (PDB ID: 6b5e), que se alinha com a D da EpaE, foi escolhida para exposição dos resíduos catalíticos, isto devido a presença da dTDP-glicose no sítio onde se esperava encontrar apenas dTDP, indicando uma característica de inespecificidade deste sítio. Na Figura 57f, além dos resíduos TYR112, THR118, SER119 e GLU252 destacados anteriormente, a ARG256 está mais próxima do resíduo GLU252, a uma distância de 2.74 Å, em comparação com a vista anteriormente, de 2.96 Å (Figura 54a), promovendo uma interação eletrostática mais forte entre esses dois resíduos, e estabilizando mais o GLU252, o qual está em uma ligação de hidrogênio com oxigênio do carbono C-4.



**Figura 58** – Mecanismo da reação catalisada pela RmlA. O oxigênio O1P da glicose-1-fosfato promove um ataque nucleofílico no fosfato  $\alpha$  do dTTP. A RmlA orienta essas duas moléculas de modo a facilitar essa substituição nucleofílica bimolecular.

Fonte: Adaptada de BLANKENFELDT *et al.*<sup>132</sup> no ChemSketch™.

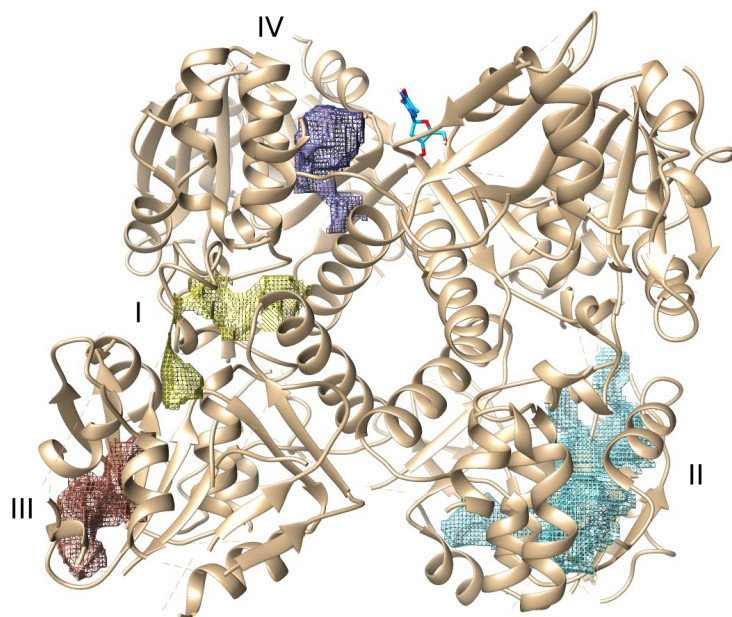
### 5.5.3 Cavidades

ProteinsPlus<sup>135</sup> é um agregador de ferramentas direcionado a analisar interação proteínas-ligantes. Dentre estas ferramentas, está a identificadora de cavidades, chamada *DoGSiteScorer*. Além de características físicas, como o volume da cavidade, o *DoGSiteScorer* fornece um parâmetro nomeado *DrugScore*, uma função de pontuação quantificadora da *drogabilidade*. A *drogabilidade* é a probabilidade de uma região da proteína alvo de uma doença ser modulada por uma molécula pequena,<sup>136</sup> isto é calculado pelos seguintes parâmetros: polaridade, hidrofobicidade, volume e a razão superfície volume da cavidade, o invólucro.<sup>137</sup> Uma cavidade com *DrugScore* maior do que 0.5 já indica a *drogabilidade* da mesma, e quanto mais próximo de 1 esse valor é, maior a confiabilidade da função de pontuação.<sup>137</sup>

Na estrutura da EpaE, foram identificadas quatro cavidades principais (Figura 59), I, II, III e IV, com *DrugScore* 0.81, 0.81, 0.79 e 0.73, respectivamente. A cavidade I intercepta uma seção do sítio catalítico da cadeia C, e a cavidade IV, da D, mas nenhuma abrange completamente as regiões alostéricas e catalíticas discutidas anteriormente. A cavidade I está presente na interface entre as cadeias D e C, a II na cadeia B, a III na cadeia C e a IV na cadeia D da EpaE. Não foi identificada uma cavidade que abrangesse por completo a molécula de timidina.

## 5.6 Triagem de Potenciais Ligantes

Uma vez que a estrutura obtida da EpaE possuía uma baixa resolução, e densidade eletrônica não permitia o traçado completo da cadeia ou mesmo a modelagem de todas as



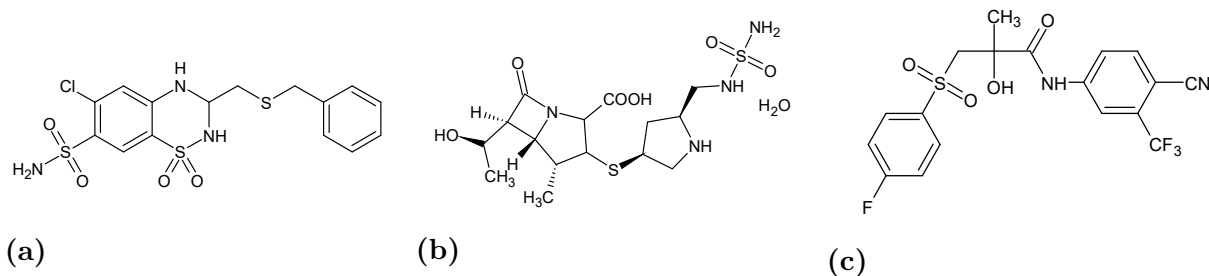
**Figura 59** – Cavidades identificadas na EpaE a partir da ferramenta *DoGSiteScorer* do *ProteinsPlus*, enumeradas em ordem decrescente do *DrugScore* de cada uma. A I se apresenta na interface entre as cadeias D e C da EpaE, a II na cadeia B, a III na C e a IV na cadeia D.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

cadeias laterais, optamos por empregar uma estratégia não baseada em estrutura para a triagem de potenciais inibidores. Neste caso, a triagem baseada na similaridade com um outro ligante bioativo foi empregada.

No total, foram selecionadas 627 moléculas a partir da triagem do programa LiBELa,<sup>102–103</sup> possuindo similaridade de pelo menos 70% com dTTP, calculada pelo índice de Hodgkin.<sup>104</sup> Dessas, apenas 33 moléculas restaram após inspeção visual. Nesta última etapa, foram consideradas: boa sobreposição com o dTTP, a presença de grupos químicos em posições equivalente às do dTTP, e moléculas que não apresentam grupos trifosfatos, pois uma eventual clivagem deles pela EpaE impediria a identificação da inibição, já que essa é a base do método de medida de atividade enzimática. Dos 33 compostos, apenas 8 estavam disponíveis para compra pelo fornecedor Sigma-Aldrich. Mesmo aplicando os protocolos de solubilização de cada molécula, não foi possível atingir uma solução aquosa homogênea de todas elas. Como o objetivo era aplicar uma relação de massas 3:1 entre ligante e EpaE, não foram realizadas diminuições nas massas dos potenciais ligantes em solução aquosa. No fim, uma molécula se solubilizou em água, o Doripenem mono hidratado, e duas precisaram ser dissolvidas em DMSO para então serem diluídas em água, a Benzotiazida e a Bicalutamida, e com essas três foram feitas os testes seguintes (Figura 60).

A Benzotiazida e Bicalutamida foram solubilizadas em DMSO. No caso destas, a medida por Fluorimetria Diferencial de Temperatura foi realizada a 0.5% e 2.5% de DMSO,



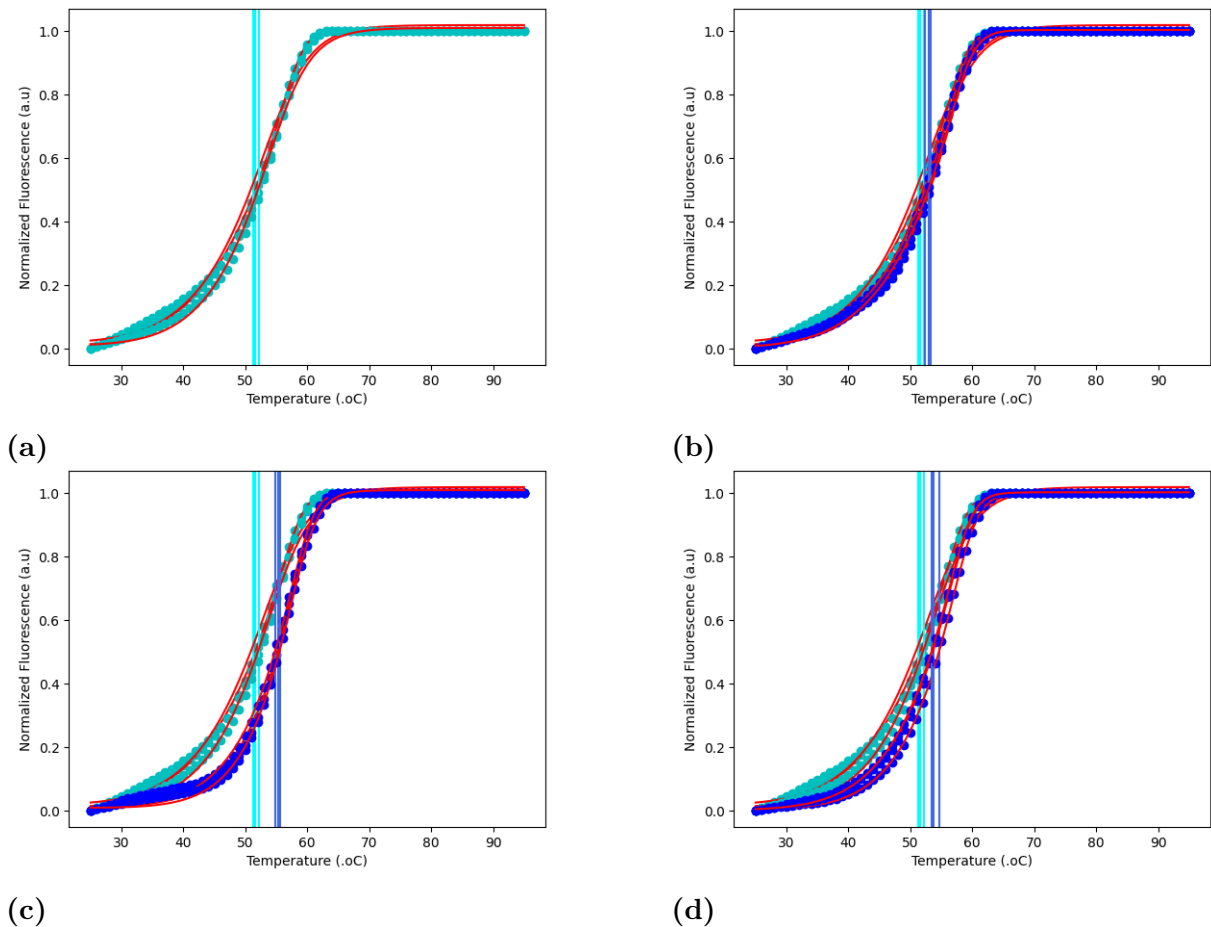
**Figura 60** – Estrutura química dos compostos (a) Benzthiazida, (b) Doripenem mono hidratado, e (c) Bicalutamida.

Fonte: Elaborada pela autora no ChemSketch <sup>TM</sup>.

esta última para aumentar a concentração delas em ambiente aquoso. Este experimento foi realizado em triplicatas (Figura 79 do Apêndice C), visando testar o efeito na EpaE na presença dessas moléculas. Em tampão 50 mM Tris pH 8.0, 100 mM NaCl e 10% Glicerol, a  $T_m$  da EpaE é  $(48.7 \pm 0.4) ^\circ\text{C}$ , quando adicionado 2.5% DMSO, este valor cai para  $(43.7 \pm 0.4) ^\circ\text{C}$ . Essas duas condições são os brancos das medidas, cujas  $T_m$  foram utilizadas para comparar com outras oriundas da adição dos potenciais ligantes. Na presença dos substratos dTTP e G-1-P e do co-fator  $\text{Mg}^{+2}$ , a  $T_m$  aumenta para  $(67.7 \pm 0.4) ^\circ\text{C}$ .

Os primeiros resultados obtidos na desnaturação térmica da enzima EpaE na presença de Doripenem, Benzthiazida e Bicalutamida não demonstraram aumento significativo na  $T_m$ . Contudo, observou-se nestes experimentos que a presença do DMSO, mesmo em concentração de 2.5% (v/v) foi suficiente para reduzir a  $T_m$  da proteína em  $5 ^\circ\text{C}$ , como dito acima. Por outro lado, a presença do dTTP aumenta a  $T_m$  da proteína em aproximadamente  $11 ^\circ\text{C}$ , como observado em experimentos preliminares anteriores. Desta forma, hipotetizamos que a presença do dTTP possa ser um fator estabilizador necessário para a enzima EpaE. Esta hipótese é também baseada na evidência experimental do nosso próprio grupo e da literatura,<sup>132</sup> a qual mostra que algumas enzimas RmlA apenas são cristalizadas na presença de dTTP. Dados iniciais obtidos pelo Mestrando Louis Oliveira mostram ainda uma mudança no estado de compactação da enzima RmlA de *Streptococcus pneumoniae* quando o dTTP é adicionado à solução. Desta forma, optamos por reavaliar a potencial interação da enzima EpaE com os ligantes identificados na triagem computacional mantendo uma concentração equimolar de dTTP em relação à proteína. O racional por trás deste desenho experimental é que o dTTP possa ocupar o sítio ativo de cada monômero da EpaE, mas que o sítio alostérico permaneça disponível para a potencial interação com os candidatos a ligantes.

No novo experimento (Figura 61), uma relação de concentrações molares de no mínimo 1:1:3 foi empregada para a enzima EpaE, dTTP e dos candidatos a ligantes. Em todas as condições, havia também presente 5 mM  $\text{MgCl}_2$ . A  $T_m$  observada para a EpaE nesta condição foi de  $(51.3 \pm 0.4) ^\circ\text{C}$ . Com o Doripenem, a  $T_m$  subiu para  $(52.3 \pm 0.4)$



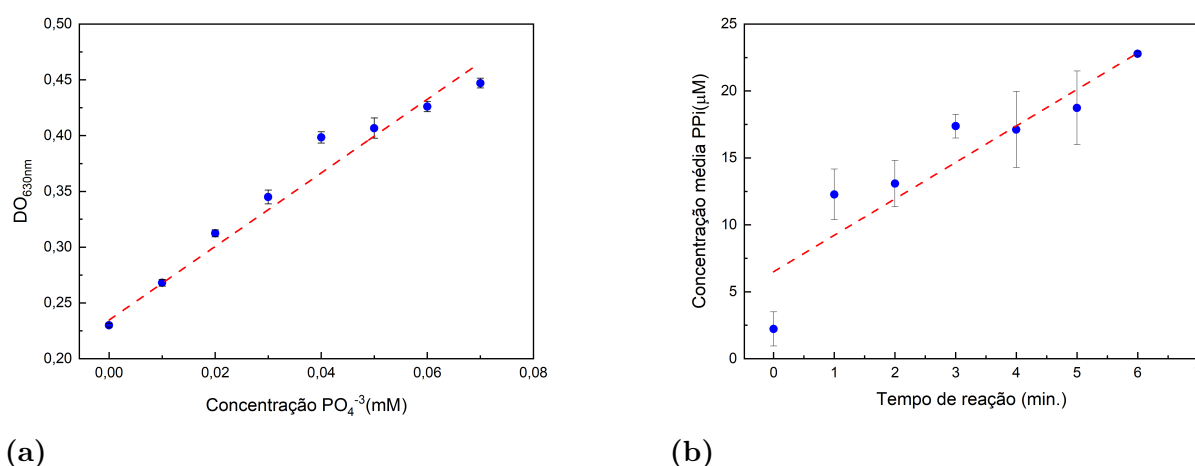
**Figura 61** – Resultados das medidas realizadas para identificação de um ligante no sítio alostérico da EpaE. Eles consistem na fluorescência normalizada da sonda JBS *ThermoFluor Dye* em função da temperatura submetida pela EpaE. Em todos está presente EpaE e dTTP ambos a 11 nM e 5 mM de MgCl<sub>2</sub>, além do tampão 50 mM Tris pH 8.0, 100 mM NaCl e 10% Glicerol. O resultado (a) é referente ao branco, sem ligantes, com T<sub>m</sub> = (51.3 ± 0.4) °C, curva em verde e reta vertical em ciano. O resultado (b) é referente a EpaE na presença de 3.2 μM de Doripenem, com T<sub>m</sub> = (52.3 ± 0.4) °C. O resultados (c) é referente a EpaE na presença de 69 nM de Bicalutamida, com T<sub>m</sub> = (54.7 ± 0.4) °C. O resultado (d) é referente a EpaE na presença de 29 nM de Benztiázida, com T<sub>m</sub> = (53.3 ± 0.4) °C. As curvas referentes aos ligantes e as retas verticais que determinam a T<sub>m</sub> estão em azul.

Fonte: Elaborada pela autora om Matplotlib.<sup>9</sup>

°C, com a Bicalutamida para (54.7 ± 0.4) °C e com a Benztiázida para (53.3 ± 0.4) °C. Estes aumentos na temperatura não são altos. Contudo, no caso da Bicalutamida, pelo teste de significância normal, aplicando  $p < 0.01$ , a diferença entre as duas temperaturas de *melting* é significativa, indicando um pequeno aumento na estabilização da EpaE.

## 5.7 Medidas de Atividade

Iniciou-se a padronização da medida de atividade da enzima RmlA com o reagente verde de malaquita, como descrito na literatura.<sup>106–109</sup> Primeiro, foi realizada a curva de calibração da  $DO_{630nm}$  por concentração de fosfato. Essa curva de calibração apresentou equação da reta  $y = (3.3 \pm 0.2)x + (0.235 \pm 0.005)$  e ajuste  $R^2 = 0.97$ . Apenas a inclinação da reta foi usada para conversão de absorbância em 630 nm para concentração de pirofosfato em solução, após a atividade de 45 ng da enzima EpaE (Figura 62b).

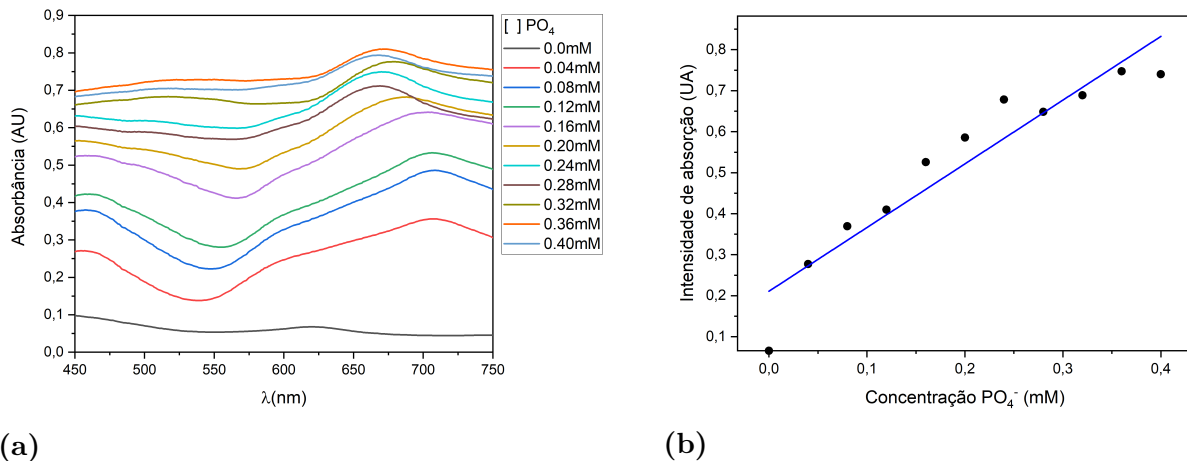


**Figura 62** – (a) Curva de calibração da  $DO_{630nm}$  pela concentração de fosfato (mM), com inclinação da reta ( $3.3 \pm 0.2$ ) e ajuste  $R^2 = 0.97$ . (b) Resultado do melhor teste atividade da enzima por tempo de reação, com inclinação da reta ( $2.7 \pm 0.3$ ) e ajuste  $R^2 = 0.91$ .

Fonte: Elaborada pela autora no OriginLab 9.5.5.

Uma das dificuldades com a aplicação do método verde de malaquita foi a minimização das barras de erro das medida de concentração de PPI, formadas em cada tempo de reação (Figura 62). Para investigar o comportamento da solução de verde de malaquita, foram feitas curvas de absorbância dessas na presença de várias concentrações de fosfato (Figura 63).

Ao analisar as curvas de absorbância do reagente verde de malaquita, com adição crescente de fosfato, observa-se o formato inconstante das curvas, a ausência de uma linha de base comum à elas, e intensidade de cada curva desproporcional a  $[PO_4^{3-}]$ , sem um pico evidente. Ao analisar o comportamento da absorbância em 630nm dessas amostras, percebe-se uma linearidade, com ajuste  $R^2 = 0.89$ . Se as curvas geradas pelo método verde de malaquita variam tanto a cada concentração, um pequeno erro de medida ou preparação de amostra acarretariam em um erro relativamente grande da medida, o que ajuda a explicar a alta barra de erro visualizada no gráfico de produto gerado por tempo de reação (Figura 62b). A dissolução do verde de malaquita em solução aquosa exigia incubação no ultrassom e agitação constante, a fim de evitar precipitação da solução.



**Figura 63** – (a) Comportamento aleatório da curva de absorção do reagente verde de malaquita na presença de fosfato. (b) Ajuste da intensidade de absorbância do reagente verde de malaquita em 630nm na presença de fosfato,  $R^2 = 0.89$ .

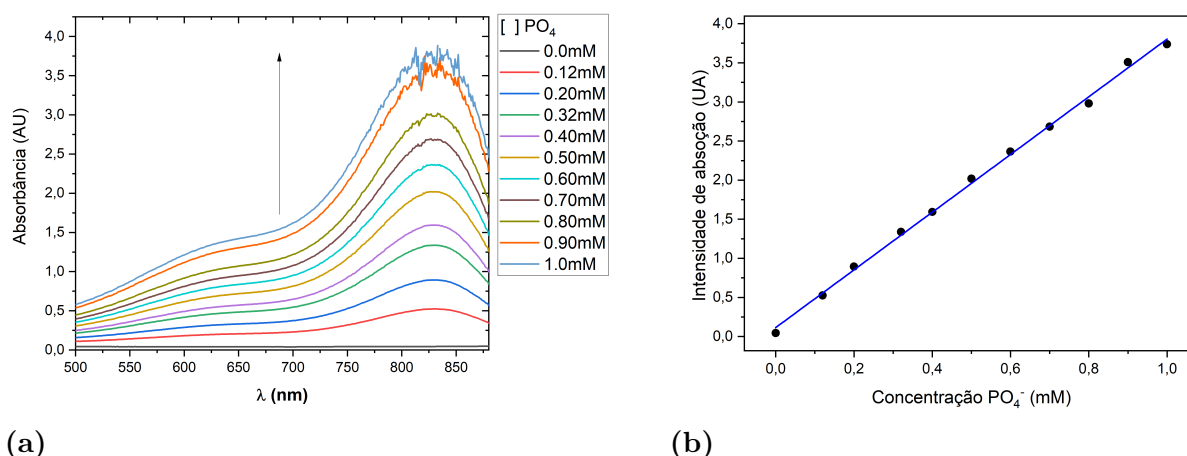
Fonte: Elaborada pela autora no OriginLab 9.5.5.

A solubilidade em água do verde de malaquita aumenta com a concentração de ácido, entretanto, o ácido em excesso, além de poder impactar a atividade da EpaE, facilita a hidrólise da glicose-1-fosfato, afetando duplamente a medida. A impossibilidade de dissolver completamente o verde de malaquita, e de manter uma solução estável deste na concentração determinada pela literatura pode também estar envolvida na falta de precisão da técnica aqui aplicada. Outra dificuldade desta metodologia era impedir a formação de bolhas geradas a pipetagem, visto que Triton X-100 é um componente do reagente verde de malaquita.

O mesmo procedimento de analisar as curvas de absorbância na presença de fosfato foi realizado com a solução colorimétrica do método de análise de fosfato (Figura 64). Diferentemente das curvas do reagente de malaquita, as intensidades das curvas desse reagente colorimétrico é diretamente proporcional a concentração de  $PO_4^{4-}$ , além de apresentar formato conservado, pico constante em torno de 830 nm e a presença de uma linha de base bem definida em cada caso. Quando analisada a absorbância em 830nm de cada curva, o ajuste linear  $R^2 = 0.99$  foi obtido. Pela maior precisão de medida, a solução colorimétrica do métodos de análise de fosfato foi escolhida para prosseguimento experimental.

Foram realizadas duas curvas de calibração, uma até 0.2 mM (Figura 65a) e outra até 0.5 mM (Figura 65b) para verificação da linearidade da curva, ambas obtiveram  $R^2 = 0.99$ , e, respectivamente,  $(3.61 \pm 0.04)$  e  $(3.66 \pm 0.03)$  de inclinação de reta. Com estes resultados iniciais demonstrando uma boa linearidade da curva padrão, foram iniciados os primeiros experimentos de avaliação da atividade enzimática. Nestes experimentos, concentrações iniciais de  $(41 \pm 6)$  nM,  $(56 \pm 4)$  nM e  $(79 \pm 2)$  nM foram empregadas para a

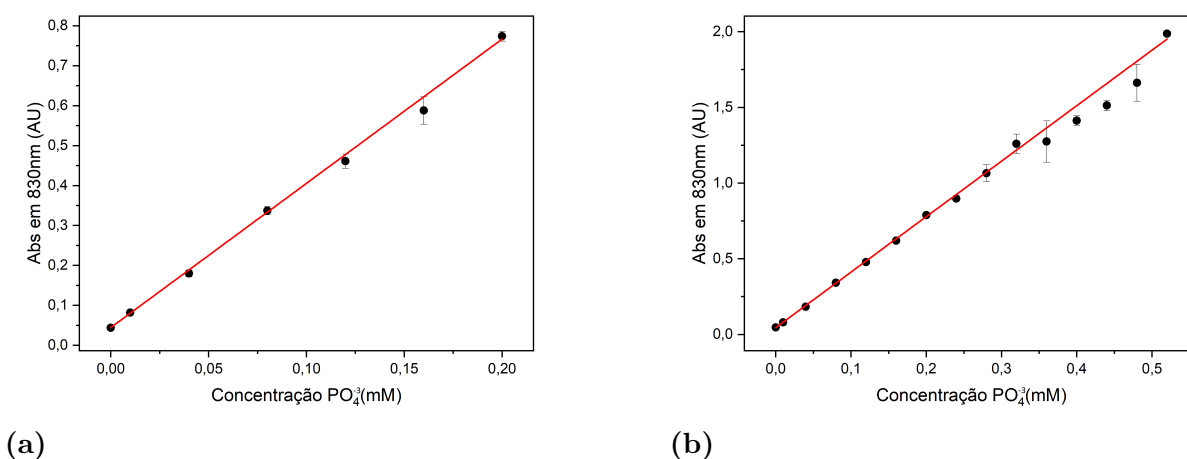




**Figura 64** – (a) Comportamento da curva de absorção do reagente colorimétrico do método de análise de fosfato, na presença de várias concentrações de fosfato. (b) Ajuste da intensidade de absorvância do azul de molibdênio em 830nm, na presença de fosfato,  $R^2 = 0.99$ .

Fonte: Elaborada pela autora no OriginLab 9.5.5.

enzima EpaE. A formação de fosfato em função do tempo nestes experimentos foi analisada e os resultados podem ser vistos na Figura 66. Curiosamente, um perfil de aumento linear da concentração de produto em função do tempo não foi observado nestes experimentos preliminares.

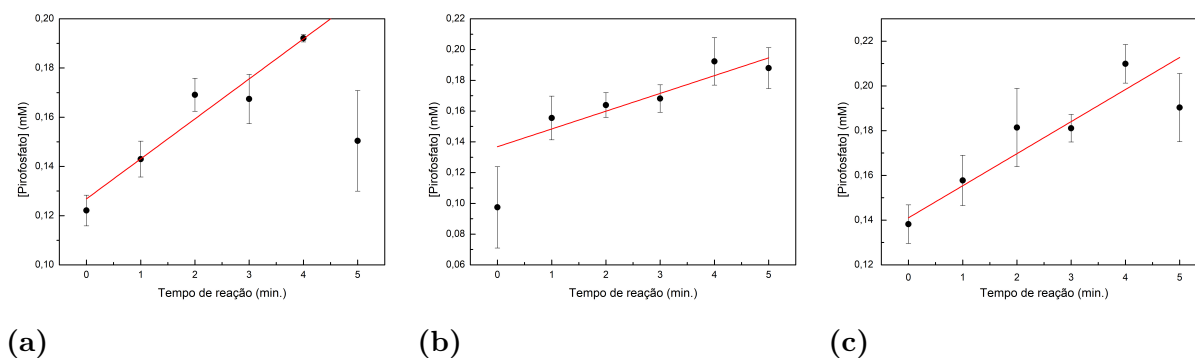


**Figura 65** – (a) Curva de calibração da intensidade absorvida em 830 nm por concentração de fosfato em solução, a partir de azul de molibdênio, com inclinação  $(3.61x \pm 0.04)$  e ajuste  $R^2 = 0.99$ . (b) Curva de calibração com mesmo método, em um intervalo maior de concentração de fosfato, inclinação  $(3.66x \pm 0.03)$ . Realizados em triplicatas.

Fonte: Elaborada pela autora no OriginLab 9.5.5.

Dados os pontos erráticos nos gráficos, as medidas foram refeitas em conjunto com outras três amostras: solução sem EpaE, solução sem dTTP e solução sem glicose-1-fosfato. Este experimento visava captar a possível interferência de algum dos componentes da

reação na leitura da concentração de fosfato na reação. Dessas três, a reação sem a presença do dTTP foi a única amostra a reagir com a solução colorimétrica do método de análise de fosfato, gerando o azul de molibdênio. Isto indica que fosfato estava sendo liberado em solução a partir da glicose-1-fosfato.



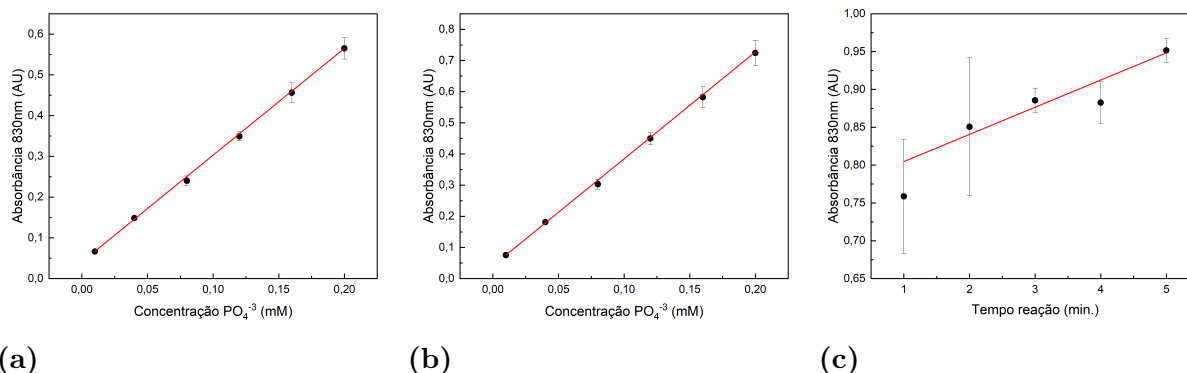
**Figura 66** – Gráficos da produção de pirofosfato por tempo de reação da EpaE nas concentrações (a)  $(41 \pm 6)$ nM (b)  $(56 \pm 4)$ nM (c)  $(79 \pm 2)$  nM.

Fonte: Elaborada pela autora no OriginLab 9.5.5.

A solução colorimétrica contém  $\text{H}_2\text{SO}_4$ , e, antes da medida ser realizada, a amostra com as enzimas e reagentes é incubada com a solução colorimétrica a  $50^\circ\text{C}$  durante 10 minutos. A acidez combinada com esse aquecimento provavelmente induz uma hidrólise na glicose-1-fosfato, afetando a medida. Para contornar este problema, a temperatura de incubação foi reduzida para  $37^\circ\text{C}$ .

Uma nova curva de calibração foi realizada aplicando a nova temperatura de incubação, a de  $37^\circ\text{C}$  (Figura 67a). Para testar a estabilidade dessa mudança, as amostras ficaram em repouso em temperatura ambiente durante 1 h, e as medidas de absorbância em 830 nm foram refeitas, como mostrado na Figura 67b. O mesmo padrão linear foi observado, apenas a inclinação da reta sofreu alteração, de  $(2.62 \pm 0.04)$  foi para  $(3.43 \pm 0.05)$ , indicando que a incubação a  $37^\circ\text{C}$  não promove uma solução perfeitamente estável. Mesmo assim, novas medidas de atividade enzimática foram medidas com esta adaptação (Figura 67c), visto que o tempo de medida era em torno de 1 minuto. Em seguida, mediram-se as intensidades de absorção em 830 nm por tempo de reação de  $(206 \pm 6)$ nM de EpaE. Estas absorbâncias não contemplam as medidas na curva de calibração, sendo assim, não foi feita a conversão entre a intensidade de absorção em 830 nm e a concentração de pirofosfato. As medidas continuaram com alta barra de erro, exceto no caso das absorbâncias medidas nos tempos de reação 3 min., 4 min. e 5 minutos.

Para investigar se as altas barras de erro eram decorrentes da baixa formação de produtos, resultando em ruídos grandes nas medidas, uma nova curva de calibração foi realizada em um intervalo maior de concentração de fosfato, assim como uma nova medida da atividade específica da EpaE.



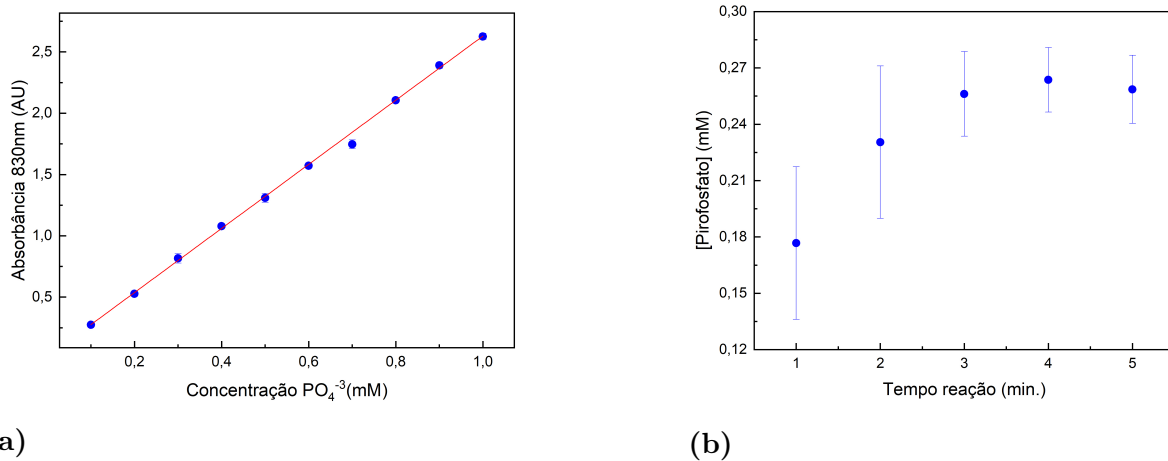
**Figura 67** – (a) Curva de calibração da intensidade absorvida em 830 nm por concentração de fosfato em solução, a partir de azul de molibdênio, ao diminuir a temperatura de incubação para 37 °C, inclinação de  $(2.62 \pm 0.04)$  e ajuste  $R^2 = 0.99$ . (b) Mesma curva, com medidas realizadas 1h depois da incubação a 37 °C, inclinação de  $(3.43 \pm 0.05)$  e ajuste  $R^2 = 0.99$ . (c) Gráfico da absorbância em 830 nm por tempo de reação de  $(206 \pm 6)$ nM de EpaE.

Fonte: Elaborada pela autora no OriginLab 9.5.5.

Uma nova curva de calibração foi padronizada. Esta apresentou o mesmo coeficiente da mostrada na Figura 67a,  $(2.61 \pm 0.01)$ , com ajuste da reta de  $R^2 = 0.99$  (Figura 68a). Com esta nova curva, um novo experimento de atividade enzimática foi realizado, desta vez usando uma concentração de  $(266 \pm 1)$ nM de EpaE, 10 mM de  $MgCl_2$ , 0.4 mM de dTTP e 2 mM D-Glc-1-P. Neste caso, as medidas de absorbância em 830 nm foram realizadas em cinco replicatas e o resultado pode ser observado na Figura 68b. As barras de erro de cada medida continuaram altas, e a linearidade é perdida a partir de 3 min. de reação enzimática. Isto pode indicar que (i) o produto formado está inibindo a catálise da EpaE ou ainda que (ii) 0.1U de pirofosfatase não é mais suficiente para converter pirofosfato em fosfato, (iii) a enzima consumiu todo substrato a partir de 4 min. de reação, uma alternativa seria fazer os experimentos em uma concentração da EpaE, evitando a condição de saturação da reação.

Apenas em nível de curiosidade, após a remoção do último ponto do gráfico da Figura 68b, a atividade específica calculada para a enzima EpaE com os dados mostrados na Figura 68b é de  $2.4 \mu\text{mol}/\text{min mg}$ . Já após a remoção dos últimos dois pontos do mesmo gráfico, a atividade específica da enzima é de  $3.8 \mu\text{mol}/\text{min mg}$ . Visto que o erro desses dois valores seriam oriundos do ajuste da reta, o qual não reflete o erro da medida em si, neste caso, eles foram omitidos. Considerando que o coeficiente linear da reta de calibração é o mesmo da realizada em um intervalo menor de concentrações de fosfato, a atividade específica resultante da catálise de  $(206 \pm 6)$ nM de EpaE (Figura 67c) é  $1 \mu\text{mol}/\text{min mg}$ .

O reagente colorimétrico da técnica de análise de fosfato demonstrou ser o mais adequado para a medida, a intensidade das curvas de absorção deste é diretamente proporcional à concentração de fosfato no meio, além de apresentar acurácia na detecção



**Figura 68** – (a) Curva de calibração da intensidade absorvida em 830 nm por concentração de fosfato em solução, a partir de azul de molibdênio em uma faixa maior de concentração de fosfato, inclinação de  $(2.61 \pm 0.01)$  e ajuste  $R^2 = 0.99$ . (b) Gráfico de produção de pirofosfato por tempo de reação de  $(266 \pm 1)$  nM de EpaE.

Fonte: Elaborada pela autora no OriginLab 9.5.5.

de fosfato, visto pelas baixas barras de erros presentes nas curvas de calibração. Entretanto, para as medidas da atividade específica da EpaE, as barras de erros das medidas eram bem maiores dos que as observadas nas curvas de calibração. As medidas também apresentaram menor linearidade quando comparada com as das curvas de calibração. Uma possibilidade desse erro sistemático durante as medidas é a hidrólise ácida em meio ácido e aquecido, apesar da tentativa de minimizar esse fenômeno.

Em resumo, as medidas de atividade da enzima EpaE se mostraram bastante difíceis, uma vez que as técnicas de medida de fosfato envolvem, em geral, ambientes ácidos e requerem o aquecimento da amostra. A acidez e a temperatura favorecem a perda do fosfato de um dos substratos, o G1P, contribuindo para ruído na própria medida.

Como perspectivas para a solução do problema, consideramos técnicas não colorimétricas. O grupo de pesquisa do prof. Dr. Alessandro Nascimento está bastante interessado, por exemplo, na determinação da atividade enzimática por calorimetria de titulação isotérmica (ITC). Esta ferramenta poderia, em tese, contornar os problemas enfrentados. Contudo, a indisponibilidade do equipamento no momento final da tese não permitiu avançar neste sentido.

## Capítulo 6

### CONSIDERAÇÕES FINAIS

---

A resistência a antimicrobianos é um problema global, com tendência a se agravar gradualmente. No cenário atual, já é estimado que 13.6% do número de mortes em escala mundial é decorrente de infecções de bactérias resistentes a antibióticos. Uma abordagem eficaz a este problema é a prevenção, realizada por meio da ampliação de saneamento básico, aplicação de testes clínicos acurados, acompanhamento rigoroso na prescrição a antibióticos em animais e humanos e desenvolvimento de novos fármacos.

Com o objetivo de identificar novos alvos de inibição, foram escolhidas proteínas do *cluster epa* de *Enterococcus faecalis*. Algumas enzimas virulentas de *Enterococcus faecalis*, também envolvidas na estruturação e biomassa de biofilmes, foram identificadas como sendo codificadas pelo *cluster epa*. Este *cluster* é responsável pela produção de proteínas responsáveis pela síntese e transporte de polissacarídeos decoradores da parede celular de *Enterococcus*. Esses polissacarídeos constituem biofilmes, um facilitador de dessensibilização a antibióticos. Isto, aliado ao pouco conhecimento disponível sobre o *cluster epa*, e, de *Enterococcus faecalis*, motivaram ter as enzimas EpaB, EpaE, EpaI e EpaOX como objeto de estudo.

As proteínas EpaB, EpaI e EpaOX não apresentaram rendimento e estabilidade suficientes para prosseguir experimentalmente com os estudos estruturais. Das tentativas de viabilizar essa finalidade estão: **(i)** expressão em vetores diferentes, **(ii)** expressão das versões com resíduos C-terminais deletados das proteínas originais, **(iii)** aplicação de cepas *E. coli* de expressão distintas e **(iv)** *refolding* com SDS. A enzima EpaE foi a que se mostrou mais solúvel, viabilizando a resolução da estrutura desta. As outras enzimas do *cluster epa* de *E. faecalis* foram modeladas a partir do *AlphaFold2*.

A modelagem das enzimas do *cluster epa* evidenciou alta similaridade estrutural entre algumas enzimas, sugerindo uma redundância biológica destas. Confirmou também a homologia entre EpaJ e GacJ e EpaI e GacI. A EpaI é predita como sendo uma proteína solúvel com boa probabilidade de cristalização, entretanto, assim, como sua homóloga, GacI, uma explicação para sua alta insolubilidade, além desta característica ser intrínseca de GTs, está na associação dela com a EpaJ, uma proteína de membrana, para otimização da atividade enzimática da EpaI. Ainda através da modelagem, foi confirmado que duas proteínas, EpaL e EpaM, constituem um transportador ABC da cadeia de ramnose, e que EpaN e EpaO possuem mais do que um domínio GT2 para realizar a catálise enzimática, apesar desses domínios adicionais não possuírem região catalítica tão conservada quanto dos domínios principais com as EpaBCDI.

A EpaE é uma RmlA, a primeira enzima atuante na síntese de dTDP-L-ramnose, a qual resultará em L-ramnose, sintetizada apenas em bactérias e plantas. A EpaE foi cristalizada em 0.1 M Tris pH 8.5, 0.2 M NaCl e 20% w/v PEG 4000, e aditivo BICINE. O grupo espacial do cristal é o  $P22_12_1$ , apresentando *twinning* pseudomeroedral, com regra de *twinning* “-h, -k, l”. A resolução da estrutura refinada é de 2.85 Å. Tentaram-se cristalizações com EpaE na presença dos substratos dTTP, D-Glc-1-P e o co-fator  $Mg^{+2}$ , visando melhora da resolução estrutural desta. Entretanto, mesmo aplicando as mesmas condições de cristalização descritas anteriormente, não foi observado nenhum cristal de proteína. A partir de sobreposição da estrutura da EpaE com outras RmlA de maior resolução, foi possível identificar o sítio catalítico da EpaE. Uma molécula de timidina teve sua densidade eletrônica evidenciada durante o refinamento, indicando um possível sítio alostérico da EpaE.

Testes enzimáticos foram feitos por medida indireta, o produto da reação catalisada pela EpaE, pirofosfato, foi hidrolisado por pirofosfatase, originando fosfato em solução. Testaram-se dois métodos de detecção de fosfato na solução, o do reagente verde de malaquita e o método de análise de fosfato. A metodologia do reagente verde de malaquita foi inicialmente aplicada, contudo, a solução colorimétrica não se apresentou adequada para os testes enzimáticos, mesmo replicando a técnica descrita na literatura. A absorvância do reagente de malaquita apresentou comportamento aleatório na presença de diversas concentrações de fosfato, diferentemente do que foi observado com a solução colorimétrica do método de análise de fosfato. Esta, sim, apta a detectar fosfato em solução. O método de análise de fosfato foi adaptado a fim de minimizar hidrólise ácida da glicose-1-fosfato. No entanto, as medidas realizadas para determinar a atividade específica da EpaE apresentaram altas barras de erro. Algum erro sistemático ocorreu durante as medidas, visto que a solução colorimétrica do método de análise de fosfato é precisa na identificação de fosfato no meio, observado pelas pequenas barras de erro presentes nas curvas de calibração realizadas com esta metodologia.

Realizaram triagens com potenciais ligantes da EpaE a partir da similaridade das moléculas com o dTTP. Com Benztiázida, Doripenem mono hidratado e Bicalutamida foram realizados experimentos de Fluorimetria Diferencial de Temperatura para verificar o efeito destas moléculas na EpaE. Inicialmente, apenas nos casos da Bicalutamida e Benztiázida foram observadas alterações do  $T_m$  da EpaE. Para evitar que as moléculas se ligassem não especificamente à EpaE em algum sítio alostérico, a Fluorimetria Diferencial de Temperatura foi realizada novamente, na relação 1:1:3 das concentrações molares de EpaE, dTTP e as três moléculas. Neste caso, foi observado aumento da  $T_m$  ao adicionar as três moléculas em solução. A Bicalutamida foi a que apresentou a maior  $T_m$ , ocorreu um aumento de 3 °C em comparação com o valor do branco, indicando uma leve estabilização da EpaE na presença dessas desta molécula.

Por fim, com este trabalho foi possível modelar a maioria das proteínas codificadas pelo *cluster* *epa* de *E. faecalis* e resolver a estrutura de uma delas, a primeira resolvida dentro desse *cluster* até o momento.

- 1 WORLD HEALTH ORGANIZATION. **Antimicrobial resistance**. 2020. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance>. Acesso em: 17 maio 2021.
- 2 PERRY, J.; WAGLECHNER, N.; WRIGHT, G. The Prehistory of Antibiotic Resistance. **Cold Spring Harbor Perspectives in Medicine**, v. 6, n. 6, p. a025197, 6 2016. ISSN 2157-1422. DOI:10.1101/cshperspect.a025197.
- 3 PAWLOWSKI, A. C. *et al.* A diverse intrinsic antibiotic resistome from a cave bacterium. **Nature Communications**, v. 7, p. 1–10, 2016. ISSN 20411723. DOI: 10.1038/ncomms13803.
- 4 TORTORELLA, E. *et al.* Antibiotics from deep-sea microorganisms: current discoveries and perspectives. **Marine Drugs**, v. 16, n. 10, p. 1–16, 2018. ISSN 16603397. DOI: 10.3390/md16100355.
- 5 PERRY, J. A.; WESTMAN, E. L.; WRIGHT, G. D. The antibiotic resistome: what's new? **Current Opinion in Microbiology**, v. 21, p. 45–50, 2014. ISSN 18790364. DOI: 10.1016/j.mib.2014.09.002.
- 6 WRIGHT, G. D. The antibiotic resistome: the nexus of chemical and genetic diversity. **Nature Reviews Microbiology**, v. 5, n. 3, p. 175–186, 2007. ISSN 17401526. DOI: 10.1038/nrmicro1614.
- 7 CENTERS FOR DISEASE CONTROL AND PREVENTION. **How antimicrobial resistance happens**. 2020. Disponível em: <https://www.cdc.gov/drugresistance/about/how-resistance-happens.html>. Acesso em: 5 maio 2023.
- 8 CENTERS FOR DISEASE CONTROL AND PREVENTION. **About antibiotic resistance**. 2020. Disponível em: <https://www.cdc.gov/drugresistance/about.html>. Acesso em: 10 fevereiro 2021.
- 9 HUNTER, J. D. Matplotlib: a 2D graphics environment. **Computing in Science & Engineering**, IEEE COMPUTER SOC, v. 9, n. 3, p. 90–95, 2007.
- 10 KOLÁŘ, M.; URBÁNEK, K.; LÁTAL, T. Antibiotic selective pressure and development of bacterial resistance. **International Journal of Antimicrobial Agents**, v. 17, n. 5, p. 357–363, 2001. ISSN 09248579. DOI: 10.1016/S0924-8579(01)00317-X.
- 11 WORLD HEALTH ORGANIZATION. **Infographics: Antibiotic resistance**. 2016. Disponível em: <https://www.who.int/mediacentre/events/2015/world-antibiotic-awareness-week/infographics/en/>. Acesso em: 17 novembro 2020.
- 12 KLEIN, E. Y. *et al.* Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. **Proceedings of the National Academy of Sciences**, v. 115, n. 15, p. E3463–E3470, 4 2018. ISSN 0027-8424. DOI: 10.1073/pnas.1717295115.



- 13 WORLD HEALTH ORGANIZATION. **WHO Report on surveillance of antibiotic consumption: 2016-2018 early implementation.** 2018. Disponível em: <https://apps.who.int/iris/bitstream/handle/10665/277359/9789241514880-eng.pdf>. Acesso em: 18 maio 2021.
- 14 BARTOLLETTI, F. *et al.* Polymyxin B resistance in carbapenem-resistant *Klebsiella pneumoniae*, São Paulo, Brazil. **Emerging Infectious Diseases**, v. 22, n. 10, p. 1849–1851, 2016. ISSN 1080-6040. DOI: 10.3201/eid2210.160695.
- 15 TORRES, R. S. L. de A. *et al.* Group A *Streptococcus* antibiotic resistance in southern Brazil: a 17-Year Surveillance Study. **Microbial Drug Resistance**, v. 17, n. 2, p. 313–319, 2011. ISSN 1076-6294. DOI: 10.1089/mdr.2010.0162.
- 16 ANVISA. **Relatório do monitoramento da prevalência e do perfil de suscetibilidade aos antimicrobianos em enterococos e salmonelas isolados de carcaças de frango congeladas comercializadas no Brasil.** 2008. Disponível em: <http://antigo.anvisa.gov.br/documents/33916/395481/Relat%CC80rio+Prebaf+-+Programa+Nacional+de+Monitoramento+da+Preval%CC80ncia+e+da+Resist%CC80ncia+Bacteriana+em+Frango+-+2008/04658e9f-7ca6-4e4b-b4fe-0fb7ef54a04e>. Acesso em: 20 maio 2021.
- 17 ROSSI, F. The challenges of antimicrobial resistance in Brazil. **Clinical Infectious Diseases**, v. 52, n. 9, p. 1138–1143, 5 2011. ISSN 1058-4838. DOI: 10.1093/cid/cir120.
- 18 MULANI, M. S. *et al.* Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: a review. **Frontiers in Microbiology**, v. 10, n. APR, 2019. ISSN 1664-302X. DOI: 10.3389/fmicb.2019.00539.
- 19 LANGFORD, B. J. *et al.* Bacterial co-infection and secondary infection in patients with COVID-19: a living rapid review and meta-analysis. **Clinical Microbiology and Infection**, v. 26, n. 12, p. 1622–1629, 2020. ISSN 14690691. DOI: 10.1016/j.cmi.2020.07.016.
- 20 MIRZAEI, R. *et al.* Bacterial co-infections with SARS-CoV-2. **IUBMB Life**, v. 72, n. 10, p. 2097–2111, 2020. ISSN 15216551. DOI: 10.1002/iub.2356.
- 21 MCCULLERS, J. A. The co-pathogenesis of influenza viruses with bacteria in the lung. **Nature Reviews Microbiology**, v. 12, n. 4, p. 252–262, 2014. ISSN 17401534. DOI: 10.1038/nrmicro3231.
- 22 AVADHANULA, V. *et al.* Respiratory viruses augment the adhesion of bacterial pathogens to respiratory epithelium in a viral species- and cell type-dependent manner. **Journal of Virology**, v. 80, n. 4, p. 1629–1636, 2006. ISSN 0022-538X. DOI: 10.1128/jvi.80.4.1629-1636.2006.
- 23 BALLINGER, M. N.; STANDIFORD, T. J. Postinfluenza bacterial pneumonia: host defenses gone awry. **Journal of Interferon and Cytokine Research**, v. 30, n. 9, p. 643–652, 2010. ISSN 10799907. DOI: 10.1089/jir.2010.0049.
- 24 DAMASIO, G. A. *et al.* Does virus-bacteria coinfection increase the clinical severity of acute respiratory infection? **Journal of Medical Virology**, v. 87, n. 9, p. 1456–1461, 2015. ISSN 01466615. DOI: 10.1002/jmv.24210.

- 25 WORLD HEALTH ORGANIZATION. **Antimicrobial resistance** - preventing the COVID-19 pandemic from causing an antibiotic resistance catastrophe. 2020. Disponível em: <https://www.who.int/europe/news/item/18-11-2020-preventing-the-covid-19-pandemic-from-causing-an-antibiotic-resistance-catastrophe>. Acesso em: 14 janeiro 2021.
- 26 LAI, C. C.; WANG, C. Y.; HSUEH, P. R. Co-infections among patients with COVID-19: the need for combination therapy with non-anti-SARS-CoV-2 agents? **Journal of Microbiology, Immunology and Infection**, v. 53, n. 4, p. 505–512, 2020. ISSN 19959133. DOI: 10.1016/j.jmii.2020.05.013.
- 27 THORPE, K. E.; JOSKI, P.; JOHNSTON, K. J. Antibiotic-resistant infection treatment costs have doubled since 2002, now exceeding \$2 billion annually. **Health Affairs**, v. 37, n. 4, p. 662–669, 2018. ISSN 0278-2715. DOI: 10.1377/hlthaff.2017.1153.
- 28 BUSH, K. *et al.* Tackling antibiotic resistance. **Nature Reviews Microbiology**, v. 9, n. 12, p. 894–896, 2011. ISSN 1740-1526. DOI: 10.1038/nrmicro2693.
- 29 FRIEDMAN, N. D.; TEMKIN, E.; CARMELI, Y. The negative impact of antibiotic resistance. **Clinical Microbiology and Infection**, v. 22, n. 5, p. 416–422, 2016. ISSN 14690691. DOI: 10.1016/j.cmi.2015.12.002.
- 30 KRAKER, M. E. A. de; DAVEY, P. G.; GRUNDMANN, H. 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, 10 2011. ISSN 1549-1676. DOI: 10.1371/journal.pmed.1001104.
- 31 CENTERS FOR DISEASE CONTROL AND PREVENTION. **Antibiotic resistance threats in the United States**. 2019. Disponível em: [https://www.cdc.gov/drugresistance/biggest\\_threats.html](https://www.cdc.gov/drugresistance/biggest_threats.html). Acesso em: 16 maio 2021.
- 32 IKUTA, K. S. *et al.* Global mortality associated with 33 bacterial pathogens in 2019: a systematic analysis for the global burden of disease study 2019. **The Lancet**, v. 400, n. 10369, p. 2221–2248, 12 2022. ISSN 01406736. DOI: 10.1016/S0140-6736(22)02185-7.
- 33 WORLD HEALTH ORGANIZATION. **WHO Coronavirus (COVID-19) Dashboard**. 2023. Disponível em: <https://covid19.who.int/>. Acesso em: 7 junho 2023.
- 34 VENTOLA, C. L. The antibiotic resistance crisis: part 1: causes and threats. **P & T**, v. 40, n. 4, p. 277–83, 4 2015. ISSN 1052-1372.
- 35 PAULSEN, I. T. *et al.* Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. **Science**, v. 299, n. 5615, p. 2071–2074, 2003. ISSN 00368075. DOI: 10.1126/science.1080613.
- 36 AARESTRUP, F. *et al.* Comparison of antimicrobial resistance phenotypes and resistance genes in *Enterococcus faecalis* and *Enterococcus faecium* from humans in the community, broilers, and pigs in Denmark. **Diagnostic Microbiology and Infectious Disease**, v. 37, n. 2, p. 127–137, 6 2000. ISSN 07328893. DOI: 10.1016/S0732-8893(00)00130-9.
- 37 XU, Y. *et al.* Analysis of a gene cluster of *Enterococcus faecalis* involved in polysaccharide biosynthesis. **Infection and Immunity**, v. 68, n. 2, p. 815–823, 2000. ISSN 00199567. DOI: 10.1128/IAI.68.2.815-823.2000.

- 38 GAO, W.; HOWDEN, B. P.; STINEAR, T. P. Evolution of virulence in *Enterococcus faecium*, a hospital-adapted opportunistic pathogen. **Current Opinion in Microbiology**, v. 41, p. 76–82, 2 2018. ISSN 13695274. DOI: 10.1016/j.mib.2017.11.030.
- 39 DALE, J. L. *et al.* Multiple roles for *Enterococcus faecalis* glycosyltransferases in biofilm-associated antibiotic resistance, cell envelope integrity, and conjugative transfer. **Antimicrobial Agents and Chemotherapy**, v. 59, n. 7, p. 4094–4105, 2015. ISSN 10986596. DOI: 10.1128/AAC.00344-15.
- 40 GÓMEZ-GIL, R. *et al.* Nosocomial outbreak of linezolid-resistant *Enterococcus faecalis* infection in a tertiary care hospital. **Diagnostic Microbiology and Infectious Disease**, Elsevier Inc., v. 65, n. 2, p. 175–179, 10 2009. ISSN 07328893. DOI: 10.1016/j.diagmicrobio.2009.06.010.
- 41 WEINER-LASTINGER, L. M. *et al.* Antimicrobial-resistant pathogens associated with adult healthcare-associated infections: summary of data reported to the National Healthcare Safety Network, 2015–2017. **Infection Control & Hospital Epidemiology**, v. 41, n. 1, p. 1–18, 1 2020. ISSN 0899-823X. DOI: 10.1017/ice.2019.296.
- 42 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**, Centers for Disease Control, v. 37, n. 11, p. 1288–1301, 2016. ISSN 15596834. DOI: 10.1017/ice.2016.174.
- 43 FLEMMING, H.-C.; WINGENDER, J. The biofilm matrix. **Nature Reviews Microbiology**, v. 8, n. 9, p. 623–633, 9 2010. ISSN 1740-1526. DOI: 10.1038/nrmicro2415.
- 44 NAGARATHINAM, S. *et al.* Effect of triple antibiotic loaded apatitic nanocarriers on *Enterococcus faecalis* biofilm – an in vitro study. **Journal of Drug Delivery Science and Technology**, v. 51, n. November 2018, p. 499–505, 2019. ISSN 17732247. DOI: 10.1016/j.jddst.2019.03.035.
- 45 KORIR, M. L.; DALE, J. L.; DUNNY, G. M. Role of *epaQ*, a previously uncharacterized *Enterococcus faecalis* gene, in biofilm development and antimicrobial resistance. **Journal of Bacteriology**, v. 201, n. 18, p. 1–15, 3 2019. ISSN 0021-9193. DOI: 10.1128/JB.00078-19.
- 46 PAGANELLI, F. L.; WILLEMS, R. J.; LEAVIS, H. L. Optimizing future treatment of enterococcal infections: attacking the biofilm? **Trends in Microbiology**, v. 20, n. 1, p. 40–49, 2012. ISSN 0966842X. DOI: 10.1016/j.tim.2011.11.001.
- 47 POTERA, C. Microbiology: forging a link between biofilms and disease. **Science**, v. 283, n. 5409, p. 1837–1839, 1999. ISSN 00368075. DOI: 10.1126/science.283.5409.1837.
- 48 CH'NG, J. H. *et al.* Biofilm-associated infection by enterococci. **Nature Reviews Microbiology**, v. 17, n. february, 2018. ISSN 17401534, DOI: 10.1038/s41579-018-0107-z.
- 49 CENTERS FOR DISEASE CONTROL AND PREVENTION. **Covid-19 U.S.:** Impact on antimicrobial resistance. 2022. Disponível em: <https://www.cdc.gov/drugresistance/covid19.html?fbclid=IwAR0TV3eq4lxFTTrjOk9tJBLjNAaxJWvwuS5OGnVcZ1V81-gbUrouOmI8fsk%0Ahttps://www.cdc.gov/drugresistance/pdf/covid19-impact-report-508.pdf?fbclid=>

IwAR2G5DyjJHcyjfWvUMYJRaeHBgr1Qd0wIYebh83VmcCBfLskK7P6MRont1. Acesso em: 19 maio 2021.

50 ERLANDSEN, S. L.; KRISTICH, C. J.; DUNNY, G. M. Ultrastructure of *Enterococcus faecalis* biofilms. **Biofilms**, v. 1, n. 2, p. 131–137, 2004. ISSN 1479-0505. DOI: 10.1017/S1479050504001206.

51 MADIGAN, M.T.; STAHL, A.; MARTINKO, J.M.; CLARK, D.P. **Brock: biology of microorganisms**. 13<sup>a</sup> ed. San Francisco: Pearson, 2012. Cap. 23. p. 669.

52 DALE, J. L. *et al.* Restructuring of *Enterococcus faecalis* biofilm architecture in response to antibiotic-induced stress. **npj Biofilms and Microbiomes**, Springer US, v. 3, n. 1, p. 1–9, 2017. ISSN 20555008. DOI: 10.1038/s41522-017-0023-4.

53 TENG, F. *et al.* Evidence that the enterococcal polysaccharide antigen gene (*epa*) cluster is widespread in *Enterococcus faecalis* and influences resistance to phagocytic killing of *E. faecalis*. **Infection and Immunity**, v. 70, n. 4, p. 2010–2015, 2002. ISSN 00199567. DOI: 10.1128/IAI.70.4.2010-2015.2002.

54 GILMORE, M.S.; CLEWELL, D.B.; IKE, Y.; *et al.* **Enterococci from commensals to leading causes of drug resistant infection**. Boston: Massachusetts Eye and Ear Infirmary, 2014. p. 375.

55 GUERARDEL, Y. *et al.* Complete structure of the Enterococcal Polysaccharide Antigen (EPA) of vancomycin-resistant *Enterococcus faecalis* V583 reveals that EPA decorations are teichoic acids covalently linked to a rhamnopolysaccharide backbone. **mBio**, v. 11, n. 2, p. 1–15, 4 2020. ISSN 2161-2129. DOI: 10.1128/mBio.00277-20.

56 QIN, X. *et al.* Complete genome sequence of *Enterococcus faecium* strain TX16 and comparative genomic analysis of *Enterococcus faecium* genomes. **BMC Microbiology**, v. 12, n. 1, p. 135, 2012. ISSN 1471-2180. DOI: 10.1186/1471-2180-12-135.

57 TENG, F. *et al.* Further characterization of the *epa* gene cluster and *epa* polysaccharides of *Enterococcus faecalis*. **Infection and Immunity**, v. 77, n. 9, p. 3759–3767, 2009. ISSN 00199567. DOI: 10.1128/IAI.00149-09.

58 SINGH, K. V.; LEWIS, R. J.; MURRAY, B. E. Importance of the *epa* locus of *Enterococcus faecalis* OG1RF in a mouse model of ascending urinary tract infection. **The Journal of Infectious Diseases**, v. 200, n. 3, p. 417–420, 2009. ISSN 0022-1899. DOI: 10.1086/600124.

59 SMITH, R. E. *et al.* Decoration of the enterococcal polysaccharide antigen EPA is essential for virulence, cell surface charge and interaction with effectors of the innate immune system. **PLOS Pathogens**, v. 15, 2019. ISSN 1553-7374. DOI: 10.1371/journal.ppat.1007730.

60 SINGH, K. V.; MURRAY, B. E. Loss of a major Enterococcal polysaccharide antigen (*Epa*) by *Enterococcus faecalis* is associated with increased resistance to ceftriaxone and carbapenems. **Antimicrobial Agents and Chemotherapy**, v. 63, n. 5, p. 1–3, 2019. ISSN 10986596. DOI: 10.1128/AAC.00481-19.

- 61 MISTOU, M.-Y.; SUTCLIFFE, I. C.; SORGE, N. M. van. Bacterial glycobiology: rhamnose-containing cell wall polysaccharides in Gram-positive bacteria. **FEMS Microbiology Reviews**, v. 40, n. 4, p. 464–479, 2016. ISSN 1574-6976. DOI: 10.1093/femsre/fuw006.
- 62 DRULA, E. *et al.* The carbohydrate-active enzyme database: functions and literature. **Nucleic Acids Research**, v. 50, n. D1, p. D571–D577, 2022. ISSN 13624962. DOI: 10.1093/nar/gkab1045.
- 63 GIRAUD, M.-F.; NAISMITH, J. H. The rhamnose pathway. **Current Opinion in Structural Biology**, v. 10, n. 6, p. 687–696, 12 2000. ISSN 0959440X. DOI: 10.1016/S0959-440X(00)00145-7.
- 64 MA, Y. *et al.* Determination of the pathway for rhamnose biosynthesis in mycobacteria: cloning, sequencing and expression of the Mycobacterium tuberculosis gene encoding  $\alpha$ -D-glucose-1-phosphate thymidyltransferase. **Microbiology**, v. 143, n. 3, p. 937–945, 1997. ISSN 13500872. DOI: 10.1099/00221287-143-3-937.
- 65 BEEK, S. L. van der *et al.* Streptococcal dTDP-L-rhamnose biosynthesis enzymes: functional characterization and lead compound identification. **Molecular Microbiology**, v. 111, n. 4, p. 951–964, 2019. ISSN 13652958. DOI: 10.1111/mmi.14197.
- 66 BISCHER, A. P. *et al.* Disruption of L-rhamnose biosynthesis results in severe growth defects in Streptococcus mutans. **Journal of Bacteriology**, v. 202, n. 6, p. 1–14, 2019. ISSN 0021-9193. DOI: 10.1128/JB.00728-19.
- 67 SIVARAMAN, J. *et al.* Crystal structure of Escherichia coli glucose-1-phosphate thymidyltransferase (RffH) complexed with dTTP and Mg<sup>2+</sup>. **Journal of Biological Chemistry**, v. 277, n. 46, p. 44214–44219, 2002. ISSN 00219258. DOI: 10.1074/jbc.M206932200.
- 68 JIANG, N. *et al.* Rhamnose in plants - from biosynthesis to diverse functions. **Plant Science**, Elsevier B.V., v. 302, n. October 2020, p. 110687, 2021. ISSN 18732259. DOI: 10.1016/j.plantsci.2020.110687.
- 69 HSU, C. Y. *et al.* Genomic differences between listeria monocytogenes EGDe isolates reveal crucial roles for sigB and wall rhamnosylation in biofilm formation. **Journal of Bacteriology**, v. 202, n. 7, p. 1–13, 2020. ISSN 10985530. DOI: 10.1128/JB.00692-19.
- 70 BROWN, H. A. *et al.* The structure of glucose-1-phosphate thymidyltransferase from Mycobacterium tuberculosis reveals the location of an essential magnesium ion in the RmlA-type enzymes. **Protein Science**, v. 27, n. 2, p. 441–450, 2018. ISSN 1469896X. DOI: 10.1002/pro.3333.
- 71 AZIMI, A. *et al.* Distribution of smf-1, rmlA, spgM and rpff genes among Stenotrophomonas maltophilia isolates in relation to biofilm-forming capacity. **Journal of Global Antimicrobial Resistance**, v. 23, p. 321–326, 2020. ISSN 22137173. DOI: 10.1016/j.jgar.2020.10.011.
- 72 MOHAGHEGHZADEH, N. *et al.* High occurrence of antibiotic resistance and biofilm-formation among Stenotrophomonas maltophilia isolated from a tertiary hospital in southwest of Iran. **Gene Reports**, v. 21, n. July, p. 100827, 2020. ISSN 24520144. DOI: 10.1016/j.genrep.2020.100827.

- 73 DUAN, Z. *et al.* Clinical and molecular epidemiology of *Stenotrophomonas maltophilia* in pediatric patients from a chinese teaching hospital. **Frontiers in Cellular and Infection Microbiology**, v. 10, n. August, p. 1–8, 2020. ISSN 22352988. DOI: 10.3389/fcimb.2020.00411.
- 74 WANG, Y. *et al.* ITRAQ-based proteomics reveals potential anti-virulence targets for ESBL-producing *Klebsiella pneumoniae*. **Infection and Drug Resistance**, v. 13, p. 2891–2899, 2020. ISSN 11786973. DOI: 10.2147/IDR.S259894.
- 75 YANG, S. *et al.* Novel dTDP-L-rhamnose synthetic enzymes (RmlABCD) from *Saccharothrix syringae* CGMCC 4.1716 for one-pot four-enzyme synthesis of dTDP-L-rhamnose. **Frontiers in Microbiology**, v. 12, n. November, 2021. ISSN 1664-302X. DOI: 10.3389/fmicb.2021.772839.
- 76 ALPHEY, M. S. *et al.* Allosteric competitive inhibitors of the glucose-1-phosphate thymidyltransferase (RmlA) from *Pseudomonas aeruginosa*. **ACS Chemical Biology**, v. 8, n. 2, p. 387–396, 2013. ISSN 15548937. DOI: 10.1021/cb300426u.
- 77 JUMPER, J. *et al.* Highly accurate protein structure prediction with AlphaFold. **Nature**, v. 596, n. 7873, p. 583–589, 2021. ISSN 0028-0836. DOI: 10.1038/s41586-021-03819-2.
- 78 MARIANI, V. *et al.* IDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. **Bioinformatics**, v. 29, n. 21, p. 2722–2728, 2013. ISSN 13674803. DOI: 10.1093/bioinformatics/btt473.
- 79 HOLM, L. DALI and the persistence of protein shape. **Protein Science**, v. 29, p. 128–140, 1 2020. ISSN 0961-8368. DOI: 10.1002/pro.3749.
- 80 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. ISSN 0192-8651. DOI: 10.1002/jcc.20084.
- 81 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. ISSN 10465928. DOI: 10.1016/j.pep.2014.03.008.
- 82 CAMILO, C. M. *et al.* HTP-OligoDesigner: an online primer design tool for high-throughput gene cloning and site-directed mutagenesis. **Journal of Computational Biology**, v. 23, n. 1, p. 27–29, 2016. ISSN 10665277. DOI: 10.1089/cmb.2015.0148.
- 83 PROMEGA corporation. **pGEM(R)-T and pGEM(R)-T**: easy vector systems technical manual TM042 - pgem-t and pgem-t easy vector systems protocol. 2010. Disponível em: <https://www.promega.co.uk/~~/media/files/resources/protocols/technicalmanuals/0/pgem-tandpgem-teasyvectorsystemsprotocol.pdf>. Acesso em: 06 junho 2019.
- 84 BERGMANS, H. E.; VAN DIE, I. M.; HOEKSTRA, W. P. Transformation in *Escherichia coli*: stages in the process. **Journal of Bacteriology**, v. 146, n. 2, p. 564–570, 1981. DOI: 10.1128/jb.146.2.564-570.1981.

- 
- 85 RAHIMZADEH, M. *et al.* Impact of heat shock step on bacterial transformation efficiency. **Molecular biology research communications**, v. 5, n. 4, p. 257–261, 2016. ISSN 2322-181X. DOI: 10.22099/mbrc.2016.3915.
- 86 RUSH, J. S. *et al.* The molecular mechanism of N-acetylglucosamine side-chain attachment to the Lancefield group A carbohydrate in *Streptococcus pyogenes*. **Journal of Biological Chemistry**, v. 292, n. 47, p. 19441–19457, 2017. ISSN 1083351X. DOI: 10.1074/jbc.M117.815910.
- 87 WAGNER, S. *et al.* Tuning *Escherichia coli* for membrane protein overexpression. **Proceedings of the National Academy of Sciences**, v. 105, n. 38, p. 14371–14376, 2008. ISSN 0027-8424. DOI: 10.1073/pnas.0804090105.
- 88 DUMON-SEIGNOVERT, L.; CARIOT, G.; VUILLARD, L. The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). **Protein Expression and Purification**, v. 37, n. 1, p. 203–206, 2004. ISSN 10465928. DOI: 10.1016/j.pep.2004.04.025.
- 89 CARVALHO, H.; MENEHINI, R. Increased expression and purification of soluble iron-regulatory protein 1 from *Escherichia coli* co-expressing chaperonins GroES and GroEL. **Brazilian Journal of Medical and Biological Research**, v. 41, n. 4, p. 270–276, 2008. ISSN 16784510. DOI: 10.1590/S0100-879X2008005000009.
- 90 Agilent Technologies. **ArcticExpress™ Competent Cells and ArcticExpress™ (DE3) Competent Cells**. 2010. Disponível em: [https://www.agilent.com/en/product/protein-expression/competent-cells-for-difficult-protein-expression/insoluble-proteins/arcticexpress-\(de3\)-competent-cells-232939](https://www.agilent.com/en/product/protein-expression/competent-cells-for-difficult-protein-expression/insoluble-proteins/arcticexpress-(de3)-competent-cells-232939). Acesso em: 20 junho 2020.
- 91 HE, C.; OHNISHI, K. Efficient renaturation of inclusion body proteins denatured by SDS. **Biochemical and Biophysical Research Communications**, v. 490, n. 4, p. 1250–1253, 2017. ISSN 10902104. DOI: 10.1016/j.bbrc.2017.07.003.
- 92 GASTEIGER, E. *et al.* **The Proteomics Protocols Handbook**. Totowa, NJ: Humana Press, 2005. 571–608 p. ISBN 978-1-58829-343-5.
- 93 KABSCH, W. XDS. **Acta Crystallographica Section D**, v. 66, n. 2, p. 125–132, 2010. ISSN 0907-4449. DOI: 10.1107/S0907444909047337.
- 94 EVANS, P. Scaling and assessment of data quality. **Acta Crystallographica Section D**, v. 62, n. 1, p. 72–82, 2006. ISSN 0907-4449. DOI: 10.1107/S0907444905036693.
- 95 MCCOY, A. J. *et al.* Phaser crystallographic software. **Journal of Applied Crystallography**, v. 40, n. 4, p. 658–674, 2007. ISSN 00218898. DOI: 10.1107/S0021889807021206.
- 96 WEBB, B.; SALI, A. Comparative protein structure modeling using MODELLER. **Current Protocols in Bioinformatics**, v. 54, n. 1, p. 1–18, 2016.
- 97 TERWILLIGER, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. **Acta Crystallographica Section D**, v. 64, n. 1, p. 61–69, 2008. ISSN 0907-4449. DOI: 10.1107/S090744490705024X.

- 98 AFONINE, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. **Acta Crystallographica Section D**, v. 68, n. 4, p. 352–367, 2012. ISSN 0907-4449. DOI: 10.1107/S0907444912001308.
- 99 MURSHUDOV, G. N.; VAGIN, A. A.; DODSON, E. J. Refinement of macromolecular structures by the maximum-likelihood method. **Acta Crystallographica Section D**, v. 53, n. 3, p. 240–255, 1997. ISSN 0907-4449. DOI: 10.1107/S0907444996012255.
- 100 EMSLEY, P. *et al.* Features and development of Coot. **Acta Crystallographica Section D**, v. 66, n. 4, p. 486–501, 2010. ISSN 0907-4449. DOI: 10.1107/S0907444910007493.
- 101 WILLIAMS, C. J. *et al.* MolProbity: more and better reference data for improved all-atom structure validation. **Protein Science**, v. 27, n. 1, p. 293–315, 2018. ISSN 09618368. DOI: 10.1002/pro.3330.
- 102 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, 2017. ISSN 1932-6203. DOI: 10.1371/journal.pone.0174336.
- 103 MUNIZ, H. dos 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. ISSN 0920-654X. DOI: 10.1007/s10822-015-9856-1.
- 104 REYNOLDS, C. A.; BURT, C.; RICHARDS, W. G. A linear molecular similarity index. **Quantitative Structure-Activity Relationships**, v. 11, n. 1, p. 34–35, 1992. ISSN 09318771. DOI: 10.1002/qsar.19920110106.
- 105 NIESEN, F. H.; BERGLUND, H.; VEDADI, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. **Nature Protocols**, v. 2, n. 9, p. 2212–2221, 2007. ISSN 1754-2189. DOI: 10.1038/nprot.2007.321.
- 106 SHA, S. *et al.* Development of a colorimetric assay and kinetic analysis for Mycobacterium tuberculosis D-glucose-1-phosphate thymidyltransferase. **Journal of Biomolecular Screening**, v. 17, n. 2, p. 252–257, 2012. ISSN 10870571. DOI: 10.1177/1087057111421373.
- 107 SHERWOOD, A. R. *et al.* A malachite green-based assay to assess glucan phosphatase activity. **Analytical Biochemistry**, v. 435, n. 1, p. 54–56, 2013. ISSN 10960309. DOI: 10.1016/j.ab.2012.10.044.
- 108 FISHER, D. K.; HIGGINS, T. J. A sensitive, high-volume, colorimetric assay for protein phosphatases. **Pharmaceutical research**, v. 11, n. 5, p. 759–63, 1994. ISSN 0724-8741. DOI: 10.1023/a:1018996817529.
- 109 BAYKOV, A. A.; EVTUSHENKO, O. A.; AVAEVA, S. M. A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay. **Analytical Biochemistry**, v. 171, n. 2, p. 266–270, 1988. ISSN 10960309. DOI: 10.1016/0003-2697(88)90484-8.
- 110 CHEN, P. S.; TORIBARA, T. Y.; WARNER, H. Microdetermination of phosphorus. **Analytical Chemistry**, v. 28, n. 11, p. 1756–1758, 1956. ISSN 0003-2700. DOI: 10.1021/ac60119a033.



- 
- 111 KIM, T. W.; LEI, X. G. An improved method for a rapid determination of phytase activity in animal feed1. **Journal of Animal Science**, v. 83, n. 5, p. 1062–1067, 5 2005. ISSN 0021-8812.
- 112 DIVYA, S. *et al.* Specific H<sup>+</sup> level is crucial for accurate phosphate quantification using ascorbate as a reductant. **Protoplasma**, v. 257, n. 1, p. 319–330, 2020. ISSN 0033-183X. DOI: 10.1007/s00709-019-01424-9.
- 113 NAGUL, E. A. *et al.* The molybdenum blue reaction for the determination of orthophosphate revisited: opening the black box. **Analytica Chimica Acta**, v. 890, p. 60–82, 2015. ISSN 00032670. DOI: 10.1016/j.aca.2015.07.030.
- 114 HALLGREN, J. *et al.* DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. **bioRxiv**, 2022.
- 115 DONG, Y. Y. *et al.* Structures of DPAGT1 explain glycosylation disease mechanisms and advance TB antibiotic design. **Cell**, v. 175, n. 4, 2018. ISSN 00928674. DOI: 10.1016/j.cell.2018.10.037.
- 116 GUÉRIN, H.; KULAKAUSKAS, S.; CHAPOT-CHARTIER, M.-P. Structural variations and roles of rhamnose-rich cell wall polysaccharides in Gram-positive bacteria. **Journal of Biological Chemistry**, v. 298, n. 10, p. 102488, 2022. ISSN 00219258. DOI: 10.1016/j.jbc.2022.102488.
- 117 LIU, F. *et al.* The small molecule luteolin inhibits N-acetyl- $\alpha$ -galactosaminyltransferases and reduces mucin-type O-glycosylation of amyloid precursor protein. **Journal of Biological Chemistry**, v. 292, n. 52, p. 21304–21319, 12 2017. ISSN 00219258. DOI: 10.1074/jbc.M117.814202.
- 118 WHEATLEY, R. W. *et al.* Tetrameric structure of the GlfT2 galactofuranosyltransferase reveals a scaffold for the assembly of mycobacterial arabinogalactan. **Journal of Biological Chemistry**, v. 287, n. 33, p. 28132–28143, 8 2012. ISSN 00219258. DOI: 10.1074/jbc.M112.347484.
- 119 LIRA-NAVARRETE, E. *et al.* Substrate-guided front-face reaction revealed by combined structural snapshots and metadynamics for the polypeptide N-acetylgalactosaminyltransferase 2. **Angewandte Chemie**, v. 126, n. 31, p. 8345–8349, 7 2014. ISSN 00448249. DOI: 10.1002/ange.201402781.
- 120 OSAWA, T. *et al.* Crystal structure of chondroitin polymerase from *Escherichia coli* K4. **Biochemical and Biophysical Research Communications**, v. 378, n. 1, p. 10–14, 2009. ISSN 0006291X. DOI: 10.1016/j.bbrc.2008.08.121.
- 121 ZHANG, H. *et al.* New helical binding domain mediates a glycosyltransferase activity of a bifunctional protein. **Journal of Biological Chemistry**, v. 291, n. 42, p. 22106–22117, 2016. ISSN 00219258. DOI: 10.1074/jbc.M116.731695.
- 122 ALLARD, S. T. *et al.* Toward a structural understanding of the dehydratase mechanism. **Structure**, v. 10, n. 1, p. 81–92, 2002. ISSN 09692126. DOI: 10.1016/S0969-2126(01)00694-3.
- 123 GANDINI, R. *et al.* Structural basis for dolichylphosphate mannose biosynthesis. **Nature Communications**, v. 8, n. 1, p. 120, 2017. ISSN 2041-1723. DOI: 10.1038/s41467-017-00187-2.

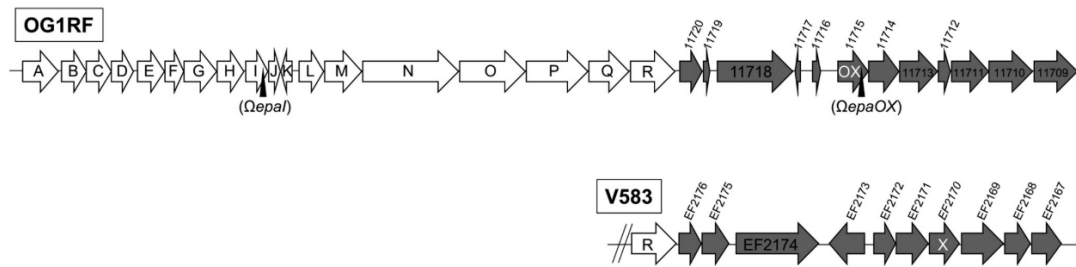
- 124 BRETON, C. *et al.* Structures and mechanisms of glycosyltransferases. **Glycobiology**, v. 16, n. 2, p. 29R–37R, 2006. ISSN 1460-2423. DOI: 10.1093/glycob/cwj016.
- 125 BUCHAN, D. W. A.; JONES, D. T. The PSIPRED protein analysis workbench: 20 years on. **Nucleic Acids Research**, v. 47, n. W1, p. W402–W407, 2019. ISSN 0305-1048. DOI: 10.1093/nar/gkz297.
- 126 SLABINSKI, L. *et al.* XtalPred: a web server for prediction of protein crystallizability. **Bioinformatics**, v. 23, n. 24, p. 3403–3405, 2007. ISSN 1460-2059. DOI: 10.1093/bioinformatics/btm477.
- 127 ALTSCHUL, S. F. *et al.* Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403–410, 1990. ISSN 00222836. DOI: 10.1016/S0022-2836(05)80360-2.
- 128 BERMAN, H. M. *et al.* The Protein Data Bank. **Acta Crystallographica Section D**, v. 58, n. 6, p. 899–907, 2002. ISSN 0907-4449. DOI: 10.1107/S0907444902003451.
- 129 Schrödinger, LLC. The PyMOL molecular graphics system, version 1.8. 2015.
- 130 ZWART, P. H., AFININA, P. V., *et al.* Automated structure solution with the phenix suite. **Methods in molecular biology (Clifton, N.J.)**, v. 426, p. 419–435, 2008. ISSN 1064-3745. DOI: 10.1007/978-1-60327-058-8\_28.
- 131 KRISSINEL, E.; HENRICK, K. Inference of macromolecular assemblies from crystalline state. **Journal of Molecular Biology**, v. 372, p. 774–797, 2007. DOI: 10.1016/j.jmb.2007.05.022.
- 132 BLANKENFELDT, W. *et al.* The purification, crystallization and preliminary structural characterization of glucose-1-phosphate thymidyltransferase (RmlA), the first enzyme of the dTDP-L-rhamnose synthesis pathway from *Pseudomonas aeruginosa*. **Acta Crystallographica Section D**, v. 56, n. 11, p. 1501–1504, 2000. ISSN 09074449. DOI: 10.1107/S0907444900010040.
- 133 ZUCCOTTI, S. *et al.* Kinetic and crystallographic analyses support a sequential-ordered bi bi catalytic mechanism for *Escherichia coli* glucose-1-phosphate thymidyltransferase. **Journal of Molecular Biology**, v. 313, n. 4, p. 831–843, 2001. ISSN 00222836. DOI: 10.1006/jmbi.2001.5073.
- 134 BLANKENFELDT, W. The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidyltransferase (RmlA). **The EMBO Journal**, v. 19, n. 24, p. 6652–6663, 12 2000. ISSN 14602075. DOI: 10.1093/emboj/19.24.6652.
- 135 SCHÖNING-STIERAND, K. *et al.* ProteinsPlus: interactive analysis of protein–ligand binding interfaces. **Nucleic Acids Research**, v. 48, n. W1, p. W48–W53, 2020. ISSN 13624962. DOI: 10.1093/NAR/GKAA235.
- 136 VOLKAMER, A. *et al.* Analyzing the topology of active sites: on the prediction of pockets and subpockets. **Journal of Chemical Information and Modeling**, v. 50, n. 11, p. 2041–2052, 2010. ISSN 1549-9596. DOI: 10.1021/ci100241y.
- 137 VOLKAMER, A. *et al.* Combining global and local measures for structure-based druggability predictions. **Journal of Chemical Information and Modeling**, v. 52, n. 2, p. 360–372, 2012. ISSN 1549-9596. DOI: 10.1021/ci200454v.

138 BOURGOGNE, A. *et al.* Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. **Genome Biology**, v. 9, n. 7, p. R110, 2008. ISSN 1465-6906. DOI: 10.1186/gb-2008-9-7-r110.

139 PÉDELACQ, J.-D. *et al.* Engineering and characterization of a superfolder green fluorescent protein. **Nature Biotechnology**, v. 24, n. 1, p. 79–88, 2006. ISSN 1087-0156. DOI: 10.1038/nbt1172.

140 PINTO JUNIOR, F. **Interação e internalização de curcumina em culturas bacterianas**. Dissertação (Mestrado em Ciências). Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2019.

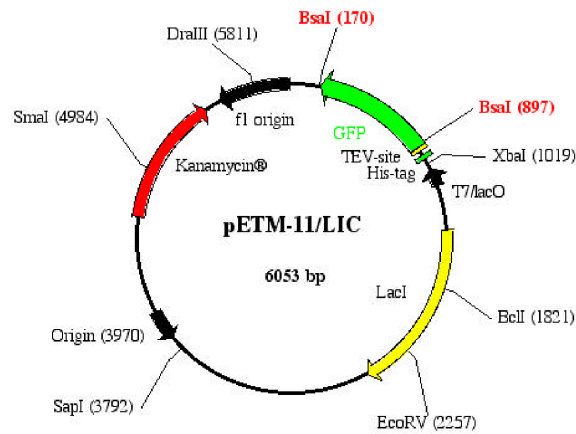
### A.1 Região variável do *cluster epa*



**Figura 69** – Comparação da região variável no locus *epa* das cepas OG1RF e V583 de *E. faecalis*.

Fonte: Dale.<sup>39</sup>

### A.2 Vetor pETM11

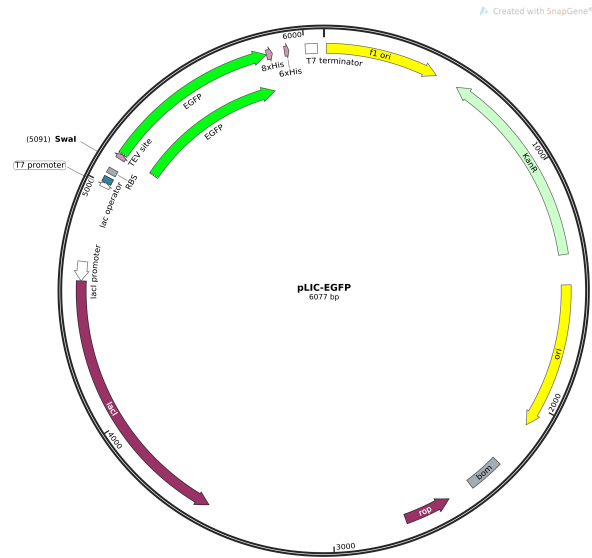


**Figura 70** – Mapa do vetor pETM11.

Fonte: Arie Geerlof - geerlof@embl-hamburg.



## A.5 Vetor pLIC-EGFP



**Figura 73** – Mapa do vetor pLIC-EGFP visualizado a partir do programa SnapGene, disponível em [snapgene.com](http://snapgene.com).

Fonte: Elaborada pela autora.

## A.6 Anotação dos genes do *cluster epa*

**Tabela 3** – Genes *epa* catalogados nas cepas OG1RF e V583. <sup>a</sup> A partir de Xu *et al.*, 1999.<sup>37</sup> <sup>b</sup> orfde12 e orfde13 constituem *epaM*. <sup>c</sup> orfde14 e orfde15 constituem *epaN*. <sup>d</sup> Designado por Bourgoigne *et al.*, 2008.<sup>138</sup> <sup>e</sup> EF2175 é um pseudogene em OG1RF.

Nomeação anterior <sup>a</sup>	OG1RF ID	V583 ID	Anotação Biológica
orfde1	orfde1	EF2200	Metionina aminopeptidase
orfde2	orfde2	EF2199	Possível RNase
orfde3	<i>epaA</i>	EF2198	Glicosiltransferase família 4
orfde4	<i>epaB</i>	EF2197	Glicosiltransferase família 2
orfde5	<i>epaC</i>	EF2196	Glicosiltransferase família 2
orfde5-6	<i>epaD</i>	EF2195	Glicosiltransferase família 2
orfde6	<i>epaE</i>	EF2194	Glicose-1-fosfato timidiltransferase
orfde7	<i>epaF</i>	EF2193	dTDP-4-dehidroramnose 3,5-epimerase
orfde8	<i>epaG</i>	EF2192	dTDP-glicose 4,6-dehidratase
orfde9	<i>epaH</i>	EF2191	dTDP-4-dehidroramnose redutase
orfde10	<i>epaI</i>	EF2190	Glicosiltransferase família 2
Não identificado	<i>epaJ</i>	EF2189	Proteína de membrana hipotética
Não identificado	<i>epaK</i>	EF2184	Proteína de membrana hipotética
orfde11	<i>epaL</i>	EF2183	Transportador ABC, permease
orfde12	<i>epaM</i> <sup>b</sup>	EF2182	Transportador ABC, ligante ATP
orfde13			
orfde14	<i>epaN</i> <sup>c</sup>	EF2181	Glicosiltransferase família 2
orfde15			
orfde16	<i>epaO</i>	EF2180	Glicosiltransferase família 2
	<i>epaP</i>	EF2179	Proteína de Membrana hipotética
	<i>epaQ</i>	EF2178	Proteína de Membrana hipotética
	<i>epaR</i>	EF2177	Açúcar transferase
	OG1RF_0164 <sup>d</sup>	EF2176	Glicosiltransferase família 2
		EF2175 <sup>e</sup>	Proteína do tipo LicD
	OG1RF_0163	EF2174	Glicosiltransferase família 25

Fonte: Adaptada de TENG *et al.*<sup>57</sup>

## A.7 Primers

**Tabela 4** – *Primers* utilizados para a clonagem dos gene *epaI*, *epaOX* e suas construções, *epaB* e *epaE* de *E. faecalis* em vetor pETTrx-1a e pETTM11. O *epaB\** e *epaE\** indicam os *primers* da clonagem dos gene *epaB* e *epaE*, respectivamente, em vetor pLIC-EGFP. Orientação 5' a 3'.

Gene	<i>Primer forward</i>	<i>Primer reverse</i>
<i>epaI</i>	CAGGGCGCCATGAAGGT ACTTTTAATCATTCCTG	GACCCGACGCGGTTATTAA TCTCCCTCCTTCATAAAGG
<i>epaI</i> <sup>130</sup>	CAGGGCGCCATGAAGGT ACTTTTAATCATTCCTG	GACCCGACGCGGTTATT AAGACGCTTCGTTTCCA
<i>epaI</i> <sup>210</sup>	CAGGGCGCCATGAAGGT ACTTTTAATCATTCCTG	GACCCGACGCGGTTAT TAACCAAGCCGTTCCA
<i>epaOX</i>	CAGGGCGCCATGGAAAA GAAAATTTAAAGTTAAGC	GACCCGACGCGGTTATTACT TCCCTAAAATTTTTCGATAC
<i>epaOX</i> <sup>220</sup>	CAGGGCGCCATGGAAA AGGAAAACCTGAAAC	GACCCGACGCGGTTAGA ATTTCTGATGCACGATG
<i>epaOX</i> <sup>277</sup>	CAGGGCGCCATGGAA AAGGAAAACCTGAAAC	GACCCGACGCGGTTACAGAT ATTTCTTATTAATTTTCGCCA
<i>epaB*</i>	TACTTCCAATCCAATGCA AGAAATGGTAACAATCTC	ATTAATTAATCCTTAAAA GAACCTCCAGCCC
<i>epaE*</i>	TACTTCCAATCCAATGAA AGGAATTATTTTAGCAGG	ATTAATTAATCCCTATTC TGCTGCTAAACGTAGC
<i>epaE</i>	CAGGGCGCCATGAAAGGA ATTATTTTAGCAGG	GACCCGACGCGGTTACTA TTCTGCTGCTAAACGTAGC

Fonte: Elaborada pela autora.

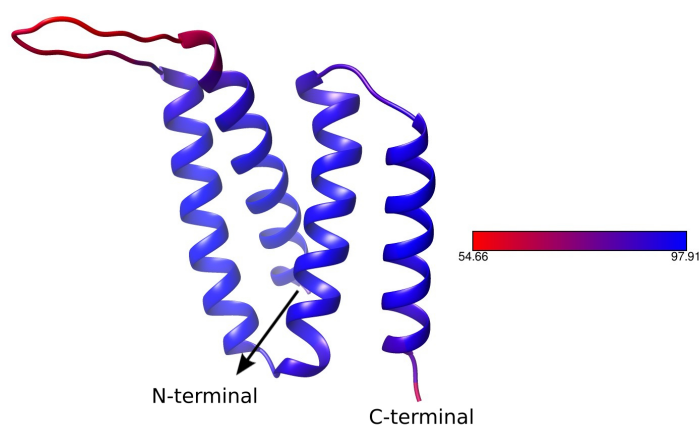


## APÊNDICE B

## MODELAGEM - CONTINUAÇÃO

## B.1 EpaK

O modelo estrutural predito da EpaK (NCBI ID: EOL33612) tem pLDDT médio de 88, o pLDDT de cada resíduo está apresentado na Figura 74. A estrutura predita da EpaK é constituída por quatro feixes de hélice e um longo *loop*. A sequência da EpaK é predita como sendo de uma proteína de membrana pelo programa DeepTMHMM.<sup>114</sup>



**Figura 74** – Modelo estrutural da EpaK, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

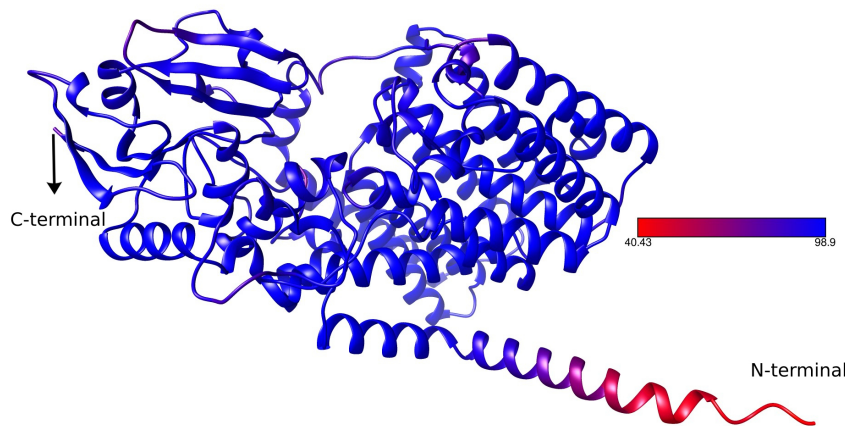
Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

O modelo predito da EpaK é similar a vários transportadores de antibióticos, apesar dos baixos valores dos indicadores de correlação entre as estruturas. A mais similar, por exemplo, do transportador de *Clostridiales bacterium*, PDB ID: 6wk5, apresenta apenas *Z-score* 11.1, RMSD 3.0 e 15% de identidade com a predição estrutural da proteína EpaK, uma proteína pequena constituída por 113 aminoácidos. No trabalho de Guerardel e colaboradores, 2020,<sup>55</sup> é proposto que o homólogo da EpaK na cepa V538, EF2189, seja uma das proteínas envolvidas no transporte de Glc e GlcNAc vinculados ao carreador lipídico undecaprenol fosfato e na transferência desses açúcares ao polímero de ramnose.

## B.2 EpaP

O modelo estrutural predito da EpaP (NCBI ID: EOL33607) tem pLDDT médio de 94, o padrão de pLDDT ao longo da estrutura predita está apresentado na Figura 75. O enovelamento da EpaP é constituído por feixes de hélice e folhas  $\beta$ . O programa DeepTMHMM<sup>114</sup> identificou a sequência da EpaP como sendo de uma proteína de membrana,

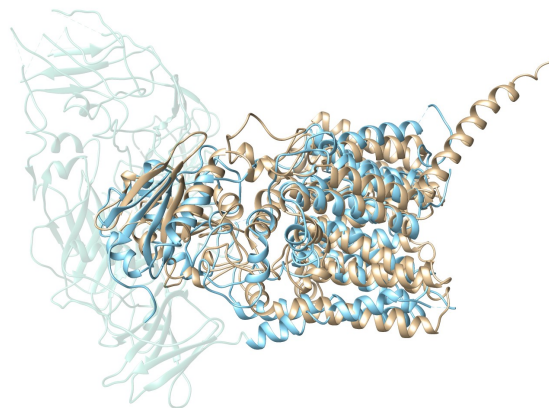
além disso, as estruturas mais semelhantes ao modelo estrutural da EpaP também são anotadas como proteínas de membrana.



**Figura 75** – Modelo estrutural da EpaP, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

A estrutura mais similar ao modelo predito da EpaP é a da arabino-furanose transferase de *Mycobacteroides abscessus*, PDB ID: 6wby, apresentando *Z-score* 12.8, RMSD 4.3 e 9% de identidade com a estrutura predita da EpaP. Essas duas não compartilham semelhança estrutural significativa, mas possuem um enovelamento parecido, como pode ser visto na Figura 76.

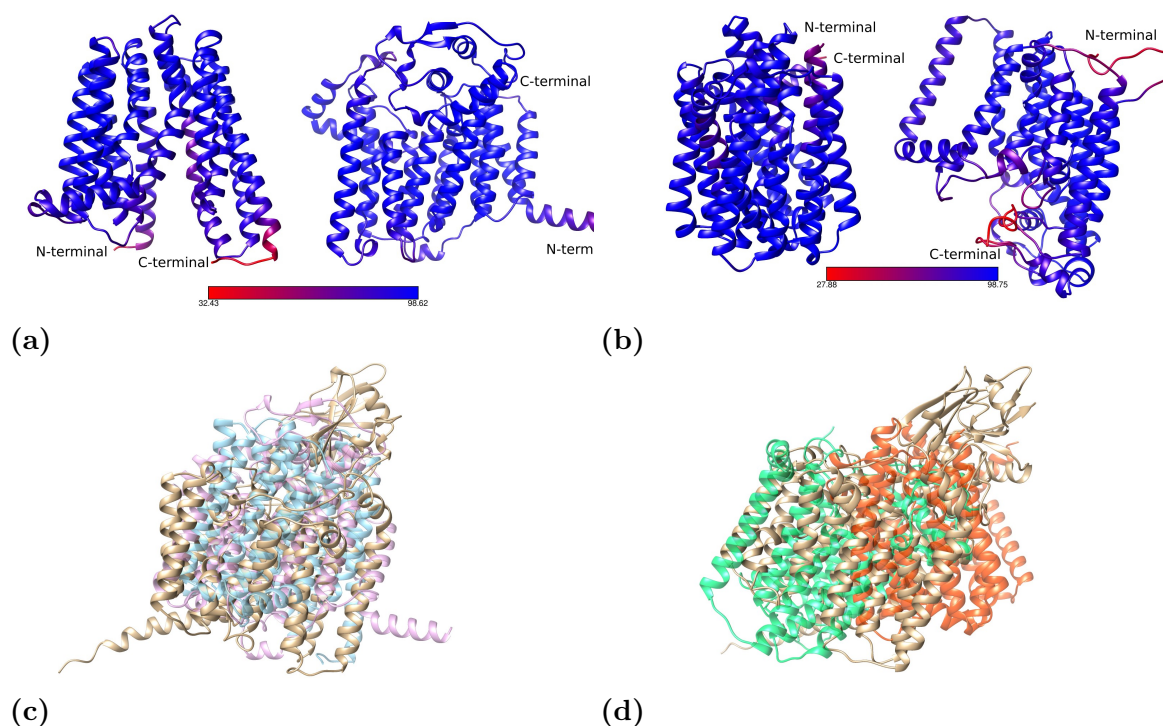


**Figura 76** – Modelo estrutural da EpaP, em dourado, alinhado com a estrutura da arabino-furanose transferase de *M. abscessus* (PDB ID: 6wby), em azul.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

Ao analisar dois *clusters* análogos ao epa, e bem caracterizados, das espécies *S. pyogenes* e *Lactococcus lactis*, *gac* e *wsp*, respectivamente, há duas proteínas de membrana em cada que também são transportadoras e glicosiltransferases, Gack e GacL de *S. pyogenes*, e WpsG e WpsJ de *Lactococcus lactis*. Nenhuma sequência dessas enzimas apresentam alinhamento significativo com a EpaP. As predições estruturais das enzimas Gack (NCBI ID: AAZ51230), GacL (NCBI ID: AAZ51231), WpsG (NCBI ID: ADJ59250) e WpsJ

(ADJ59240) possuem o enovelamento parecido com o da predição estrutural da EpaP, constituído majoritariamente por feixes de hélice. No caso da GacL, há ainda sequências de hélices e fitas  $\beta$  na estrutura predita pelo *AlphaFold2* (Figura 77a). A GacK pertence à família de flipases Wzx, ela reorienta o carreador lipídico ligado a oligossacarídeos para o exterior celular, e a GacL é uma glicosiltransferase que retira o oligossacarídeo do intermediário lipídico e adiciona-o à cadeia lateral do polímero de ramnose.<sup>116</sup> Ambas enzimas são descritas como sendo ancoradas na membrana plasmática. Apesar do modelo estrutural da EpaP não possuir um bom alinhamento com as predições estruturais da GacK e GacL, ao analisar o enovelamento em si das enzimas, o da predição estrutural da EpaP é mais semelhante à estrutura preditas da GacL de *S. pyogenes*. Guerardel e colaboradores<sup>55</sup> sugerem que o homólogo EF2179 da EpaP seja um dos transportadores dos açúcares Glc e GlcNAc para o exterior celular e os adicionam na lateral da cadeia de ramnose. A EpaP provavelmente é uma proteína transmembrana, mas nada específico pode ser dito sobre sua atividade como glicosiltransferase.

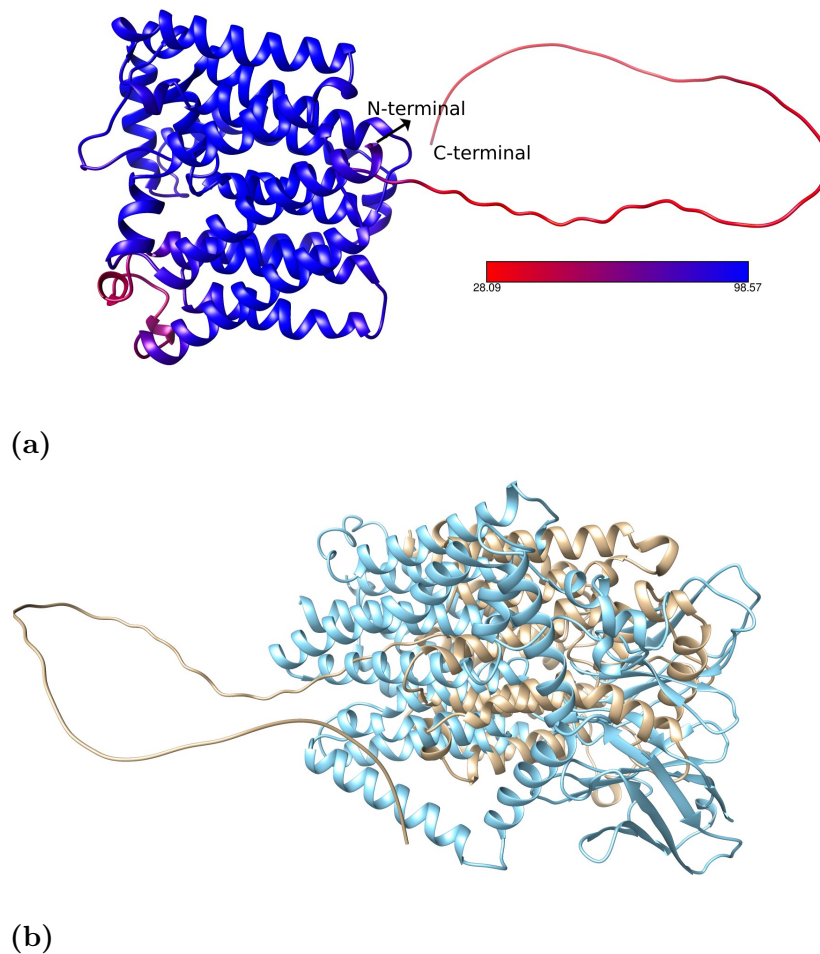


**Figura 77** – (a) Modelo estrutural da GacK, à esquerda, e da GacL, à direita. (b) Modelo estrutural da WpsG, à esquerda, e da WpsG, à direita. A coloração indica do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. Sobreposição da predição estrutural da EpaP, em dourado, com (c) as da GacK, em azul e da GacL, em roxo, e (d) WpsG, em laranja, e WpsJ, em ciano.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

### B.3 EpaQ

O modelo estrutural predito da EpaQ (NCBI ID: EOL33606) tem pLDDT médio de 83, o padrão de pLDDT ao longo da estrutura predita está apresentado na Figura 78a. O enovelamento da EpaQ é constituído por feixes de hélice, há ainda uma longa região desta enzima que não foi possível ser predita pelo *AlphaFold2*. A partir da análise da sequência da EpaQ, o programa DeepTMHMM<sup>114</sup> prediz que esta proteína seja ancorada na membrana plasmática.



**Figura 78** – (a) Modelo estrutural da EpaQ, com coloração indicando do menor (em vermelho) para o maior (em azul) valor do pLDDT de cada resíduo. (b) Modelo estrutural da EpaQ, em dourado, sobreposto com a estrutura da proteína de membrana polissacarídeo polimerase/ligase (PDB ID: 7tpg), em azul.

Fonte: Elaborada pela autora no UCSF-Chimera.<sup>80</sup>

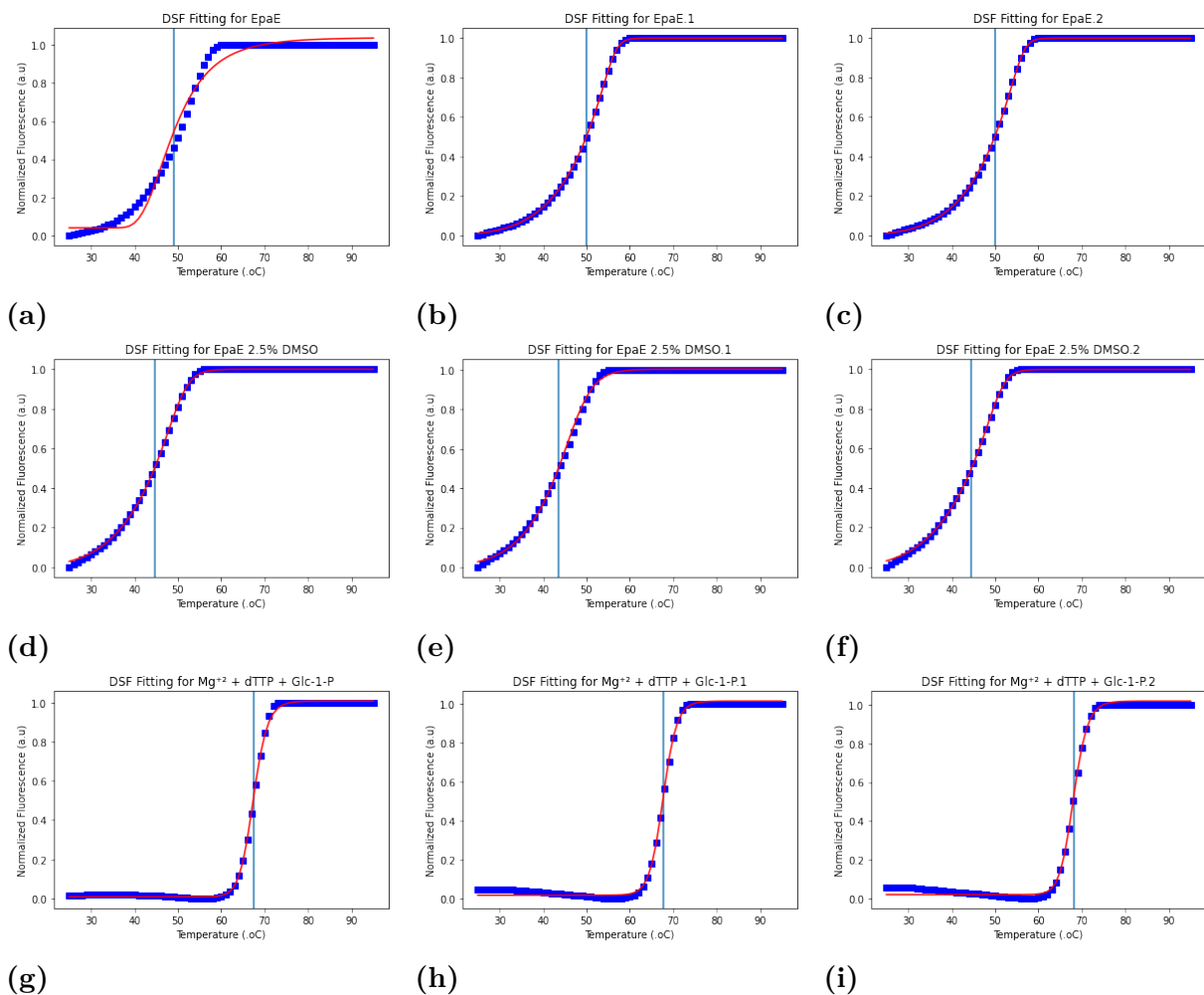
A estrutura mais similar ao modelo predito da EpaQ é a da proteína de membrana polissacarídeo polimerase/ligase de *Cupriavidus metallidurans*, PDB ID: 7tpg, apresentando *Z-score* 17.0, RMSD 4.5 e 13% de identidade com a estrutura predita da EpaQ. Essas duas não compartilham semelhança estrutural significativa, apenas um enovelamento

parecido em uma região da estrutura da polissacarídeo polimerase/ligase (PDB ID: 7tpg), como visualizado na Figura 78b. Assim como no caso do homólogo da EpaP, Guerardel e colaboradores<sup>55</sup> propõem que o homólogo EF2178 da EpaQ seja um dos transportadores dos açúcares Glc e GlcNAc para o exterior celular e os adicionam na lateral da cadeia de ramnose. Os dados obtidos da predição estrutural da EpaQ não produziu nenhuma evidência sobre a provável reação catalisada por esta enzima.

# APÊNDICE C

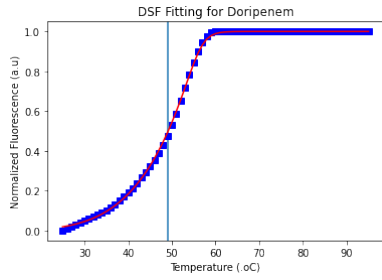
## GRÁFICOS INDIVIDUAIS DAS MEDIDAS DE FLUORIMETRIA DIFERENCIAL DE TEMPERATURA

As triplicatas das medidas dos testes de interação entre os potenciais ligantes com a EpaE são mostrados a seguir (Figura 79). Pela técnica de fluorimetria diferencial de temperatura, isto foi feito analisando as  $T_m$  das soluções branco, EpaE em tampão e EpaE em tampão e 2.5 % DMSO, com as soluções com EpaE em tampão na presença de cada molécula, Doripenem, Benztiázida e Bicalutamida. Apenas nos casos 8.1  $\mu\text{M}$  Bicalutamida em 2.5% DMSO, 24 nM Benztiázida em 0.5% DMSO e 0.2 mM Benztiázida em 2.5% DMSO foi observado aumento da  $T_m$  em relação aos dois brancos,  $(44.3 \pm 0.9)$ ,  $(49.7 \pm 0.4)$  e  $(44.7 \pm 0.4)$ , respectivamente.

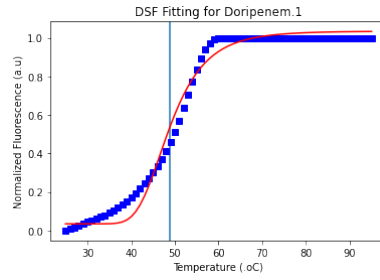


(continua)

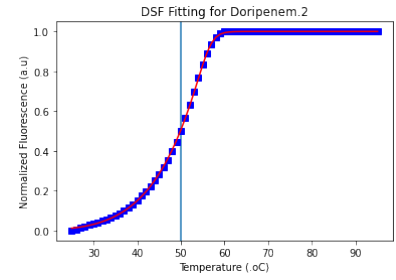
(continuação)



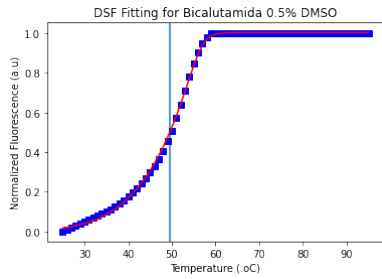
(j)



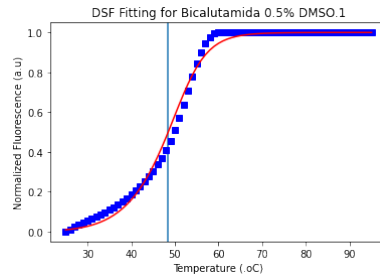
(k)



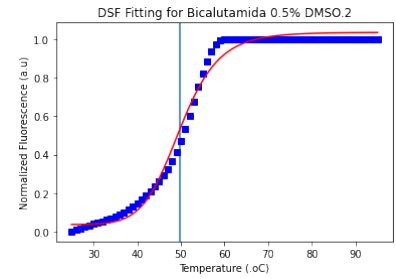
(l)



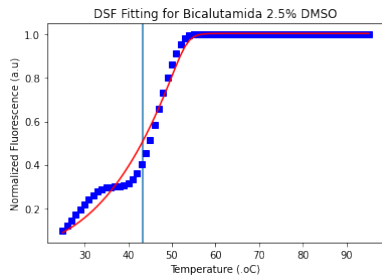
(m)



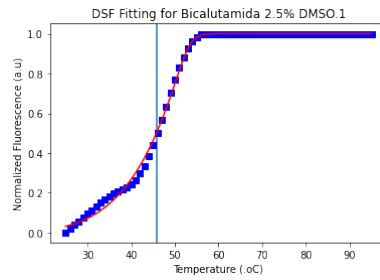
(n)



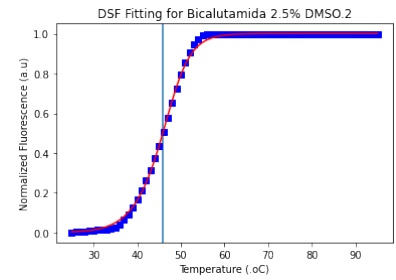
(o)



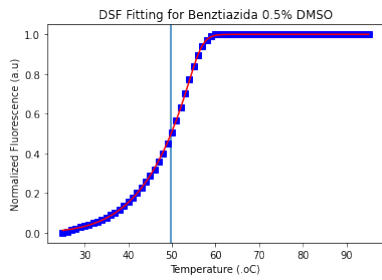
(p)



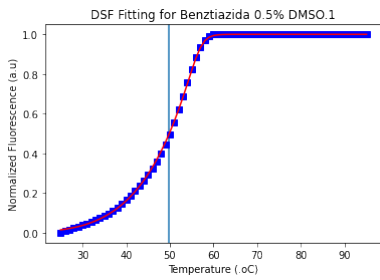
(q)



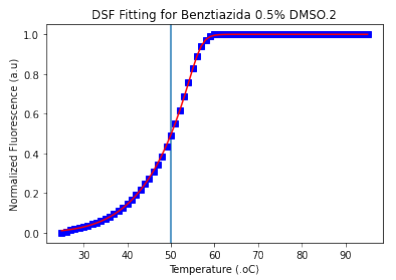
(r)



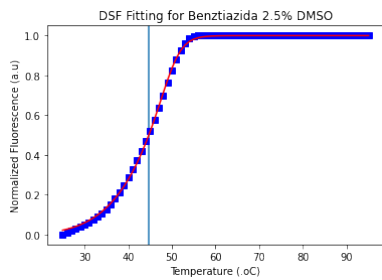
(s)



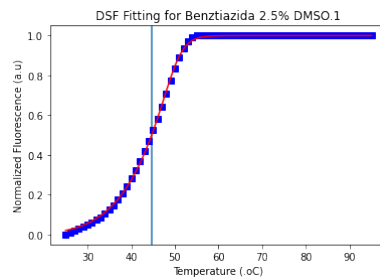
(t)



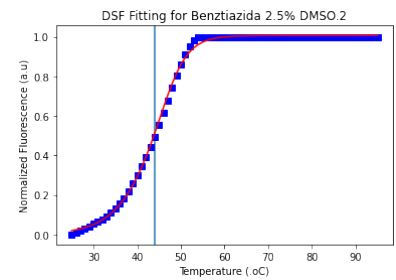
(u)



(v)



(w)



(x)

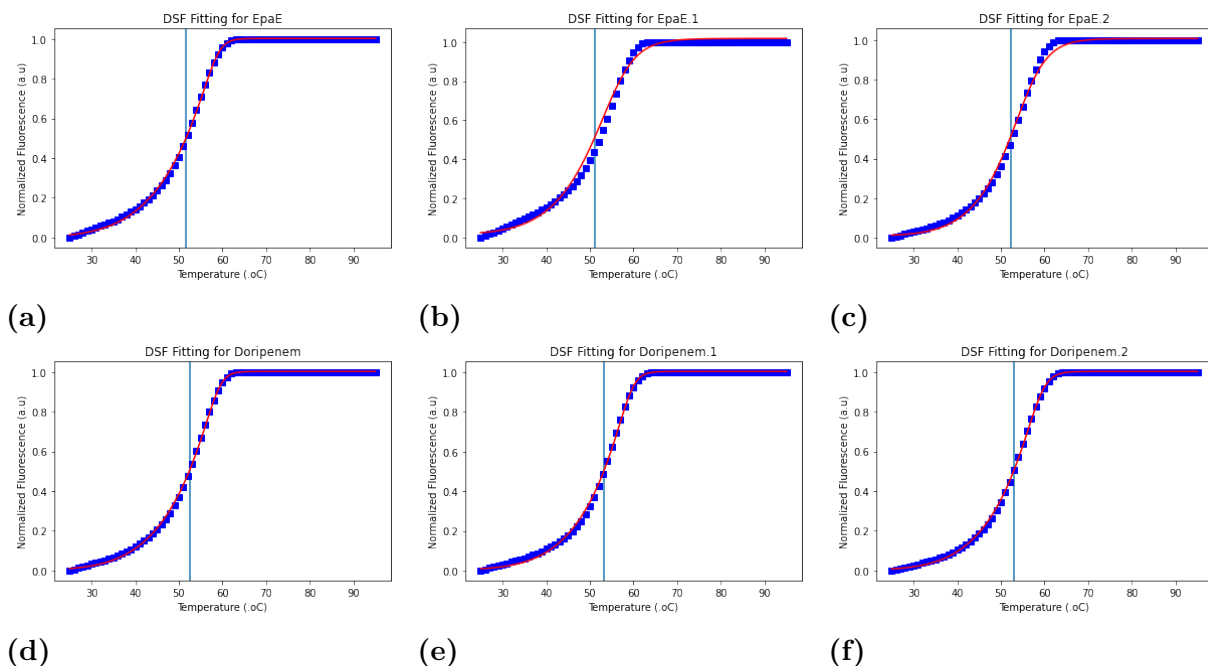
(continua)

(continuação)

**Figura 79** – Resultados das medidas realizadas para identificação de interações entre a EpaE e um potencial ligante. Eles consistem na fluorescência normalizada da sonda JBS *Thermofluor Dye* em função da temperatura submetida pela EpaE. Em todos os casos, está presente tampão 50 mM Tris pH 8.0, 100 mM NaCl e 10% Glicerol e 11 nM de EpaE. Os resultados (a), (b) e (c) são referentes ao branco, sem ligantes ou DMSO, apenas EpaE, com  $T_m = (48.7 \pm 0.4) ^\circ\text{C}$ . Os resultados (d), (e) e (f) são referentes à EpaE na presença de 2.5% DMSO em solução, com  $T_m = (43.7 \pm 0.4) ^\circ\text{C}$ . Os resultados (g), (h) e (i) são referentes à EpaE na presença de 5 mM  $\text{MgCl}_2$ , 0.2 mM dTTP e 1 mM Glicose-1-fosfato, com  $T_m = (67.7 \pm 0.4) ^\circ\text{C}$ . Os resultados (j), (k) e (l) são referentes à 3.1 mM de Doripenem, com  $T_m = (48.7 \pm 0.4) ^\circ\text{C}$ . Os resultados (m), (n) e (o) são 81 nM de Bicalutamida em 0.5% DMSO, com  $T_m = (48.7 \pm 0.4) ^\circ\text{C}$ . Os resultados (p), (q) e (r) são referentes à 810  $\mu\text{M}$  Bicalutamida e 2.5% DMSO, com  $T_m = (44.3 \pm 0.9) ^\circ\text{C}$ . Os resultados (s), (t) e (u) são referentes à 25  $\mu\text{M}$  Benztiázida e 0.5% DMSO, com  $T_m = (49.7 \pm 0.4) ^\circ\text{C}$ . Os resultados (v), (w) e (x) são referentes à 24 nM Benztiázida e 2.5% DMSO, com  $T_m = (44.7 \pm 0.4) ^\circ\text{C}$ .

Fonte: Elaborada pela autora com Matplotlib.<sup>9</sup>

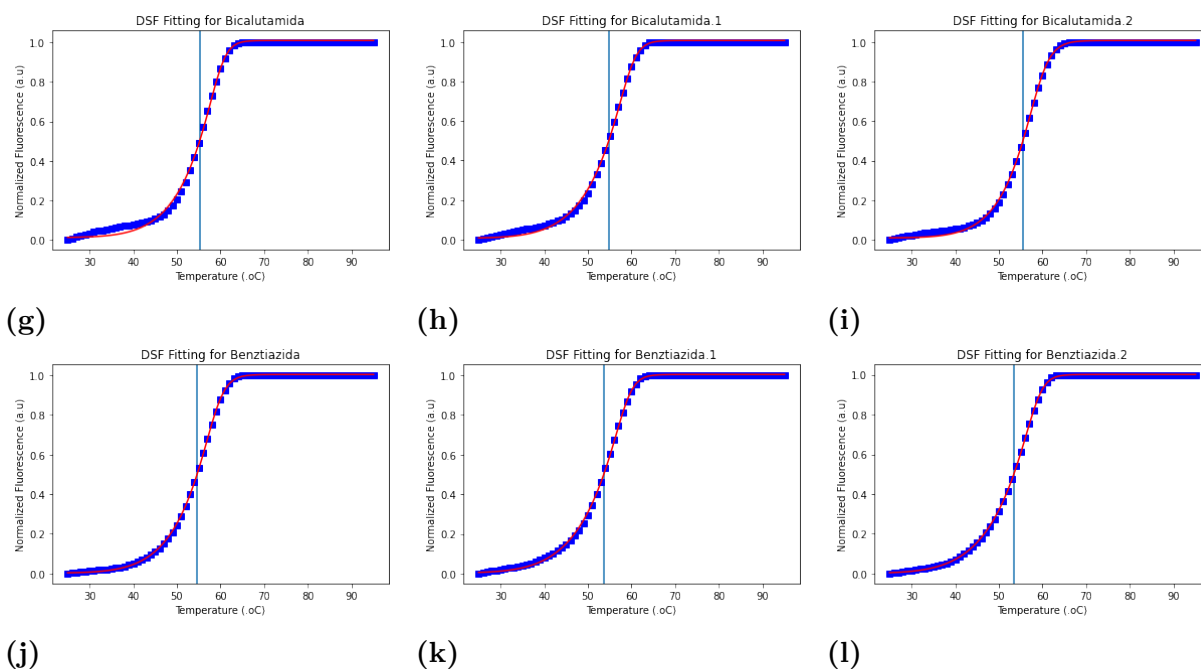
Para impedir que os potenciais ligantes interagissem com a EpaE pelo sítio catalítico, o experimento foi repetido mantendo a relação 1:1 entre a concentração molar de proteína e do dTTP, na presença de  $\text{Mg}^{+2}$ , cujo  $T_m$  é  $(51.3 \pm 0.4) ^\circ\text{C}$  (Figuras 80a a 80c), enquanto que no caso dos ligantes, a relação das concentrações molares da EpaE com as moléculas pequenas se mantivesse em, no mínimo, 1:3 (Figura 80).



(continua)



(continuação)



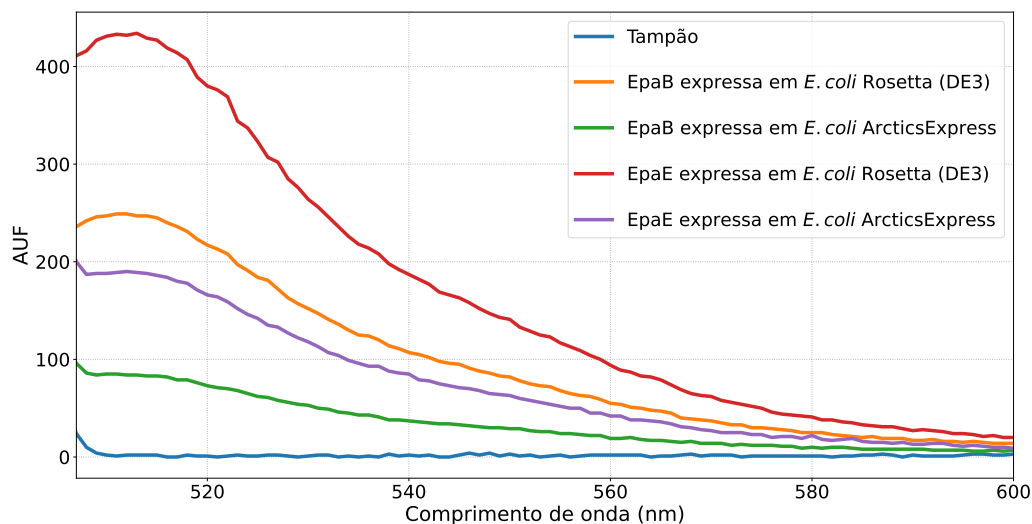
**Figura 80** – Resultados das medidas realizadas para identificação de um ligante no sítio alostérico da EpaE. Eles consistem na fluorescência normalizada da sonda JBS *ThermoFluor Dye* em função da temperatura submetida pela EpaE. Em todos os casos, estão presentes EpaE e dTTP, ambos a 11 nM e 5 mM de MgCl<sub>2</sub>, além do tampão 50 mM Tris pH 8.0, 100 mM NaCl e 10% Glicerol. Os resultados (a), (b) e (c) são referentes ao branco, sem ligantes, com T<sub>m</sub> = (51.3 ± 0.4) °C. Os resultados (d), (e) e (f) são referentes a 3.2 μM de Doripenem, com T<sub>m</sub> = (52.3 ± 0.4) °C. Os resultados (g), (h) e (i) são referentes a 69 nM de Bicalutamida, com T<sub>m</sub> = (54.7 ± 0.4) °C. Os resultados (j), (k) e (l) são referentes a 29 nM de Benztiazida, com T<sub>m</sub> = (53.3 ± 0.4) °C.

Fonte: Elaborada pela autora com Matplotlib.<sup>9</sup>

### D.1 EpaB e EpaE

Os genes *epaB* e *epaE* foram clonados no vetor pLIC-EGFP a partir da biblioteca de gDNA de *E. faecalis*, e, consecutivamente, as enzimas codificadas por eles foram expressas nas cepas Rosetta (DE3) e ArcticExpress. A proteína EpaB manifestou menor solubilidade do que a EpaE, a qual se apresentou mais viável para experimentos cristalográficos e enzimáticos.

Na construção do vetor pLIC-EGFP, a eGFP (*Enhanced green fluorescent protein*) é expressa na região C-terminal do que é codificado após a clonagem molecular neste vetor. A eGFP e a proteína a ela fusionada interferem mutuamente com o enovelamento uma da outra. Caso a eGFP não esteja enovelada, a intensidade de fluorescência por ela emitida é menor.<sup>139</sup> Pela fluorescência das células após expressão da fusão de eGFP com as enzimas EpaE e EpaB, é possível ver um pico de intensidade maior em torno de 512 nm nas células contendo EpaE em relação com as que contêm a EpaB (Figura 81).

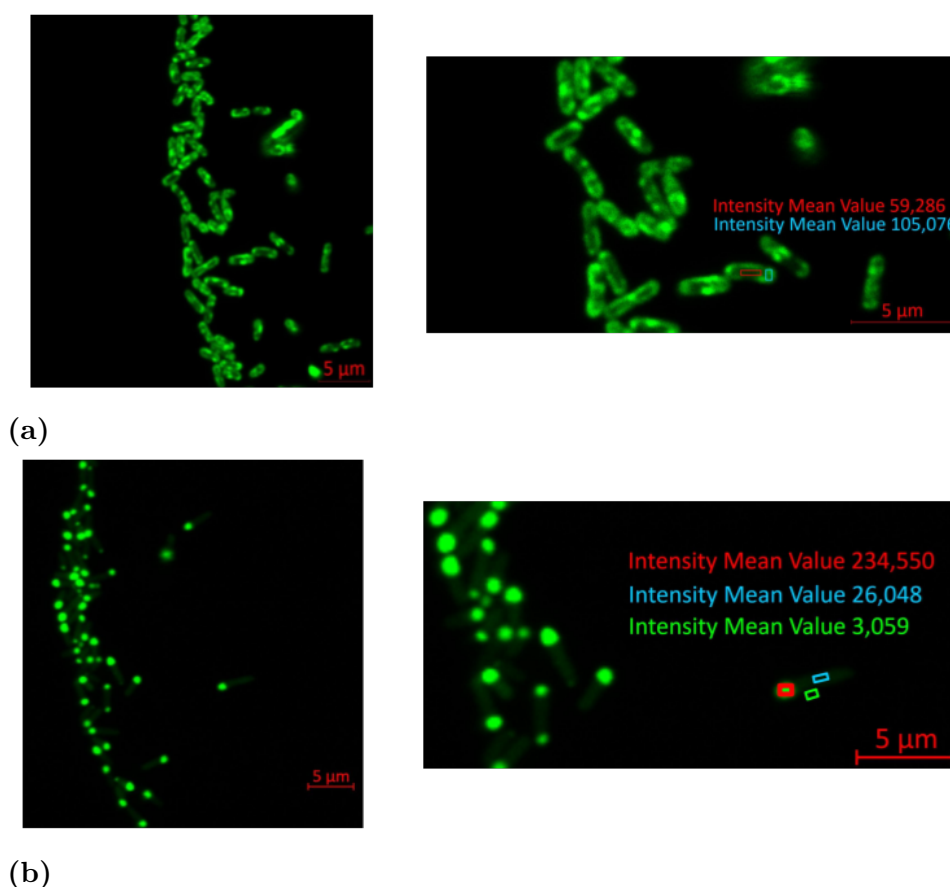


**Figura 81** – Espectro de fluorescência das células pós expressão no vetor pLIC-EGFP com IPTG, em *E. coli* Rosetta (DE3) e ArcticExpress. Medição realizada no espectrofotômetro *infinite* M200 PRO (TECAN).

Fonte: Elaborada pela autora com Matplotlib.<sup>9</sup>

O preparo da amostra para visualização no microscópio confocal aconteceu após a co-expressão das proteínas EpaB e EpaE com a eGFP, com volume final de expressão

de 50 mL em meio LB. Após a expressão, 1 mL do meio de cultura foi diluído 10x para atingir  $DO_{600nm} = 0.1$ ,<sup>140</sup> de modo a evitar sobreposição entre as células na lamínula. O meio foi então centrifugado a  $10k \times g$  durante 10 min, o sobrenadante foi descartado e as células foram lavadas com tampão PBS pH 7.4. Este último processo foi repetido três vezes, utilizando o mesmo volume do sobrenadante para cada lavagem. Por fim, 10  $\mu$ L da amostra foram depositados em uma lamínula, e levadas ao microscópio óptico de fluorescência confocal Zeiss (modelo LSM 780 invertido), obtido pelo processo FAPESP 2009/54035-4, e comandado pelo Prof. Dr. Francisco Guimarães. A lente utilizada foi uma de 63x; 1.2NA; wd = 0.2 mm (óleo), com excitação em 488 nm do laser de Argônio. O programa ZEN® foi aplicado para analisar as medidas de fluorescência no citosol e na membrana celular das células. Esse procedimento foi realizado conjuntamente com a então aluna de iniciação científica Maria Luiza Ferreira Vicente.



**Figura 82** – Microscopia confocal de fluorescência a partir da excitação em 488nm da eGFP fusionada com as proteínas (a) EpaB e (b) EpaE.

Fonte: Elaborada pela aluna de iniciação científica na época, Maria Luiza.

Após a expressão das proteínas EpaB e EpaE em *E. coli* Rosetta (DE3), e preparação das amostras, depositadas em lamínulas de vidro, as eGFP fusionadas com EpaB/E dentro das células foram excitadas em 488 nm pelo laser de Argônio no microscópio confocal, e a emissão intensa da eGFP entre 504 nm-548 nm foi observada (Figura 82). Com aproximação

64x e resolução 1024 x 1024, é possível perceber a fluorescência da construção EpaB-eGFP se localizando ao redor da célula, concentrada próximo ao invólucro celular (Figura 82a). Para afirmação da localização exata da EpaB, seria necessário um marcador de localização celular, em específico, na membrana celular, o que não foi realizado neste trabalho. Já para o caso da eGFP fusionada à EpaE (Figura 82b), a emissão é mais intensa no interior da célula Rosetta (DE3) *E. coli*, sugerindo uma localização citosólica para esta fusão. Ensaio de solubilidade anteriores mostraram que a EpaB permanecia, em sua maioria, na fração insolúvel, enquanto que a EpaE se mostrava presente tanto na fração solúvel quanto insolúvel. Sendo assim, os experimentos da EpaB também foram encerrados.