+UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

ANDREW ALBERT DE OLIVEIRA

Estudos estruturais e descoberta de ligantes da enzima diidropteroato sintase de
Xanthomonas albilineans para o combate da escaldadura das folhas

São Carlos

2019

ANDREW ALBERT DE OLIVEIRA

Estudos estruturais e descoberta de ligantes da enzima diidropteroato sintase de
Xanthomonas albilineans para o combate da escaaldadura das folhas

Tese apresentada ao Programa de Pós-Graduação em Física no Instituto de Física de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Doutor em Ciências.

**Área de Concentração: Física Aplicada
Opção: Física Biomolecular
Orientador: Prof. Dr. Rafael V. C. Guido**

Versão Corrigida

(Versão original disponível na Unidade que aloja o Programa)

São Carlos

2019

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.

Oliveira, A. A. de

Estudos estruturais e descoberta de ligantes da enzima diidropteroato sintase de *Xanthomonas albilineans* para o combate da escaldadura das folhas / A. A. de Oliveira; orientador Rafael Victorio Carvalho Guido - versão corrigida -- São Carlos, 2019.
165 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Aplicada Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2019.

1. Escaldadura das folhas. 2. *Xanthomonas albilineans*. 3. DHPS. I. Victorio Carvalho Guido, Rafael, orient. II. Título.

AGRADECIMENTOS

Às minhas três mulheres: Maurícia, Fátima e Maria. Obrigado pelo carinho, dedicação, inspiração, zelo, amor e exemplos de luta e garra para vencer as batalhas diárias. À matriarca, onde tudo começou, deixo aqui minha eterna saudade e o resultado de que nosso sonho foi realizado. Sem vocês nada disso seria possível!

Ao meu irmão, Jorge, pelas conversas, amizade e companheirismo por toda essa vida.

À minha eterna namorada, Elisa, que nesses anos tem sido meu alicerce para lidar com leveza com os desafios e obstáculos. Obrigado pelo amor, carinho, pela presença incansável nas dificuldades e nas celebrações das vitórias atingidas.

Ao Prof. Dr. Rafael V.C. Guido, pela oportunidade de ingresso no mundo acadêmico da pós-graduação, pelas discussões construtivas e por ter me orientado nesses anos todos. Obrigado também pela paciência, amizade, ideias sugeridas e por celebrar comigo nossas conquistas.

Aos meus contemporâneos de doutorado, Fernando, Renata, Gustavo e Larissa. Da mesma forma, obrigado às pós-doutorandas Fernanda, Maria Amélia e Carol. Obrigado pela incrível convivência, pelos risos, pelas discussões, pelas críticas e o espaço cedido para poder aprender com vocês.

Ao Dr. Marko Hyvönen, pela oportunidade única de explorar e conhecer um ambiente científico com pessoas tão incríveis. Obrigado por me orientar e ter sido sempre solícito. Agradeço também a todos do “MH_group”: Theo, Tom, Xuelu, Paul, Beata e Michael, Alexei, Tônia, Joe, Matt, Katharina, Mariana, Monique, Nathália e Luana que fizeram as minhas experiências em Cambridge ainda mais inesquecíveis. Obrigado, Ivan, por me receber no exterior, pela ajuda, conversas, consultorias e pela amizade neste último ano.

Aos pesquisadores Dr. Márcio V. B. Dias, Dra. Silvana C. D. Souza, Dr. Henri Schrekker, agradeço pelas colaborações que incrementaram e auxiliaram este projeto. Agradeço, ainda, a todos os integrantes da linha MX2 do LNLS e funcionários das diversas linhas de luz Diamond por serem tão atenciosos e prestativos nas coletas realizadas.

Às técnicas da biofísica, Andressa e Bel, por serem o ponto de apoio na realização de alguns dos experimentos citados nessa obra.

Ao Instituto de Física de São Carlos e à Universidade de São Paulo, pela oportunidade de realização do curso de Doutorado.

À Fundação de Amparo à Pesquisa do Estado de São Paulo, pela concessão da bolsa de Doutorado Direto (Processos 2013/17093-1 e 2015/07005-3) e BEPE (2017/17816-4), tornando o projeto possível de realização através do aporte financeiro.

Finalmente, aos amigos, familiares e pessoas próximas, obrigado por estarem comigo até aqui.

“O conhecimento serve para encantar as pessoas, não para humilhá-las”.

Mario Sérgio Cortella

RESUMO

OLIVEIRA, A.A. **Estudos estruturais e descoberta de ligantes da enzima diidropteroato sintase de *Xanthomonas albilineans* para o combate da escaldadura das folhas**. 2019. 165p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2019.

A bactéria gram-negativa *Xanthomonas albilineans*, invasora de xilema, é causadora de uma das principais doenças encontradas em culturas de cana-de-açúcar: a escaldadura das folhas. Atualmente, não é conhecido qualquer produto, químico ou biológico, que controle essa praga. Portanto, o grave panorama torna urgente o desenvolvimento de novas alternativas para o controle da fitodoença. Para tanto, foram traçadas estratégias da descoberta e desenvolvimento de novas moléculas bioativas através da Triagem Biológica, Química Medicinal e Biologia Estrutural. Foram testados os potenciais inibitórios de duas séries imidazólica, totalizando 41 compostos frente à cultura de *X. albilineans*, que resultaram em 11 compostos com potência intermediária no crescimento da bactéria. Em outra vertente, selecionamos dois alvos moleculares baseados em duas vias essenciais para o organismo: i. via dos folatos e ii. síntese de albicidina. A enzima diidropteroato sintase (*XaDHPS*) está relacionada à produção do tetrahidrofolato, enquanto a enzima benzoato coenzima A ligase (*XaBCL*), ligada ao fator de virulência. Com a enzima *XaDHPS* avanços significativos foram obtidos, dentre eles destacam-se: i. Purificação; ii. Cristalização; iii. Coleta de mais de 300 conjuntos de dados iv. Resolução da estrutura 3D inédita (na forma apoenzima e em complexos com substrato nativo e produto) e v. triagem, identificação, validação e detalhamento estrutural de fragmentos. No total foram elucidados os modos de ligação de 43 fragmentos moleculares em diferentes regiões proteica. Identificamos a presença de um sítio único, $\beta 1$, bem como a caracterização de bolsões auxiliares e estados de transição na movimentação das alças catalíticas. A presença de compostos próximos facilita a ligação em uma nova molécula com poder inibitório da enzima *XaDHPS*. Os resultados deste trabalho permitirão que métodos de planejamento baseado na estrutura do alvo receptor (SBDD) e do fragmento (FBDD) possam ser aplicados para a construção e desenvolvimento de novos compostos bioativos como candidatos à agroquímicos para cultura de cana-de-açúcar. Ademais, o mapeamento desta enzima contribui para o desenvolvimento de novos agentes antimicrobianos no combate de doenças causadas por bactérias gram-negativa.

Palavras-chave: Escaldadura das folhas. *Xanthomonas albilineans*. DHPS.

ABSTRACT

OLIVEIRA, A.A. **Structural studies and ligand discovery of the enzyme dihydropteroate synthase from *Xanthomonas albilineans* for the leaf scald treatment.** 2019. 165p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2019.

The gram-negative bacterium *Xanthomonas albilineans* causes one of the substantial diseases found in sugarcane crops: leaf scald. Currently, no chemical or biological product is known to control this plague. Therefore, this severe panorama raises the urgent need for developing new alternatives for pathogen control. Therefore, strategies were drawn for the discovery and development of new bioactive molecules through Phenotypic Screening, Medicinal Chemistry and Structural Biology. The inhibitory potentials of two imidazole series were assessed through bacterial growth analysis, totalizing 41 compounds, that resulted in 11 intermediate capacity inhibitors to stop the bacterium growth. Simultaneously, we have selected two molecular targets based on two *X. albilineans* essential pathways: i. Folate biosynthesis and ii. Synthesis of albicidin. The enzyme dihydropteroate synthase (*XaDHPS*) is related to the production of tetrahydrofolate, while the enzyme benzoate coenzyme A ligase (*XaBCL*) linked to the virulence factor. Over the enzyme *XaDHPS*, significant advances were obtained, among them: i. Purification; ii. Crystallisation; iii. Collection of more than 300 datasets iv. Resolution of the unpublished 3D structure (in the apoenzyme form and complexes with native substrate and product) and v. Screening, identification, validation and structural detailing of complexed fragments. In detail, the binding modes of 43 molecular fragments bound to diverse protein regions were elucidated. We recognised the presence of a single site, $\beta 1$, as well as the characterisation of auxiliary pockets and transition states in the movement of the catalytic loops. The presence of nearby compounds facilitates binding into a new molecule with inhibitory potential against *XaDHPS*. The results of this work will allow planning methods based on the receptor target structure (SBDD) and the fragment (FBDD) to be applied for the construction and development of new bioactive compounds as agrochemicals candidates for sugarcane culture. In addition, the mapping of this enzyme contributes to the development of new antimicrobial agents in the fight against diseases caused by gram-negative bacteria.

Keywords: Leaf scald. *Xanthomonas albilineans*. DHPS.

Lista de Figuras

Figura 1	- Número de agentes etiológicos resistentes da lavoura reportados por tipo de organismo. O crescimento do número de pragas do campo foi analisado entre 1930 e 2015.	28
Figura 2	- Comparativo entre o processo de descoberta e desenvolvimento de novos fármacos e agroquímicos. As moléculas resultantes de ambos processos são similares, no entanto, o tempo e custo da liberação de um novo agente no mercado é diferente. O comparativo engloba também a quantidade de compostos utilizada em cada fase e os valores desembolsados. NTBC = nitrosinona.	29
Figura 3	- Matriz energética brasileira. Ano base 2018 – Empresa de Pesquisa Energética.	31
Figura 4	- Produtividade mundial de cana-de-açúcar por país. O índice térmico se refere em vermelho à maior produtividade, ou seja, o maior rendimento em área plantada pela área colhida. O Brasil configura entre os países com valores médios de produtividade (74 %).	32
Figura 5	- Fases sintomáticas da doença escaldadura das folhas. Fase aguda caracterizada pela presença de estrias brancas da bainha ao limbo foliar. Fase crônica, marcada pelo embranquecimento completo das folhas das plantas infectadas.	35
Figura 6	- Cladograma das bactérias do gênero Xanthomonadaceae. As características da bactéria Xanthomonas albilineans se aproxima do gênero Xylella.	37
Figura 7	- Comparativo da afinidade dos inibidores de DNA girase (Topoisomerase II) GyrA. A molécula de simociclinona (magenta) apresenta valores de $IC_{50} = 1,45 \mu M$ contra <i>Staphylococcus aureus</i> . Para o mesmo alvo, a molécula de albicidina se mostra mais potente que outros inibidores descritos.	38
Figura 8	- Esquema ilustrativo da via de produção de albicidinas. O complexo de produção entre policetídeos e peptídeos não-ribossomais conferem a produção da fitotoxina. Esta molécula é composta por ácido metil p-coumárico (MCA-1), dois ácidos 4-aminobenzóico (pABA-2 e pABA-4), a parte proteica não usual α -aminoácido β -ciano-L-alanina no centro da estrutura e dois ácidos 4-amino-2-hidroxi-3-metoxibenzoico (pMBA-5 e pMBA-6) na extremidade C-terminal. A enzima benzoato coenzima A ligase tem papel fundamental no fornecimento de substratos.	39
Figura 9	- Ligação do composto albicidina na enzima albA (PDB 6ET8). A ligação é apresentada com a proteína em cartoon e em superfície. O ligante ocupa toda a extensão proteica em mecanismo de detoxificação celular.	40
Figura 10	- Representação 2D da molécula de tetrahidrofolato. As caixas hachuradas indicam as três porções da molécula: anel pterina, ácido 4-aminobenzóico (pABA) e cauda glutamilada, que pode ser estendida (n). A formação do tetrahidrofolato é iniciada pelo fornecimento dos subprodutos HMDHP e pABA das vias GTP-ciclodrolase e shikimato sintase, respectivamente. Dentro da mitocôndria, há a fosforilação do substrato HMDHP (pela enzima HPPK), com a subsequente transferência do produto HMDHP- P_2 ao pABA pela enzima DHPS. Especificamente, ocorre a liberação do pirofosfato e o diidropteroato. As enzimas em sucessoras utilizam a molécula de	

diidropteroato para a adição da cauda de ácido glutâmico (DHFS/FPGS) ou a redução do anel pterina (DHFR).....	41
Figura 11 - Esquema representativo das reações catalisadas pelas enzimas diidropterina pirofosfoquinase (HPPK) e diidropteroato sintase (DHPS). A enzima HPPK que fosforila o 6HMPH2, prepara o substrato (DHPP) para o uso na enzima diidropteroato sintase (DHPS). Nesta última, a enzima DHPS catalisa a ligação do pABA ao DHPP formado pela HPPK, liberando pirofosfato e o diidropteroato	43
Figura 12 - Reação bioquímica catalisada pela enzima diidrofolato redutase (DHFR). A enzima reduz o anel pterina para a formação do tetrahydrofolato.....	44
Figura 13 - Reação bioquímica catalisada pela enzima benzoato coenzima A ligase (BCL). Com o consumo de ATP há a ligação do benzoato à coenzima A. A enzima funciona como carreadora dos substratos necessários à formação da albicidina.	45
Figura 14 - Esquema indicativo da interação entre a cauda hexa-histidina expressa conjuntamente com as proteínas de interesse e a coluna carregada com o metal níquel (Ni^{2+}) na cromatografia de afinidade. As proteínas e debris celulares sem afinidade pela coluna é eluído, restando apenas a proteína de interesse.	58
Figura 15 - Quadro expositivo da configuração das placas de 96 poços com as condições triadas em experimento de estabilidade térmica.	60
Figura 16 - Estrutura 2D da série de compostos testada frente a cultura de <i>Xanthomonas albilineans</i>	65
Figura 17 - Compostos testados contra o crescimento de <i>Xanthomonas albilineans</i> . A segunda série aumenta a cadeia carbônica dos derivados imidazólicos. Em alguns casos há a inclusão de um segundo anel imidazólico (33), ou a troca do imidazol por uma piridina (31). Os compostos foram diluídos em 100 % DMSO.	70
Figura 18 - Gel 0,8 % agarose contendo resultado da amplificação da região específica PGBL e das proteínas não ribossomais NRPS1/3,2,5 de <i>Xanthomonas albilineans</i> . MK = marcador de peso molecular	70
Figura 19 - Representação 2D da molécula de imidazol. As possíveis substituições estão indicadas em R1-5. Alguns derivados podem atuar como antibacteriano, como por exemplo o Metronidazol.	712
Figura 20 - Esquema ilustrativo do resultado do teste de crescimento bacteriano de <i>Xanthomonas albilineans</i> , em placa de 100 poços, na presença de ligantes. As placas em colméia pertencem ao equipamento Bioscreen C MBR.....	72
Figura 21 - Exemplo de resultados da atividade dos compostos com efeito no crescimento bacteriano da <i>Xanthomonas albilineans</i> , em diferentes concentrações. O cálculo dos valores destas concentrações efetivas é feito após titulação em diluição seriada do ligante que vão desde 2 mM até 2,5 nM.	72
Figura 22 - Resultado da triagem fenotípica de compostos imidazólicos frente a cultura de <i>Xanthomonas albilineans</i> . Os compostos organizados no grupo em verde causaram inibição de 100 % do crescimento bacteriano à 125 μ M. Os compostos no grupo em vermelho não influenciaram em taxas maiores que 50 % no crescimento bacteriano à 2 mM.	74

Figura 23 - Representação das curvas de EC ₅₀ obtidas no ensaio de triagem fenotípica de compostos. Foram utilizadas 20 concentrações diferentes para a determinação dos valores de concentração efetiva. Os valores de EC ₅₀ determinados estão apresentados em cada estrutura bidimensional dos compostos de todas as séries testadas.....	Error! Bookmark not defined.
Figura 24 - Amplificação gênica das ORF <i>XafolP</i> e <i>XafolA</i> resultante dos experimentos de PCR de colônia. A amplificação destas regiões confirma que os vetores trabalhados contém os genes inseridos.....	77
Figura 25 - Sequência da região gênica responsável pela produção da enzima diidropteroato sintase de <i>Xanthomonas albilineans</i> . A região sequenciada pertence ao plasmídeo pETTrx, o qual contém uma tiorredoxina como proteína de fusão na extremidade N-terminal, além de uma cauda hexa-histidina conjugada a esta última proteína.	78
Figura 26 - Disposição dos códons raros presentes nos segmentos gênicos <i>folA</i> , <i>folK</i> e <i>folP</i>	79
Figura 27 - A. Gel 10 % SDS-PAGE referente a avaliação do perfil de expressão proteico XaDHPS nos vetores do sistema LIC. B. Purificação da enzima XaDHPS (31,8 kDa). S = fração solúvel, I = fração insolúvel, TEV = clivagem com a enzima TEV protease, GF = filtração em gel, por exclusão molecular e MK = marcador de peso molecular.	80
Figura 28 - Cromatogramas de purificação da enzima XaDHPS nas diversas etapas realizadas. A. cromatografia de afinidade. B. remoção de excesso de sais e imidazol (desalting). C. cromatografia de afinidade e D. exclusão molecular. A remoção de cada enzima nas etapas de purificação está indicada nos retângulos em cores.	81
Figura 29 - Resultado do experimento de triagem de soluções tamponante via DSF. A enzima XaDHPS é estabilizada em diferentes configurações de tampões e aditivos. Cada sigmoide representa uma condição testada durante o experimento	82
Figura 30 - Gel 10 % SDS contendo padrão de expressão da enzima XaDHFR nos vetores pET-M11, pETTrx-1a e pETNus-1a. Esquerda = Expressão em Rosetta 2 (DE3); Sol. = fração solúvel, Ins = fração insolúvel e MK = marcador de peso molecular.	84
Figura 31 - Etapa final de purificação da enzima XaDHFR. (O gel SDS-PAGE 15 % confirma a pureza das frações recolhidas após utilização da coluna de exclusão molecular. O cromatograma indica a fração majoritária como sendo a proteína de interesse isolada.	85
Figura 32 - Testes de estabilidade térmica (DSF) frente a proteína XaDHFR. Os resultados indicam preferência e estabilidade em tampão fosfato pH 7,5 e 200 mM NaCl com adição de 10 % glicerol.....	86
Figura 33 - Medidas do consumo de NADPH ao longo do tempo. As soluções contêm quantidades equimolar do substrato diidropteroato e o cofator NADPH, acrescidos da enzima XaDHFR A. A enzima XaDHFR se mostrou não ativa no ensaio realizado. B. Na varredura do substrato e cofator observou-se que ambos absorvem luz no mesmo comprimento de onda 340 nm.....	86

- Figura 34 - Gel SDS-PAGE 15 % com a purificação da enzima humana diidrofolato redutase de Homo sapiens por cromatografia de exclusão molecular. Os diversos poços dizem respeito às frações eluídas no experimento de gel filtração. MK - Marcador molecular, S - fração solúvel.....87
- Figura 35 - Estrutura tridimensional da enzima humana diidrofolato redutase de Homo sapiens (PDB 1U70). A atividade enzimática desta enzima foi mensurada por absorbância através do consumo de NADPH e DHP ao longo do tempo. Quantidades iguais de cofator e substrato foram inseridos e, na presença da HsDHFR, consumidos durante os 4 minutos de experimento.....88
- Figura 36 - Clonagem da enzima XaHPPK utilizando o sistema SLIC (Sequence-and-ligation-independent cloning). A. Duas construções foram exploradas na extremidade N-terminal, sendo: 8xHis e 8xHis + ‘maltose binding protein’ (MBP) como proteína de fusão. B. A integridade do material foi assegurada por sequenciamento.....89
- Figura 37 - Gel SDS-PAGE 15 % com o resultado da purificação por exclusão molecular da enzima XaHPPK. Mk = marcador molecular, S = fração solúvel e T = lise total. Comparativo entre a atividade da enzima XaHPPK e o controle comparado no teste de atividade.90
- Figura 38 - Representação das 96 imagens que representam os poços de triagem das condições de cristalização. As triagens de soluções de cristalização foram feitas contra as enzimas XaDHFR e XaHPPK. Para a proteína XaHPPK também foram construídas as condições de adição 10 mM ATP e/ou 10 mM MgCl₂, previamente ao início do ensaio.91
- Figura 39 - A. Diagrama de solubilidade proteica em diferentes concentrações de agente precipitante. B. Esquema descritivo da técnica de difusão de vapor através da metodologia de gota suspensa para a triagem e produção de cristais proteicos.....92
- Figura 40 - Esquema ilustrativo dos experimentos de difração de raios X. A amostra utilizada (cristal) sofre a incidência dos raios X, difratando fótons que serão coletados pelo detector. O cálculo da densidade eletrônica, bem como a construção do modelo representativo seguem cálculos computacionais.92
- Figura 41 - Padrão de difração da enzima XaDHPS. O poderio de difração ultrapassa as resoluções de canto (1,5 Å). Destacamos o primeiro cristal obtido após triagem de condições de cristalização, os cristais comumente obtidos e o resultado do experimento de semeadura.....94
- Figura 42 - Diferentes métodos de coleta de dados pela técnica de difração de raios X. O método de fatiamento angular por fine ϕ -slicing auxilia no processo de captação de pontos únicos de difração. Outros métodos de fatiamento grosseiro (coarse ϕ -slicing) é indicado para conjuntos de dados que não apresentem problemas de pontos de difração próximos.....94
- Figura 43 - Resolução estrutural do cristal monoclinico de XaDHPS. A densidade eletrônica e disposição residual é correspondente à faixa da interface dimérica proteica.....96
- Figura 44 - Alinhamento de estruturas secundárias e comparativo entre o enovelamento de DHPS singular (PDB 1EYE) e do enovelamento na fusão com a enzima HPPK (2BMB). A enzima SchPPK está destacada em cinza, enquanto a porção

ScDHPS em azul. A apresentação da enzima DHPS de *M. tuberculosis* é feita em laranja. 98

- Figura 45 - Alinhamento sequencial das estruturas de Diidropteroato sintase disponíveis no banco de dados PDB. Os resíduos conservados estão indicados em vermelho, enquanto os resíduos idênticos em bastões verdes. Estruturas alinhadas: *Francisella tularensis* (PDB 3MCM, 4PVZ), *Escherichia coli* (PDB 3V5O, 1AJ0, 5U0V), *Burkholderia cenocepacia* (PDB 2Y5J), *Bacillus anthracis* (PDB 1TWS), *Mycobacterium tuberculosis* (PDB 1EYE, 2VP8), *Staphylococcus aureus* (PDB 4HB7, 1AD1, 6CLU), *Streptococcus pneumoniae* (PDB 2VEF), *Francisella tularensis* (PDB 3MCM), *Coxiella burnetii* (PDB 3TR9), *Mycobacterium smegmatis* (PDB 6OMZ), *Vibrio fischeri* (PDB 5U5W), *Klebsiella pneumoniae* (PDB 5UMG), *Thermus thermophilus* (PDB 2DQW) e *Yersinia pestis* (PDB 5JQ9, 3TYU)..... 101
- Figura 46 - Divisão das estruturas secundárias da enzima XaDHPS. O enovelamento de um barril 'TIM' compõe o complexo (α/β)₈. Em vários casos, principalmente em fungos, o grampo β não faz parte da estrutura de DHPS..... 101
- Figura 47 - Estrutura 3D dimérica da enzima XaDHPS. A estrutura é formada por oito hélices α (1-8) que circundam um barril β (1-8). Estas estruturas secundária são conectadas por oito alças L(1-8), das quais as alças L1, L2 e L5 participam da catálise enzimática. A interação recíproca das hélices α 7 e α 8 de cada monômero formam uma cavidade nesta interface..... 102
- Figura 48 - Estrutura 3D da enzima Diidropteroato sintase de *Xanthomonas albilineans* (XaDHPS) cristalizada na solução de tartrato de amônio e PEG 3350. São 8 moléculas na cela unitária organizada no grupo P21. O potencial eletrostático é descrito indicando as regiões carregadas negativamente (branco) e positivamente (marrom). 104
- Figura 49 - Caracterização estrutural do sítio catalítico da enzima diidropteroato sintase de *Xanthomonas albilineans*. A ligação do anel pterina segue a mesma orientação tanto isolado quanto no aduto formado com o ácido 4-aminobenzóico. O sítio de ligação do pirofosfato é mimetizado pelos íons sulfato (SO₄). 105
- Figura 50 - Detalhamento estrutural do sítio catalítico da enzima XaDHPS. A. Ligação do produto diidropteroato e do substrato 6HMPH₂. B. Interação dos compostos DHP e 6HMPH₂ no sítio ativo. 106
- Figura 51 - Resultado do experimento de ITC quantificando a afinidade da enzima XaDHPS pelo produto da reação catalítica, o ácido diidropteróico (DHP). O valor de K_d = 1.8 μ M foi definido para dois sítios de ligação. 107
- Figura 52 - Representação da estrutura 3D em complexo com o respectivo substrato nativo, ácido 4-aminobenzóico. Na estrutura DHPS:pABA o reconhecimento das alças nas hélices α 1, α 2 e α 5 são apresentados. 108
- Figura 53 - Movimentação da alça 1 (Phe33) para a entrada e acomodação de ligantes na interface dimérica. Na ausência de ligantes (ciano), a alça 1 preenche o sítio com a F33. Na entrada de um ligante ocorre o giro da alça em sentido ao sítio ativo (amarelo), passando pelo estado de transição (magenta)..... 109
- Figura 54 - Representação da molécula de pABA em complexo com o sítio da interface da XaDHPS. Enzimas homólogas não dispõem da mesma característica do bolsão hidrofóbico. Em outras enzimas que ligam pABA há ocupação da cadeia lateral

de resíduos mais volumosos que a A243 em <i>Xanthomonas albilineans</i> . Destacamos a presença de histidina em <i>T. thermophilus</i> (2DQW) e isoleucina em <i>Y. pestis</i> (5JQ9).....	110
Figura 55 - Resíduos disponíveis para reação com as sondas RED (amina primária - laranja) e MALEIMIDE (cisteína - amarelo). A exposição dos resíduos em direção ao solvente auxilia na marcação proteica. O sítio ativo está destacado pela superfície em verde.....	113
Figura 56 - Compostos e fragmentos identificados por afinidade a enzima XaDHPS por experimentos de DSF. A determinação da constante de afinidade contra a enzima XaDHPS realizada por MST.	115
Figura 57 - Complexos formados entre os fragmentos identificados por DSF com a enzima XaDHPS. As moléculas estão atracadas na interface dimérica. As densidades eletrônicas estão representadas pelos mapas de contorno 2Fo-Fc (1,0 σ).	117
Figura 58 - Representação da densidade eletrônica resultado do experimento de difração de raios X da enzima XaDHPS. Em amarelo está representado o fragmento em complexo com a proteína, que se liga tanto no sítio catalítico como na interface dimérica. As densidades eletrônicas estão representadas pelos mapas de contorno 2Fo-Fc (2,0 σ).....	120
Figura 59 - Resultado da triagem de fragmentos contra a enzima XaDHPS. Os compostos atingiram, majoritariamente, quatro regiões enzimática: sítio ativo, interface dimérica, alça L1 e contatos cristalográficos. A biblioteca contém 352 compostos organizados em 88 coquetéis.	121
Figura 60 - Fragmentos complexados na molécula de XaDHPS após a triagem de uma biblioteca de compostos por cristalografia de raios X. Os dados estão organizados por região de ligação na molécula: i. pABA (PB1-3), ii. contatos cristalográficos (CC1-3), iii. sítio da interface (IT1-20) e sítio do anel pterina (PT1-8).....	122
Figura 61 - Sobreposição dos fragmentos PB1 (verde) e PB2 (ciano) ligados no sítio de interação do ácido 4-aminobenzóico (magenta). Os compostos PB1, PB2 se assemelham no modo de ligação com o substrato natural. A ligação da molécula PB3 favorece a interação com os compostos na posição pterina do sítio catalítico. A cadeia proteica é destacada pelos carbonos em verde.	123
Figura 62 - A enzima XaDHPS em complexo com quatro moléculas ao mesmo tempo. No sítio catalítico PT4 e PB3 se localizam de maneira próxima, enquanto CC1-2 se mantém nas cavidades cristalográficas. As densidades eletrônicas estão representadas pelos mapas de contorno 2Fo-Fc (1,8 σ).....	124
Figura 63 - A. Interações realizadas entre os hits complexados no sítio do anel pterina. B. A possibilidade de expansão dos diversos fragmentos está indicada pelas setas. As densidades eletrônicas estão representadas pelos mapas de contorno 2Fo-Fc (1,5 σ).....	125
Figura 64 - A. Fragmentos complexados na interface dimérica da enzima XaDHPS. B. O sítio tem profundidade de 14 Å e características hidrofóbicas. O fragmento IT14 percorre toda a cavidade do sítio. A extremidade fenol está enterrada no bolsão, enquanto a outra extremidade faz face com o solvente e a alça 1. As densidades eletrônicas estão representadas pelos mapas de contorno 2fo-fc (1,8 σ).	126

Figura 65 - Análogos adquiridos com base nos compostos cloropirimidina (interface) e piridina (sítio ativo). As densidades eletrônicas estão representadas pelos mapas de contorno 2Fo-Fc (1,5 σ).....	128
Figura 66 - Variações nos valores da curva de fusão (T _m) pelo método de DSF. Os compostos foram divididos entre variações positivas e negativas.	128
Figura 67 - Empilhamento dos resultados dos complexos com os análogos de piridina identificadas pela triagem de fragmentos por cristalografia de raios X. Nd = não determinado.	129
Figura 68 - Proximidade de sítios na enzima XaDHPS. Um novo bolsão capaz de acomodar novos ligantes está presente entre a fita β 1 e a hélice α 8. A cavidade presente em cada monômero facilita a comunicação entre o bolsão e o sítio da interface.	130
Figura 69 - Representação do modelo 3D final gerado por homologia para a enzima XaBCL.....	152
Figura 70 - Gel SDS-Page 10 % cotendo padrão de expressão da enzima XaBCL nos vetores pET/LIC e pET/Sumo. Todas as construções mantidas nos debris celulares. Direita = Adição de aditivos no tampão de lise celular para expressão em vetor pET-M11: i. 5 % sarcosil, ii. 1 % Triton X-100, iii. 5 % glicose, iv. 100 mM L-arginina, v. 5 % sacarose. Sol. = fração solúvel, Ins = fração insolúvel, TEV = clivagem com a enzima TEV protease, 2 ^a afinidade, GF = filtração em gel, por exclusão molecular e MK = marcador de peso molecular.	154
Figura 71 - Esquema representativo da remoção dos 200 resíduos iniciais da enzima benzoato coenzima A ligase de Xanthomonas albilineans. Em cinza o domínio removido na versão truncada da enzima. O sítio de ligação do AMP está colorido por cilindros verdes.	155
Figura 72 - Esquema com as diferentes construções da enzima Benzoato coenzima A ligase de Xanthomonas albilineans (XaBCL).....	156

Lista de Tabelas

Tabela 1 - Oligonucleotídeos iniciadores utilizados em reação de PCR na confirmação do DNA genômico de <i>X. albilineans</i> . As regiões amplificadas são únicas nesta bactéria gram-negativa.	54
Tabela 2 - Parâmetros utilizados na reação de PCR para amplificação dos genes específicos de <i>X. albilineans</i>	54
Tabela 3 - Pares senso (Fw) e antisenso (Rv) 5' → 3' dos oligonucleotídeos iniciadores utilizados nas amplificações das ORF folA, folK e folP.....	55
Tabela 4 - Parâmetros utilizados para reação de PCR na amplificação dos fragmentos gênicos folA, folK e folP.....	56
Tabela 5 - Condições promissoras com surgimento de cristais da enzima XaDHPS.....	93
Tabela 6 - Dados estatísticos do processamento e refinamento das estruturas de XaDHPS, na forma apo e em complexo com ácido 4-aminobenzóico (pABA), o substrato 6HMPH2 e o produto diidropteroato (DHP). Estatísticas na máxima resolução estão em parênteses.	97
Tabela 7 - Parâmetros estatísticos do refinamento e resolução 'Table 1' dos complexos entre os compostos (B9, C7 e E5) e a enzima XaDHPS.	116
Tabela 8 - Lista com empresas que fornecem a infraestrutura para o desenvolvimento e triagem de fragmentos, bem como as técnicas utilizadas. Métodos abordados: X-ray (Cristalografia de raios X), ITC (Calorimetria de titulação isotérmica), NMR (Ressonância magnética nuclear), SPR (Ressonância plasmônica de superfície) e TR-FRET (Fluorescência resolvida no tempo).....	119

SUMÁRIO

1	Introdução	27
1.1	Desenvolvimento de novos agroquímicos.....	27
1.2	Desenvolvimento Sustentável e Agroenergia.....	30
1.3	Cana-de-açúcar como fonte de energia	32
1.4	Fitodoeças da cana-de-açúcar.....	33
1.5	Escaldadura das folhas	34
1.6	A bactéria <i>Xanthomonas albilineans</i>	35
1.7	Alvos Moleculares	40
1.7.1	Biossíntese dos Folatos	41
1.7.1.1	Enzimas 6-hidroximetil-7,8-diidropterina pirofosfoquinase / Diidropteroato sintase.....	42
1.7.1.2	Enzima Diidrofolato redutase.....	43
1.7.2	Via de Albicidinas - Benzoato coenzima A ligase	44
2	Objetivos.....	49
3	Material e Métodos	53
3.1	Isolamento e Cultura de <i>Xanthomonas albilineans</i>	53
3.2	Extração DNA genômico e Caracterização de X. albilineans	53
3.3	Amplificação e clonagem dos segmentos gênicos <i>folA</i> , <i>folK</i> , <i>folP</i>	54
3.4	Testes de Expressão Heteróloga.....	57
3.5	Purificação	58
3.6	Triagem de Condições de Estabilidade Térmica	59
3.7	Identificação de ligantes por DSF	61
3.8	Ensaio de Cristalização.....	61
3.9	Coletas de Dados.....	62
3.10	Processamento, Resolução e Refinamento	62
3.11	Termoforese em microescala (MST).....	63
3.12	Calorimetria de Titulação Isotérmica (ITC).....	63
3.13	Cinética enzimática <i>XaDHPS</i> , <i>XaDHFR</i> , <i>XaHPPK</i>	64
3.14	Triagem Fenotípica de Compostos.....	64
4	Resultados e Discussão (<i>Xanthomonas albilineans</i>).....	69
4.1	Isolamento, Cultura e Caracterização de <i>Xanthomonas albilineans</i>	69
4.2	Triagem de Compostos	70
5	Resultados e Discussão	77
5.1	Amplificação e Clonagem.....	77
5.2	Sequenciamento	77
5.3	Expressão e Purificação	79
5.3.1	Diidropteroato sintase <i>XaDHPS</i>	80
5.4	Ensaio Cinético e de atividade	82
5.4.1	Diidrofolato redutase <i>XaDHFR</i>	83
5.4.2	6-hidroximetil-7,8-diidropterina pirofosfoquinase (<i>XaHPPK</i>).....	88

5.5	Cristalização, Coleta de Dados, Resolução e Refinamento (<i>Xa</i> DHPS).....	91
5.6	Caracterização Estrutural.....	98
5.7	Análise estrutural da enzima <i>Xa</i> DHPS em complexo com pABA, 6HMPH2 ou DHP.....	104
5.8	Planejamento de Moléculas baseado na Estrutura do Fragmento (FBDD)	111
5.9	Termoforese em microescala.....	112
5.10	Identificação de Ligantes por DSF para <i>Xa</i> DHPS.....	114
5.11	Triagem de Inibidores por cristalografia de raios X.....	118
6	Considerações Finais.....	135
6.1	Conclusões	135
6.2	Perspectivas.....	137
	Referências.....	141
	Apêndice I.....	151
	Apêndice II	157

INTRODUÇÃO

1 Introdução

1.1 Desenvolvimento de novos agroquímicos

A domesticação e o cultivo de espécies de plantas e animais para o consumo humano, iniciados há cerca de 10.000 anos, têm sido significativos para a mudança ambiental global. Atividades de agricultura, como a fertilização e controle de doenças, vêm desencadeando mudanças fundamentais nos fluxos de carbono (C), nitrogênio (N) e fósforo (P) que originalmente compunham o ecossistema.¹

O crescimento contínuo da população mundial requer esforços para o aumento da produção de alimentos, bem como dos produtos de consumo e bem-estar. É estimado que, ao menos, 10 % da produção mundial de alimentos é perdida devido a ocorrência de doenças no cultivo (**Figura 1**). Os maiores patógenos são plantas parasitas, oomicetos, nematoides, vírus, fungos e bactérias. Dentro do reino Monera, existem mais de 200 espécies patogênicas, que compreendem os gêneros: *Pseudomonas*, *Ralstonia*, *Agrobacterium*, *Xanthomonas*, *Erwinia*, *Xylella*, *Pectobacterium* e *Dickeya*.²⁻³

Com a expansão da produção agrícola, o controle de pragas do campo sempre esteve atrelado à plantação de alimentos. A proteção química da lavoura provavelmente se iniciou com os Romanos por meio da aplicação de enxofre elementar. Posteriormente, o processo foi aprimorado com o uso de materiais trabalhados no dia-a-dia: esterco, mercúrio, tabaco, entre outros. Alguns anos mais tarde, a era dos agroquímicos se iniciou com a introdução do DDT (diclorodifeniltricloroetano) e dos inseticidas do tipo carbamato e 2,4-diclorofenoxiacético.⁴

No entanto, o uso de compostos organoclorados (e.g. DDT) foi banido de circulação em alguns países. Os efeitos nocivos destes compostos vão da sua volatilidade à persistência em se manter no meio ambiente, passando pela contaminação da cadeia ambiental subsequente e pela associação com doenças humanas (e.g. câncer).⁵⁻⁶

Figura 1 - Número de agentes etiológicos resistentes da lavoura reportados por tipo de organismo. O crescimento do número de pragas do campo foi analisado entre 1930 e 2015.

Fonte: Adaptada de SPARKS; LORSBACH³

Mudanças nos sistemas de cultivo e o uso de espécies geneticamente modificadas são algumas das técnicas de manejo atuais. No entanto, a resistência patogênica e a manifestação de novos agentes etiológicos, somados à restrição de uso de agroquímicos mencionada, alertam sobre a importância do desenvolvimento de novas alternativas. Portanto, o surgimento de espécies resistentes faz necessária a busca de novas entidades moleculares conjuntamente com seus específicos modos de ação (MoA).

Tradicionalmente, em um programa de descoberta de novos compostos bioativos, a geração de compostos líderes faz uso de estratégias de triagem em larga escala (HTS, do inglês, *high throughput screening*). A decisão de qual estratégia adotar deve levar em consideração os fatores custo, tempo e acesso a bibliotecas de compostos. Apesar das diversas similaridades no processo de descoberta e desenvolvimento de fármacos com agroquímicos, existem várias divergências nas moléculas utilizadas, por exemplo: a taxa de persistência no meio ambiente que assegure a proteção da lavoura por mais tempo, a quantidade distribuída no tratamento, valores desembolsados na elaboração e venda, dentre outros (**Figura 2**).⁷

corresponderia a uma colher de chá para proteger uma área similar a um campo de futebol.⁸⁻⁹ Em contraponto, no combate de cupins e coleópteros na lavoura de cana-de-açúcar com o Endossulfan (Thiodan ® - Bayer), são administrados um mínimo de 2100 g por hectare.

Assim, suprir as crescentes metas do desenvolvimento humano e, concomitantemente, proteger a saúde e o meio ambiente desafia toda a comunidade científica. A investigação de moléculas mais seletivas, construídas através da relação estrutura-atividade, do modo de ação e do planejamento racional contribui no manejo integrado e consciente.

1.2 Desenvolvimento Sustentável e Agroenergia

A definição de desenvolvimento sustentável passa por atender as necessidades humanas atuais sem comprometer as gerações futuras, meta delineada no acordo firmado pelos países componentes das Nações Unidas em 2015. O acordo descreve uma agenda universal que deve ser implementada e respeitada a fim de obter o máximo do desenvolvimento humano com a preservação global.

De forma a garantir equilíbrio nas ações tomadas, deve-se respeitar o contexto: i. Social, que compreende o desenvolvimento humano, a equidade na distribuição de renda, a estabilidade das instituições públicas e, a consequente redução dos conflitos entre as classes; ii. Econômico, com a capacidade de produção, distribuição e utilização equitativa das riquezas produzidas pelo homem; iii. Ambiental, com conjuntos de condutas que favoreçam a alteração saudável do ambiente, impactando positivamente a curto, médio e a longo prazo.

Um dos temas desenvolvidos neste sentido é a evolução dos sistemas energéticos. Em todos os países, com maior ou menor grau de sucesso, o plano sobre este setor atua para assegurar o suprimento de energia limpa e a preços acessíveis, o que é essencial para o bem-estar da sociedade e para o incremento econômico.¹⁰⁻¹¹ Por energia limpa, entende-se o desenvolvimento energético que subsidie o crescimento das atividades produtivas e da geração de empregos, ao mesmo tempo em que diminua e se responsabilize pelos impactos causados à qualidade ambiental.

Encontramos no etanol de cana-de-açúcar uma alternativa economicamente plausível e com significativo potencial de expansão. A produção e uso de etanol como combustível veicular são praticados regularmente no Brasil desde 1931, com notável evolução durante as últimas décadas. A atual estagnação da produção e aumento da importação e do preço de derivados do petróleo (e.g., gasolina), amplia a oportunidade de desenvolvimento interno do setor sucroenergético.

Cultivada desde o período colonial, a cana-de-açúcar se transformou em um dos principais produtos brasileiros. O Brasil é um dos maiores produtores de açúcar e etanol do mundo e estimula, cada vez mais, o uso de biocombustíveis como alternativa energética no mercado. O país é destaque, também, no uso de energias renováveis, que representam mais de 44 % da matriz energética nacional. O setor sucroenergético é peça-chave neste montante, uma vez que os produtos da cana-de-açúcar são responsáveis por 17 % de toda a oferta de energia limpa do país. Este valor já ultrapassa o fornecido pelas usinas hidroelétricas (**Figura 3**).¹²

Figura 3 - Matriz energética brasileira. Ano base 2018 – Empresa de Pesquisa Energética.*
Fonte: Elaborada pelo autor.

A Empresa de Pesquisa Energética (EPE) prevê que os derivados da cana-de-açúcar irão crescer 8,1 % ao ano na próxima década, ficando atrás apenas do gás natural, com 9,1 %. Com este crescimento, a cana-de-açúcar aumentará sua participação na matriz energética brasileira dos atuais 17 % em 2018 para 21,2 % em 2021. O etanol de primeira geração e a biomassa de cana ascenderão mais rápido do que a média das fontes renováveis (hidroelétrica, lixívia, dentre outras) na próxima década: em média 8,1 % ao ano, contra 5,1 % anuais, respectivamente.¹³ Para tanto é necessário aumentar a produtividade, ou seja, produzir maiores quantidades com o mesmo espaço dedicado à cultura. Atualmente a Austrália lidera o ranking dos países com maior produtividade em cana-de-açúcar (**Figura 4**).

* Empresa de Pesquisa Energética. Disponível em <http://www.epe.gov.br/pt/abcedenergia/matriz-energetica-e-eletrica>. Acessado em: 20 de março de 2019.

Figura 4 - Produtividade mundial de cana-de-açúcar por país. O índice térmico se refere em vermelho à maior produtividade, ou seja, o maior rendimento em área plantada pela área colhida. O Brasil configura entre os países com valores médios de produtividade (74 %).*

Fonte: Elaborada pelo autor.

Os benefícios da utilização do etanol como fonte alternativa de energia vão além da produção. Segundo dados da Agência Internacional de Energia (IEA), a utilização de etanol produzido através da cana-de-açúcar reduz em média 89 % a emissão de gases responsáveis pelo efeito estufa – como dióxido de carbono (CO₂), metano (CH₄) e óxido nitroso (NO₂) – quando comparado com a gasolina.¹⁴

1.3 Cana-de-açúcar como fonte de energia

Originária da Nova Guiné, o cultivo da ‘planta que brotou do mel’ foi introduzido no Brasil por Martim Affonso de Souza, em 1532. Localizadas no Nordeste, as capitanias de Pernambuco e da Bahia foram as pioneiras a multiplicarem a produção da monocotiledônea. A intenção do plantio era de firmar os portugueses em solo nacional, através do monopólio internacional da lavoura da cana-de-açúcar, o que foi frustrado pela atuação dos holandeses na competição de produção, extração e refino do açúcar no caribe.¹⁵

* União da Indústria de cana-de-açúcar. Disponível em: www.unicadata.com.br. Acesso em: 18 de março 2019.

A cultura extensiva dessa planta se manteve de maneira amena nos anos seguintes até que fosse assegurado o uso de um dos seus subprodutos: o etanol. Desde a sua estreia comercial no Brasil, nos anos 1920, a cadeia do etanol alterna graus de maior e menor intervenção do Estado. Porém, somente com o advento do Programa Nacional do Álcool (Proálcool), na década de 1970, é que se definiu claramente uma estratégia de médio e longo prazo, permitindo que o setor público intervisse positivamente na melhora da produção e no desenvolvimento da cadeia. A motivação governamental para lançar o Proálcool foi o peso do uso de derivados do petróleo na balança de pagamentos do país e na segurança energética, uma vez que importava, na época, mais de 80 % do que produzia. Contribuiu, também, o fato de que a indústria canavieira apresentava capacidade instalada ociosa.¹⁶

Considerando que as condições de produção canavieira no país são extremamente favoráveis (climáticas, solo, área), o programa visava a implementação do seguinte: a. uso compulsório de 10 % de etanol anidro como aditivo a gasolina e b. disponibilização em bombas do etanol 95 % como combustível (apena para motores modificados à época). Após 20 anos do programa em funcionamento, a indústria automobilística lançou carros bicompostíveis, que são confeccionados no país até hoje.

Atualmente, o Brasil é o maior produtor de cana-de-açúcar do mundo. Entretanto, há um alto custo ambiental: o desmatamento conjunto com a disputa pela produção de alimentos no cerrado faz com que a cultura canavieira descampe extensivamente as regiões tropicais e subtropicais do sul e sudeste brasileiro. Paralelamente, para atingir a demanda mundial de biocompostíveis, os líderes na produção de etanol deverão aumentar suas produções dos atuais 80 para aproximadamente 200 milhões de litros até 2021, o que demandará a disponibilização de áreas maiores para cultivo.¹⁷

Essa demanda futura acende o alerta para o tratamento das áreas já utilizadas, de modo que se atinja a produtividade necessária e se atendam os objetivos do desenvolvimento sustentável mundial.

1.4 Fitodoeças da cana-de-açúcar

Apesar de ser altamente eficiente na produção de açúcar, a cana-de-açúcar, assim como outras culturas, é alvo de vários patógenos, como vírus, fungos e bactérias, que causam prejuízos significativos, reduzindo a produtividade e, conseqüentemente, impactando o desenvolvimento do setor. Assim, são necessárias algumas práticas de controle e manejo para que esses prejuízos causados pelas fitodoeças sejam minimizados. Já foram identificadas mais

de cem doenças que atingem o cultivo da cana-de-açúcar, algumas causando danos substanciais. Entre elas, cinco são consideradas as mais importantes: i. carvão – causado pelo fungo *Ustilago scitaminea*; ii. ferrugem – causada pelos fungos *Puccinia melanocephala* e *Puccinia kuehnii*; iii. raquitismo das soqueiras – causado pela bactéria *Leifsonia xyli*; iv. mosaico – causado pelo vírus do mosaico; e v. escaldadura das folhas – causada pela bactéria *Xanthomonas albilineans*, sendo esta última o foco do presente projeto.¹⁸

1.5 Escaldadura das folhas

A “escaldadura das folhas” é um distúrbio vascular causada pela bactéria gram-negativa *Xanthomonas albilineans* (Ashby) Dowson em culturas de cana-de-açúcar (*Saccharum officinarum*). A doença está distribuída mundialmente e causa prejuízos econômicos expressivos devido à diminuição da produtividade, à necessidade de reforma precoce dos canaviais e à queda da qualidade do caldo extraído.¹⁹

O patógeno penetra na planta através de fissuras presentes no colmo e sua produção de albicidina age diretamente na enzima DNA girase (Topoisomerase II) dos cloroplastos, comprometendo o enovelamento dos ácidos nucleicos, e resultando em um déficit na geração de energia por parte do vegetal. Os sintomas e agressividade da doença dependem da fase de desenvolvimento do patógeno e das condições ambientais, sendo descritos três estágios: i. latente, caracterizada por ser assintomática; ii. crônica, caracterizada pelo aparecimento de estrias brancas paralelas por todo o limbo e que podem se estender até a bainha (fase semelhante ao raquitismo das soqueiras); iii. aguda, caracterizada pelo agravamento dos sintomas e aparência de folhas escaldadas em água quente, daí o surgimento do nome (**Figura 5**).^{20,21}

Figura 5 - Fases sintomáticas da doença escaldadura das folhas. Fase aguda caracterizada pela presença de estrias brancas da bainha ao limbo foliar. Fase crônica, marcada pelo embranquecimento completo das folhas das plantas infectadas.*

Fonte: Elaborada pelo autor.

Atualmente, as estratégias de controle da doença consistem na substituição de variedades suscetíveis (B69566) por resistentes (B8008). Assim, variedades de cana-de-açúcar com características agronômicas desejáveis e com alto potencial de produção, mas que sejam suscetíveis à doença, são frequentemente descartadas durante o processo de seleção de novas variedades comerciais. Esse cenário é agravado pelo aumento da agressividade do patógeno em condições de tempo seco e pela propensão em sofrer variações genéticas que favoreçam a quebra de resistência.

Esse grave panorama torna urgente o desenvolvimento de novas alternativas para o controle da escaldadura das folhas. Nesse sentido, a descoberta de novas moléculas bioativas como defensivos agrícolas pode contribuir significativamente para o controle e manejo eficaz da doença.

1.6 A bactéria *Xanthomonas albilineans*

O gênero *Xanthomonas* está classificado na divisão gama das Proteobacterias, o qual é constituído por agentes etiológicos que atingem plantas consideradas economicamente importantes. Podemos citar algumas culturas afetadas: Citrus spp. (limão, laranja, lima, entre outras), Oryzae spp. (arroz), crucíferas (repolho, brócolis, couve-flor), com algumas espécies

* Disponível em: www.agencia.cnptia.embrapa.br/gestor/cana-de-acucar/arvore/contag01_78_22122006154841.html. Acessado em: 25 de abril de 2019.

em alto grau de especificidade com o hospedeiro (e.g., *Xanthomonas albilineans*), totalizando 27 espécies com mais de 120 patovares (**Figura 6**).²²

As bactérias patogênicas podem ser distinguidas em dois grupos, de acordo com as características do genoma e os modos de obtenção de alimento: A. força bruta, as quais causam podridão e B. patógenos parasitas, os quais acarretam manchas foliares, necrose ou murcha. A estratégia bacteriana tipo A consiste em quebrar a parede celular do hospedeiro pela secreção de enzimas através do sistema de excreção tipo II (T2SS). Para o segundo caso, a estratégia é modular as funções moleculares do hospedeiro, geralmente do sistema imune, usando fatores que são diretamente injetados nas células através da secreção pelo sistema tipo III (T3SS).²³

No entanto, a bactéria *Xanthomonas albilineans* difere de ambos sistemas e pode ser encontrada dentro do hospedeiro nas diversas partes do feixe vascular (xilema e floema). Nesse caso, o organismo não possui maquinário para quebra da parede celular e similarmente não possui sistema de secreção do tipo III. Esse tipo específico de colonização, não classificado nos parâmetros supramencionados, pode ser justificado pela erosão do genoma e consecutivamente pela redução da artilharia responsável pela invasão e manutenção no hospedeiro.²³

As características únicas presentes em *Xanthomonas albilineans* a fazem especial dentro do seu gênero. A redução gênica, por exemplo, representa um genoma 25 % menor quando comparado às outras Xanthomonadaceae. Esta peculiaridade pode estar relacionada com a presença de duas regiões palindrômicas curtas agrupadas e regularmente interespaçadas (CRISPR – do inglês, *Clustered Regularly Interspaced Short Palindromic Repeats*) e que não aparecem nos demais. Ausentes, também, são os fatores de virulência associados à invasão e manutenção no hospedeiro, como a produção de biofilme, goma xantana e um sistema de secreção do tipo VI (T6SS).²⁴

Figura 6 - Cladograma das bactérias do gênero Xanthomonadaceae. As características da bactéria *Xanthomonas albilineans* se aproxima do gênero *Xylella*.

Fonte: Adaptado de PIERETTI²⁴.

No genoma de *Xanthomonas albilineans* é descrito a presença de enzimas específicas responsáveis pela degradação de parede celular e a síntese de policetídeos e peptídeos não ribossomais (PKS/NPRS).²⁵ É destacada, neste último, a produção de albicidina, molécula que inibe a DNA girase em concentrações nanomolares (**Figura 7**). Seu modo de ação difere de todos os inibidores conhecidos por afetar a subunidade A da DNA girase, atuando em bactérias gram-positivas e gram-negativas, dificultando a replicação celular. A deficiência de características invasoras presentes na família Xanthomonadaceae faz com que *X. albilineans* seja específica à colonização em xilemas de plantas da espécie *Saccharum officinarum*.²⁶

Albicidina é uma fitotoxina que bloqueia a enzima DNA girase dos cloroplastos, a qual é essencial para a replicação e regulação gênica na sua ação de modular as atividades de superrenovamento. A atividade desta molécula pode ser comparada a duas classes de moléculas com o mesmo modo de ação: coumarinas e quinolonas. No entanto, os compostos moleculares diferem no local de atuação, sendo que a fitotoxina produzida pela *Xanthomonas albilineans* atinge a subunidade GyrA, diretamente ligada na atividade catalítica de quebra e ligação, e coumarinas/quinolonas atacam a subunidade GyrB, responsável pela atividade ATPase. Adicionalmente, a albicidina, em concentrações nanomolares, atua como bactericida contra uma gama de microorganismos: *Enterobacter aerogenes*, *Escherichia coli*, *Haemophilus influenza*, *Klebsiella pneumonia*, *Shigella sonnei* e *Staphylococcus aureus*.²⁶

Figura 7 - Comparativo da afinidade dos inibidores de DNA girase (Topoisomerase II) GyrA. A molécula de simociclinona (magenta) apresenta valores de $IC_{50} = 1,45 \mu M$ contra *Staphylococcus aureus*. Para o mesmo alvo, a molécula de albicidina se mostra mais potente que outros inibidores descritos.

Fonte: Elaborada pelo autor.

A síntese de albicidina é baseada no misto da ação de síntese de policetídeos e peptídeos não ribossomais e consiste em seis fragmentos: ácido metil p-coumárico (MCA-1), dois ácidos 4-aminobenzóico (pABA-2 e pABA-4), a parte proteica não usual α -aminoácido β -ciano-L-alanina no centro da estrutura e dois ácidos 4-amino-2-hidroxi-3-metoxibenzoico (pMBA-5 e pMBA-6) na extremidade C-terminal (**Figura 8**).²⁷⁻²⁹

De outra parte, alguns organismos são resistentes à ação desta fitotoxina. Em 1999, Robert Birch, maior pesquisador sobre *Xanthomonas albilineans*, e colaboradores publicaram um artigo no periódico *Nature Biotechnology* informando a produção de cana-de-açúcar transgênica, a qual possuía o gene de resistência a albicidina - albD. Como resultado, plantas que carregam a enzima não apresentaram nenhuma doença nos cloroplastos, todavia, os exemplares sem o gene de detoxificação desenvolveram sintomas severos.³⁰

Figura 8 - Esquema ilustrativo da via de produção de albicidinas. O complexo de produção entre policetídeos e peptídeos não-ribossomais conferem a produção da fitotoxina. Esta molécula é composta por ácido metil p-coumárico (MCA-1), dois ácidos 4-aminobenzóico (pABA-2 e pABA-4), a parte proteica não usual α -aminoácido β -ciano-L-alanina no centro da estrutura e dois ácidos 4-amino-2-hidroxi-3-metoxibenzoico (pMBA-5 e pMBA-6) na extremidade C-terminal. A enzima benzoato coenzima A ligase tem papel fundamental no fornecimento de substratos.

Fonte: Adaptado de COCIANCH²⁶.

AlbD é um novo tipo de endopeptidase que catalisa a quebra da albicidina na ligação peptídica do ciano alanina. Atualmente, albD é o único exemplo da família de peptidases capaz de clivar uma ligação amídica entre os δ -aminoácidos, no qual o átomo de carbono- α do resíduo n é parte integral de um anel benzênico.³¹

Similarmente, mas de forma promíscua, a enzima albA de *Klebsiella oxytoca* trabalha na detoxificação e confere resistência ao organismo. A proteína carreadora de fármacos liga a fitotoxina para que não haja interação com o alvo específico do composto (Figura 9).²⁷

Figura 9 - Ligação do composto albicidina na enzima albA (PDB 6ET8). A ligação é apresentada com a proteína em cartoon e em superfície. O ligante ocupa toda a extensão proteica em mecanismo de detoxificação celular.

Fonte: Adaptado de ROSTOCK.²⁷

Portanto, os desafios apresentados em combater a bactéria gram-negativa apresenta sincronicamente o trunfo de compor novos agentes microbianos para o combate de doenças seculares, como *Staphylococcus aureus*. O aumento no número de organismos resistentes a antibióticos, aliás, tornou-se um dos maiores problemas de saúde pública. Nesse sentido, a descoberta de novas moléculas bioativas como defensivos agrícolas pode contribuir significativamente para o controle e manejo eficaz da doença, bem como para o desenvolvimento de alternativas para o tratamento de doenças causadas por bactérias gram-negativas.

1.7 Alvos Moleculares

O planejamento de novas moléculas bioativas baseia-se, principalmente, nas vias bioquímicas vitais do patógeno em estudo. A recente descrição completa do genoma de *X. albilineans* auxilia significativamente o estudo das rotas metabólicas essenciais e exclusivas da bactéria.³² Assim, selecionamos a via exclusiva de virulência e uma via essencial para o desenvolvimento da bactéria, ou seja, via de albicidina e via de folatos, respectivamente.

1.7.1 Biossíntese dos Folatos

Todas as células, procarióticas e eucarióticas, necessitam de folatos reduzidos para a biossíntese de diversos componentes celulares (e.g., síntese de purinas, timinas, aminoácidos, vitamina B5, dentre outros) (**Figura 10**). Folatos são fornecedores de grupos metil nas diversas reações de metilação diretamente relacionada à regulação gênica, e às sínteses de lipídeos, clorofila e lignina.³³

Enquanto o requerimento por folatos é universal, métodos para a sua obtenção diferem significativamente entre os organismos. Por exemplo, mamíferos possuem transporte ativo associados à membrana para sua obtenção e a maioria dos micro-organismos necessitam sintetizar via método *de novo*.^{34,35}

Figura 10 - Representação 2D da molécula de tetrahydrofolato. As caixas hachuradas indicam as três porções da molécula: anel pterina, ácido 4-aminobenzóico (pABA) e cauda glutamilada, que pode ser estendida (n). A formação do tetrahydrofolato é iniciada pelo fornecimento dos subprodutos HMDHP e pABA das vias GTP-ciclohidrolase e shikimato sintase, respectivamente. Dentro da mitocôndria, há a fosforilação do substrato HMDHP (pela enzima HPPK), com a subsequente transferência do produto HMDHP-P₂ ao pABA pela enzima DHPS. Especificamente, ocorre a liberação do pirofosfato e o diidropteróato. As enzimas em sucessoras utilizam a molécula de diidropteróato para a adição da cauda de ácido glutâmico (DHFS/FPGS) ou a redução do anel pterina (DHFR).

Fonte: Adaptada de GORELOVA³⁵

A via de produção está presente na maioria dos organismos patogênicos e difere em pontos específicos em eucariotos, o que torna a via dos folatos extremamente atrativa para o desenvolvimento de novos agentes antimicrobianos. Os substratos necessários para a produção do tetrahydrofolato são oriundos de vias associadas que ocorrem no citoplasma celular. São essenciais ao início da formação os produtos diidroneopterina e o ácido 4-aminobenzóico

(pABA), subprodutos da enzima GTP ciclohidrolase I (GCH) e via de síntese do shikimato, respectivamente.

1.7.1.1 Enzimas 6-hidroximetil-7,8-diidropterina pirofosfoquinase / Diidropteroato sintase

Dentro da mitocôndria a atuação da enzima HPPK (6-hidroximetil-7,8-diidropterina pirofosfoquinase), que fosforila o 6HMPH₂, prepara o substrato (DHPP) para o uso na enzima diidropteroato sintase (DHPS) (E.C. 2.7.6.3). A enzima DHPS (E.C. 2.5.1.15), por sua vez, catalisa a ligação do pABA ao DHPP formado pela HPPK, liberando pirofosfato e o diidropteroato (**Figura 11**).³⁶

Tanto a enzima HPPK quanto a DHPS se apresentam como monofuncionais em algumas bactérias, como por exemplo, em *Xanthomonas albilineans* (PDB 6DAY). No entanto, em alguns eucariotos (plantas), protozoários e fungos essas enzimas são comumente encontradas na forma bifuncional (e.g. *Saccharomyces cerevisiae*. PDB: 2BMB),³⁷ ligadas por meio de um conector proteico. É descrito que, em plantas, a enzima DHPS é retro inibida pelos substratos DHP, DHF e THF.³³

Em especial, a enzima diidropteroato sintase tem sido estudada como alvo para o desenvolvimento de novos fármacos (e.g., sulfonamidas), pois quando inibida impede o fornecimento do diidropteroato para produção do tetrahidrofolato.^{38,39} A descoberta das sulfonamidas foi um marco na história da medicina e do tratamento de infecções microbianas. Prontosil, desenvolvido pela Bayer em 1937, é um pró-fármaco que ao ser metabolizado libera a sulfanilamida. Foi o primeiro antibiótico a ser utilizado no tratamento de bactérias gram-positivas durante a Segunda Guerra Mundial.³⁶

Os índices de infecções bacterianas, após a introdução das sulfonamidas no mercado, diminuíram drasticamente, entre 20 e 80 %. Por outro lado, o uso de algumas sulfonamidas declinou instantaneamente devido ao enaltecimento de organismos resistentes, bem como o desenvolvimento de novos agentes antimicrobianos mais potentes, como exemplo o análogo sulfametoxazol. Atualmente, ainda, são administrados coquetéis contendo sulfametoxazol e trimetopim, os quais atuam de maneira sinérgica inibindo a enzima diidropteroato sintase e diidrofolato redutase (DHFR), respectivamente.⁴⁰

Figura 11 - Esquema representativo das reações catalisadas pelas enzimas diidropterina pirofosfoquinase (HPPK) e diidropteroato sintase (DHPS). A enzima HPPK que fosforila o 6HMPH2, prepara o substrato (DHPP) para o uso na enzima diidropteroato sintase (DHPS). Nesta última, a enzima DHPS catalisa a ligação do pABA ao DHPP formado pela HPPK, liberando pirofosfato e o diidropteroato

Fonte: Elaborada pelo autor.

Pela notória importância dentro da via de biossíntese de folatos, a enzima DHPS é extensivamente estudada. Como exemplo é possível acessar, no banco de dados proteico PDB, diversas entradas provenientes de diferentes organismos: *Bacillus anthracis*,⁴¹ *Burkholderia cenocepacia*,⁴² *Yersinia pestis*,⁴³ *Staphylococcus aureus*,⁴⁴ *Escherichia coli*,³⁶ dentre outros. Diversos são os estudos em desenvolvimento: mutações, complexos com substratos, fatores de resistência, análogos de inibidores e diferentes empregos de planejamento de novos fármacos.

Dada a ausência da DHPS em humanos, o estudo desta enzima se torna cada vez mais interessante do ponto de vista do planejamento de novos agentes antimicrobianos. Novos inibidores acarretarão a paralisação do funcionamento do sistema patogênico enquanto o hospedeiro não sofrerá com a ação do fármaco.

1.7.1.2 Enzima Diidrofolato redutase

Consecutivamente na via, a enzima diidrofolato redutase (*XaDHFR*, E.C. 1.5.1.3), primeira a reduzir o anel pterina, trabalha em sinergismo com a enzima DHPS, aumentando sua própria atividade na diminuição da produção do diidropteroato (**Figura 12**).⁴⁵ DHFR é uma enzima amplamente estudada a partir da perspectiva de função até da evolução nos organismos. A enzima catalisa a redução do 7,8-diidrofolato (DHF) para 5,6,7,8,-tetrahydrofolato (THF) na presença do cofator, o fosfato de dinucleotídeo de nicotinamida e adenina (NADPH).⁴⁶

Figura 13 - Reação bioquímica catalisada pela enzima benzoato coenzima A ligase (BCL). Com o consumo de ATP há a ligação do benzoato à coenzima A. A enzima funciona como carreadora dos substratos necessários à formação da albicidina.

Fonte: Elaborada pelo autor.

O gene *xabE* codifica 686 aminoácidos com 30-50 % de similaridade com a família de acil-coenzima A ligases na região do resíduo 240, incluindo um sítio de ligação AMP típico destas enzimas. A função enzimática desta proteína inclui o metabolismo de benzoato, aminobenzoato e hidroxibenzoato de várias fontes e organismos, os quais estejam envolvidos nos processos anaeróbicos e aeróbicos do metabolismo de benzoato e síntese de metabólitos secundários. Além disso, há estudos que indicam que o gene é essencial para a produção da albicidina.⁵⁰

Por fim, a descoberta e planejamento de inibidores seletivos para as enzimas *XaBCL*, *XaDHPS* e *XaDHFR* impediria a produção de albicidina e o crescimento de *X. albilineans*, respectivamente, levando a uma condição metabólica na qual o patógeno seria incapaz de se desenvolver e, conseqüentemente, causar prejuízos às lavouras de cana-de-açúcar.

OBJETIVOS

2 Objetivos

O objetivo central desta tese de Doutorado foi a descoberta de moléculas bioativas candidatas à novos agroquímicos para o combate da escaldadura das folhas. Aplicamos no processo de descoberta a Triagem Biológica Fenotípica, Biologia Estrutural e Química Medicinal. Portanto, estabelecemos os seguintes objetivos específicos:

- Clonagem das regiões abertas de leitura (ORF) dos segmentos gênicos específicos responsáveis pelas enzimas chave: *folP*, *folA* e *xabE*;
- Ensaios de expressão e purificação proteica em sistema bacteriano;
- Padronização dos ensaios de atividade enzimática e ensaios para a determinação de parâmetros cinéticos e de inibição;
- Triagem biológica de candidatos a inibidores;
- Determinação da propriedade biológica dos inibidores identificados nos ensaios enzimáticos;
- Cristalização na forma nativa e co-cristalização com os inibidores identificados;
- Coleta de dados cristalográficos e determinação da estrutura 3D das enzimas na forma nativa e em complexos com inibidores;
- Triagem de biblioteca de fragmentos moleculares para mapear regiões de ligação na molécula alvo;
- Planejamento baseado na estrutura do alvo molecular (SBDD) e do fragmento (FBDD);
- Testes com análogos dos *hits* iniciais e determinação da relação entre a estrutura e atividade das séries mais promissoras;
- Triagem biológica de inibidores do crescimento de *X. albilineans*.

MATERIAL E MÉTODOS

3 Material e Métodos

3.1 Isolamento e Cultura de *Xanthomonas albilineans*

Culturas isoladas de *X. albilineans* (Xa3, Xa5, Xa11 e Xa37) foram gentilmente cedidas pela Dra. Silvana Aparecida Creste Dias de Souza, pesquisadora do Centro de Cana - Instituto Agronômico de Campinas. As cepas diferem conforme a agressividade. Por conseguinte, selecionamos a cepa de maior grau patogênico para cultivo e realização de estudos.

Uma colônia única foi cultivada em meio líquido XAS (do inglês, *Xanthomonas albilineans selective media*).⁵¹ Suplementamos o caldo com os antimicrobianos: ciclohexamida (1 mg/mL), cefalexina (0,25 mg/mL), novobiocina (0,30 mg/mL) e kasugamicina (0,5 mg/mL) após esterilização em autoclave. Após incubação por 7 dias a 28 °C o cultivo foi semeado em placas de Petri contendo meio XAS ágar 1,5 %.

O passo anterior foi repetido com o intuito de produzir as amostras de trabalho e estocar as alíquotas mãe. Para tanto, o novo cultivo bacteriano foi dividido em três alíquotas, sendo uma delas direcionada para a produção de DNA genômico, uma para triagem fenotípica de inibidores e a última utilizada para produção de estoque glicerinado 25 % glicerol, resfriado em nitrogênio líquido e armazenado em ultra freezer (-80 °C).

3.2 Extração DNA genômico e Caracterização de *X. albilineans*

As células da cepa isolada (Xa11) de *X. albilineans*, cultivada em meio líquido XAS a 28 °C até a saturação, foram obtidas por centrifugação a 4.000 x g durante 20 min. Em seguida, o aglomerado celular contendo o conjunto bacteriano foi ressuspenso em tampão de lise CTAB (2 % brometo de cetil trimetil amônio, 1,4 M NaCl, 100 mM Tris-HCl pH 8 e 20 mM EDTA) pré-aquecido e mantido a 65 °C por 1 h. Subsequentemente, adicionou-se 1 volume de solução clorofórmio/álcool isoamílico (24:1) e as fases resultantes foram separadas por centrifugação (16.000 x g durante 10 min). Sobre a fase superior isolada, repetiu-se o processo anterior em dupla lavagem.

A precipitação dos ácidos nucleicos presentes na fase superior foi realizada pela adição de 1 volume de isopropanol (-20 °C) seguido de incubação a -20 °C por 30 min. A fração de DNA após centrifugação a 16.000 x g durante 10 min passou por duas etapas de lavagem, consecutivas, com etanol 70 % e 96 %, respectivamente, e seco à temperatura ambiente. O

material foi ressuspensão em água ultrapura e a confirmação realizada através de eletroforese em gel de agarose 0,8 %. O RNA residual foi removido da amostra pela digestão com 100 µg/mL RNase A (Invitrogen) por 1 h a 37 °C.

A amplificação das regiões de DNA específicas (**Tabela 1**) de *X. albilineans* foi realizada através da Reação em Cadeia da Polimerase (PCR, em inglês, *Polymerase Chain Reaction*).⁵² Os detalhes da ciclagem térmica utilizada estão apresentados na **Tabela 2**.

Tabela 1 - Oligonucleotídeos iniciadores utilizados em reação de PCR na confirmação do DNA genômico de *X. albilineans*. As regiões amplificadas são únicas nesta bactéria gram-negativa.

Gene, módulo	Senso 5' → 3'	Antisenso 5' → 3'
PGBL	CTTTGGGTCTGTAGCTCAGG	GCCTCAAGGTCATATTCAGC
albl, NRPS-1/3	TATGCGCAGCTGTACAACTG	AGCACGTATACCTCGGTGTT
albl, NRPS-2	TTCTGCCGAAGTGGCGAATA	GGCTGTGTGATCCACAATTC
alblV, NRPS-5	ATTGCATCTATACCTCCGGC	GGCTTGTTACATAGCCGAT

Fonte: Elaborada pelo autor.

A enzima Taq DNA polimerase de alta fidelidade (Cellco) foi utilizada para a amplificação, em ciclagem térmica usando o DNA genômico extraído como molde. O produto da reação foi analisado através de eletroforese em gel de agarose 1 %.

Tabela 2 - Parâmetros utilizados na reação de PCR para amplificação dos genes específicos de *X. albilineans*.

Reagente	Concentração Final	Ciclagem térmica
Tampão Taq HF 10X	1X	
dNTP	0,2 mM	95 °C – 30 s
Oligonucleotídeo senso	0,5 µM	95 °C – 30 s
Oligonucleotídeo anti-senso	0,5 µM	55 °C – 30 s
Molde	DNA genômico 50-250 ng	72 °C – 45 s
	DNA plasmidial 1 - 100 ng	
Taq DNA polimerase	1 U	72 °C – 10 min
H ₂ O	qsp. 50 µL	

Fonte: Elaborada pelo autor.

3.3 Amplificação e clonagem dos segmentos gênicos *folA*, *folK*, *folP*

Os fragmentos gênicos codificantes das enzimas *XaDHPS* (*folP*), *XaDHFR* (*folA*) e *XaHPPK* (*folK*) foram amplificados a partir de DNA genômico. A estratégia de clonagem adotada compõe o sistema de clonagem independente de ligação (LIC, em inglês, *ligation-independent cloning*). O sistema baseia-se na atividade exonuclease da enzima T4 DNA polimerase (3' → 5') para criar extremidades coesivas nos vetores após remoção das regiões

originais eGFP. Os produtos de PCR, gerados com oligonucleotídeos iniciadores (em inglês, *primer*) específicos (**Tabela 3**), formam as extremidades complementares no inserto e, assim, os fragmentos se unem independentemente da enzima DNA ligase em reação de hibridização. Esse sistema alternativo fornece a possibilidade de trabalhar com vetores em paralelo com um mesmo par de oligonucleotídeos iniciadores (e.g., pETM-11, pETTrx-1a e pETNus-1a).⁵³

Tabela 3 - Pares senso (Fw) e antisenso (Rv) 5' → 3' dos oligonucleotídeos iniciadores utilizados nas amplificações das ORF *folA*, *folK* e *folP*.

Enzima	Oligonucleotídeos
<i>folA</i>	CAGGGCGCCAT GGTCGCGTTGACGC (Fw)
	GACCCGACGCGGT TACACGGCGCTCGG (Rv)
<i>folK</i>	CAGGGCGCCAT GACCACCGTGCTACTGA (Fw)
	GACCCGACGCGGT TACCCTGGAGTCCCGG (Rv)
<i>folP</i>	CAGGGCGCCAT GTTCGATACCTCGCCTC (Fw)
	GACCCGACGCGGT AGTCCTCGTCCGGCC (Rv)

Fonte: Elaborada pelo autor.

A síntese do par (senso e antisenso) de cada oligonucleotídeos iniciador foi adquirida da empresa Extend® respeitando os seguintes termos protocolais:

- Senso (5' → 3') **CAGGGCGCCA**TG-gene de interesse
- Antisenso (5' → 3') **GACCCGACGCGGT**TA-gene de interesse (reverso complementar)

A preparação do inserto foi feita através de reação de PCR. Especificamente, para o volume contendo 50 µL utilizamos 5 µL tampão 10X Phusion, 0,5 µM de cada oligonucleotídeos iniciador senso e antisenso, 200 µM dNTP, 100 ng DNA genômico e 1U enzima Phusion DNA polimerase e água ultrapura q.s.p. Os reagentes foram transferidos para tubos 200 µL e o termociclador ajustado de acordo com os parâmetros descritos na **Tabela 4**. As condições ótimas de temperatura e ciclagem térmica foram utilizadas conforme sugeridas pelo fabricante. Os produtos das reações foram analisados em eletroforese em gel de agarose 0,8 % e a quantificação dos produtos da PCR realizada com NanoVue Plus (GE Healthcare) após a extração do material com o kit *Wizard® SV Gel and PCR Clean-Up System* (Promega).

Tabela 4 - Parâmetros utilizados para reação de PCR na amplificação dos fragmentos gênicos *folA*, *folK* e *folP*.

Reagente		Concentração Final	Ciclagem térmica
Tampão <i>Phusion</i> HF 10X		1X	
dNTP		0,2 mM	95 °C – 30 s
Oligonucleotídeo senso		0,5 µM	95 °C – 30 s } 55 °C – 30 s } 72 °C – 45 s } 35 X
Oligonucleotídeo anti-senso		0,5 µM	
Molde	DNA genômico	50 - 250 ng	
	DNA plasmidial	1 - 100 ng	
<i>Phusion</i> DNA polimerase		1 U	72 °C – 10 min
H ₂ O		qsp. 50 µL	

Fonte: Elaborada pelo autor.

A produção dos vetores desejados foi alcançada pela inserção dos plasmídeos originais LIC (pET-M11, pETTrx-1a e pETNus-1a) em cepas bacterianas *E. coli* DH-5α atingidas competência de transformação pelo método INOUE.⁵⁴ A propagação plasmidial foi feita com o cultivo de uma colônia carreadora do plasmídeo, isolada a 37 °C por 16 h em meio líquido Lysogenic broth (LB - 10 g.L⁻¹ triptona, 5 g.L⁻¹ extrato de levedura, 10 g.L⁻¹ NaCl) suplementado com canamicina 50 µg/mL.⁵⁵

O DNA plasmidial extraído, obtido pelo kit Wizard® *Plus Minipreps DNA Purification System* (Promega), foi digerido com a endonuclease BsaI (New England Biolabs) em reação (1 X New England Biolabs buffer 3, 25 U BsaI) para abertura do plasmídeo e remoção do sítio eGFP.

Os produtos das reações foram analisados em eletroforese em gel de agarose 0,8 % e os vetores digeridos foram recuperados do gel com o kit de extração supramencionado. Ao final do procedimento, vetores e insertos foram tratados, especificamente, com T4 DNA polimerase (Fermentas) em reações de 20 µL (2 µL - 10X Tampão T4 DNA polimerase, 1 µL - 5 mM DTT, 1U T4 DNA polimerase) adicionado de 2,5 mM dATP / 0,2 pmol insertos e 2,5 mM dTTP / 600 ng vetores por 30 min a 22 °C e 20 min a 75 °C para a formação das extremidades coesivas de ligação.

Os fragmentos tratados, contendo as sequências codificantes para as enzimas *XaDHFR*, *XaHPPK* e *XaDHPS*, foram hibridizados aos vetores na proporção 3:1, em temperatura ambiente por 30 min. As reações de hibridização foram utilizadas para a transformação em células competentes *E. coli* DH-5α.

Assim sendo, 100 µL de células foram incubados com a mistura reacional em gelo por 30 min. O choque térmico a 42 °C por 45 s seguido de manutenção em gelo por 5 min garantiu a incorporação dos plasmídeos desejados. Em seguida, adicionou-se 1 mL de meio líquido SOC

(20 g.L⁻¹ triptona, 5 g.L⁻¹ extrato de levedura, 0,5 g.L⁻¹ NaCl, 2,5 mM KCl, 10 mM MgCl₂, 20 mM glicose) e manteve-se por 50 min a 37 °C sob agitação constante de 200 rpm. Ao término, o cultivo foi semeado em placas de Petri contendo meio sólido LB suplementado com canamicina (50 µg/mL) e foi mantido em 37 °C por 16 h.

As colônias positivas foram identificadas via PCR de colônia, cujo protocolo reproduz o método de amplificação mencionado, alterando o molde para uma colônia isolada. O material positivo foi cultivado em meio líquido LB por 16 h a 37 °C, tendo plasmídeo recombinante obtido com o auxílio do kit comercial de extração. Todas as construções tiveram sua integridade verificada por sequenciamento e as colônias positivas foram mantidas em estoque glicerinado 25 % em freezer -80 °C.

3.4 Testes de Expressão Heteróloga

Colônias únicas foram cultivadas em paralelo a 37 °C em meio líquido LB e meio auto indutor ZYM-5052,⁵⁶ para cada proteína alvo, até atingirem a densidade óptica (D.O., em 600 nm) entre 0,6-0,8. Em seguida, induziu-se a expressão proteica com a adição de 0,4 e 1 mM de isopropil-β-D-galactosídeo (IPTG) para o meio LB. Dois protocolos distintos foram empregados para avaliação da condição de expressão: i. expressão por 3 h a 37 °C e ii. expressão por 20 h a 18 °C. Após o período de expressão, as células foram obtidas por centrifugação a 4.000 x g por 15 min e ressuspensas em tampão de lise (50 mM Tris pH 7,0, 100 mM NaCl e 10 mM imidazol) em um décimo do volume da cultura inicial, suplementado com 3 unid./mL DNaseI e 1 mg/mL lisozima.

As células foram submetidas à lise celular por congelamento e descongelamento em nitrogênio líquido e mantidas em gelo por 30 min. As frações solúvel e insolúvel foram separadas por centrifugação a 16.000 x g por 20 min e submetidas à análise de eletroforese em gel de poliacrilamida 10 % (SDS-PAGE), corado com 0,01 % de Coomassie Brilliant Blue R-250, para avaliação do perfil de expressão.

Com base nos resultados obtidos, novos protocolos foram empregados visando o máximo de produção na fração solúvel: i. alteração da faixa pH e concentração de agente indutor; ii. aditivos para o tampão de lise celular; iii. modificação da temperatura de expressão; iv. aditivos para o meio de expressão.

3.5 Purificação

Um concentrado celular foi obtido pela centrifugação do meio de expressão. O *pellet* resultante foi ressuspensão em solução-tampão para lise celular (50 mM Tris-HCl pH 8, 200 mM NaCl, 20 mM Imidazol e 10 % Glicerol), suplementado com 1 mM do inibidor de protease fluoreto de metilfenilsulfonil (PMSF) e 4 mM do agente redutor ditioneitol (DTT). As células foram lisadas por sonicação, com os debris celulares sendo separados da fração solúvel por centrifugação a 12.000 x g por 1 h.

A purificação consistiu em métodos cromatográficos utilizando o sistema semi automatizado AKTA™ (GE Healthcare) para cromatografia líquida de alta eficiência (CLAE), que incluíram quatro passos: i. cromatografia de afinidade a metal imobilizado (IMAC); ii. remoção da cauda hexa-histidina; iii. cromatografia de afinidade e iv. cromatografia por exclusão molecular. Na primeira etapa, foi explorado a cauda 6x-His para cromatografia de afinidade utilizando coluna HisTrap™ HP 5 mL (GE Healthcare) pré-carregada com Ni^{2+} (Figura 14).

Figura 14 - Esquema indicativo da interação entre a cauda hexa-histidina expressa conjuntamente com as proteínas de interesse e a coluna carregada com o metal níquel (Ni^{2+}) na cromatografia de afinidade. As proteínas e debris celulares sem afinidade pela coluna é eluído, restando apenas a proteína de interesse.

Fonte: Elaborada pelo autor.

Na etapa de afinidade, foram utilizados dois tampões: a. 50 mM Tris-HCl pH 8, 500 mM NaCl, 20 mM Imidazol e 10 % glicerol; b. 50 mM Tris-HCl pH 8, 500 mM NaCl, 500 mM Imidazol. Assim, na eluição formamos um gradiente imidazólico que compete pela ligação do metal imobilizado (Ni^{2+}). Em um segundo momento, todo conteúdo eluído em pico único da coluna contendo níquel foi injetado em coluna de *desalting* (HiLoad® XK 26/40 – Sephadex

g25 fine – GE Healthcare) com tampão 50 mM Tris-HCl, 200 mM NaCl, para retirada do excesso de imidazol, etapa indispensável para o pleno funcionamento da protease TEV. A clivagem da cauda hexa-histidina e da proteína de fusão das frações de interesse ocorreu pela incubação com 4 mM DTT e 10 % v/v TEV protease por 16 h.

Em seguida, a proteína livre foi obtida em fração pura através da segunda etapa cromatográfica de afinidade. Por fim, para garantir a elevada pureza das amostras, passamos o purificado por coluna de exclusão molecular (HiLoad® XK 26/100 – Superdex® 200 - GE Healthcare) a fim de garantir o isolamento (> 95 %), que foi confirmado em gel de eletroforese desnaturante 10 % (SDS-PAGE).

3.6 Triagem de Condições de Estabilidade Térmica

Utilizamos a técnica de fluorimetria de varredura diferencial – DSF (em inglês, *Differential Scanning Fluorimetry*) para triar condições promissoras de estabilização proteica.⁵⁷ A temperatura na qual ocorre a desnaturação é aferida com o incremento da fluorescência, uma vez que as porções hidrofóbicas são expostas ao solvente e facilitam a ligação de uma sonda fluorescente que possui afinidade por grupos hidrofóbicos.

Inicialmente foram montadas reações de triagem das condições de estabilidade frente ao DMSO (1, 2,5, 5, 10 e 20 %) e, em outro ensaio, as concentrações de sonda hidrofóbica (1X, 2,5X, 5X, 10X) que forneceriam os melhores valores sinal-ruído para a detecção no ensaio. Ao final do ensaio de calibração obtivemos o melhor padrão de reação para as diferentes aplicações que propusemos.

Brevemente, para uma reação de 25 µL, adicionamos 30 µM de proteína pura, 5x SYPRO Orange Protein e 12,5 µL de cada condição descrita na **Figura 15**, em placa de 96 poços. O monitoramento foi realizado em termociclador CFX Connect Real-Time PCR Detection System (Bio Rad) a cada grau, até a temperatura final 95 °C, iniciando à temperatura ambiente 25 °C. Os dados foram tratados seguindo recomendações do fabricante. Testamos a estabilidade térmica proteica em diferentes condições de soluções que englobaram pH, sais, tampões e glicerol (**Figura 15**). Os resultados desta triagem indicam condições de estabilidade proteica para os devidos fins de manutenção por longos períodos e a necessidade de íons ou outros agentes na estocagem do material.

	1	2	3	4	5	6	7	8	9	10	11	12
A	Citrate pH 4.0 100 mM	Citrate pH 4.0 100 mM NaCl	Citrate pH 4.0 250 mM NaCl	Citrate pH 4.0 500 mM NaCl	Citrate pH 4.0 100 mM KCl	Citrate pH 4.0 250 mM KCl	Citrate pH 4.0 500 mM KCl	Citrate pH 4.0 5% DMSO 200 mM NaCl	Citrate pH 4.0 10% DMSO 200 mM NaCl	Citrate pH 4.0 10 % glycerol 200 mM NaCl	Hepes pH 7.0 100 mM MgCl2	Hepes pH 7.0 200 mM L-Arg
B	Acetate pH 5.0 100 mM	Acetate pH 5.0 100 mM NaCl	Acetate pH 5.0 250 mM NaCl	Acetate pH 5.0 500 mM NaCl	Acetate pH 5.0 100 mM KCl	Acetate pH 5.0 250 mM KCl	Acetate pH 5.0 500 mM KCl	Acetate pH 5.0 5% DMSO 200 mM NaCl	Acetate pH 5.0 10% DMSO 200 mM NaCl	Acetate pH 5.0 10 % glycerol 200 mM NaCl	Hepes pH 7.0 250 mM MgCl2	Hepes pH 7.0 400 mM L-Arg
C	MES pH 6.0 100 mM	MES pH 6.0 100 mM NaCl	MES pH 6.0 250 mM NaCl	MES pH 6.0 500 mM NaCl	MES pH 6.0 100 mM KCl	MES pH 6.0 250 mM KCl	MES pH 6.0 500 mM KCl	MES pH 6.0 5% DMSO 200 mM NaCl	MES pH 6.0 10% DMSO 200 mM NaCl	MES pH 6.0 10 % glycerol 200 mM NaCl	Hepes pH 7.0 500 mM MgCl2	Hepes pH 7.0 50 mM Arg/Glu
D	Hepes pH 7.0 100 mM	Hepes pH 7.0 100 mM NaCl	Hepes pH 7.0 250 mM NaCl	Hepes pH 7.0 500 mM NaCl	Hepes pH 7.0 100 mM KCl	Hepes pH 7.0 250 mM KCl	Hepes pH 7.0 500 mM KCl	Hepes pH 7.0 5% DMSO 200 mM NaCl	Hepes pH 7.0 10% DMSO 200 mM NaCl	Hepes pH 7.0 10 % glycerol 200 mM NaCl	Hepes pH 7.0 100 mM NaAcetate	Hepes pH 7.0 100 mM NaAcetate
E	Phosphate pH 7.5 100 mM	Phosphate pH 7.5 100 mM NaCl	Phosphate pH 7.5 250 mM NaCl	Phosphate pH 7.5 500 mM NaCl	Phosphate pH 7.5 100 mM KCl	Phosphate pH 7.5 250 mM KCl	Phosphate pH 7.5 500 mM KCl	Phosphate pH 7.5 5% DMSO 200 mM NaCl	Phosphate pH 7.5 10% DMSO 200 mM NaCl	Phosphate pH 7.5 10 % glycerol 200 mM NaCl	Hepes pH 7.0 100 mM NaAcetate	Hepes pH 7.0 100 mM NaAcetate
F	Tris pH 8.0 100 mM	Tris pH 8.0 100 mM NaCl	Tris pH 8.0 250 mM NaCl	Tris pH 8.0 500 mM NaCl	Tris pH 8.0 100 mM KCl	Tris pH 8.0 250 mM KCl	Tris pH 8.0 500 mM KCl	Tris pH 8.0 5% DMSO 200 mM NaCl	Tris pH 8.0 10% DMSO 200 mM NaCl	Tris pH 8.0 10 % glycerol 200 mM NaCl	Hepes pH 7.0 100 mM (NH4)2SO4	Hepes pH 7.0 100 mM NaCitrate
G	CHES pH 9.0 100 mM	CHES pH 9.0 100 mM NaCl	CHES pH 9.0 250 mM NaCl	CHES pH 9.0 500 mM NaCl	CHES pH 9.0 100 mM KCl	CHES pH 9.0 250 mM KCl	CHES pH 9.0 500 mM KCl	CHES pH 9.0 5% DMSO 200 mM NaCl	CHES pH 9.0 10% DMSO 200 mM NaCl	CHES pH 9.0 10 % glycerol 200 mM NaCl	Hepes pH 7.0 100 mM NH4Acetate	Hepes pH 7.0 5% PEG 4,000
H	CAPS pH 10.0 100 mM	CAPS pH 10.0 100 mM NaCl	CAPS pH 10.0 250 mM NaCl	CAPS pH 10.0 500 mM NaCl	CAPS pH 10.0 100 mM KCl	CAPS pH 10.0 250 mM KCl	CAPS pH 10.0 500 mM KCl	CAPS pH 10.0 5% DMSO 200 mM NaCl	CAPS pH 10.0 10% DMSO 200 mM NaCl	CAPS pH 10.0 10 % glycerol 200 mM NaCl	Hepes pH 7.0 100 mM MgAcetate	Hepes pH 7.0 10% PEG 4000

Figura 15 - Quadro expositivo da configuração das placas de 96 poços com as condições triadas em experimento de estabilidade térmica.
Fonte: Elaborada pelo autor.

3.7 Identificação de ligantes por DSF

A diferença na temperatura entre a estabilidade térmica na presença e na ausência de ligante é relacionada com a afinidade do composto com o alvo molecular. Resumidamente, uma reação de 25 μ L contém 30 μ M de proteína pura, 5x SYPRO Orange Protein e 2 mM de cada fragmento, em placa de 96 poços. As medidas foram realizadas em duplicatas. O monitoramento foi realizado em termociclador CFX Connect Real-Time PCR Detection System (Bio Rad) a cada grau, por 70 min, iniciando à temperatura ambiente 25 °C. Os dados foram tratados seguindo recomendações do fabricante, considerando ligação o $\Delta T_m > 0,5$ °C.

3.8 Ensaios de Cristalização

Submetemos a proteína pura ao processo de triagem de condições de cristalização utilizando kits comerciais. As estratégias de matriz esparsa, grade, triagem primária e secundária foram empregadas com os seguintes kits comerciais: Crystal Screen, Classic, Classic II, PEG, PEG II, PEG III, JCSG Cor Suite I-IV. Os experimentos foram conduzidos a 18 °C e concentrações de 5 e 10 mg/mL de proteína. O método gota assentada (do inglês, *sitting drop*), variação da técnica de difusão de vapor, foi empregado em placa de 96 poços. As gotas foram inspecionadas a cada dois dias, partindo do primeiro dia de montagem das caixas.

As misturas que favoreceram a obtenção de cristais (dimensão e qualidade) foram expandidas utilizando-se a técnica de gota suspensa (do inglês, *hanging drop*) em placa de 24 poços na proporção 1:1. Os cristais foram retirados da solução de cristalização com auxílio de alça de nylon e imersos rapidamente em solução crioprotetora (solução matricial adicionada de 20 % etilenoglicol) anteriormente à difração de raios X.

Ensaios de semeadura (*seeding*) e imersão (*soaking*) foram conduzidos com o objetivo de obter cristais com tamanho adequados para os experimentos de difração de raios X e em complexo com os ligantes, respectivamente. Semeadura: cristais formados são submetidos ao agitador vortex em solução contendo mistura da solução mãe e 20 % etilenoglicol e, com auxílio de ferramenta específica (*seeding tool* - HAMPTON), introduzimos pequenas quantidades às gotas pré equilibradas. Após 24 h, novos cristais foram observados nas gotas frequentadas. Para a obtenção de cristais da XaDHPS em complexo com ligantes, mergulhamos os mesmos em solução contendo a condição matricial e um concentrado dos substratos pABA, 6HMPH2 e do

produto DHP, por 10 min. Em seguida, os cristais foram resfriados anteriormente ao experimento de difração de raios X.

3.9 Coletas de Dados

A estratégia de coleta foi traçada mediante a captura de algumas imagens por poucos segundos de exposição, em varredura ortogonal. Especificamente, com uma varredura do cristal avalia-se o seu padrão de difração, resolução máxima atingida, mosaicidade, grupo espacial e tempo de exposição necessário, com comprimento de onda $\lambda = 1,4587$ nm na linha MX-2 do Laboratório Nacional de Luz Síncrotron (LNLS – Campinas/SP). Foram obtidas imagens a cada 90° dos cristais, as quais foram processadas com a interface iMosflm.⁵⁸ Conforme as características de simetria, traçamos a estratégia de coleta baseada em *fine ϕ -slicing* que atende a coleta do maior número de imagens para um melhor detalhamento dos fótons captados pelo detector por uma pequena varredura angular. Submetemos a coleta total em 180° , obtendo uma imagem a cada $0,2^\circ$ com 1,2 s de exposição para cada imagem a uma distância de 120 mm (linha MX2).

Dadas as características dos sincrotrons de última geração, nossa estratégia de coleta foi alterada para atender as necessidades dos cristais e explorar ao máximo as linhas de luz das quais utilizamos (SSRL, Diamond e ESRF). Nas linhas de luz de terceira e quarta geração utilizamos o comprimento de onda $\lambda = 0,97$ Å, menor tempo de exposição e varredura de 360° a fim de se obter o melhor conjunto de dados possível.

3.10 Processamento, Resolução e Refinamento

Todas as imagens obtidas foram processadas manualmente utilizando o programa XDS⁵⁹, ou em rotinas de processamento automático com autoPROC⁶⁰. Para os conjuntos manualmente processados implementamos os tratamentos de escalonamento, definição do grupo pontual e corte de resolução com o programa Aimless.⁶¹ Os problemas das fases foram resolvidos por substituição molecular com PHASER,⁶² utilizando a proteína homóloga de *E. coli* (PDB 1AJ0).⁶² Utilizamos o conjunto de software Phenix⁶³ para refinamento com o protocolo *phenix.refine*. Em outra vertente, empregamos o refinamento no espaço recíproco com o programa BUSTER.⁶⁴ A qualidade dos modelos gerados foram avaliadas pelo servidor Molprobity⁶⁵, que aplica a inspeção dos diagramas de *Ramachandran*, impedimentos estéricos,

rotâmetros, dentre outros parâmetros. As etapas de refinamento em espaço real foi performeda com o programa Coot⁶⁶.

3.11 Termoforese em microescala (MST)

O experimento de termoforese em microescala (em inglês, *Microscale thermophoresis*) foi conduzido no equipamento Monolith NT115 (Nanotemper). Esta versão carece a marcação do alvo molecular com uma sonda fluorescente. Para tanto, reagimos a proteína *Xa*DHPS com a sonda reativa de amina primária (NT-647-NHS) e a sonda formadora de ligações covalente em cisteínas (MALEIMIDE). Foram utilizados 300 μ L de proteína a 20 μ M, inicialmente, para troca de solução tamponante anterior à marcação. Consecutivamente, a sonda liofilizada foi solubilizada em 100 % DMSO para concentração final 470 μ M. Incubamos a mistura da proteína com a sonda na proporção 5:1 de sonda e proteína, mantido por 1 h longe da exposição luminosa. O excesso de sonda foi removido em coluna específica de exclusão molecular, fornecida pela empresa (Nanotemper), tendo como resultante a proteína marcada singularmente. À coleta foram adicionados 10 % glicerol e esta mistura, após alíquotagem, foi mantida em freezer -80 °C.

Para os ensaios foram acondicionados 100 nM *Xa*DHPS-RED-NHS, tampão 50 mM fosfato de sódio pH 7, 200 mM NaCl, 10 mM MgCl₂ e 0,05 % Tween-20. Em outra frente, empregamos 10 nM *Xa*DHPS-MALEIMIDE, tampão 50 mM HEPES pH 7,2, 200 mM NaCl, e 0,05 % Tween-20. Foram testados compostos solubilizados em 100 % DMSO, com concentração variável e dependente da solubilidade em solução aquosa resultante da mistura com a proteína. Os capilares utilizados possuíam revestimento padrão e não apresentaram dificuldades para a realização do experimento.

3.12 Calorimetria de Titulação Isotérmica (ITC)

Os experimentos foram realizados em equipamento Microcal VP-ITC (em inglês, *Isothermal Titration Calorimetry*). As propriedades termodinâmicas da enzima pura *Xa*DHPS na concentração 30 μ M foram analisadas quando titulados os tampões utilizados na purificação proteica (calibração), o solvente DMSO (estabilidade) e os ligantes (teste de afinidade). Realizamos o experimento com o ligante ácido diidropteróico (DHP) que foi diluído na concentração 500 μ M em tampão teste. Os dados foram analisados com o software ORIGIN® v9.6 seguindo recomendação do fabricante.

3.13 Cinética enzimática *XaDHPS*, *XaDHFR*, *XaHPPK*

Os parâmetros cinéticos e de atividade são aferidos via ensaio colorimétrico (absorbância) através da taxa de liberação de fosfato inorgânico no meio reacional. A enzima *XaDHPS* catalisa a conversão do ácido 4-aminobenzóico (pABA) e 6-hidroxi-metil-7,8-diidropterina pirofosfato (DHPP) em diidropteroato (DHP) (**Figura 11**). Nesta reação há liberação do pirofosfato que é convertido em dois íons fosfatos por uma pirofosfoquinase inorgânica de levedura (Sigma-Aldrich). Os fosfatos inorgânicos são, então, detectados utilizando o kit PiColorlock Gold (Innova Biosciences) medindo absorbância em leitor de placa Spectramax Plus 384.

A mistura reacional consistiu em: 50 mM Tris-HCl pH 8, 10 mM MgCl₂, 2 % DMSO, 20 µM pABA, 20 µM DHPP e 5 nM *XaDHPS*. Para a conversão em fosfatos inorgânicos adicionamos 0,01 U pirofosfoquinase inorgânica. O volume final foi de 100 µL e a reação foi incubada a 37 °C por 30 min anteriormente a leitura.

3.14 Triagem Fenotípica de Compostos

Os compostos sintéticos avaliados neste trabalho foram fornecidos pelo Prof. Dr. Henri Stephan Schrekker (UFRGS). No total, a capacidade inibitória de 21 derivados imidazólicos foram adicionadas no crescimento de culturas de *X. albilineans* (**Figura 16**). De maneira geral, a base do set testado foi construída de maneira aleatória mediante a disponibilidade de compostos no laboratório do docente. Brevemente, os compostos foram diluídos em 100% dimetilsulfóxido (DMSO) na concentração final 200 mM, sendo os ensaios conduzidos utilizando a incubadora/leitor de placa/agitador Bioscreen C. Os compostos foram dispostos, em meio XAS, em placas de 100 poços e distribuídos em diluições seriada de 2 mM a 250 µM, com as respectivas duplicatas.

Monitoramos a densidade óptica ($\lambda = 600$ nm) do crescimento bacteriano por 4 dias, com leituras realizadas a cada 10 min. Os dados foram ajustados pelo método de regressão não linear com o programa GraphPad Prism 6, baseado na equação de Gompertz, para cálculos da mínima concentração inibitória (MIC, do inglês, *minimal inhibitory concentration*).

Figura 16 - Estrutura 2D da série de compostos testada frente a cultura de *Xanthomonas albilineans*.
Fonte: Elaborada pelo autor.

Por conseguinte, após os resultados obtidos com o primeiro conjunto de moléculas, novos compostos foram submetidos ao teste de atividade contra a bactéria *X. albilineans* (**Figura 17**).

A atividade inibitória para cada concentração avaliada foi determinada através da área fracional (Af) do crescimento, que corresponde à área sob a curva do crescimento na presença do composto testado (At) dividido pela área sob a curva do crescimento do controle negativo (Ac). Este índice varia entre 0 e 1, sendo 1: o composto testado não apresenta nenhum efeito no crescimento bacteriano e 0: representa inibição completa do crescimento.

Figura 17 - Compostos testados contra o crescimento de *Xanthomonas albilineans*. A segunda série aumenta a cadeia carbônica dos derivados imidazólicos. Em alguns casos há a inclusão de um segundo anel imidazólico (33), ou a troca do imidazol por uma piridina (31). Os compostos foram diluídos em 100 % DMSO.

Fonte: Elaborada pelo autor.

A porcentagem de inibição foi calculada pela **Equação 1** para cada composto avaliado. Valores de potência (EC₅₀, do inglês, *effective concentration*) serão determinados através da concentração na qual o composto induz 50% do efeito máximo no crescimento celular.

$$\% \text{ Inibição} = \left(1 - \frac{A_t}{A_c}\right) \times 100 \quad (1)$$

Xanthomonas albilineans

4 Resultados e Discussão (*Xanthomonas albilineans*)

4.1 Isolamento, Cultura e Caracterização de *Xanthomonas albilineans*

A bactéria gram-negativa *Xanthomonas albilineans* possui características específicas de cultivo *in vitro*: i. crescimento lento (4-7 dias), ii. coloração amarelada não mucoidal e iii. temperatura ótima 28 °C de incubação e iv. meio seletivo suplementado com antibióticos e fungicidas.

Todas as nossas cepas foram adquiridas, inicialmente, do Instituto Biológico de Campinas (IBSBF1374, 740, 326 e 654), recuperadas e submetidas a várias tentativas de isolamento sem sucesso. Como as observações sobre o crescimento bacteriano em estudo não condiziam com o perfil de crescimento reportado na literatura (e.g. saturação no crescimento a 37 °C em 16 h), realizou-se vários testes que identificaram a presença de contaminantes nas amostras enviadas.

Frente a este problema, contatamos a Dra. Silvana Aparecida Creste Dias de Souza, pesquisadora do Instituto Agrônomo de Campinas (IAC). Cepas coletadas diretamente de canaviais infectados, isoladas e administradas por ela nos foram gentilmente fornecidas. As amostras diferem na taxa de agressividade observada em culturas de cana-de-açúcar. Todos os exemplares foram recuperados em meio XAS, suplementado com os antibióticos de seleção. As novas cepas apresentaram taxa de crescimento e características morfológicas compatíveis àquelas descritas.

Para caracterização molecular das cepas, utilizamos a amplificação por PCR da região ribossomal específica. O método identificou um fragmento gênico único de *X. albilineans* de 288 pares de base (bp, do inglês base pairs) localizado na região espaçadora entre as subunidades ribossomais 16S-23S. Concomitantemente, amplificamos outras regiões específicas de proteínas não ribossomais exclusivas da bactéria (NPRS - 1/3, 2 e 5) (**Figura 18**). Por consequência, houve confirmação de que as cepas em estudo são *X. albilineans*.

Figura 18 - Gel 0,8 % agarose contendo resultado da amplificação da região específica PGBL e das proteínas não ribossomais NRPS1/3,2,5 de *Xanthomonas albilineans*. MK = marcador de peso molecular.
Fonte: Elaborada pelo autor.

4.2 Triagem de Compostos

Imidazol é uma molécula planar, altamente polar, anfótero, composta de um anel de cinco membros de três átomos de carbono e dois de nitrogênios. As características multifacetadas são ocasionadas pelas diferentes possibilidades de interações: interação de hidrogênio, interação íon-dipolo, cátion- π , π - π *stacking*, efeitos hidrofóbicos, forças de van der Waals. Alguns derivados são de uso farmacológico nas múltiplas frentes: antifúngicos e antibacterianos, antivirais, antidiabéticos e antimaláricos (**Figura 19**).⁶⁷

Figura 19 - Representação 2D da molécula de imidazol. As possíveis substituições estão indicadas em R1-5.
Um exemplo de molécula que possa atuar como antibacteriano (Metronidazol).
Fonte: Elaborada pelo autor.

Apesar da alta solubilidade em soluções aquosas, quando adicionadas longas cadeias alifáticas os derivados imidazólicos tendem a diminuir a solubilidade em água. Assim, todos os

compostos foram diluídos em 100 % DMSO. Dada a não solubilização em 200 mM, os compostos foram acondicionados em concentrações menores para que o ensaio fosse realizado. Por conseguinte, a concentração máxima testada neste caso específico foi menor para que atendesse às normas protocolares. Todos os compostos indicados na **Figura 16** foram submetidos ao ensaio de inibição. Os resultados foram divididos em dois grupos conforme a atividade inibitória observada (**Figura 20**). Sendo:

- **Grupo 1:** Inibição > 90 % na concentração 125 μ M
- **Grupo 2:** Inibição < 50 % na concentração 2 mM

Figura 20 - Esquema ilustrativo do resultado do teste de crescimento bacteriano de *Xanthomonas albilineans*, em placas de 100 poços, na presença de ligantes. As placas em colmeia pertencem ao equipamento Bioscreener C MBR.

Fonte: Elaborada pelo autor.

Os valores de potência são estimados conforme a atividade na faixa de concentração resposta dos compostos testados (2 mM – 2,5 nM). Os gráficos, como exemplo demonstrado na **Figura 21**, mostram as curvas de crescimento pelas medidas de densidade óptica $\lambda = 600$ nm durante os quatro dias de experimento. Após tratamento dos dados, normalização e cálculo da área sob a curva, foi possível determinar a atividade inibitória a partir da concentração de 125 μ M. Escolhemos esta faixa de limite mínimo de atividade como filtro e que corrobora com a concentração controle da tetraciclina.

Figura 21 - Exemplo de resultados da atividade dos compostos com efeito no crescimento bacteriano da *Xanthomonas albilineans*, em diferentes concentrações. O cálculo dos valores destas concentrações efetivas é feito após titulação em diluição seriada do ligante que vão desde 2 mM até 2,5 nM.

Fonte: Elaborada pelo autor.

Os resultados da primeira triagem estão organizados na **Figura 22**. Todos os compostos que tiveram atividade inibitória satisfatória foram alocados no grupo em verde e, por contraste, a baixa atividade contra o crescimento bacteriano foi mantida no grupo vermelho. No total, 12 compostos do grupo verde foram selecionados para determinação da potência.

Figura 22 - Resultado da triagem fenotípica de compostos imidazólicos frente a cultura de *Xanthomonas albilineans*. Os compostos organizados no grupo em verde causaram inibição de 100 % do crescimento bacteriano à 125 μ M. Os compostos no grupo em vermelho não influenciaram em taxas maiores que 50 % no crescimento bacteriano à 2 mM.

Fonte: Elaborada pelo autor.

Com o primeiro resultado de inibição em mãos, novos derivados foram enviados para teste. Na nova série, há um aumento na cadeia alifática e alteração do contra íon na formação salina. As novas moléculas estão representadas na **Figura 17**. Os resultados foram novamente divididos em dois grupos conforme a atividade inibitória observada, tendo o valor corte escolhido quando a atividade inibitória de 50 % na concentração 125 μM .

Problemas de realização dos experimentos foram vivenciados quando testada a solubilidade em 100 % DMSO nos compostos com dupla cadeia alifática (**32**), ou dupla inserção de anel imidazólico na molécula (**33**). Portanto os testes de atividade nestas moléculas ficaram prejudicados (e.g. os compostos com dupla cadeia alifática de 16 membros).

As moléculas **32** e **36** não foram ensaiados por não se apresentarem solúveis em 15 μM 100 % DMSO. Contudo, o experimento com os compostos restantes foi realizado com sucesso e as porcentagens de inibição são modestas, porém totalmente consideráveis para o planejamento de novas moléculas.

Os valores de EC_{50} (**Figura 23**) quantificaram a potência dos compostos que diminuem o crescimento de *X. albilineans*. Foi possível observar que a adição grupos metil **18** ($\text{EC}_{50} = 94 \pm 6 \mu\text{M}$) $\rightarrow$ **30** ($\text{EC}_{50} = 4,4 \pm 2 \mu\text{M}$) $\rightarrow$ **34** ($\text{EC}_{50} = 8,4 \pm 2 \mu\text{M}$) ao anel imidazol aumentaram a faixa de inibição. Por outro lado, a mudança no íon cloreto Cl^- (**15** - $\text{EC}_{50} = 16 \pm 2 \mu\text{M}$) para brometo Br^- (**25** - $\text{EC}_{50} = 29 \pm 4 \mu\text{M}$) diminuem a potência inibitória. A variação no anel imidazólico (**15**, $\text{EC}_{50} = 16 \pm 2 \mu\text{M}$) para piridínico (**31**, $\text{EC}_{50} = 49 \pm 4 \mu\text{M}$) carbonos também reduzem os efeitos causados no crescimento (**Figura 23**).

Figura 23 - Representação das curvas de EC_{50} obtidas no ensaio de triagem fenotípica de compostos. Foram utilizadas 20 concentrações diferentes para a determinação dos valores de concentração efetiva. Os valores de EC_{50} determinados estão apresentados em cada estrutura bidimensional dos compostos de todas as séries testadas.

Fonte: Elaborada pelo autor.

O processo de planejamento de novos inibidores passa pela etapa de busca por novos esqueletos moleculares e que podem ser caracterizados frente ao alvo molecular ou de maneira fenotípica. Atualmente, não existem compostos comerciais que atuem como inibidores do crescimento de *Xanthomonas albilineans*. Os derivados imidazólicos testados contribuem essencialmente na busca por estas moléculas.

Podemos agrupar os resultados das 41 moléculas testadas em suas respectivas atividades e potência. Alguns compostos tiveram modesta atividade inibitória, mas são necessários para definirmos a relação estrutura-atividade (SAR) desta série e, investigarmos modificações moleculares favoráveis para o aumento da potência inibitória. Dadas as características moleculares dos derivados imidazólicos, sugerimos que estas possam exercer sua atividade inibitória em membrana plasmática, no entanto, o alvo molecular não está elucidado.

A atividade de compostos derivados imidazólicos foi observada nos organismos *Staphylococcus aureus*, *Bacillus subtilis*, *Escherichia coli*, *Pseudomonas aeruginosa*, entre outros. No entanto, não encontramos descrição de atividade dessas moléculas em *Xanthomonas albilineans*. Os compostos com atividade bactericida diferem dos aqui testados, pois não apresentam cadeia carbônica entre 4-18 átomos, são cíclicos e com outros grupos substituintes que inferimos diferenciar no alvo molecular de atuação. Moléculas maiores, com atividade antibacteriana, aparecem conjugadas a peptídeos.⁶⁸ A não-descrição da atividade biológica dessa série em outros microrganismos aumenta nosso interesse por essas moléculas.

VIA DE FOLATOS

5 Resultados e Discussão

5.1 Amplificação e Clonagem

Os resultados obtidos para a proteína da via de produção de albicidina (*XaBCL*) será abordado no **Apêndice I**. Os fragmentos gênicos *folA* e *folP* foram amplificados utilizando técnica de PCR (**Figura 24**). Os oligonucleotídeos iniciadores específicos foram planejados para expressão em vetores pET-LIC em *E. coli*. Utilizamos, nas etapas de clonagens, 3 vetores: i. cauda 6x-His (pET-M11); ii. cauda 6x-His + tiorredoxina (pETTrx-1a); iii. cauda 6x-His + NusA (pETNus-1a). O objetivo da adição das proteínas de fusão tiorredoxina e NusA foi aumentar a solubilidade. Essas construções também auxiliam o estágio de purificação, uma vez que exploramos a afinidade da cauda 6x-His ao metal imobilizado presente em nossas colunas cromatográficas de afinidade. Os sítios marcadores são removidos com auxílio da enzima TEV (em inglês, *Tobacco Etch Virus*), uma cisteína protease, produzida de maneira recombinante em nosso laboratório.

Figura 24 - Amplificação gênica das ORF *XafolP* e *XafolA* resultante dos experimentos de PCR de colônia. A amplificação destas regiões confirma que os vetores trabalhados contém os genes inseridos.

Fonte: Elaborada pelo autor.

5.2 Sequenciamento

Exploramos a instalação de sequenciamento disponível no Departamento de Bioquímica da Universidade de Cambridge para mapear todas as sequências utilizadas nesta tese. A infraestrutura disponível emprega a nova geração de sequenciamento (em inglês, *Next Generation Sequencing – NGS*) com três Illumina MiSeqs e um NextSeq500. Com isso, é possível realizar a identificação de regiões gênicas da noite para o dia e com alta precisão. Em

Mensuramos a porcentagem de códons raros pelo Índice de Adaptação de Códons (em inglês, *Codon Adaptation Index*). Os valores vão de 0-1, sendo considerados ótimo os números entre 0,8-1, que se referem à porcentagem de presença da trinca de nucleotídeo não usual (**Figura 26**). Os genes trabalhados representam, também, a alta porcentagem de bases G/C (~70%), a qual é característica intrínseca do organismo como um todo. O alto conteúdo destas bases ($\uparrow 70\%$) afeta, inevitavelmente, a expressão proteica pela transcrição e tradução das ORF's devido à alta estabilidade da ligação e/ou formação de estruturas secundárias no RNA mensageiro (mRNA).

Resíduo	Códon	folP	folK	folA
Arginina	CGA	2	0	3
	CGG	3	3	5
	AGG	0	0	0
	AGA	0	0	0
Glicina	GGA	1	3	2
	GGG	2	4	3
Isoleucina	AUA	0	1	0
Leucina	CUA	0	2	1
Prolina	CCC	3	1	0
Treonina	ACG	7	3	4
IAC		0,75	0,70	0,68

Xanthomonas albilineans

Figura 26 - Disposição dos códons raros presentes nos segmentos gênicos *folA*, *folK* e *folP*.
Fonte: Elaborada pelo autor.

5.3 Expressão e Purificação

Protocolos similares para as enzimas investigadas em paralelo (*XaDHFR*, *XaHPPK* e *XaDHPS*) foram empregados. Todas as construções foram transformadas em cepas de expressão padrão *E. coli* BL-21 (DE3) e Rosetta 2 (DE3). Avaliamos o impacto da temperatura (18 °C e 37 °C) na expressão proteica em todas as construções.

Os meios LB, 2xYT ou auto indutor ZYM-5052 foram produzidos em faixa neutra de pH = 7. Esse tipo de expressão bacteriana (pET) baseia-se na atividade da enzima T7 RNA polimerase, no operon lac, para produzir proteínas recombinantes. Quando utilizado o meio LB ou 2xYT houve adição do agente indutor de expressão isopropil β -D-1-tiogalactopiranosídeo (IPTG), molécula mimética à alolactose. Para ambos os casos (IPTG ou autoindução), há a interação da alolactose com o repressor lac através de ligação de hidrogênio com os resíduos Asp149 e Ser193, o que impede o acesso ao operador, consequentemente, permitindo a

produção proteica. Os avanços obtidos para cada um dos alvos moleculares selecionados estão detalhados nas seções seguintes:

5.3.1 Diidropteroato sintase *XaDHPS*

Todas as construções permitiram a expressão recombinante da proteína alvo, contudo, somente para os vetores pETTrx-1a e pETNus-1a houve manutenção da mesma na fração solúvel (**Figura 27**). Diante disso, selecionamos as construções em pETTrx-1a e pETNus-1a para escalonamento e estudos de purificação.

Figura 27 – **A.** Gel 10 % SDS-PAGE referente a avaliação do perfil de expressão proteico *XaDHPS* nos vetores do sistema LIC. **B.** Purificação da enzima *XaDHPS* (31,8 kDa). S = fração solúvel, I = fração insolúvel, TEV = clivagem com a enzima TEV protease, GF = filtração em gel, por exclusão molecular e MK = marcador de peso molecular.

Fonte: Elaborada pelo autor.

A primeira etapa de purificação por cromatografia de afinidade ocorreu com sucesso, liberando proteínas e debris celulares não específicos à coluna. O gradiente imidazólico desvencilha a 6xHis-Trx-DHPS da coluna entre as concentrações 40-50 %, o que corresponde a 200-250 mM Imidazol. Os testes de remoção da proteína de fusão Nus-1a indicaram que não houve clivagem. Sugerimos que o problema ocorra pela obstrução do sítio de reconhecimento da TEV protease, que ficaria escondido entre as duas proteínas.

Por outro lado, observamos a separação da *XaDHPS* da proteína de fusão tiorredoxina em gel 10 % SDS (**Figura 27**). Diante disso, uma nova injeção em coluna de cromatografia por afinidade foi utilizada para separar as proteínas 6xHis-TEV e 6xHis-Trx da proteína de interesse. Por fim, a *XaDHPS* foi submetida a uma etapa adicional de purificação através de cromatografia de exclusão molecular (**Figura 28**). Os rendimentos finais foram de 30 mg/L a partir do meio ZYM-5052, 15 mg/L meio LB e 5 mg/L meio 2xYT.

O crescimento celular em meio auto indutor ocorre com o consumo do glicerol como fonte de carbono, glicose como fonte energética, e com a suplementação com metais traço até que ocorra o crescimento celular e atinja o processamento da lactose para iniciar a expressão proteica. O baixo e diferente rendimento entre a implementação dos meios de cultura, nas proteínas da via dos folatos, está na formação de debris celulares e no aumento da proteína de interesse na fração insolúvel.

Figura 28 - Cromatogramas de purificação da enzima *XaDHPS* nas diversas etapas realizadas. **A.** cromatografia de afinidade. **B.** remoção de excesso de sais e imidazol (*desalting*). **C.** cromatografia de afinidade e **D.** exclusão molecular. A remoção de cada enzima nas etapas de purificação está indicada nos retângulos em cores.

Fonte: Elaborada pelo autor.

Experimentos de DSF monitoram a desnaturação proteica em função do aumento de temperatura na presença de uma sonda hidrofóbica fluorescente. A ligação de um fragmento ou ligante gera um salto térmico que é percebido pela diferença na temperatura de desenovelamento. Essa metodologia é uma estratégia bastante atrativa para triar condições de estabilidade na presença de diferentes tampões e bibliotecas de ligantes, pois pode ser realizada com baixas concentrações tanto de proteína, quanto de compostos.

A busca pela melhor condição de ensaio e manutenção proteica foi alcançada pela triagem de uma biblioteca de soluções tampão. Triamos 96 condições diferentes que potencialmente estabilizariam a enzima *XaDHPS*. O melhor resultado indicou a utilização do

tampão 100 mM Hepes pH 7,0, 250 mM NaCl (**Figura 29**). Os experimentos consecutivos seguiram os resultados apresentados nesta triagem.

Figura 29 - Resultado do experimento de triagem de soluções tamponante via DSF. A enzima *XaDHPS* é estabilizada em diferentes configurações de tampões e aditivos. Cada sigmoide representa uma condição testada durante o experimento.

Fonte: Elaborada pelo autor.

5.4 Ensaio Cinético e de atividade

O composto sal de lítio 6-hidroximetilpterina-difosfato foi adquirido da empresa *Schircks Laboratories*. A companhia envia esse material apenas entre os meses de outubro e março, devido à instabilidade em temperatura ambiente. A única fornecedora encontrada é referência na síntese de pterinas, no entanto, o substrato nativo da DHPS não é mais sintetizado pela empresa. O tempo total da compra do produto à chegada ao laboratório foram de, aproximadamente, 2 meses. A fornecedora indica que há conversão de 0,3 % do composto de difosfato para monofosfato por dia em freezer -20 °C. A taxa de mudança aumenta quando mantido em temperatura ambiente.

Quantidades iguais (m/m) do composto foram adicionados em mistura com diselenito de sódio. A junção foi mantida em solução 0,4 M ácido ascórbico / cloreto de potássio. O resultante foi incubado a 50 °C por 5 min, dividido em alíquotas e armazenado em freezer -80 °C. A produção do substrato foi confirmada pela diminuição da absorbância em 340 nm (forma oxidada) para o aumento da absorbância em 315 nm (forma reduzida).

O ensaio colorimétrico pretendido faz uso do kit PiColorlock Gold, no qual é necessário a adição da enzima Pirofosfoquinase e, assim, geram-se tonalidades da cor verde (presença de fosfato inorgânico). A calibração das leituras foi feita no comprimento de onda 635 nm pela diluição de solução de fosfato inorgânico presente no kit. Valores da concentração de fosfato

no meio reacional podem ser diretamente correlacionadas através da estequiometria 1:2, uma vez que a produção de uma molécula de diidropteroato geram 2 moléculas de Pi.

Essa conversão impactou na padronização do ensaio cinético da enzima *XaDHPS*, pois o kit PiColorlock Gold reage com os fosfatos livres no meio reacional. A presença de monofosfato nativo confirma a instabilidade e aumento do ruído, dificultando a análise dos dados.

A alta taxa de conversão do substrato impossibilitou a realização do ensaio de atividade, uma vez que o ruído presente nas medidas iniciais mascarou qualquer funcionamento enzimático. Diferentes concentrações testadas apresentaram o mesmo problema, mesmo através da alteração de pH, concentrações salinas e do substrato.

Por conseguinte, os valores de leitura que foram comparados a uma curva de calibração pré-definida até o momento não indicaram qualquer comportamento enzimático satisfatório. Outras opções descritas na literatura para ensaio da DHPS são: i. ensaio via fluorescência polarizada,⁷¹ ii. ensaio com carbono marcado [C^{14}]pABA⁷²⁻⁷⁴ e iii. ensaio acoplado à outras enzimas da via.

Nas duas primeiras opções, ambos necessitavam de vários compostos que dependeriam de importação, não viável à época. Adicionalmente, os ensaios desenvolvidos purificam alguns compostos bem como trabalham com material radioativo, o que exigem níveis de segurança laboratoriais e uma infraestrutura dedicada, que não estavam disponíveis.

Isto posto, selecionamos o ensaio acoplado às duas enzimas, antecessora e sucessora na via dos folatos, para a produção do substrato DHPP com a enzima *XaHPPK* e a futura medida do consumo do NADPH pela enzima *XaDHFR*, respectivamente.

Assim sendo, trataremos nas seções seguintes os métodos e a produção das novas enzimas citadas, bem como o teste de atividade anterior ao acoplamento de todas no experimento citado. E devido às dificuldades encontradas na padronização do ensaio cinético, métodos biofísicos foram empregados para a determinação da constante de afinidade da enzima por seus substratos naturais e futuros inibidores.

5.4.1 Diidrofolato redutase *XaDHFR*

A expressão na fração solúvel foi observada para todas as construções com a *XaDHFR* (18,3 kDa) na temperatura de 18 °C (**Figura 30**). A primeira etapa de purificação por cromatografia de afinidade removeu a maioria das proteínas não desejadas, entretanto, na etapa seguinte de clivagem para remoção da proteína de fusão e cauda de histidinas, não obtivemos

sucesso. Possivelmente, devido ao tamanho significativo da enzima NusA (55 kDa), houve oclusão do sítio de reconhecimento específico da protease TEV, impedindo dessa forma o acesso e, conseqüentemente, não ocorrendo a hidrólise para separação das proteínas.

Figura 30 - Gel 10 % SDS contendo padrão de expressão da enzima *XaDHFR* nos vetores pET-M11, pETTrx-1a e pETNus-1a. Esquerda = Expressão em Rosetta 2 (DE3); Sol. = fração solúvel, Ins = fração insolúvel e MK = marcador de peso molecular.

Fonte: Elaborada pelo autor.

Dada a incompatibilidade momentânea na remoção da proteína de fusão migramos para o vetor pETTrx-1a. Para o caso específico da enzima *XaDHFR*, a qual possui um valor de pI igual a 8 (predito pelo servidor ExPASy)⁷⁵, dividimos o material celular expresso em duas partes, as quais foram submetidos à lise celular com solução tamponante em pH = 7,0 (Tris-HCl) e pH = 9,5 (CHES). A diferença serviu para observar o comportamento proteico na cromatografia de afinidade e na tentativa de remoção da proteína de fusão com a protease TEV.

Surpreendentemente, e não corroborando com o gel SDS-PAGE apresentado nos testes de expressão, a proteína aumentou exponencialmente o rendimento quando escalonada a expressão. A primeira tentativa de eluição, com gradiente imidazólico, precipitou a solução proteica em pH = 7 e 500 mM de NaCl. O mesmo efeito não foi observado quando utilizamos o tampão CHES em pH = 9,5 e 500 mM, uma vez que não encontramos problemas em eluir a proteína fusionada à tiorredoxina.

A concentração salina pode mimetizar o ambiente químico de neutralização das cargas dos resíduos da superfície proteica, condição similar encontrada no ponto zwitteriônico. Com isso, a perda de cargas leva a proteína a se agregar no início da injeção do gradiente imidazólico para o desligamento da cauda hexa-histidina (6xHis-Trx-*XaDHFR*) com a coluna de afinidade por níquel. Após a redução da concentração de sais, na etapa seguinte, não foi observado precipitação do material.

Testamos, seguidamente, a atividade da protease TEV nos diferentes tampões. É descrito que a atividade desta protease é inibida quando em soluções de alta concentração imidazólica, alta concentração salina (NaCl) e faixa tamponante acima de pH 9,0.⁷⁶ A cauda marcadora, e a proteína de fusão, não foram excisadas quando em solução pH = 9,5, o que nos obrigou a alterar a faixa tamponante da solução de trabalho (**Figura 30**).

Com o ajuste protocolar até a remoção da proteína de fusão, os passos de segunda coluna cromatográfica de afinidade e polimento com coluna de exclusão molecular ocorreram sem maiores problemas em pH = 7 e 200 mM NaCl (**Figura 31**). Observou-se que não foi possível concentrar a proteína em valores acima de 5 mg/mL (pH = 7), no entanto, quando alteramos a faixa tamponante (pH = 9), elevamos a concentração de proteína na solução sem dificuldades.

Figura 31 - Etapa final de purificação da enzima *XaDHFR* (18,3 kDa). O gel SDS-PAGE 15 % confirma a pureza das frações recolhidas após utilização da coluna de exclusão molecular. O cromatograma indica a fração majoritária como sendo a proteína de interesse isolada.

Fonte: Elaborada pelo autor.

De acordo com o levantamento dos métodos de purificação descritos na literatura, os passos desenvolvidos neste caso específico isolaram a proteína *XaDHFR* sem a utilização de metotrexato (composto tóxico) ou técnicas de re-novelamento proteico.

Empregamos teste de estabilidade térmica para definir a melhor condição de trabalho com esta proteína. Os valores alcançados na máxima térmica indicam preferência por tampão 100 mM Fosfato de sódio pH 7,5 e 200 mM NaCl (**Figura 32**). Encontramos a mesma estabilidade térmica no tampão HEPES, em faixa tamponante mais distante do valor de pI. É indispensável, entretanto, a presença de glicerol ou DMSO. Optamos em trabalhar, para os próximos ensaios, com o tampão HEPES, uma vez que as outras proteínas da via ligam fosfato

e pela fácil formação de arranjos cristalinos de sal com tampão fosfato, o que dificulta a identificação de cristais proteicos de alta qualidade nos ensaios de cristalização.

Figura 32 - Testes de estabilidade térmica (DSF) frente a proteína *Xa*DHFR. Os resultados indicam preferência e estabilidade em tampão fosfato pH 7,5 e 200 mM NaCl com adição de 10 % glicerol.

Fonte: Elaborada pelo autor.

A atividade enzimática foi avaliada utilizando as medidas de absorção de radiação UV-VIS no comprimento de onda (λ) de 340 nm, o que corresponde à formação do NADPH ao longo do tempo. Estando a proteína ativa, ocorreria a oxidação do cofator e consequente formação do NADP^+ , que não absorve luz no mesmo comprimento de onda indicado anteriormente.

Não notamos alteração nos valores de absorbância para a proteína *Xa*DHFR. Foram aplicadas enzimas provenientes de diferentes purificações, variou-se a concentração dos componentes da solução reacional, o tempo de incubação, bem como a temperatura utilizada (Figura 33).

Figura 33 - Medidas do consume de NADPH ao longo do tempo. As soluções contêm quantidades equimolar do substrato diidropteroato e o cofator NADPH, acrescidos da enzima *Xa*DHFR. **A.** A enzima *Xa*DHFR se mostrou não ativa no ensaio performedo. **B.** Na varredura do substrato e cofator observou-se que ambos absorvem luz no mesmo comprimento de onda 340 nm.

Fonte: Elaborada pelo autor.

Outro ponto negativo encontrado neste ensaio é que o cofator absorve luz no mesmo comprimento de onda que o substrato enzimático, o DHP. Desta forma, os ensaios desenvolvidos para esta enzima utilizam concentrações equimolares de ambos os reagentes e o cálculo da concentração final emprega a soma dos coeficientes de extinção molar. Na tentativa de evitar ensaios enzimáticos radioativos, este é o único experimento de aferição da atividade enzimática para DHFR descrito.

Durante o período de intercâmbio no Departamento de Bioquímica da Universidade de Cambridge, foi possível expressar a proteína ortóloga de humano (*Homo sapiens* DHFR, *HsDHFR*). O processo de produção desta proteína seguiu o protocolo anteriormente descrito para a enzima *XaDHPS*. O ótimo rendimento e a facilidade de isolamento tornaram simples o processo de obtenção da *HsDHFR* (Figura 34).

Figura 34 - Gel SDS-PAGE 15 % com a purificação da enzima humana diidrofolato redutase de *Homo sapiens* por cromatografia de exclusão molecular. Os diversos poços dizem respeito às frações eluídas no experimento de gel filtração. MK - Marcador molecular, S - fração solúvel.

Fonte: Elaborada pelo autor.

Os testes de atividade seguiram protocolo descrito em ensaios desta enzima, comparando-se à atividade homóloga descrita na literatura.⁷⁷ As condições utilizadas são as mesma empregadas no ensaio com a *XaDHFR*. O decaimento na absorbância foi observado, tendo obtido o 100 μ M de substrato totalmente processado em 3 min (Figura 35). Dependendo da conformação, do sítio de ligação do NADPH, a proteína pode assumir dois estados: aberto e fechado.⁷⁸ Analisando os resultados em paralelo, sugerimos que a proteína purificada de *X. albilineans* pode não estar enovelada corretamente, ou apresentar conformação incompatível com a ligação do cofator e substrato para que haja catálise.

Figura 35 - Estrutura tridimensional da enzima humana diidrofolato redutase de *Homo sapiens* (PDB 1U70). A atividade enzimática desta enzima foi mensurada por absorvância através do consumo de NADPH e DHP ao longo do tempo. Quantidades iguais de cofator e substrato foram inseridos e, na presença da HsDHFR, consumidos durante os 4 minutos de experimento.

Fonte: Elaborada pelo autor.

5.4.2 6-hidroximetil-7,8-diidropterina pirofosfoquinase (*XaHPPK*)

Como mencionado na Introdução, a enzima HPPK catalisa a formação do substrato DHPP para a função catalítica exercida pela DHPS. Os planos de purificação traçados seguiram os protocolos implementados para as outras enzimas, no entanto, com o vetor contendo tiorredoxina, não foi possível separar a proteína de interesse da proteína de fusão após a clivagem.

As características intrínsecas da enzima tiorredoxina são semelhantes à proteína *XaHPPK*. A proteína de fusão, utilizada para aumentar solubilidade, tem o mesmo peso molecular aproximado, a mesma faixa de pH do ponto isoelétrico e grau de lipofilicidade. Adicionalmente, a enzima tiorredoxina liga o cofator NADPH e, neste caso, interferiria no ensaio cinético acoplado com a enzima *XaDHFR*.

Dada a problemática encontrada, optamos por alterar o vetor de expressão proteica. Realizamos sub-clonagem com vetores do sistema SLIC (em inglês, *Sequence-and-ligation-independent cloning*) para otimizar o tempo e os custos de uma nova construção.⁷⁹ Foram escolhidos, inicialmente, um vetor que contém na extremidade N-terminal uma cauda octa-histidina (8x-His) e outro vetor com a cauda octa-histidina atrelada à proteína de fusão MBP (em inglês, *maltose binding protein*). A repetição da amplificação padrão por PCR do inserto *folK* foi suficiente para a clonagem. O aumento no número de histidina auxilia na especificidade de ligação da proteína com a coluna de afinidade, eliminando espaços para ligações de baixa afinidade que geram impurezas no produto.

Empregamos como molde da PCR o vetor previamente clonado em pETTrx-1a, o qual gerou o fragmento necessário a ser transferido para os novos vetores. Para as extremidades coesivas, de maneira similar ao sistema LIC, adicionamos dCTP à reação após 5 min de atividade da enzima T4 DNA polimerase (exonuclease). A mistura reacional foi hibridizada e transformada em células DH-5 α . A integridade do inserto foi confirmada por sequenciamento (Figura 36).

Figura 36 - Clonagem da enzima *XaHPPK* utilizando o sistema SLIC (*Sequence-and-ligation-independent cloning*). **A.** Duas construções foram exploradas na extremidade N-terminal, sendo: 8xHis e 8xHis + 'maltose binding protein' (MBP) como proteína de fusão. **B.** A integridade do material foi assegurada por sequenciamento.

Fonte: Elaborada pelo autor.

Todas as construções auxiliaram a expressão da quinase *XaHPPK*. No entanto, com a proteína de fusão MBP, a enzima de interesse foi direcionada por completo para corpos de inclusão e não faz parte da fração solúvel. O comparativo de rendimento final dos vetores de expressão indicou um aumento de duas vezes no rendimento proteico com o vetor com cauda octa-histidina (8xHis), quando comparado ao vetor pETTrx. Os métodos cromatográficos de purificação são os mesmos utilizados na obtenção da proteína *XaDHPS* (Figura 37).

Figura 37 - Gel SDS-PAGE 15 % com o resultado da purificação por exclusão molecular da enzima *XaHPPK*.

Mk = marcador molecular, S = fração solúvel e T = lise total. Comparativo entre a atividade da enzima *XaHPPK* e o controle comparado no teste de atividade.

Fonte: Elaborada pelo autor.

A caracterização de atividade da enzima iniciou com a medida do ATP remanescente após período de incubação da proteína no meio reacional. O ensaio luminescente compara a atividade enzimática com uma curva padrão de ATP, pré-construída sem enzima. Podemos destacar que a enzima se encontrou ativa após a purificação e foi utilizada no processo de produção do DHPP para atividade da enzima *XaDHPS*.

Os próximos passos para a enzima incluem: a. determinação da taxa de produção, valores de K_m e V_{max} e o rendimento final do produto e b. ajustes de pH, concentração salina e temperatura para que possam ser adicionadas outras duas enzimas no ensaio acoplado.

Paralelamente, testamos a cristalização das enzimas *XaHPPK* e *XaDHFR*. Reduzimos os kits na triagem para dois: Index e JCSG+ (HAMPTON). Avaliamos o comportamento das enzimas frente às diferentes soluções nas seguintes concentrações: 10 e 5 mg/mL para ambas enzimas. No caso da proteína *XaHPPK*, adicionamos 10 mM de ATP, 10 mM $MgCl_2$ ou um mix de ambos compostos anteriormente à mistura com a solução de cristalização. Até a presente data, não existe nenhuma estrutura que se assemelhe a um arranjo cristalino. A presença de cristais proteicos foi monitorada por incubadora robotizada que contém leituras de UV. Além disso, inspeções visuais foram realizadas periodicamente. (Figura 38).

Figura 38 - Representação das 96 imagens que representam os poços de triagem das condições de cristalização. As triagens de soluções de cristalização foram feitas contra as enzimas *XaDHFR* e *XaHPPK*. Para a proteína *XaHPPK* também foram construídas as condições de adição 10 mM ATP e/ou 10 mM $MgCl_2$, previamente ao início do ensaio.

Fonte: Elaborada pelo autor.

5.5 Cristalização, Coleta de Dados, Resolução e Refinamento (*XaDHPS*)

A técnica de cristalografia de macromoléculas é um dos principais métodos de identificação e caracterização da estrutura tridimensional de proteínas. Em altas concentrações proteicas, as macromoléculas tendem a se agrupar para esconder as regiões hidrofóbicas e expor as partes polares ao solvente. O tamanho e a complexidade proteica refletem na habilidade em formar os mais diferentes contatos intermoleculares, mesmo que estes contatos existam apenas em condições cristalinas. A diversidade dos grupos químicos envolvidos nestes contatos implica em pequenas alterações de pH, força iônica, temperatura ou concentração de um determinado íon ou molécula que seja crucial para a cristalização.⁸⁰

O diagrama de solubilidade traduz as áreas nas quais a alta concentração proteica, conjuntamente com agente precipitante, pode formar cristais, precipitar ou estar solúvel a todo momento (**Figura 39**). Portanto, a cristalização é a resposta única de uma macromolécula com a respectiva solução.

Figura 39 – **A.** Diagrama de solubilidade proteica em diferentes concentrações de agente precipitante. **B.** Esquema descritivo da técnica de difusão de vapor através da metodologia de gota suspensa para a triagem e produção de cristais proteicos.

Fonte: Elaborada pelo autor.

Os cristais resultantes do experimento são enviados para análise do poder de difração dos raios X. Várias são as fontes de raios X úteis para esta finalidade, no entanto, destacaremos neste trabalho apenas a radiação Síncrotron, e que é construído em tecnologia suficiente para coleta de dados em alta resolução (e.g. 0.7 Å).

Uma das maiores vantagens da cristalografia de raios X é a possibilidade de determinação estrutural em alta resolução (resolução > 2,5 Å). A disposição atômica sobre a organização estrutural do alvo molecular pode ser utilizada para a melhor compreensão da função da proteína em estudo, bem como elucidar o modo de ligação com outras moléculas (**Figura 40**).

Figura 40 - Esquema ilustrativo dos experimentos de difração de raios X. A amostra utilizada (cristal) sofre a incidência dos raios X, difratando fótons que serão coletados pelo detector. O cálculo da densidade eletrônica, bem como a construção do modelo representativo seguem cálculos computacionais.

Fonte: Elaborada pelo autor.

Nesta tese empregamos a cristalografia de raios X para a determinação do modo de ligação dos candidatos a inibidores da enzima *XaDHPS*. Para tanto, com o método de cristalização por difusão de vapor, utilizando a técnica de gota assentada, triamos manualmente 11 conjuntos comerciais nas concentrações de 5 e 10 mg/mL. As gotas foram analisadas a cada dois dias, partindo do primeiro dia de montagem das caixinhas de 96 poços. Os experimentos foram conduzidos a 18 °C. Houve precipitação imediata em algumas condições. Após duas semanas alguns poços apresentaram sítios de nucleação ou cristais bem definidos (**Tabela 5**).

Tabela 5 - Condições promissoras com surgimento de cristais da enzima *XaDHPS*.

SAL	PEG	Solução Tampão
Sulfato de lítio	25 % 3350	Bis-Tris pH 6,5
Tartrato de Amônio	20 % 3350	-
Formato de amônio	20 % 3350	-
Fluoreto de sódio	20 % 6000	-
Fosfato de amônio dibásico	20 % 3350	-
-	25 % 3350	HEPES pH 7,5

Fonte: Elaborada pelo autor.

As condições mais promissoras foram expandidas com a técnica de gota suspensa (**Figura 39**) em placa de 24 poços. Neste experimento, variou-se pH, agente tamponante e concentração de polietilenoglicol (PEG) 3350. O método de gota suspensa permite trabalhar com volumes maiores de proteína e solução, o que melhora a formação dos cristais. Expandimos as condições de sulfato de lítio, tartrato de amônio e hepes. Apenas na condição de sulfato de lítio houve reprodução dos arranjos iniciais.

A melhor condição encontrada para obtenção de cristais apropriados para experimentos de difração de raios X da enzima *XaDHPS* foi: 0,1 M Bis-Tris pH 6,5, 25 % PEG 3350, 0,2 M Sulfato de Lítio a 18 °C (**Figura 41**). Os cristais cresceram em 15 dias e, com a adição de 1 mM de MgCl₂, houve o adiantamento para 7 dias. A aplicação da técnica de semeadura (em inglês, *seeding*) permitiu a formação de cristais após 24 h (**Figura 41**). Experimentos de imersão (em inglês, *soaking*) foram conduzidos com os cristais para a obtenção de complexos cristalográficos entre a *XaDHPS* e os substratos ácido p-aminobenzóico (pABA), 6-hidroximetil-7,8-diidropterina (6HMPH2) e produto diidropteroato (DHP).

Os dados cristalográficos dos cristais nativos e em complexo foram coletados na linha MX-2 do Laboratório Nacional de Luz Síncrotron (LNLS) com o detector Pilatus 2M. com exceção do conjunto referente ao 6HMPH2 que foi coletado na linha de luz Diamond I04.

Figura 41 - Padrão de difração da enzima XaDHPS. O poderio de difração ultrapassa as resoluções de canto (1,5 Å). Destacamos o primeiro cristal obtido após triagem de condições de cristalização, os cristais comumente obtidos e o resultado do experimento de semeadura.

Fonte: Elaborada pelo autor.

No processamento das imagens coletadas empregamos o programa XDS ou programas derivados que realizam o processamento automaticamente, como autoPROC (Global Phasing) e DIALS. A utilização destes softwares é de fundamental importância, uma vez que os dados foram coletados utilizando a técnica de *fine ϕ -slicing* (0,2 °/imagem) (**Figura 42**). O fatiamento dos *spots* pelo método de coleta só é devidamente identificado com o perfil de integração 3D (em inglês, *profile 3D fitting*), encontrado no XDS e programas derivados. O método considera as direções angulares, bem com a superfície do detector utilizado (1+2D).

Figura 42 - Diferentes métodos de coleta de dados pela técnica de difração de raios X. O método de fatiamento angular por *fine ϕ -slicing* auxilia no processo de captação de pontos únicos de difração. Outros métodos de fatiamento grosseiro (*coarse ϕ -slicing*) é indicado para conjuntos de dados que não apresentem problemas de pontos de difração próximos.

Fonte: Elaborada pelo autor.

A integração e indexação das diversas reflexões geradas pelos diferentes planos é produzida inicialmente. O software corrige as distorções do detector, estima o nível de sinal de

fundo (*background*), seleciona os picos de intensidade e indica as dimensões de cela e sua orientação no espaço. Em seguida, integra os picos coletados e corrige, ao final, os dados empíricos aplicados (decaimento, absorção, vibrações do goniômetro, dentre outros parâmetros). A indicação do grupo espacial é feita por tentativa e erro para se adequar o grupo cristalino ao espaço da cela unitária pela menor penalidade possível. Este índice de correlação indicou duas possíveis soluções para grupos monoclinicos: mP (P 1 2 1 ou P 1 2₁ 1) e mC (C 1 2 1).

No total, quatro conjuntos de dados com estatísticas satisfatórias foram coletados (*XaDHPS*, *XaDHPS*:pABA, *XaDHPS*:6HMPH2 e *XaDHPS*:DHP). Os conjuntos *XaDHPS*, *XaDHPS*:6HMPH2 e *XaDHPS*:DHP foram processados no grupo espacial C2 (nº 5), enquanto o conjunto *XaDHPS*:pABA foi processado no grupo espacial P2₁ (nº 4).

Em seguida, utilizamos o programa Aimless (associação dos pacotes Pointless, Scala e cTruncate) para determinação de grupo pontual, avaliação estatística completa dos dados agrupados e conversão das intensidades em amplitude, respectivamente. A resolução de cada conjunto foi definida mediante os parâmetros estatísticos individuais, uma vez que levamos em consideração os valores da razão sinal-ruído, completeza e outros. As faixas de resolução que não se enquadram nos valores definidos foram automaticamente descartadas.

As estatísticas do processamento estão indicadas na **Tabela 6**. De acordo com o coeficiente de Matthews há probabilidade de 85 %, aproximadamente, de obtermos quatro moléculas por unidade assimétrica quando o grupo espacial é P2₁ (*XaDHPS*:pABA) e 99 % de existir duas moléculas no grupo quando é C2 (apo e *XaDHPS*:DHP). A estimativa de solvente na cela unitária fica entre 40-50 % para todos os conjuntos trabalhados.

A substituição molecular encontrou as fases iniciais das estruturas cristalográficas com o programa Phaser. O sequenciamento previamente realizado nos auxiliou no cálculo de massa molecular para estimar o número de moléculas dentro da cela unitária predita. Para a busca da estrutura molde na densidade eletrônica, empregamos a enzima homóloga de *Escherichia coli* (*EcDHPS* – PDB 1AJ0). A identidade sequencial entre *XaDHPS* e *EcDHPS* é de 51 %. Uma vez que o molde é um monômero, buscamos o mesmo número de cópias indicado pelo coeficiente de Matthews na célula unitária.

O programa *Phaser* busca verossimilhança pelo molde na densidade eletrônica e pontua a cada solução encontrada. Os parâmetros de verossimilhança para a estrutura nativa foram: LLG > 10,000 e TFZ > 15, validando o processo de resolução. Ao final da substituição molecular obtivemos duas estruturas com duas moléculas por unidade assimétrica (C2) e uma

com quatro moléculas por unidade assimétrica (P2₁). Empregamos como molde a enzima nativa resolvida para todos os outros de dados trabalhados.

Os modelos foram manualmente construídos, com diversos ciclos de refinamento entre o espaço real (Coot) e o espaço recíproco (Buster / phenix.refine). A interpretação dos mapas de densidade eletrônica seguiu os contornos 1.0 σ (2F_o-F_c) e 4.0 σ (F_o-F_c). As estruturas foram refinadas, no espaço real, através do ajuste das coordenadas xyz, encaixe dos resíduos na densidade eletrônica, adição de águas, adequação dos ângulos diedrais phi (ϕ) e psi (Ψ), os quais compõem o diagrama de Ramachandran (**Figura 43**).

Figura 43 – Resolução estrutural do cristal monoclinico de XaDHPS. A densidade eletrônica e disposição residual é correspondente à faixa da interface dimérica proteica.

Fonte: Elaborada pelo autor.

Foram enquadrados, também, os ângulos das ligações (χ - χ^{1-4}) entre os carbonos C β -C γ -C δ -C ϵ , que integram as cadeias laterais dos resíduos a partir da ligação peptídica. No espaço recíproco foram ajustadas as imperfeições pela anisotropia da difração, as distorções da simetria não cristalográfica (em inglês, *non crystallographic symmetry*) e o refino dos ligantes mediante as restrições geradas pelo servidor Mogul. As estatísticas do refinamento estão apresentadas na **Tabela 6**.

Tabela 6 - Dados estatísticos do processamento e refinamento das estruturas de XaDHPS, na forma apoenzima e em complexo com ácido 4-aminobenzóico (pABA), o substrato 6HMPH2 e o produto diidropteroato (DHP). Estatísticas na máxima resolução estão em parênteses.

	XaDHPS	XaDHPS:pABA	XaDHPS:DHP	XaDHPS:6HMPH2
Comprimento de onda (Å)	0,97	1,45	1,45	0,97
Resolução (Å)	43,65 - 1,76 (1,823 - 1,76)	32,75 - 1,65 (1,709 - 1,65)	31,08 - 1,96 (2,03 - 1,96)	40,41 - 2,033 (2,106 - 2,033)
Grupo Espacial	C 1 2 1	P 1 21 1	C 1 2 1	C 1 2 1
Cela - Dimensão (Å)	109,30 73,98 82,22 90 127,0	82,589 75,387 89,106 90	109,176 74,405 82,152 90	109,414 76,321 82,378 90 126,427
Ângulos - cela α, β, γ (°)	90	101,102 90	126,414 90	90
Reflexões (total)	354071 (33709)	645102 (45277)	214765 (13522)	241071 (23476)
Reflexões (únicas)	51853 (5109)	128121 (12427)	36725 (3288)	35167 (3456)
Multiplicidade	6,8 (6,6)	5,0 (3,6)	5,8 (4,1)	6,9 (6,7)
Completeza (%)	99,13 (98,89)	98,95 (96,86)	95,94 (86,59)	98,24 (97,24)
I/sigma(I) médio	15,49 (2,66)	14,45 (2,57)	29,47 (4,25)	13,38 (2,49)
Wilson B-factor	27,88	23,07	31,59	42,79
R-merge	0,075 (0,6605)	0,069 (0,5455)	0,041 (0,3035)	0,097 (0,7046)
R-meas	0,082 (0,7168)	0,076 (0,6479)	0,044 (0,346)	0,106 (0,7647)
R-pim	0,031 (0,2757)	0,031 (0,3401)	0,017 (0,1634)	0,040 (0,2939)
CC 1/2	0,997 (0,945)	0,98 (0,0556)	0,99 (0,976)	0,98 (0,966)
R-work	0,24 (0,4883)	0,21 (0,3128)	0,19 (0,2452)	0,24 (0,4903)
R-free	0,28 (0,5157)	0,23 (0,3401)	0,22 (0,3083)	0,28 (0,5946)
RMS (ligação)	0,014	0,013	0,014	0,013
RMS (ângulos)	1,66	1,55	1,62	1,68
Ramachandran favorável (%)	97,26	98,35	97,83	97,30
Ramachandran permitido (%)	2,74	1,55	1,98	2,70
Ramachandran outliers (%)	0,00	0,10	0,20	0,00
B-fator médio	38,45	33,50	38,36	61,12

$$R = \frac{\sum hkl |F_{obs} - F_{calc}|}{\sum hkl F_{obs}}$$

$$R\text{-merge} = \frac{\sum hkl \sum j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum hkl \sum j I_{hklj}}$$

$$R\text{-pim} = \frac{\sum hkl \sqrt{\frac{n}{n-1}} \sum j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum hkl \sum j I_{hklj}}$$

Fonte: Elaborada pelo autor.

5.6 Caracterização Estrutural

A enzima *Xa*DHPS está presente em um vasto número de organismos, incluindo procariotos, eucariotos primitivos e plantas, no entanto, está ausente em mamíferos. O enovelamento da proteína é conservado entre as espécies e os resíduos que compõem o sítio catalítico apresentam elevada similaridade sequencial (> 90 %). Até julho de 2019, existiam estruturas de DHPS de mais de 15 organismos diferentes depositadas no banco de dados proteico PDB.

O cálculo do desvio quadrático médio (RMSD do inglês, *root mean square deviation*) entre a *Xa*DHPS e as estruturas ortólogas de *Francisella tularensis* (PDB 3MCM, 4PVZ), *Escherichia coli* (PDB 3V5O, 1AJ0, 5U0V), *Burkholderia cenocepacia* (PDB 2Y5J), *Bacillus anthracis* (PDB 1TWS), *Mycobacterium tuberculosis* (PDB 1EYE, 2VP8), *Staphylococcus aureus* (PDB 4HB7, 1AD1, 6CLU), *Streptococcus pneumoniae* (PDB 2VEF), *Francisella tularensis* (PDB 3MCM), *Coxiella burnetii* (PDB 3TR9), *Mycobacterium smegmatis* (PDB 6OMZ), *Vibrio fischer* (PDB 5U5W), *Klebsiella pneumoniae* (PDB 5UMG), *Thermus thermophilus* (PDB 2DQW) e *Yersinia pestis* (PDB 5JQ9, 3TYU) sugerem elevada similaridade sequencial (r.m.s.d < 2,0 Å) (**Figura 45**).

Em alguns eucariotos a enzima DHPS se encontra ligada à enzima 6-hidroximetil-7,8-diidropterina pirofosfoquinase (HPPK). Os complexos DHPS:HPPK de *Saccharomyces cerevisiae* (PDB 2BMB) e *Plasmodium vivax* (PDB 2Z79) estão disponíveis e mostram a ligação das enzimas por uma dupla de fitas β flexível (**Figura 44**).

Figura 44 - Alinhamento de estruturas secundárias e comparativo entre o enovelamento de DHPS singular (PDB 1EYE) e do enovelamento na fusão com a enzima HPPK (2BMB). A enzima ScHPPK está destacada em cinza, enquanto a porção ScDHPS em azul. A apresentação da enzima DHPS de *M. tuberculosis* é feita em laranja.

Fonte: Elaborada pelo autor.

O alinhamento sequencial das estruturas encontradas no PDB está apresentado na **Figura 45**. A alta conservação entre os resíduos de aminoácidos (em vermelho), principalmente

aqueles presentes nos sítios catalíticos (bastões verdes), sugerem que a descoberta de inibidores seletivos para a enzima de bactéria é uma tarefa desafiadora.

Com base na elevada identidade sequencial, principalmente na cavidade catalítica, e no fato de vários inibidores de DHPSs terem sido desenvolvidos, a estratégia que empregamos buscou evitar a investigação de novos derivados do anel pterina.⁴¹ Na ausência de um análogo de pterina, o sítio catalítico é ocupado pela cadeia lateral da Arg261, resíduo conservado em todos os organismos comparados. Este mecanismo garante estabilidade e regulação da atividade enzimática.⁸¹

O monômero da *Xa*DHPS possui enovelamento de um clássico barril triose fosfato isomerase “TIM” (α_8/β_8). Este enovelamento é caracterizado por oito hélices alfa (α_1 - α_8) que circundam oito fitas betas (β_1 - β_2) intercaladas, dispostas de maneira que um barril seja formado por fitas no centro e hélices que o envolvam (**Figura 46**).

Figura 45 - Alinhamento sequencial das estruturas de Diidropteroato sintase disponíveis no banco de dados PDB. Os resíduos conservados estão indicados em vermelho, enquanto os resíduos idênticos em bastões verdes. Estruturas alinhadas: *Francisella tularensis* (PDB 3MCM, 4PVZ), *Escherichia coli* (PDB 3V5O, 1AJ0, 5U0V), *Burkholderia cenocepacia* (PDB 2Y5J), *Bacillus anthracis* (PDB 1TWS), *Mycobacterium tuberculosis* (PDB 1EYE, 2VP8), *Staphylococcus aureus* (PDB 4HB7, 1AD1, 6CLU), *Streptococcus pneumoniae* (PDB 2VEF), *Francisella tularensis* (PDB 3MCM), *Coxiella burnetii* (PDB 3TR9), *Mycobacterium smegmatis* (PDB 6OMZ), *Vibrio fischeri* (PDB 5U5W), *Klebsiella pneumoniae* (PDB 5UMG), *Thermus thermophilus* (PDB 2DQW) e *Yersinia pestis* (PDB 5JQ9, 3TYU).

Fonte: Elaborada pelo autor.

Na extremidade N-terminal há um grampo beta (do inglês, *beta hairpin*), que auxilia o fechamento do barril e caracteriza a sua polaridade (resíduos 6-16). Os resíduos iniciais das fitas N β 1 e N β 2 alternam sua polaridade na face exposta ao solvente enquanto, do outro lado, há formação do sítio catalítico. A presença deste grampo beta ampara a estabilização do barril, mas não é totalmente conservado entre os organismos, como por exemplo em *M. tuberculosis* e *S. aureus*. Outro ponto que merece destaque é a presença de volta α no início da hélice 7 (resíduos 226-235). Esta estrutura secundária auxilia o arranjo do ácido 4-aminobenzóico no sítio catalítico.⁸²

Figura 46 - Divisão das estruturas secundárias da enzima XaDHPS. O enovelamento de um barril 'TIM' compõe o complexo $(\alpha/\beta)_8$. Em vários casos, principalmente em fungos, o grampo β não faz parte da estrutura de DHPS.

Fonte: Elaborada pelo autor.

Como observado em outras estruturas, a enzima se apresenta como um dímero funcional, com área de contato de aproximadamente 1200 Å² entre os monômeros. A interface é formada pela interação recíproca das hélices α 7 e α 8 dos monômeros, composta basicamente por resíduos hidrofóbicos, como valina e leucina. Adicionalmente, a interface dimérica é

mantida pela interação eletrostática recíproca dos resíduos Asp270⁻ e Arg235⁺ de cada monômero e que auxiliam na estabilização das hélices 7 e 8.⁸³ O sítio está localizado na porção C-terminal, enterrado no centro do barril, e depende da interação conjunta das alças $\alpha 1$ - $\alpha 2$ e $\alpha 5$ para estabilização do substrato pABA e das fitas $\beta 2$ - $\beta 7$ para ligação do DHP (**Figura 47**).

As maiores variações sequenciais estão nas extremidades C-terminal das estruturas secundárias. Os resíduos que compõem as alças 1 e 2 são altamente conservados entre os organismos ortólogos (**Figura 45**), no entanto, eles não estão evidenciados na maioria das estruturas depositadas no PDB.⁴² Este fato é corroborado pela elevada flexibilidade dessas alças, o que dificulta a determinação precisa das coordenadas cristalográficas devido à baixa ou ausência de densidade eletrônica nesta região. Por outro lado, os contatos cristalográficos formados no arranjo cristalino trabalhado permitiram a determinação da conformação dos resíduos das alças 1 e 2 na estrutura *XaDHPS*.

Figura 47 - Estrutura 3D dimérica da enzima *XaDHPS*. A estrutura é formada por oito hélices $\alpha(1-8)$ que circundam um barril $\beta(1-8)$. Estas estruturas secundária são conectadas por oito alças L(1-8), das quais as alças L1, L2 e L5 participam da catálise enzimática. A interação recíproca das hélices $\alpha 7$ e $\alpha 8$ de cada subunidade formam uma cavidade nesta interface.

Fonte: Elaborada pelo autor.

Existem várias proposições da função catalítica deste arranjo de alças. Podemos ressaltar que a conservação do aminoácido Pro30 na extremidade C-terminal e a dupla de glicinas na extremidade N-terminal da alça 1 auxiliam na mobilidade do conjunto, tanto para reconhecimento de substratos, quanto para a proteção do sítio catalítico contra potenciais efeitos do solvente.

Em casos específicos, são encontradas estas alças completas. Comumente, podemos elucidar que em uma mesma estrutura a presença e ordenamento das alças descritas em uma subunidade não são simétricos a outra, ou seja, enquanto um monômero se apresenta ordenado, a molécula-espelho está descolada. Isto indicaria o funcionamento cooperativo de cada monômero, um de cada vez, na atividade catalítica de reconhecimento de substrato, direcionamento ao sítio ativo e liberação do produto.

O contato cristalino formado por estas moléculas diferem dependendo do grupo espacial. No caso dos cristais da proteína nativa, são frequentemente encontrados dois grupos espaciais na mesma gota de cristalização: C 1 2 1 e P 1 2₁ 1.

Nas etapas iniciais de triagens das condições de cristalização obtivemos uma estrutura cristalina na condição com tartrato de amônio e PEG 3350. Após a coleta de dados, a solução encontrada apresentou 8 moléculas na unidade assimétrica no grupo P2₁. O octâmero cristalográfico é formado por três dímeros dentro da unidade mais dois monômeros que se completam pela junção dos monômeros da cela vizinha. O potencial eletrostático indica regiões carregadas positivamente dentro do sítio catalítico, enquanto as áreas de contato entre os dímeros são formadas por resíduos carregados negativamente (**Figura 48**).

Figura 48 - Estrutura 3D da enzima Dihidropterotato sintase de *Xanthomonas albilineans* (XaDHPS) cristalizada na solução de tartrato de amônio e PEG 3350. São 8 moléculas na cela unitária organizada no grupo P21. O potencial eletrostático é descrito indicando as regiões carregadas negativamente (branco) e positivamente (marrom).

Fonte: Elaborada pelo autor.

O que distingue essa condição de cristalização das demais resolvidas são os contatos formados entre as unidades diméricas. A área de contato entre os dímeros nesta condição de tartrato de amônio é duas vezes menor ($\sim 350 \text{ \AA}^2$) que o contato encontrado na solução de sulfato de lítio ($\sim 700 \text{ \AA}^2$). A proximidade maior se deve ao movimento da alça 1, que interage com a entrada da interface dimérica. Esse processo de abertura e fechamento auxilia na catálise com a comunicação direta com o sítio catalítico.

5.7 Análise estrutural da enzima XaDHPS em complexo com pABA, 6HMPH2 ou DHP

Considerando a alta especificidade do sítio catalítico e a natureza dos substratos que se ligam à DHPSs, destacamos a dificuldade em se descobrir novos inibidores solúveis e que sejam seletivos para a homóloga de *X. albilineans*. Estas características dificultam a identificação de candidatos a novos inibidores, uma vez que as concentrações iniciais avaliadas são entre 250 μM e 500 μM e necessitam que as moléculas candidatas a ligantes apresentem estrutura de no máximo dois anéis fundidos.⁷⁴ Isto posto, identificamos os possíveis sítios de ligação para que o planejamento busque vias alternativas de ligação e inibição da atividade bacteriana.

O sítio catalítico da XaDHPS é composto por três regiões distinta de interação com o substrato DHPP: a. anel pterina, ii. ácido 4-aminobenzoico e iii. pirofosfato (**Figura 49**). Identificamos as características moleculares responsáveis por essas regiões de interação através

de experimentos de imersão em solução concentrada de cada substrato separado. Não obtivemos complexos quando formamos coquetéis para o ensaio de imersão. Os complexos *Xa*DHPS:pABA, *Xa*DHPS:6HMPH2 e *Xa*DHPS:DHP foram obtidos com solução aquosa preparada no máximo da solubilidade dos ligantes neste solvente.

A triagem inicial das concentrações de DMSO indicaram que os cristais podem suportar até 25 % DMSO na diluição dos compostos. No entanto, os cristais de *Xa*DHPS são extremamente sensíveis a presença de agentes crioprotetores.

Figura 49 - Caracterização estrutural do sítio catalítico da enzima diidropteroato sintase de *Xanthomonas albilineans*. A ligação do anel pterina segue a mesma orientação tanto isolado quanto no aduto formado com o ácido 4-aminobenzóico. O sítio de ligação do pirofosfato é mimetizado pelos íons sulfato (SO4).

Fonte: Elaborada pelo autor.

Outro ponto considerado foi a necessidade dos íons sulfato na formação dos cristais de *Xa*DHPS. Presente na condição de cristalização, decidimos variar as concentrações de sulfato de lítio até não fazer parte da composição da solução de cristalização. Com o tempo de 2-3x maior no crescimento dos cristais, percebemos que o íon SO_4^{2-} não é vital para a formação do cristal. Nos cristais obtidos na ausência de íons SO_4^{2-} , identificamos moléculas de etilenoglicol ligados na posição equivalente no sítio catalítico. Além disso, testamos a importância deste íon na manutenção do sítio catalítico e na estrutura proteica como um todo. Para tanto, embebemos alguns cristais em solução sem sulfato de lítio e em água, confiando que o fator osmótico fosse capaz de remover os íons presentes no sítio.

Não identificamos o íon SO_4^{2-} no sítio catalítico após o experimento de remoção dos íons. Além disso, não houve mudança conformacional dos resíduos do sítio ativo. Na posição equivalente ao íon verificamos a presença de moléculas de água. Testamos, em um terceiro

momento, embeber os cristais em solução contendo ligantes e na ausência do sulfato de lítio. Infelizmente, esta tentativa não gerou uma melhora na identificação de complexos.

A análise estrutural do complexo *Xa*DHPS:DHP e *Xa*DHPS:6HMPH2 mostra que o anel pterina é estabilizado por ligação de hidrogênio entre os átomos polares do anel e as cadeias laterais dos resíduos Asp101, Asn120, Asp190 e Lys226. Em ambos os complexos o anel pterina assume a mesma orientação, interação e espaço ocupado dentro do sítio ativo. Além disso, observa-se interações de elétrons π do anel pterina com a cadeia lateral dos resíduos Phe195 e Arg261 e interações hidrofóbicas com as cadeias laterais dos resíduos Met144 e Val122 (**Figura 50**). Esse modo de ligação permite a liberação do pirofosfato pelo carbocátion formado no carbono C9 durante catálise enzimática. É possível visualizar uma representação do estado de transição no complexo com a molécula de 6HMPH2, a qual não possui o pirofosfato que seria deslocado após a entrada na cavidade.

Figura 50 – Detalhamento estrutural do sítio catalítico da enzima *Xa*DHPS. **A.** Ligação do produto diidropteroato e do substrato 6HMPH2. **B.** Interação dos compostos DHP e 6HMPH2 no sítio ativo.

Fonte: Elaborada pelo autor.

Dessa maneira, é proposto que após a ligação do DHPP (anel pterina) os resíduos His263, Arg68, Asn27 estabilizam o modo de ligação do substrato e auxiliam na hidrólise do grupo pirofosfato. Essa tríade de resíduos interage, nas estruturas resolvidas, com moléculas de SO_4^{2-} que mimetizam o modo de ligação do fosfato ou com agente crioprotetor etilenoglicol. Na etapa seguinte da catálise, ocorre o movimento das alças $\alpha1$, $\alpha2$ e $\alpha5$, direcionando a entrada do pABA para o ataque nucleofílico ao carbocátion C9. É proposto que a liberação do produto

só ocorra na presença do íon Mg^{2+} .⁴³ No entanto, não encontramos densidade correspondente a este íon nas estruturas analisadas. A suposição deste mecanismo SN_1 de catálise foi feito para a homóloga de *B. anthracis*. Foram testados outros mecanismos de reação, mas com a elucidação de um estado de transição na homóloga de *Yersinia pestis* exclui-se a viabilidade do mecanismo SN_2 por não encontrarem indícios de hidrólise.⁴³

O composto ácido diidropteróico (DHP) foi avaliado no experimento de ITC (em inglês, *isothermal titration calorimetry*) com o objetivo de quantificar a afinidade do ligante, ou produto da catálise, com a enzima *XaDHPS*. Através da titulação de 500 μM foi possível traçar isoterma resultante dessa técnica. O valor de $K_d = 1,8 \mu M$ foi determinado com o ajuste no número de sítios igual a dois (**Figura 51**).

Figura 51 - Resultado do experimento de ITC quantificando a afinidade da enzima *XaDHPS* pelo produto da reação catalítica, o ácido diidropteróico (DHP). O valor de $K_d = 1.8 \mu M$ foi definido para dois sítios de ligação. Fonte: Elaborada pelo autor.

Os valores de afinidade obtidos são essenciais para confirmar a interação do composto DHP com a enzima, bem como a identificação de dois sítios de ligação. O número de sítios definidos corrobora com os dados cristalográficos e de exclusão molecular encontrado na purificação da enzima *XaDHPS*. Em plantas, o excesso de DHP faz a inibição pela retroalimentação. O mecanismo corrobora com os dados aqui encontrados: o preenchimento de ambos os sítios ativos, eliminando o número de sítios livres para a produção do DHP.⁸⁴ Isto posto, indicamos que a enzima funcione alternadamente em cada monômero por vez.

Na estrutura obtida em complexo com o pABA, notamos a presença da molécula do substrato em apenas um dos dois sítios ativos do dímero, mantido em interação a uma molécula de íon SO_4^{2-} , dadas as polaridades opostas. No total, foram posicionadas oito moléculas do ácido 4-aminobenzóico na estrutura composta por quatro monômeros de *XaDHPS*.

Especificamente, uma molécula está no sítio catalítico, uma na interface dimérica e duas na interface cristalográfica. O modo de ligação das outras quatro moléculas de pABA foi simétrico (eixo de ordem 2) ao modo de ligação descrito (**Figura 53**). A presença de duas moléculas de pABA na interface cristalográfica é, provavelmente, resultado de um artefato do experimento de imersão, e desta forma, consideramos na análise do modo de ligação as moléculas de pABA ligadas ao sítio catalítico e da interface dimérica.

Os resíduos Lys226, Phe195 e Pro69 são responsáveis pela estabilização do modo de ligação do substrato no sítio ativo. Além disso, a conformação da Phe195 altera para que haja a interação cátion- π com o substrato. O substituinte carboxilato do pABA interage com o NH da cadeia principal da R227 e com uma molécula de água (**Figura 52**).

Figura 52 - Representação da estrutura 3D em complexo com o respectivo substrato nativo, ácido 4-aminobenzóico. Na estrutura DHPS:pABA o reconhecimento das alças nas hélices $\alpha 1$, $\alpha 2$ e $\alpha 5$ são apresentados.

Fonte: Elaborada pelo autor.

A molécula de pABA identificada na interface entre os monômeros A e B interage também com resíduos da alça 1 da cadeia D. A cavidade é formada entre as hélices $\alpha 7$ e $\alpha 8$ das cadeias A e B. Nesta cavidade, o pABA é estabilizado por interações hidrofóbicas com os átomos da cadeia principal da Gly37D e átomos da cadeia lateral da V242A/B, A243A/B, A267A/B, V242A/B, A243A/B, A267A/B, bem como por contatos polares com moléculas de solvente (**Figura 54**).

Na forma nativa da enzima, o espaço equivalente ao sítio de ligação do pABA é ocupado pela cadeia lateral da Phe33, que pela alta flexibilidade gerada pela dupla possibilidade de conformações da Pro30 e a dupla de glicinas (Gly36 e Gly37), na extremidade C-terminal da alça 1, conseguem se movimentar para a acomodação do ligante (**Figura 53**). Essa conformação da alça 1 não foi observada nas estruturas homólogas depositadas no PDB, casos nos quais a alça está voltada para o topo do sítio ativo.

Figura 53 - Movimentação da alça 1 (Phe33) para a entrada e acomodação de ligantes na interface dimérica. Na ausência de ligantes (ciano), a alça 1 preenche o sítio com a F33. Na entrada de um ligante ocorre o giro da alça em sentido ao sítio ativo (amarelo), passando pelo estado de transição (magenta).

Fonte: Elaborada pelo autor.

A função inibitória e de regulação alostérica na interface foi proposta na estrutura da proteína homóloga de *Bacillus anthracis*.⁴¹ Mudanças na atividade foram observadas após testes com 36 ligantes diversos, os quais tiveram os modos de ligação confirmados por cristalografia de raios X e espectroscopia RMN. Após determinação dos valores cinéticos, estudos computacionais auxiliaram os autores a sugerir que a ligação de inibidores não impede a entrada dos substratos naturais e sucessiva catálise, mas compromete a saída do produto diidropteroato.

Como identificado em outras enzimas (Ribonuclease A), a enzima DHPS apresenta flexibilidade estrutural que lembra o movimento respiratório. Especificamente, as alças 1, 2 e 5 conjuntamente com as respectivas hélices atuam na abertura da interface (inspiração "∩") e fechamento da mesma (expiração "∪"). Uma vez que um ligante interaja com os resíduos deste bolsão, a função de movimentação está limitada, logo, não há liberação do produto diidropteroato.

A comparação das estruturas de DHPs atualmente disponíveis no PDB em complexo com o pABA indica que a composição dos resíduos difere na região equivalente ao sítio de ligação na interface dimérica. Por exemplo, as estruturas de *T. thermophilus* e *Y. pestis* não possuem moléculas do substrato ligadas na interface (**Figura 54**). A presença de cadeias laterais de grupos mais volumosos (A243H *Tt*DHPS e A243I *Yp*DHPS) nestas enzimas bloqueiam a cavidade que está totalmente livre em *X. albilineans*.

O ordenamento das alças com o preenchimento do sítio alostérico pelo substrato pABA é um fenômeno raro que foi observado somente em nossa estrutura. As enzimas ortólogas de *T. thermophilus*, *Yersinia pestis* e, mais recentemente, *P. vivax* apresentam as moléculas em complexo ternário no sítio catalítico. Interessantemente, mutações que conferem resistência à classe de antibióticos sulfonamidas são observadas nas regiões de alça catalítica e na extremidade C-terminal do sítio catalítico (face dimérica).

Figura 54 - Representação da molécula de pABA em complexo com o sítio da interface da *Xa*DHPS. Enzimas homólogas não dispõem da mesma característica do bolsão hidrofóbico. Em outras enzimas que ligam pABA há ocupação da cadeia lateral de resíduos mais volumosos que a A243 em *Xanthomonas albilineans*. Destacamos a presença de histidina em *T. thermophilus* (2DQW) e isoleucina em *Y. pestis* (5JQ9).

Fonte: Elaborada pelo autor.

Nossos dados reforçam a viabilidade da proposição de um sítio alostérico de inibição. Tanto a composição dos resíduos de aminoácidos quanto a conformação do sítio fazem com que a enzima *Xa*DHPS tenha pré-disposição em formar novos complexos nos ensaios de imersão de cristais. Por se tratar de uma região não conservada entre as enzimas homólogas, diferenças estruturais podem ser exploradas para o planejamento de inibidores seletivos para o

controle da escaldadura das folhas. Além disso, a estrutura de *XaDHPS* pode ser utilizada como modelo para a descoberta de novos agentes antimicrobianos com atividade seletiva para bactérias gram-negativa.

5.8 Planejamento de Moléculas baseado na Estrutura do Fragmento (FBDD)

A identificação de regiões e alvos moleculares são de suma importância no processo de planejamento de novas moléculas. Uma vez definidas, os métodos modernos de planejamento racional utilizam o seguinte regime: a. triagem de compostos; b. caracterização biofísica (elucidação do modo de ligação e caracterização dos complexos); c. planejamento e otimização de novas séries ou análogos. O objetivo é atingido mediante o atendimento das características de seletividade, absorção, distribuição, metabolismo, excreção e toxicidade (ADME-Tox).

O conceito de fragmento molecular engloba moléculas que possuam até 25 átomos (desconsiderando hidrogênio). De maneira ampla, bibliotecas destes compostos englobam e são classificadas por características similares denominada ‘regra dos três’: i. massa molecular ≤ 300 Da; ii. número de átomos aceptores e doadores de ligação de hidrogênio ≤ 3 ; iii. número de ligações rotacionáveis ≤ 3 ; iv. área de superfície polar $\leq 60 \text{ \AA}^2$ e $\text{cLogP} \leq 3$.^{85,86} A regra, consequentemente, seleciona compostos mais solúveis em meio aquoso, com menor lipofilicidade, o que torna favoráveis os testes de triagem contra alvos moleculares.⁸⁷

De maneira a adequar os fragmentos resultantes de baixa afinidade (e.g., 100-1000 μM) são necessárias medidas técnicas ortogonais da interação proteína-ligante. Métodos com nível de sensibilidade alto são indicados para aferir a relação do complexo. Nesse contexto, métodos como espectroscopia de ressonância magnética nuclear (RMN),⁸⁸ cristalografia de raios X (X-tal), ressonância plasmônica de superfície (SPR) e calorimetria de titulação isotérmica (ITC) são amplamente empregados para a triagem e caracterização de fragmentos como compostos bioativos. Nos últimos anos, outros métodos biofísicos têm sido aplicados no âmbito de FBDD, dentre os quais se destacam interferometria de bicamada (BLI), fluorimetria diferencial de varredura (DSF) e termoforese em microescala (MST).^{89,90}

São performados dupla confirmação da ligação dos fragmentos ao alvo molecular, aplicada em paralelo no planejamento de novas moléculas, a fim de evoluir a entidade molecular encontrada com conjuntamente com os melhores parâmetros de administração, seletividade, solubilidade, toxicidade, dentre outros.⁹¹

O método de desenvolvimento de novos fármacos baseado na estrutura do ligante está em constante evolução e aprimoramento. O desenvolvimento científico, sobretudo do

entendimento dos alvos moleculares, contribui diariamente para o desenvolvimento de novas terapias e controle de patógenos, sejam eles de cunho humano ou vegetal.

5.9 Termoforese em microescala

O experimento de termoforese em microescala baseia-se na análise da fluorescência quando um gradiente térmico é aplicado. Especificamente, a termoforese descreve o deslocamento proteico no gradiente térmico. A relação entre o movimento das moléculas da região de baixo calor para temperaturas maiores, e vice-versa, que no estado de equilíbrio estão distribuídas igualmente, podem ser medidos pela **Equação 2**.

$$\frac{c(calor)}{c(frio)} = \exp [-St * (T - T_0)] \quad (2)$$

A concentração relativa depende intrinsicamente do coeficiente de Soret (St) e da variação de temperatura empregada. O coeficiente de Soret é a medida da força da movimentação proteica em eletroforese quando comparado à difusão ordinária. Experimentalmente o coeficiente pode ser descrito por:

$$St = \frac{A}{kT} (-\Delta s_{hid}(T) + \frac{\beta \sigma_{eff}^2}{4\epsilon \epsilon_0 T} * \lambda_{DH}) \quad (3)$$

Esta equação traz os descritores envolvidos na variação da área de superfície molecular (A), da carga efetiva σ_{eff} e da variação entrópica causada pela interface molécula-solvente. Adicionalmente, o coeficiente depende da força iônica λ_{DH} , da constante dielétrica (ϵ) bem como a derivação da temperatura β . Portanto, o ensaio de termoforese mede as alterações causadas pela interação proteína-ligante, como por exemplo pelo rearranjo das águas Δ_{hid} (hidratação) presente no sítio ativo proteico para a acomodação do novo ligante.⁹² Pela sensibilidade da técnica é possível aferir qualquer transformação nos parâmetros descritos acima (e.g. afinidade de ligação).

Na realização dos experimentos, a marcação da proteína foi realizada com sucesso, entretanto, a eficiência de marcação foi baixa quando exploradas as aminas primárias. Considera-se de alta eficiência quando a quantidade de proteína necessária para realização do experimento esteja na escala baixo nanomolar (< 100 nM), mas para a enzima *XaDHPS* foi necessária uma quantidade mínima de 100 nM na realização experimental. Utilizamos tampão

carbonato, fornecido com o kit, uma vez que tampões à base de amina (Tris-HCl) são reativos com a sonda fluorescente.

O tipo de sonda utilizado (NT-647 RED-NHS) reage com aminas primárias (Lisina). A enzima *XaDHPS*, em sua composição residual, apresenta 4 lisinas por monômero. Deste modo, a taxa de eficiência da marcação depende diretamente do número de resíduos expostos ao solvente, ou que estejam próximos a cavidades que facilitem a acomodação da sonda fluorescente. Com isso, apenas a K226, catalítica, está alocada em região de fácil acesso e que possibilita a ligação da sonda de interesse. As outras três (K104, K197 e K184) estão voltadas para o centro proteico ou interagindo com a cadeia principal.

Da mesma maneira, há a perspectiva de marcação com sonda reagente aos resíduos de cisteína (MALEIMIDE), o qual está presente em duas unidades na *XaDHPS* (C213 e C10) por unidade monomérica **Figura 55**.

Similarmente às lisinas descritas, uma das duas cadeias laterais está voltada para o centro do barril e não faz face com o solvente. No entanto, a cisteína C213 está situada na parte inferior do barril e se mostra livre para a interação com a sonda. Ademais, o pH trabalhado na marcação garante aproximadamente 90 % das cisteínas na forma de tiolato (S^-) e que aumentam a chance de reação. Como resultado, a marcação da *XaDHPS* foi de alta eficiência e rendimento.

A titulação da proteína contra diferentes concentrações de tampão e DMSO foi inicialmente realizada para determinação da melhor condição de ensaio para a determinação da afinidade de ligação. O *software* do equipamento oferece a checagem da quantidade de fluorescência necessária a realização do experimento, bem como a informação sobre adsorção proteica aos capilares ou se há agregação da macromolécula no experimento.

Figura 55 - Resíduos disponíveis para reação com as sondas RED (amina primária - laranja) e MALEIMIDE (cisteína - amarelo). A exposição dos resíduos em direção ao solvente auxilia na marcação proteica. O sítio ativo está destacado pela superfície em verde.

Fonte: Elaborada pelo autor.

O segundo passo congrega a checagem da afinidade dos fragmentos com a proteína em valor único de alta concentração do ligante. Nesse ponto, a elevada solubilidade é essencial nos testes de interação e determinação da constante de afinidade entre proteína-ligante. Os candidatos a ligantes selecionados pela técnica de DSF foram avaliados por MST.

Todos os fragmentos foram testados, entretanto, algumas moléculas induzem a alteração na quantidade de fluorescência. A marcação utilizada neste experimento está baseada na ligação de amina primária. Sugerimos que a diminuição na fluorescência ocorre pela competição do ligante com a sonda fluorescente na entrada no sítio ativo, identificado pela lisina K226. Por conseguinte, ao alterar o lote de proteína marcada na Lys (*Xa*DHPS-lis) para a marcada na Cys (*Xa*DHPS-cis), notamos a interação proteína-ligante. A titulação do ligante de interesse, em diluições seriadas (16 concentrações), geraram os pontos de concentração-reposta necessários à determinação da constante de afinidade (K_d). A constante é estimada pelo cálculo das concentrações dos estados ligado (LM) e não ligado da proteína (M) e o ligante (L) em solução (**Equação 4**).⁹³

$$K_d = [L] \times \frac{[M]}{[LM]} \quad (4)$$

Não identificamos nenhum valor de afinidade para o ácido 4-aminobenzóico, o que sugere a necessidade do preparo do sítio de ligação para a sua entrada nos termos de mudanças conformacionais e alinhamentos de alças após a entrada do anel pterina. Nos derivados triazólicos, o composto com maior afinidade foi o **M9** ($K_d = 25 \mu\text{M}$). Apesar da baixa afinidade do fragmento **M7** ($K_d = 5 \text{ mM}$), podemos observar que a substituição ou o aumento da molécula com um grupo triazol resultou em uma afinidade 200 vezes maior no composto **M9**. Dado o arranjo cristalino e o posicionamento da alça 1 nos cristais de *Xa*DHPS, ocorre o impedimento estérico da entrada dos derivados triazólicos no sítio da interface. Por isso não obtivemos dados cristalográficos dos complexos com os triazóis.

5.10 Identificação de Ligantes por DSF para *Xa*DHPS

O objetivo deste experimento foi a identificação de fragmentos moleculares promissores para o desenvolvimento de inibidores para a enzima *Xa*DHPS. Em parceria com o Prof. Dr. Marcio V. B. Dias (ICB – USP/SP), realizamos a triagem de uma biblioteca de fragmentos

contendo, aproximadamente, 400 compostos. A biblioteca é bastante diversa e inclui moléculas com características desde fragmentos moleculares (e.g. massa molecular < 250 Da) até compostos fármaco-similar (e.g., massa molecular < 500, log P < 5, N+O < 10 e doador de ligação de H < 5).

No estudo de triagem por DSF foram identificados 5 fragmentos derivados de quatro classes químicas (quinolina, indol, ácido benzóico e piridinas). A interação desses fragmentos com a *XaDHPS* aumentou a temperatura média de desnaturação (ΔT_m). Além disso, identificamos 6 candidatos a ligantes (derivados triazólicos e pirazólicos) que possuem afinidade frente a enzima *XaDHPS* (aumento de ΔT_m entre 0,5-2 °C) (**Figura 56**).

Figura 56 - Compostos e fragmentos identificados por afinidade a enzima *XaDHPS* por experimentos de DSF.

A determinação da constante de afinidade contra a enzima *XaDHPS* realizada por MST.

Fonte: Elaborada pelo autor.

A ortogonalidade na confirmação biofísica dos hits foi testada com a imersão de cristais da *XaDHPS* em soluções contendo os ligantes identificados. Os cristais provenientes de semeadura foram acondicionados em solução concentrada dos ligantes a 28 mM por 1 h. Os dados cristalográficos foram coletados na linha de luz síncrotron (SSRL), da Universidade de Stanford – CA/USA. No total foram coletados 20 conjuntos de dados na linha BL 9-2. Os experimentos foram realizados em triplicatas, e apenas os arranjos cristalinos que visivelmente não sofreram com a adição do solvente orgânico e o acréscimo dos fragmentos foram enviados para coleta.

Os conjuntos foram processados com o programa XDS e a redução dos dados através do pacote de programas Aimless. Utilizamos o critério de $I/\sigma \sim 2$ para cortes de resolução

e a estrutura *XaDHPS* nativa (PDB 5ULQ) como molde para substituição molecular. Os parâmetros estatísticos da coleta e refinamento estão apresentados na **Tabela 7**.

Tabela 7 - Parâmetros estatísticos da coleta e do refinamento dos complexos entre os compostos (B9, C7 e E5) e a enzima *XaDHPS*.

	M2: <i>XaDHPS</i>	M7: <i>XaDHPS</i>	M6: <i>XaDHPS</i>
Comprimento de onda (Å)	0,97	0,97	0,97
Resolução (Å)	26,01 - 1,84 (1,906 - 1,84)	33,09 - 1,982 (2,053 - 1,982)	35,17 - 1,81 (1,875 - 1,81)
Grupo Espacial	C 1 2 1	C 1 2 1	C 1 2 1
Cela - Dimensão (Å)	109,687 75,797	109,89 76,345	109,631 75,624
Ângulos - cela α, β, γ (°)	82,568 90 126,832 90	82,744 90 126,695 90	82,515 90 126,871 90
Reflexões (total)	291655 (31664)	151588 (15093)	169410 (16794)
Reflexões (únicas)	42792 (4519)	37626 (3745)	48347 (4780)
Multiplicidade	6,8 (7,0)	4,0 (4,0)	3,5 (3,5)
Completeza (%)	90,23 (95,04)	94,22 (97,88)	97,24 (96,91)
I/sigma(I) médio	22,03 (2,88)	10,13 (2,81)	15,80 (2,90)
Wilson B-factor	34,74	37,23	32,30
R-merge	0,05774 (0,6676)	0,1531 (0,3273)	0,04274 (0,4199)
R-meas	0,06236 (0,721)	0,1773 (0,3772)	0,05051 (0,4976)
R-pim	0,02331 (0,2695)	0,08796 (0,1842)	0,02659 (0,2633)
CC 1/2	0,999 (0,973)	0,977 (0,977)	0,999 (0,964)
R-work	0,2159 (0,3569)	0,2556 (0,4070)	0,2177 (0,3424)
R-free	0,2402 (0,4296)	0,2709 (0,4342)	0,2388 (0,3761)
RMS (ligação)	0,013	0,013	0,013
RMS (ângulos)	1,65	1,66	1,62
Ramachandran favorável (%)	97,87	97,65	98,61
Ramachandran permitido (%)	1,94	2,35	1,39
Ramachandran outliers (%)	0,19	0,00	0,00
Clashscore	45,50	47,97	42,30

$$R = \frac{\sum_{hkl} |F_{obs} - F_{calc}|}{\sum_{hkl} F_{obs}}$$

$$R\text{-merge} = \frac{\sum_{hkl} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

$$R\text{-pim} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

Fonte: Elaborada pelo autor.

A ligação dos fragmentos identificados por DSF ocorreu na interface dimérica. O modo de ligação dos três compostos se assemelha ao descrito para o substrato pABA. Especificamente, as moléculas são mantidas por interações hidrofóbicas onde não foram encontradas moléculas de água ocupando o espaço na forma nativa.

Provavelmente, a baixa afinidade dos ligantes pelo sítio de acomodação resulte na ocupância demonstrada pela densidade eletrônica dos ensaios de imersão (**Figura 57**). Como observado nos compostos **M7** ($K_d = 5$ mM) e **M6**, a ocupância é menor quando comparada ao fragmento **M2** ($K_d = 374$ μ M). É desejável, para que se resolva complexos entre proteína-ligante o uso de, pelo menos, 10X a concentração do K_d .⁹⁴ Para estes compostos, os valores de concentração ideais ficaram comprometidos pela baixa solubilidade e limite de DMSO na interação com os cristais de *Xa*DHPS. Além disso, as imagens coletadas resultantes da interação com os compostos fármacos-similar (triazóis) apresentaram baixa completeza, sem nenhuma variação, em múltiplas faixas de resolução. Indicamos que a interação com os compostos quebra o arranjo cristalino em níveis não perceptíveis a olho nu.

Quando analisamos as moléculas por simetria é possível observar que a alça 1 auxilia a manutenção dos compostos na interface dimérica (**Figura 57**). A presença da estrutura secundária nesta posição limita o movimento e a possibilidade de assumir diferentes conformações dentro do sítio. A disposição das moléculas M2, M6 e M7 no sítio da interface é o primeiro passo da busca pela construção de novos ligantes para a enzima *Xa*DHPS.

Figura 57 - Complexos formados entre os fragmentos identificados por DSF com a enzima *Xa*DHPS. As moléculas estão atracadas na interface dimérica. As densidades eletrônicas estão representadas pelos mapas de contorno 2Fo-Fc (1 σ).

Fonte: Elaborada pelo autor.

5.11 Triagem de Inibidores por cristalografia de raios X

Estratégias de busca por novos esqueletos moleculares empregam variados métodos biofísicos para a identificação do modo de ligação e determinação da potência/afinidade dos agentes encontrados. Dessa forma, durante o período de intercâmbio no Departamento de Bioquímica da Universidade de Cambridge/UK, realizamos a triagem de uma biblioteca de compostos por cristalografia de raios X. O robusto protocolo de produção proteica construído conjuntamente com a alta reprodutibilidade dos cristais de *Xa*DHPS auxiliaram a condução dos experimentos.

A biblioteca comercial investigada, Zenobia, contém 352 fragmentos moleculares, os quais foram dispostos em coquetéis de 4 moléculas por solução. Todos os fragmentos são constituídos de pelo menos um anel, majoritariamente aromático (70 %). Os fragmentos são baseados nos fármacos complexados em alvos depositados no banco de dados PDB.

A grande vantagem do planejamento de novas moléculas com base no fragmento molecular é a possibilidade de construção do ligante mediante as particularidades de pequenos sítios dos alvos moleculares. Inúmeras companhias oferecem o serviço de triagem de esqueletos moleculares e identificação de sítios de ligação, no entanto, utilizam paralelamente outras técnicas para a caracterização. Um dos motivos é a baixo rendimento de resultados positivos pela cristalografia de raios X, por outro lado, a fidedignidade é alta, e a resposta traz consigo o modo de ligação do complexo (**Tabela 8**).⁹⁵

Tabela 8 - Lista com empresas que fornecem a infraestrutura para o desenvolvimento e triagem de fragmentos, bem como as técnicas utilizadas. Métodos abordados: X-ray (Cristalografia de raios X), ITC (Calorimetria de titulação isotérmica), NMR (Ressonância magnética nuclear), SPR (Ressonância plasmônica de superfície) e TR-FRET (Fluorescência resolvida no tempo).

Companhias	nº de compostos	Técnicas
Astex	1,600	X-ray, NMR, ITC
Beatica	1,946	SPR
BioFocus	1,500	X-tal
Biosensor	1,500	SPR
Crown Bio.	3,400	X-ray, SPR
CrystalsFirst	1,000	X-ray
Emerald	1,500	X-ray, NMR,
Evotec	1,500	X-ray, NMR
IOTA	5,500	X-ray, NMR, SPR
Kinetic Disc.	700	SPR
Kinomed	24,000	NMR
Pharma Polyphor	1,000	SPR
Proteros	8,000	X-ray, TR-FRET
Selcia	1,000	X-ray
Vernalis	1,400	SPR, ITC, NMR, X-ray
Vive Biotech	2,000	X-ray, NMR, SPR
XChem	700 por dia	X-ray
Zenobia	1,000	X-ray, SPR
ZoBio	1,500	NMR, SPR

Fonte: Elaborada pelo autor.

Na triagem de fragmentos da enzima *Xa*DHPS, os complexos proteicos foram obtidos pela imersão dos cristais em soluções contendo 25 mM de fragmento dissolvidos em 100 %, mantidos por 1 dia. Nem todos os compostos são solúveis em solução aquosa nesta concentração. Em poucos casos nos deparamos com a presença de precipitados por parte dos fragmentos. Nos cristais, em alguns casos, notamos a dissolução e quebra, o que nos leva a crer na desestabilização do arranjo cristalino pelos compostos adicionados e pelo efeito da dessolvatação causado tanto pelo DMSO quanto pelo coquetel.

Na elucidação dos complexos moleculares, realizamos os experimentos em algumas replicatas, dependendo da condição final dos cristais. Por meio do serviço remoto de coleta nas linhas de luz Síncrotron do Diamond, exploramos as diversas linhas de luz em 21 oportunidades (e.g. microfoco). No total, foram coletados mais de 300 conjuntos de imagens correspondentes a cada coquetel testado.

A estratégia de coleta seguiu aquelas delineadas dos conjuntos obtidos anteriormente na linha de luz MX2 do LNLS. Fizemos os devidos ajustes no protocolo para a exposição e comprimento de onda, mantendo o método de *fine ϕ -slicing* (0,2° /imagem) e 360° na angulação total. Portanto, cada conjunto possui 1800 imagens à alta resolução (>1.6 Å), que inevitavelmente varia dada a especificidade dos arranjos cristalinos encontrados.

O processamento, resolução e refinamento dos conjuntos seguiram os métodos modernos de *high throughput* empregado nos centros de descoberta e desenvolvimento de fármacos (e.g. *XChem*, Astex, dentre outros). Um exemplo de um dos conjuntos em complexo obtido pode ser visualizado na **Figura 58**.

Figura 58 - Representação da densidade eletrônica resultado do experimento de difração de raios X da enzima *XaDHPS*. Em amarelo está representado o fragmento em complexo com a proteína, que se liga tanto no sítio catalítico como na interface dimérica. As densidades eletrônicas estão representadas pelos mapas de contorno 2Fo-Fc (2,0 σ).

Fonte: Elaborada pelo autor.

Os vários fragmentos identificados nos complexos estão em diversas regiões da enzima *XaDHPS*. Além dos dois sítios de ligação descritos anteriormente, foram identificadas regiões de interação formadas pelo arranjo cristalino, dado o empacotamento apresentado. Uma vez que experimentos adicionais são necessários para a investigação da relevância biológica desses sítios de ligação, os esforços foram focados nas regiões de ligação do anel pterina, do sítio formado com o ácido 4-aminobenzóico e na interface dimérica (**Figura 59**).

Os complexos obtidos indicam a diversidade química esperada para uma triagem de fragmentos e que corroboram com a limitação em substratos descrita anteriormente. Todos os fragmentos moleculares em complexo são inéditos no planejamento de inibidores para a enzima DHPS. Entre as diferentes classes químicas identificadas, cita-se: imidazol, indol, catecol, benzoato, salicilato e, na grande maioria, derivados de piridina.

Figura 59 - Resultado da triagem de fragmentos contra a enzima *XaDHPS*. Os compostos atingiram, majoritariamente, quatro regiões enzimática: sítio ativo, interface dimérica, alça L1 e contatos cristalográficos. A biblioteca contém 352 compostos organizados em 88 coquetéis.

Fonte: Elaborada pelo autor.

Dividimos os resultados mediante o local de ligação: sítio anel pterina, sítio pABA, sítio da interface e contatos cristalográficos. O mapa da diferença (Fo-Fc) na densidade eletrônica foi analisado e interpretado com o programa automatizado Rhofit (Global Phasing), seguido da inspeção visual na identificação dos ligantes. Nossa taxa de sucesso na triagem de fragmentos foi de 10 %, a qual se enquadra perfeitamente nas mais variadas campanhas que utilizaram a mesma técnica.^{90,96} Em todos os complexos desse estudo, os hits estão claramente definidos dentro dos sítios pelas suas densidades eletrônicas. Nos termos de definição da densidade eletrônica, especificamente nos casos do acoplamento de moléculas na interface dimérica, essa elucidação da interação proteína-ligante não aparece em nenhuma estrutura depositada no PDB na referida resolução.

Encontramos duas moléculas no sítio de ligação do ácido 4-aminobenzóico (**PB1-2**), três fragmentos nas interfaces cristalográfica (**CC1-3**), 20 fragmentos alocados na interface dimérica (**IT1-20**) e oito fragmentos que se ligam na posição de interação do anel pterina da cavidade catalítica (**PT1-8**) (**Figura 60**).

Figura 60 - Fragmentos complexados na molécula de *XaDHPS* após a triagem de uma biblioteca de compostos por cristalografia de raios X. Os dados estão organizados por região de ligação na molécula: i. pABA (PB1-3), ii. contatos cristalográficos (CC1-3), iii. sítio da interface (IT1-20) e sítio do anel pterina (PT1-8).

Fonte: Elaborada pelo autor.

No total, são 33 fragmentos moleculares que exploram quatro regiões de ligação da enzima *XaDHPS*. Majoritariamente, os fragmentos contêm pelo menos um átomo de nitrogênio ligado ao anel benzênico. Em outros casos, há grupos imidazóis, alguns derivados do substrato natural, nitrocompostos, bioisósteros dos substratos e benzoaminas que fazem o acoplamento ao sítio da interface. Algumas moléculas se apresentam em mais de um sítio ou em mais de uma molécula por unidade dimérica.

No sítio de ligação do ácido 4-aminobenzóico, há 2 compostos que mimetizam as interações do substrato natural no sítio. O fragmento **PB1** interage apenas com a cadeia principal da Arg227 e uma molécula de água. Já o **PB2** comunica com a F195 em interação elétrons π (π - π) do tipo T e através do halogênio Cl⁻, por fim pela interação (cátion- π) com a Lys226. Ainda próximo, no mesmo sítio, mas em outro conjunto de dados, ocorre a ocupação da molécula **CC3** na posição da Thr67, favorecendo a interação com os compostos que se acomodam na cavidade do anel pterina (**Figura 61**).

Figura 61 - Sobreposição dos fragmentos PB1 (verde) e PB2 (ciano) ligados no sítio de interação do ácido 4-aminobenzóico (magenta). Os compostos PB1, PB2 se assemelham no modo de ligação com o substrato natural. A ligação da molécula PB3 favorece a interação com os compostos na posição pterina do sítio catalítico. A cadeia proteica é destacada pelos carbonos em verde.

Fonte: Elaborada pelo autor.

Os fragmentos identificados ligados na região de contatos cristalográficos estabilizam a alça 5, formando uma volta α não reportada até o momento. Além disso, elucidamos as quatro moléculas do coquetel complexadas na enzima *Xa*DHPS. Este fato inédito indica que ambas moléculas causam efeitos na enzima, tanto na ligação do sítio catalítico, quanto no enovelamento e estabilização. O resultado da alta estabilização pode ser notado pela densidade eletrônica nos compostos **PT4**, **CC1** e **CC2**. No entanto, nenhum dos 4 agentes foram encontrados no sítio da interface (**Figura 62**).

O composto mais bem definido na posição do anel pterina é o **PT4**. Ainda encontramos a ligação de uma pirimidina, a cloropurina **PT3**, que é análogo da pterina. Isso indica que a enzima *Xa*DHPS pode ligar outros análogos de pirimidina como o ATP. Este nucleotídeo faz parte da função quinase da enzima HPPK e pode interferir, tanto nos termos do ensaio acoplado, quanto em um futuro inibidor que atue em ambas enzimas.

Figura 62 - A enzima *XaDHPS* em complexo com quatro moléculas ao mesmo tempo. No sítio catalítico PT4 e PB3 se localizam de maneira próxima, enquanto CC1-2 se mantêm nas cavidades cristalográficas. As densidades eletrônicas estão representadas pelos mapas de contorno 2fo-fc (1,8 σ).

Fonte: Elaborada pelo autor.

Ademais, existem outros fármacos que fazem uso desse esqueleto molecular, como o Trimetoprim. Esta molécula liga-se à DHFR e possui em sua estrutura uma pirimidina e um substituinte trimetiltolueno. Ambas porções moleculares tiveram análogos identificados, **PT4** e **PT1**, no sítio catalítico da *XaDHPS* como resultado da triagem de fragmentos. No entanto, não obtivemos sucesso em complexar o Trimetoprim ao sítio catalítico por cristalografia de raios X.

Todos os fragmentos identificados interagem com o sítio ativo de modo semelhante ao produto DHP, ou seja, as interações do anel pterina com os resíduos polares do bolsão são reproduzidas nas novas moléculas. Por se tratar de compostos aromáticos, todos eles também formam interações π com a Arg261 e interações hidrofóbicas com a Met144. Nesses complexos, há uma pequena diferença na posição da lisina catalítica, que pode ser observado pela diferença na orientação da interação de hidrogênio (**Figura 63**).

As diferenças aparecem na outra extremidade, do nitrogênio N9, que está voltada para a abertura do sítio e livre para junção de novas extremidades. A amina primária do composto **PT8** e o oxigênio O9 do **IT15** estão livres para atuar como doadores de ligação de hidrogênio, pois estão próximos da carbonila do fragmento **PB3** e seriam passíveis de junção para formação de um novo composto (**Figura 63**).

A direção na qual estes fragmentos poderiam ser expandidos está apresentada na **Figura 64**. Se possível, durante o desenvolvimento futuro, gostaríamos de fugir da trivialidade de

congregar sulfonamidas e a área de ligação do pABA, uma vez que todos os inibidores descritos seguem esta métrica. Uma alternativa, por exemplo, seria explorar a outra face responsável pela ligação do pirofosfato e que em nossas estruturas interage ou com íons SO_4^{2-} ou etilenoglicol (EDO). A tríade de resíduos é formada pela His263, Arg261 e Asn27. Dessa forma, esses resíduos polares poderiam ser explorados para adicionar grupos que favoreçam a solubilidade da nova molécula, pois todos os compostos testados possuem solubilidade limitada para a determinação da afinidade de ligação.

Para a expansão molecular, sugerimos três canais de solvente representados pelos três retângulos da **Figura 63**. Os dois primeiros representam áreas de saída do produto diidropteroato, a área descrita com concentração de mutações relacionadas à resistência (verde e laranja), e o último a conexão do sítio catalítico com a interface dimérica (amarelo). Dependendo do tamanho das moléculas desenvolvidas seria possível construir uma entidade que acoplasse ambas as cavidades.

Figura 63 – **A.** Interações realizadas entre os hits complexados no sítio do anel pterina. **B.** A possibilidade de expansão dos diversos fragmentos está indicada pelas setas.

Fonte: Elaborada pelo autor.

Identificamos 20 novos esqueletos moleculares posicionados na interface entre os monômeros da *XaDHPS* (**Figura 64**). As diferenças consistem na distância da entrada do sítio que estas moléculas se encontram. Por se tratar de uma cavidade altamente hidrofóbica, as interações de moléculas fortemente polares não foram encontradas no sítio. O posicionamento rígido das hélices $\alpha 7$ e $\alpha 8$ se manteve em todos os conjuntos coletados.

A profundidade do sítio é de, aproximadamente, 14 Å, que representa espaço suficiente para ser ocupado por moléculas maiores. No entanto, o aumento da complexidade via ramificações laterais bloquearia a entrada no sítio. Por isso, as moléculas testadas que

tenham dois anéis aromáticos ligados por carbonos em cadeia maior que 2 não resultaram em complexos.

Entre os complexos obtidos, destacam-se duas moléculas que fogem o padrão de interação no sítio: **IT4** e **IT14**. Ambas moléculas atingem maior profundidade no sítio quando comparado a todas as outras e ao modo de ligação do substrato pABA. A substituição da carboxila do pABA para um grupo nitro faz o aprofundamento na cavidade. No mesmo sentido, o grupo fenol preenche a cavidade, com a cadeia carbônica sendo eficiente em interagir com toda o sítio na cavidade, com o grupo carboxila interagindo com a alça L1.

A característica comum é o posicionamento do heteroátomos (F, Cl, N) em direção à cavidade e os substituintes hidroxila ou imidazol voltados para a alça catalítica. Os fragmentos são estabilizados na entrada pela mudança conformacional da Phe33, por outro lado, ao se ligarem ocorre o bloqueio de movimento da alça L1 dentro da cavidade para que se adote a posição inicial. Consideramos ambas moléculas (**IT4** e **IT14**) como ótimo ponto inicial para o planejamento de novos inibidores, uma vez que divergem de todos os outros *hits* e preenchem por completo o bolsão hidrofóbico localizado na interface dimérica.

Figura 64 – A. Fragmentos complexados na interface dimérica da enzima *XaDHPS*. B. O sítio tem profundidade de 14 Å e características hidrofóbicas. O fragmento IT14 percorre toda a cavidade do sítio. A extremidade fenol está enterrada no bolsão, enquanto a outra extremidade faz face com o solvente e a alça 1. As densidades eletrônicas estão representadas pelos mapas de contorno 2fo-fc (1,5 σ).

Fonte: Elaborada pelo autor.

Apesar de existirem vários desafios técnicos a serem superados ao utilizarmos a cristalografia de raios X para a triagem de fragmentos, como a obtenção de cristais adequados

e solubilização de compostos, também se apresentam vantagens importantes sobre outras técnicas.

Embora outros métodos identifiquem fragmentos de baixa afinidade, a sua transformação em compostos líder, mantendo a eficiência de interação, só é possibilitada pelo conhecimento das orientações e interações proteína-ligante. Estas estruturas em complexos podem identificar posições do solvente, bem como o movimento conformacional proteico, ambos aspectos extremamente importantes na descoberta de ligantes como candidatos a novos fármacos. Além disso, a presença de falsos positivos na técnica de cristalografia de raios X é rara.⁹⁵

Ao identificar os múltiplos fragmentos, existem vários métodos para otimizar os *hits* encontrados. Um dos métodos empregados compreende a triagem de análogos que possuam ramificações e que explorem as cavidades e os resíduos adjacentes do fragmento inicial. Uma vez que nosso primeiro *hit* foi uma pirimidina (**PT4**), optamos por procurar novas possibilidades de ligantes próximos em direção aos canais de solvente sugeridos e que ainda não foram explorados na literatura. Para tanto, adquirimos dez análogos de pirimidina (**PT4**) da empresa Enamine.

Os compostos foram escolhidos segundo a disponibilidade de entrega, preço (< \$30) e a possibilidade de aumentar ou bloquear as interações de hidrogênio realizadas com a R261, a K226 e outros resíduos de aminoácidos próximos ao sítio catalítico. Similarmente, escolhemos 9 compostos derivados da cloropirimidina (encontrada em complexo na interface dimérica) e adquirimos da empresa Chembridge. A escolha foi feita com base no aumento na cadeia carbônica, mudança no heteroátomos para cloro, bromo, iodo ou nitrogênio (**Figura 65**).

Como método alternativo, mas ainda não explorado, existe o conhecido '*Fragment-linking*' que consiste em ligar moléculas que estejam em áreas próximas (e.g. **PB3** e **PT8** ou **PT4**) dentro bolsão de interesse. Quando dois fragmentos são identificados, estratégias de químicas sintética são utilizadas para ligar os compostos em diferentes formas, utilizando ligantes flexíveis entre as moléculas, mas com comprimentos variados. O planejamento destas novas moléculas é baseado na topologia do sítio de interação, como por exemplo o bloqueio das cadeias laterais da macromolécula. A prova de conceito é aplicada com a nova triagem por cristalografia, ou outro método biofísico, para determinação do modo de ligação e quantificação de potência/afinidade.

Figura 65 - Análogos adquiridos com base nos compostos cloropirimidina (interface) e piridina (sítio ativo).
Fonte: Elaborada pelo autor.

Testamos a afinidade dos análogos adquiridos através da alteração da temperatura e estabilidade térmica (T_m) na técnica de DSF. As mudanças positivas, que se referem a temperaturas de fusão maiores, indicam que a presença do substrato natural e o produto nativo auxiliam na manutenção proteica quando há o gradiente térmico. Neste caso, ambas moléculas se ligam no sítio ativo, variando os valores de ΔT_m em 14,1 °C e 2,9 °C para o DHP e 6HMPH2, respectivamente (**Figura 66**).

Figura 66 - Variações nos valores da curva de fusão (T_m) pelo método de DSF. Os compostos foram divididos entre variações positivas e negativas.

Fonte: Elaborada pelo autor.

Da mesma maneira, o decréscimo nos valores de T_m (ou ΔT_m negativo) também indicam ligação dos compostos, porém, em áreas não específicas que podem representar pontos de instabilidade (e.g. enovelamento incorreto, estrutura quaternária parcial, dentre outros).

Neste sentido, trabalhamos com a possibilidade das moléculas identificadas interagindo na interface dimérica ao alterarem o estado oligomérico da enzima quando em solução. O sítio da interface, como descrito, é altamente lipofílico e possui dois resíduos de triptofano na extremidade oposta à entrada do sítio da interface, o que aumenta as chances de serem pontos de ligação da sonda. Os valores variaram de $-0,5\text{ }^{\circ}\text{C}$ (**M7**) até $-17,0\text{ }^{\circ}\text{C}$ (sulfametoxazol – inibidor conhecido), confirmando a ligação das moléculas (**Figura 66**).

Os novos testes para obter os complexos com os análogos e identificação do modo de ligação, seguiram o protocolo descrito na triagem de inibidores por cristalografia de raios X. No entanto, a baixa solubilidade dos derivados de pirimidina prejudicaram o desenvolvimento do período de incubação dos cristais em solução com os compostos.

Ainda assim, após várias tentativas e coleta de dados, a ligação de alguns análogos foram confirmadas no sítio de ligação do anel pterina. A diminuição das aminas desloca as moléculas em direção a K226, diminuem a interação com o D101, além de reduzirem a ocupância no sítio ativo. As aminas remanescentes, como no composto **E7**, se curvam em direção à lisina catalítica para manter a ligação de hidrogênio (**Figura 67**).

Figura 67 - Empilhamento dos resultados dos complexos com os análogos de piridina identificadas pela triagem de fragmentos por cristalografia de raios X. Nd = não determinado.

Fonte: Elaborada pelo autor.

Já para os compostos com potencial de ligação na interface, confirmaram-se nossas expectativas. Não existe nenhum motivo aparente para a preferência de um sítio a outro. No entanto, alguns análogos têm características de ligação nas duas cavidades (e.g. **IT17**). Os

compostos ancorados na interface reproduzem o modo de ligação das moléculas anteriormente descritas.

A presença do composto **C3** em várias partes da enzima evidenciam o caminho percorrido pelos fragmentos ao se ligarem na interface. Os valores de afinidade moderada foram estimados com a técnica de MST (K_d entre 0,1 e 500 μM). A variação no número de sítios moleculares dificulta a determinação da constante de afinidade dos análogos testados. Adicionalmente, a baixa solubilidade nas condições do ensaio de MST comprometeu a determinação da maioria dos análogos de pirimidina.

Como resultado de todas as triagens, sugerimos a presença de um terceiro sítio de ligação, não apresentado na literatura, localizado entre a fita $\beta 1$ e a hélice $\alpha 8$. A cavidade rasa é formada pelos resíduos polares voltados ao solvente Q276, R19, D58 e complementado pela cadeia lateral hidrofóbica da Val267, Leu272, Gly56 e Ala266. Nas triagens iniciais encontramos quatro moléculas que se acomodaram neste bolsão. Até então, não existia nenhuma relação entre os sítios descritos e a posição dos ligantes.

Após a resolução dos complexos com os análogos do composto **M7**, confirmado com a piridina **E3**, nos deparamos com a proximidade do composto **IT17** no sítio da interface. A distância que separa os fragmentos da interface e o composto **IT17** é de, aproximadamente, 12 Å e possibilita a criação da junção de ambas moléculas nos termos da criação de um novo ligante (**Figura 68**).

Figura 68 - Proximidade de sítios na enzima *XaDHPs*. Um novo bolsão capaz de acomodar novos ligantes está presente entre a fita $\beta 1$ e a hélice $\alpha 8$. A cavidade presente em cada monômero facilita a comunicação entre o bolsão e o sítio da interface.

Fonte: Elaborada pelo autor.

Estes compostos são promissores para a criação de fármacos potentes, em termos farmacocinético e físico-químico, da mesma maneira em potência e especificidade. A junção da tecnologia, dos resultados apresentados e com o aperfeiçoamento de instalações de novos centros (e.g. SIRIUS – Campinas/BR) ascenderão, em breve, como fator primário nas etapas de construção de fármacos e/ou agroquímicos no Brasil.

CONSIDERAÇÕES FINAIS

6 Considerações Finais

6.1 Conclusões

A bactéria *Xanthomonas albilineans* é o agente causador da escaldadura das folhas em culturas de cana-de-açúcar, um dos principais problemas do cultivo. Com a dificuldade de detecção dos sintomas, que se assemelham a outras fitodoenças, a avaliação do impacto econômico é prejudicada pela falta de controle e pela ausência de tratamento biológico ou químico. Portanto, a urgência na descoberta e desenvolvimento de métodos efetivos de combate à bactéria gram-negativa nos levou a selecionar vias essenciais ao organismo para o estudo de descoberta de novas entidades moleculares como candidatos a inibidores.

Duas abordagens foram realizadas em paralelo. Em uma frente, a triagem fenotípica de compostos contra ao crescimento da bactéria *Xanthomonas albilineans* foi realizada com sucesso. Identificamos 11 derivados imidazólicos com moderado potencial inibitório no crescimento da bactéria causadora da doença em cana-de-açúcar.

Na segunda abordagem, inicialmente selecionamos duas enzimas pertencentes às vias vitais do patógeno: diidropteroato sintase (*XaDHPS*) e benzoato coenzima A ligase (*XaBCL*), as quais fazem parte da via de produção de folatos e da via de síntese de albicidinas, respectivamente. Em função dos estudos cinéticos da enzima *XaDHPS*, incluímos no projeto outras enzimas da via de folatos (diidrofolato redutase *XaDHFR* e 6-hidroximetil-7,8-diidropterina pirofosfoquinase *XaHPPK*). No mesmo sentido, incluímos a enzima diidrofolato redutase de humano (*HsDHFR*) para testar a atividade proteica das enzimas da via em ensaio de atividade acoplado.

Todos os genes que codificam para as enzimas inicialmente selecionadas foram clonados em vetores pET, gerando expressão proteica em sistema heterólogo de *E. coli*. A presença de proteínas expressas como corpos de inclusão foi solucionada pela alteração na construção gênica ou nas condições do ensaio empregado para a expressão (*e.g.*, pH).

Ensaio de purificação foram empregados e, no caso das enzimas *XaDHPS*, *XaDHFR*, *HsDHFR* e *XaHPPK*, foi possível a determinação de um protocolo robusto para isolamento (rendimento > 30 mg/L). Além disso, foi realizada a triagem e identificação de condições promissoras para cristalização da *XaDHPS*. Esses resultados permitiram a coleta de dados cristalográficos á alta resolução (>1,6 Å), processamento e determinação da estrutura 3D

inédita da enzima *XaDHPS* em duas formas (apoenzima, e em complexos com o ácido 4-aminobenzóico, o substrato 6HMPH2 e o produto diidropteroato).

A análise estrutural revelou o clássico enovelamento barril TIM, conhecido pelo arranjo (α/β_8). A análise das estruturas cristalográficas de *XaDHPS* permitiu o detalhamento da movimentação das alças catalíticas, bem como a elucidação da orientação dessas alças quando há a entrada de ligantes. Além disso, identificamos um possível sítio alostérico de inibição localizado na interface dimérica. Neste sentido, após a triagem da primeira biblioteca de fragmentos por DSF que resultou em 11 potenciais inibidores, elucidamos o modo de ligação de três dessas moléculas no sítio localizado na interface dimérica (M2, M6 e M7).

Durante o período de intercâmbio na Universidade de Cambridge (UK) triamos, por cristalografia de raios X, 352 fragmentos moleculares diversos no sentido de mapear possíveis regiões de ligação da enzima *XaDHPS*. Foram identificados 43 novos esqueletos moleculares com promissor potencial de inibição. A vasta maioria dos compostos (28, considerando o pABA) se liga ao sítio da interface, o que aumenta o nosso interesse por estas moléculas.

Para investigar com mais detalhes os resultados obtidos no período de intercâmbio, adquirimos 20 análogos dos *hits* primariamente identificados. O objetivo foi a ramificação das ligações dentro do sítio ativo mesmo que frustrada pela baixa solubilidade dos compostos, com afinidade confirmada por DSF. Por outro lado, o crescimento da cadeia alifática dos análogos da cloropirimidina (e.g. C3) produziram valiosos resultados de afinidade que variam entre 0,1 e 500 μM . Desta forma, identificamos diversos novos ligantes do sítio localizado na interface dimérica, que ainda foi pouco explorado na literatura, e que podem servir para a descoberta de novos agentes antimicrobianos seguros e seletivos.

6.2 Perspectivas

Na área de triagem fenotípica dos derivados imidazólicos, avaliaremos se há efeitos dos íons nos compostos com atividade no crescimento de *Xanthomonas albilineans*. Espera-se que, num futuro próximo, possamos identificar o modo de ação destes compostos para que seja feito o planejamento baseado na estrutura do alvo molecular (SBDD). Paralelamente, testaremos o funcionamento do ensaio acoplado para avaliação da atividade enzimática da *Xa*DHPS. Os dados serão valiosos para compor e caracterizar os fragmentos que se ligam na interface dimérica como inibidores alostérico. No mesmo sentido, construiremos novas moléculas que explorem as cavidades da interface dimérica e bolsão $\beta 1$ para a descoberta de novas entidades químicas potentes e seletivas para a enzima alvo.

REFERÊNCIAS

Referências

- 1 MCLAUCHLAN, K. The nature and longevity of agricultural impacts on soil carbon and nutrients: a review. **Ecosystems**, v. 9, n. 8, p. 1364–1382, 2006.
- 2 BUTTIMER, C. et al. Bacteriophages and bacterial plant diseases. **Frontiers in Microbiology**, v. 8, p. 34, 2017.
- 3 SPARKS, T. C., LORSBACH, B. A. Perspectives on the agrochemical industry and agrochemical discovery. **Pest Management Science**, v. 73, n. 4, p. 672–677, 2017.
- 4 DELANEY, J. et al. Modern agrochemical research: a missed opportunity for drug discovery? **Drug Discovery Today**, v. 11, n. 17–18, p. 839–845, 2006.
- 5 SIMONICH, S. L., HITEST, R. A. Global distribution of persistente organochlorine compounds. **Science**, v. 269, n. 5232, p. 1851–1854, 1995.
- 6 SNEDEKER, S. M. Pesticides and breast cancer risk: a review of DDT, DDE, and dieldrin. **Environmental Health Perspectives**, v. 109, n. 1, p. 35–47, 2001. doi: 10.1289/ehp.01109s135.
- 7 PAUL, S. M. et al. How to improve r&d productivity: the pharmaceutical industry's grand challenge. **Nature Reviews Drug Discovery**, v. 9, n. 3, p. 203–214, 2010.
- 8 LAMBERTH, C. et al. Current challenges and trends in the discovery of agrochemicals. **Science**, v. 341, n. 6147, p. 742–6, 2013.
- 9 CANTRELL, C. L., DAYAN, F. E., DUKE, S. O. Natural products as sources for new pesticides. **Journal of Natural Products**, v. 75, n. 6, p. 1231–1242, 2012.
- 10 DALY, H. E. Sustainable development—definitions, principles, policies. In: KEINER, M.(Ed.) **The future of sustainability** New York: Springer Verlag, 200. p. 39–53, doi:10.1007/1-4020-4908-0_2.
- 11 SHARACHCHANDRA, M. Sustainable development: a critical review. **World Development**, v. 19, n. 6, p.607-621,1991.
- 12 BRASIL. MINISTÉRIO DAS MINAS E ENERGIA. **Resenha Energética Brasileira 2018**. Brasília: 2018.
- 13 BRASIL. MINISTÉRIO DE MINAS E ENERGIA/EMPRESA DE PESQUISA ENERGÉTICA: MME/EPE. **Plano decenal de expansão de energia 2024**. Brasília: 2018.
- 14 SILVEIRA, J. L. et al. The benefits of ethanol use for hydrogen production in urban transportation. **Renewable and Sustainable Energy Reviews**, v. 13, n. 9, p. 2525–2534, 2009.
- 15 UNIÃO DOS PRODUTORES DE BIOENERGIA (UDOP). **A história da cana-de-açúcar - da antiguidade aos dias atuais**. Disponível em: <<https://www.udop.com.br/index.php?item=noticias&cod=993>>. Acesso em: 15 de maio 2019.
- 16 BORGES, J. M. M. The brazilian alcohol program: foundations, results, and perspectives. **Energy Sources**, v. 12, n. 4, p. 451–461, 1990.

- 17 LOARIE, S. R. et al. Direct impacts on local climate of sugar-cane expansion in Brazil. **Nature Climate Change**, v. 1, n. 2, p. 105–109, 2011.
- 18 TOKESHI, H. Doenças da cana-de-açúcar. In: KIMARI, H. et al. **Manual de fitopatologia: doenças das plantas cultivadas**. São Paulo: Agronomica Ceres, 1997.
- 19 ROTT, P. et al. Unravelling pathogenicity of *Xanthomonas albilineans*, the causal agent of sugarcane leaf scald. **International Sugar Journal**, v.113, n.1351,p.490-496,2011.
- 20 BIRCH, R. G., PATIL, S. S. Evidence that an albicidin-like phytotoxin induces chlorosis in sugarcane leaf scald disease by blocking plastid DNA replication. **Physiological and Molecular Plant Pathology**, v. 30, n. 2, p. 207–214, 1987.
- 21 HASHIMI, S. M. et al. DNA gyrase from the albicidin producer *Xanthomonas albilineans* has multiple-antibiotic-resistance and unusual enzymatic properties. **Antimicrobial Agents and Chemotherapy**, v. 52, n. 4, p. 1382–90, 2008.
- 22 RODRIGUEZ-R, L. M. et al. Genomes-based phylogeny of the genus *Xanthomonas*. **BMC Microbiology**, v. 12, n. 1, p. 43, 2012.
- 23 MENSI, I. et al. Breaking dogmas: the plant vascular pathogen *Xanthomonas albilineans* is able to invade non-vascular tissues despite its reduced genome. **Open Biology**, v. 4, n. 2, p. 130116–130116, 2014.
- 24 PIERETTI, I. et al. What makes *Xanthomonas albilineans* unique amongst xanthomonads? **Frontiers in Plant Science**, v. 6, p. 289, 2015. doi: 10.3389/fpls.2015.00289.
- 25 PIERETTI, I. et al. The complete genome sequence of *Xanthomonas albilineans* provides new insights into the reductive genome evolution of the xylem-limited Xanthomonadaceae. **BMC Genomics**, v. 10, n. 1, p. 616, 2009.
- 26 COCIANCICH, S. et al. The gyrase inhibitor albicidin consists of p-aminobenzoic acids and cyanoalanine. **Nature Chemical Biology**, v.11, n. 3, p. 195-197, 2015. doi:10.1038/nCHeMBIO.1734.
- 27 ROSTOCK, L. et al. Molecular insights into antibiotic resistance - how a binding protein traps albicidin. **Nature Communications**, v. 9, n. 1, p. 3095, 2018.
- 28 KRETZ, J. ET AL. Total synthesis of albicidin: a lead structure from *Xanthomonas albilineans* for potent antibacterial gyrase inhibitors. **Angewandte Chemie International Edition**, v. 54, n. 6, p. 1969–1973, 2015.
- 29 KERWAT, D. et al. Synthesis of albicidin derivatives: assessing the role of n-terminal acylation on the antibacterial activity. **ChemMedChem**, v. 11, n. 17, p. 1899–1903, 2016.
- 30 ZHANG, L., XU, J., BIRCH, R. G. Engineered detoxification confers resistance against a pathogenic bacterium. **Nature Biotechnology**, v. 17, n. 10, p. 1021–1024, 1999.
- 31 VIEWEG, L. et al. The albicidin resistance factor albd is a serine endopeptidase that hydrolyzes unusual oligoaromatic-type peptides. **Journal of American Chemical Society**, v. 137, n. 24, p. 7608, 2015.

- 32 PIERETTI, I. et al. Genomic insights into strategies used by *Xanthomonas albilineans* with its reduced artillery to spread within sugarcane xylem vessels. **BMC Genomics**, v. 13, n. 1, p. 658, 2012.
- 33 BOVY, A. et al. Folates in plants: research advances and progress in crop biofortification. **Frontiers in Chemistry**, v. 5, n. 21, 2017. doi: 10.3389/fchem.2017.00021.
- 34 HANSON, A. D., GREGORY, J. F. Folate biosynthesis, turnover, and transport in plants. **Annual Review of Plant Biology**, v. 62, p 105-125, 2011. doi: 10.1146/annurev-arplant-042110-103819.
- 35 GORELOVA, V. et al. Evolution of folate biosynthesis and metabolism across algae and land plant lineages. **Scientific Reports**, v. 9, n. 1, p. 5731, 2019.
- 36 ACHARI, A. et al. Crystal structure of the anti-bacterial sulfonamide drug target dihydropteroate synthase. **Natural Structural Biology**, v. 4, n. 6, p. 490–497, 1997.
- 37 LAWRENCE, M. C. et al. The three-dimensional structure of the bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/dihydropteroate synthase of *saccharomyces cerevisiae*. **Journal of Molecular Biology**, v. 348, n. 3, p. 655–670, 2005.
- 38 NZILA, A. Inhibitors of de novo folate enzymes in *Plasmodium falciparum*. **Drug Discovery Today**, v. 11, n. 19–20, p. 939–944, 2006.
- 39 WANG, Z. et al. A highly sensitive and class-specific fluorescence polarisation assay for sulphonamides based on dihydropteroate synthase. **Biosensors and Bioelectronics**, v. 70, p. 1–4, 2015. doi: 10.1016/j.bios.2015.03.016.
- 40 TALAWANICH, Y. et al. Use of bacterial surrogates as a tool to explore antimalarial drug interaction: synergism between inhibitors of malarial dihydrofolate reductase and dihydropteroate synthase. **Acta Tropica**, v. 149, p. 64–69, 2015. doi:10.1016/j.actatropica.2015.05.011.
- 41 HAMMOUDEH, D. I. et al. Identification and characterization of an allosteric inhibitory site on dihydropteroate synthase. **ACS Chemical Biology**, v. 9, p. 55, 2014. doi: 10.1021/cb500038g.
- 42 MORGAN, R. E. et al. Crystal structures of *Burkholderia cenocepacia* dihydropteroate synthase in the apo-form and complexed with the product 7,8-dihydropteroate. **BMC Structural Biology**, v. 11, n. 1, p. 21, 2011.
- 43 YUN, M.-K. et al. Catalysis and sulfa drug resistance in dihydropteroate synthase. **Science**, v. 335, n. 6072, p. 1110–4, 2012.
- 44 HAMPELE, I. C. et al. Structure and function of the dihydropteroate synthase from *staphylococcus aureus*. **Journal of Molecular Biology**, v. 268, n. 1, p. 21–30, 1997.
- 45 BWIJO, B. et al. High prevalence of quintuple mutant dhps/dhfr genes in *Plasmodium falciparum* infections seven years after introduction of sulfadoxine and pyrimethamine as first line treatment in Malawi. **Acta Tropica**, v. 85, n. 3, p. 363–373, 2003.
- 46 MHASHAL, A. R. et al. Evolutionary effects on bound substrate pK_a in dihydrofolate

reductase. **Journal of the American Chemical Society**, v. 140, n. 48, p. 16650–16660, 2018.

47 WELSCH, M. E. et al. Discovery of potent and selective leads against toxoplasma gondii dihydrofolate reductase via structure-based design. **ACS Medicinal Chemistry Letters**, v. 7, n. 12, p. 1124–1129, 2016.

48 SONG, X. et al. Molecular characteristics and serodiagnostic potential of dihydrofolate reductase from echinococcus granulosus. **Scientific Reports**, v. 7, 2017. doi: 10.1038/S41598-017-00643-5.

49 THORNBURG, C. K. et al. Kinetically and crystallographically guided mutations of a benzoate coa ligase (bada) elucidate mechanism and expand substrate permissivity. **Biochemistry**, v. 54, n. 40, p. 6230–6242, 2015.

50 HASHIMI, S. M., BIRCH, R. G. Functional analysis of genes for benzoate metabolism in the albicidin biosynthetic region of *Xanthomonas albilineans*. **Applied Microbiology and Biotechnology**, v. 87, n. 4, p. 1475–1485, 2010.

51 DAVIS, M. J. Evaluation of selective media and immunoassays for detection of *xanthomonas albilineans*, causal agent of sugarcane leaf scald disease. **Plant Disease**, v. 78, n. 1, p. 78, 1994.

52 RENIER, A. et al. Substrate specificity-conferring regions of the nonribosomal peptide synthetase adenylation domains involved in albicidin pathotoxin biosynthesis are highly conserved within the species *Xanthomonas albilineans*. **Applied and Environmental Microbiology**, v. 73, n. 17, p. 5523–5530, 2007. doi: 10.1128/AEM.00577-07.

53 ASLANIDIS, C., DE JONG, P. J. Ligation-independent cloning of pcr products (LIC-PCR). **Nucleic Acids Research**, v. 18, n. 20, p. 6069–6074, 1990. doi: 10.1093/nar/18.20.6069.

54 INOUE, H., NOJIMA, H., OKAYAMA, H. High efficiency transformation of *Escherichia coli* with plasmids. **Gene**, v. 96, n. 1, p. 23–28, 1990.

55 BERTANI, G. Lysogeny at mid-twentieth century: p1, p2, and other experimental systems. **Journal Of Bacteriology**, v. 186, n. 3, p. 595–600, 2004.

56 STUDIER, F. W. Protein production by auto-induction in high-density shaking cultures. **Protein Expression and Purification**, v. 41, n. 1, p. 207–234, 2005.

57 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.

58 BATTYE, T. G. G. et al. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. **Acta Crystallographica D: biological crystallography**, v. 67, n. 4, p. 271–81, 2011. doi: 10.1107/S0907444910048675.

59 KABSCH, W. XDS. **Acta Crystallographica D: biological crystallography**, v. 66, n. 2, p. 125–32, 2010. doi: 10.1107/S0907444909047337.

60 VONRHEIN, C. et al. Data processing and analysis with the autoPROC toolbox. **Acta Crystallographica D: structural biology**, v. 67, p. 293–302, 2011. doi: 10.1107/S0907444911007773.

- 61 EVANS, P. R., MURSHUDOV, G. N. How good are my data and what is the resolution? **Acta Crystallographica D: biological crystallography**, v. 69, n. 7, p. 1204–14, 2013. doi: 10.1107/s0907444913000061.
- 62 MCCOY, A. J. et al. Phaser crystallographic software. **Journal of Applied Crystallography**, v. 40, n. 4, p. 658–674, 2007. doi: 10.1107/s0021889807021206.
- 63 ADAMS, P. D. et al. *PHENIX*: a comprehensive python-based system for macromolecular structure solution. **Acta Crystallographica D: biological crystallography**, v. 66, n. 2, p. 213–221, 2010. doi: 10.1107/s0907444909052925.
- 64 BLANC, E. et al. Refinement of severely incomplete structures with maximum likelihood in BUSTER±TNT. **Acta Crystallographica D: structural biology**, v. 60, p. 2210–2221, 2004. doi: 10.1107/s0907444904016427.
- 65 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. doi: 10.1002/pro.3330.
- 66 EMSLEY, P. et al. Biological crystallography features and development of coot. **Acta Crystallographica D: structural biology**, v. 66, p. 486–501, 2010. doi: 10.1107/s0907444910007493.
- 67 VERMA, A., JOSHI, S., SINGH, D. Imidazole: having versatile biological activities. **Journal of Chemistry**, v. 2013, p. 1–12, 2013. doi: 10.1155/2013/329412.
- 68 REINHARDT, A. et al. Novel imidazolium salt–peptide conjugates and their antimicrobial activity. **Bioconjugate Chemistry**, v. 25, n. 12, p. 2166–2174, 2014. doi: 10.1021/bc500510c.
- 69 SANGER, F., NICKLEN, S., COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, v. 74, n. 12, p. 5463–7, 1977. doi: 10.1073/pnas.74.12.5463.
- 70 ZHANG, S. P., ZUBAY, G., GOLDMAN, E. Low-usage codons in *Escherichia coli*, yeast, fruit fly and primates. **Gene**, v. 105, n. 1, p. 61–72, 1991.
- 71 ZHAO, Y. et al. Development of a pterin-based fluorescent probe for screening dihydropteroate synthase. **Bioconjugate Chemistry**, v. 22, n. 10, p. 2110–2117. doi: 10.1021/bc200346e.
- 72 HONG, Y-L. et al. Inhibition of recombinant pneumocystis carinii dihydropteroate synthetase by sulfa drugs. **Antimicrobial Agents And Chemotherapy** v. 39, n. 8, 1995. doi: 10.1128/aac.39.8.1756.
- 73 NOPPONPUNTH, V. et al. Cloning and expression of mycobacterium tuberculosis and mycobacterium leprae dihydropteroate synthase in escherichia coli. **Journal of Bacteriology** v. 181, n. 21, 1999.
- 74 HEVENER, K. E. et al. Structural studies of pterin-based inhibitors of dihydropteroate synthase. **Journal Medicinal Chemistry**, v. 53, p. 166–177, 2010. doi: 10.1021/jm900861d.
- 75 GASTEIGER, E. et al. Protein analysis tools on the expasy server. In: WALKER, J. M. (Ed.) **The proteomics protocols handbook**. Totowa: Humana Press, p. 571–607, 2005. doi:

10.1385/1-59259-890-0:571.

76 RARAN-KURUSSI, S. et al. Removal of affinity tags with tev protease. **Methods Molecular Biology**, v. 1586, p. 221–230, 2017.

77 BILSLAND, E. et al. Plasmodium dihydrofolate reductase is a second enzyme target for the antimalarial action of triclosan. **Scientific Reports**, v. 8, n. 1, p. 1038, 2018.

78 ANTIKAINEN, N. M. et al. Conformation coupled enzyme catalysis: single-molecule and transient kinetics investigation of dihydrofolate reductase. **Biochemistry**, v. 44, n. 51, p. 16835–16843, 2005.

79 LI, M. Z., ELLEDGE, S. J. SLIC: a method for sequence- and ligation-independent cloning. **Methods Molecular Biology**, v. 852, 2012.

80 DURBIN, S. D., FEHER, G. Protein crystallization. **Methods Molecular Biology**, v. 852, p.51-9,2012. doi: 10.1007/978-1-61779-564-0_5

81 BABAOGLU, K. et al. Crystal structure of 7,8-dihydropteroate synthase from bacillus anthracis: mechanism and novel inhibitor design. **Structure**, v. 12, p. 1705–1717, 2004.

82 PEMBLE IV, C. W. et al. Crystal structure of the 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase and dihydropteroate synthase bifunctional enzyme from francisella tularensis. **PLoS ONE**, v. 5, n. 11, p. 14165, 2010.

83 MEUZELAAR, H., VREEDE, J., WOUTERSEN, S. Influence of glu/arg, asp/arg, and glu/lys salt bridges on α -helical stability and folding kinetics. **Biophysical Journal**, v. 110, n. 11, p. 2328–2341, 2016.

84 MOUILLON, J.-M., RAVANEL, P., DOUCE, R. Folate synthesis in higher-plant mitochondria : coupling between the dihydropterin pyrophosphokinase and the dihydropteroate synthase activities. **Biochemistry Journal** v. 363, pt.2, p.313-319,2002.

85 HAJDUK, P. J. Fragment-based drug design: how big is too big? **Journal of Medicinal Chemistry**, v. 49, n. 24, p. 6972–6976, 2006.

86 CONGREVE, M. et al. A “rule of three” for fragment-based lead discovery? **Drug Discovery Today**, v. 8, n. 19, p. 876–877, 2003.

87 REES, D. C. et al. Fragment-based lead discovery. **Nature Reviews Drug Discovery**, v. 3, n. 8, p. 660–672, 2004.

88 PELLECCIA, M. et al. Perspectives on NMR in drug discovery: a technique comes of age. **Nature Reviews Drug Discovery**, v. 7, n. 9, p. 738–745, 2008.

89 ERLANSON, D. A. et al. Twenty years on: the impact of fragments on drug discovery. **Nature Reviews Drug Discovery**, v. 15, n. 9, p. 605–619, 2016.

90 SCHIEBEL, J. et al. Six biophysical screening methods miss a large proportion of crystallographically discovered fragment hits: a case study. **ACS Chemical Biology**, v.11, n. 6, p. 1693-1701, 2016.

- 91 MURRAY, C. W., REES, D. C. The rise of fragment-based drug discovery. **Nature Chemistry**, v. 1, n. 3, p. 187–192, 2009.
- 92 WIENKEN, C. J. et al. Protein-binding assays in biological liquids using microscale thermophoresis. **Nature Communications**, v. 1, n. 1, p. 100, 2010.
- 93 JERABEK-WILLEMSSEN, M. et al. Microscale thermophoresis: interaction analysis and beyond. **Journal of Molecular Structure**, v. 1077, p. 101–113, 2014.
- 94 MÜLLER, I. Guidelines for the successful generation of protein-ligand complex crystals. **Acta Crystallographica D: structural biology**, v. 73, n. 2, p. 79–92, 2017. doi: 10.1107/S2059798316020271.
- 95 KESERÜ, G. M. et al. Design principles for fragment libraries: maximizing the value of learnings from pharma fragment-based drug discovery (FBDD) programs for use in academia. **Journal of Medicinal Chemistry**, v. 59, n. 18, p. 8189–8206, 2016.
- 96 HARTSHORN, M. J. et al. Fragment-based lead discovery using x-ray crystallography. **Journal Medicinal Chemistry**, v. 48, n. 2, p. 403–413, 2005.
- 97 UniProt: a hub for protein information. **Nucleic Acids Research**, v. 43, n. 1, p. 204–212, 2015.
- 98 Activities at the universal protein resource (UniProt). **Nucleic Acids Research**, v. 42, n. 1, p. 191–198, 2014.
- 99 ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403–410, 1990.
- 100 LARKIN, M. A. et al. Clustal w and clustal x version 2.0. **Bioinformatics**, v. 23, n. 21, p. 2947–2948, 2007.
- 101 JONES, D. T. Protein secondary structure prediction based on position-specific scoring matrices. **Journal of Molecular Biology**, v. 292, n. 2, p. 195–202, 1999. doi: 10.1006/jmbi.1999.3091.
- 102 BUCHAN, D. W. A., JONES, D. T. The PSIPRED protein analysis workbench: 20 years on. **Nucleic Acids Research**, v. 47, n. 1, p. 402–407, 2019.
- 103 ŠALI, A., BLUNDELL, T. L. Comparative protein modelling by satisfaction of spatial restraints. **Journal of Molecular Biology**, v. 234, n. 3, p. 779–815, 1993.
- 104 ESWAR, N. et al. Comparative protein structure modeling using Modeller. **Current Protocols in Bioinformatics**, v. 5, 2006. doi: 10.1002/0471250953.bi0506s15.
- 105 SHEN, M.-Y., SALI, A. Statistical potential for assessment and prediction of protein structures. **Protein Science : a publication of the Protein Society**, v. 15, n. 11, p. 2507–24, 2006. doi: 10.1110/ps.062416606.
- 106 JONES, G. et al. Development and validation of a genetic algorithm for flexible docking. **Journal of Molecular Biology**, v. 267, n. 3, p. 727–748, 1997.

- 107 GASTEIGER, E. ExPASy: the proteomics server for in-depth protein knowledge and analysis. **Nucleic Acids Research**, v. 31, n. 13, p. 3784–3788, 2003.
- 108 EL-GEBALI, S. et al. The Pfam protein families database in 2019. **Nucleic Acids Research**, v. 47, n. D1, p. D427–D432, 2019.
- 109 BATEMAN, A., COGGILL, P., FINN, R. D. DUFs: families in search of function. **Acta Crystallographica F: structural biology and crystallization communications**, v. 66, n. 10, p. 1148–1152, 2010. doi: 10.1107/S1744309110001685.

APÊNDICE I

Apêndice I

Modelagem por Homologia

No passo inicial da modelagem por homologia é fundamental a identificação de estruturas primária de proteínas que pertençam à mesma família e que possuam estruturas 3D determinada experimentalmente. A sequência de aminoácidos da enzima *XaBCL* foi obtida do servidor UniProt(código A1EAJ3).^{97,98} A busca por proteínas homólogas foi realizada com auxílio do algoritmo blastp, disponível no servidor BLAST.⁹⁹ Este algoritmo compara a sequência de interesse com sequências depositadas no banco de dados de proteínas (PDB), tendo como resultado uma lista de proteínas com variados graus de similaridade.

Os alinhamentos das sequências resultantes foram realizados com o servidor ClustalW.¹⁰⁰ Esta ferramenta faz o alinhamento das sequências de aminoácidos entre as proteínas homólogas e sugere, de maneira comparativa, as regiões conservadas. Por fim, para validar os resultados, empregamos o servidor PSIPRED para predizer as estruturas secundárias da *XaBCL*.^{101,102} Os modelos iniciais foram construídos com o programa Modeller 9.11,^{103,104} através da interface VitaMIn (desenvolvida pelos Dr. Fernando V. Maluf e Dr. Gustavo M. A. Lima, em nosso laboratório, e disponível em <https://github.com/gustalima/vitamin>). A validação do modelo 3D da *XaBCL* foi realizada através da verificação do diagrama de Ramachandran, assim como do gráfico de energia DOPE (em inglês, *Discrete Optimized Protein Energy*).¹⁰⁵ Como forma adicional de validação foi realizada a docagem molecular do substrato natural, benzoato, no modelo construído de *XaBCL* com a metodologia de ancoragem molecular pelo programa GOLD.¹⁰⁶

Modelo tridimensional por homologia da enzima *XaBCL*

A enzima *XaBCL* (E.C. 6.2.1.25) possui 696 aminoácidos e pertence à família das ligases formadoras de ligação carbono-enxofre. A massa molecular é de, aproximadamente, 75 KDa e ponto isoelétrico teórico 6,54 (calculados pelo servidor ExPASy).¹⁰⁷ De acordo com a análise realizada no servidor Pfam,¹⁰⁸ o monômero de *XaBCL* apresenta dois domínios: i. domínio de ligação da adenosina monofosfato (AMP) (resíduos 223-609) e ii. domínio de função desconhecida (do inglês, DUF – *domain of unknown function*), (resíduos 617-690). Segundo o banco de dados do Pfam, este tipo de enovelamento aparece em, aproximadamente,

30,000 estruturas disponíveis no banco, o que corresponde a 20 % das proteínas conhecidas. Este domínio apresenta-se conservado tanto em eucariotos quanto procariotos, entretanto, nem sempre é essencial para o desenvolvimento de uma função definida.¹⁰⁹

Realizamos a predição da estrutura secundária nos servidores PSIPRED e PDBSum, sugerindo regiões α/β e ponte dissulfeto na região N-terminal. O banco de dados CATH (do inglês, *Class Architecture Topology Homologous*) indica que a XaBCL (CATH 3.40.50.980) pertence à classe de proteínas alfa-beta, com arquitetura sanduíche de três camadas (do inglês, *3-layer sandwich*) disposta em $\alpha\beta\alpha$, e topologia correspondente ao enovelamento de Rossman.

Os estudos de modelagem molecular comparativa foram iniciados com a identificação de proteínas ortólogas que apresentam estruturas cristalográficas depositadas no PDB. Nesse sentido, utilizamos o algoritmo blastp para alinhar a sequência alvo com àquelas de estrutura 3D conhecida, retornando estatísticas com valores de identidade sequencial e probabilidades ocorrência (E-value). O resultado da busca indicou que as enzimas benzoato CoA ligase de *Burkholderia xenovorans* (PDB – 2V7B) e *Rhodopseudomonas palustris* (PDB – 4EAT) possuem similaridade com a proteína de *Xanthomonas albilineans*.

A identidade sequencial entre as enzimas *B. xenovorans* ou *R. palustris* com *Xanthomonas albilineans* é de 30 %. As duas estruturas estão relacionadas a bactérias gram-negativas com grande capacidade de metabolizar compostos aromáticos. Os dados de similaridade apresentados sugerem alta conservação de funções destas proteínas.

Por conseguinte, tendo em vista a necessidade de explorar o espaço amostral para a seleção da conformação mais representativa da XaBCL, 200 modelos moleculares foram gerados e triados pelo menor grau energético. O melhor modelo construído apresenta valores DOPE = 1,375 e r.m.s.d = 0,57 Å em relação às estruturas molde (**Figura 69**). Como forma de validação, empregamos a metodologia de redocking que consistiu em reproduzir o modo de ligação do ligante cristalográfico na estrutura do modelo construído.

Figura 69 - Representação do modelo 3D final gerado por homologia para a enzima XaBCL.
Fonte: Elaborada pelo autor.

O modelo indica a presença de dois domínios α/β e um sítio catalítico formado pelo barril β_6 torcido, o qual através da movimentação gera flexibilidade à proteína e interliga as extremidades moleculares. É possível observar o enovelamento de Rossmann presente na extremidade C-terminal e um sanduíche $\alpha\beta\alpha$ na porção N-terminal. O diagrama de Ramachandran indica 92,5 % dos resíduos em regiões favoráveis, e os demais em regiões não permitidas estão localizados em posições de alças expostas ao solvente.

O domínio de ligação ao AMP é responsável pelo reconhecimento molecular do benzoato e a coenzima A. O processo de esterificação entre o grupo tiol da coenzima A com o benzoato é mediado pela hidrólise de uma molécula de adenosina trifosfato (ATP). O produto benzoil-coenzima A é liberado conjuntamente com difosfato e AMP com auxílio do íon magnésio. O sítio catalítico possui diâmetro de 17 Å, sendo suficientemente espaçoso para a entrada do substrato e o cofator. Ademais, a baixa conservação catalítica comparada aos ortólogos descritos auxiliam no planejamento de moléculas bioativas.

Clonagem, Expressão e Purificação da enzima *XaBCL*

No campo experimental, mediante dificuldades de amplificação do gene *xabE* pelas cepas de *X. albilineans*, realizamos a compra do gene sintético com a empresa GeneScript. O material sintético foi otimizado para expressão em *E. coli*, e de maneira a facilitar a clonagem desta ORF adicionamos dois sítios de restrição flanqueando o seguimento gênico: BamHI downstream (5' → 3') e XhoI upstream (3' → 5'). O inserto recebido liofilizado está inserido no plasmídeo pUC57. Em seguida, após reidratação com água ultrapura uma alíquota foi retirada para os trabalhos de biologia molecular e o restante armazenado em freezer -20 °C. A propagação plasmidial foi realizada em células de *E. coli* DH-5 α , conforme descrito no protocolo INOUE.

Para os mesmos vetores descritos nas clonagens das ORF de folatos, inserimos em sistema LIC, o segmento gênico da enzima *XaBCL*. Todas as construções foram expressas utilizando cepas *E. coli* Rosetta 2 (DE3), entretanto, as proteínas se encontram nos debris celulares.

Investigou-se, então, alternativas para a expressão da enzima na fração sobrenadante. Foram avaliados na expressão celular da enzima *XaBCL*: i. adição do substrato nativo

(benzoato) no meio de expressão; ii. variação da temperatura; iii. concentração do agente indutor; iv. troca das cepas de expressão *E. coli* [BL-21(DE3), Arctic Express, pLyss(DE3)].

Após a expressão das construções, relacionamos a solubilidade com a introdução de aditivos no tampão de lise celular: i. 5 % sarcosil; ii. 1 % Triton X-100; iii. 5 % glicose; iv. 100 mM L-arginina; v. 5 % sacarose (**Figura 70**). Com a introdução do detergente sarcosil foi possível solubilizar a proteína *XaBCL* no vetor pET-M11, com expressão em células Rosetta 2 (DE3).

Por outro lado, apesar de conseguirmos solubilizar o alvo, há um limite para uso do detergente sarcosil em colunas de afinidade e assim reflete diretamente na possibilidade de purificação proteica. Testamos sem êxito a separação desta proteína dos debris celulares com a introdução do aditivo, no entanto, a provável formação de micelas mesmo em titulação do composto impede a ligação da proteína à coluna.

Figura 70 - Gel SDS-Page 10 % cotendo padrão de expressão da enzima *XaBCL* nos vetores pET/LIC e pET/Sumo. Todas as construções mantidas nos debris celulares. Direita = Adição de aditivos no tampão de lise celular para expressão em vetor pET-M11: i. 5 % sarcosil, ii. 1 % Triton X-100, iii. 5 % glicose, iv. 100 mM L-arginina, v. 5 % sacarose. Sol. = fração solúvel, Ins = fração insolúvel, TEV = clivagem com a enzima TEV protease, 2^a afinidade, GF = filtração em gel, por exclusão molecular e MK = marcador de peso molecular.

Fonte: Elaborada pelo autor.

A *XaBCL* nativa possui 696 aminoácidos, o que corresponderia a uma alta demanda energética para expressão no sistema heterólogo escolhido. Adicionalmente, o domínio flexível demonstrado pela modelagem comparativa sugere ser uma barreira para o enovelamento correto e manutenção da estabilidade em solução aquosa. Diante disso, investigamos se a remoção do domínio α/β na extremidade N-terminal auxiliaria positivamente na obtenção de proteína solúvel e estável. Esse domínio de características hidrofóbicas e flexível se localiza distante do sítio catalítico (**Figura 71**).

Figura 71 - Esquema representativo da remoção dos 200 resíduos iniciais da enzima benzoato coenzima A ligase de *Xanthomonas albilineans*. Em cinza o domínio removido na versão truncada da enzima. O sítio de ligação do AMP está colorido por cilindros verdes.

Fonte: Elaborada pelo autor.

Nesse sentido propusemos seis novas construções que exploram outros vetores. Escolhemos vetores que congregassem a expressão conjunta de uma cauda de histidina maior no sentido de aumentar a afinidade do complexo pela coluna de afinidade. Abrimos como possibilidade a inserção da cauda de histidina na extremidade C-terminal.

As montagens englobam a proteína nativa com os respectivos 696 amino ácidos, a remoção dos 230 resíduos iniciais, bem como uma terceira montagem sem os primeiros 244 aminoácidos. A escolha do resíduo de partida é baseada em uma metionina ou para diminuir a alça resultante do encurtamento da construção, respectivamente. Também construímos dois arranjos distintos para tentar a expressão solúvel da porção inicial afastada do domínio AMP (**Figura 72**).

Figura 72 - Esquema com as diferentes construções da enzima Benzoato coenzima A ligase de *Xanthomonas albilineans* (*XaBCL*).

Fonte: Elaborada pelo autor.

Como resultado, todas as construções foram amplificadas e clonadas produzindo fragmento de tamanho correspondente ao esperado para cada inserto. A possibilidade de trabalhar com o sistema de clonagem em paralelo facilita a ramificação em vários segmentos no sentido de obter a melhor condição de expressão proteica. Com a simples repetição dos passos foi possível produzir oito novos arranjos da enzima *XaBCL*.

Testes preliminares de expressão indicaram proteína referente a construção criada em fração solúvel. Em um futuro, será realizado o aprofundamento do estudo de estabilidade, purificação e cristalização destas estruturas. O entendimento da disposição atômica desta proteína nos auxiliará no planejamento de novos inibidores baseados na estrutura do alvo molecular.

Apêndice II

Parâmetros estatísticos dos conjuntos de dados abordados neste documento:

	PT3/IT12	IT20	IT5
Comprimento de onda (Å)	0,97	0,97	0,97
Resolução (Å)	42,09 - 1,73 (1,792 - 1,73)	39,67 - 1,63 (1,688 - 1,63)	25,43 - 1,845 (1,911 - 1,845)
Grupo Espacial	P 1 2 1 1	P 1 2 1 1	C 1 2 1
Cela - Dimensão (Å)	82,382 74,554	82,37 74,004 89,078	109,375 75,168
Ângulos - cela α, β, γ (°)	89,299 90 101,213 90	90 100,836 90	82,579 90 126,862 90
Reflexões (total)	754442 (71913)	857088 (82712)	316659 (31041)
Reflexões (únicas)	108834 (10733)	126908 (13016)	46104 (4552)
Multiplicidade	6,9 (6,7)	6,8 (6,4)	6,9 (6,8)
Completeza (%)	97,81 (96,68)	96,54 (99,69)	99,03 (98,42)
I/sigma(I) médio	14,87 (2,65)	16,70 (2,57)	14,57 (2,70)
Wilson B-factor	24,37	21,97	31,54
R-merge	0,07982 (0,7693)	0,06558 (0,6644)	0,07542 (0,5138)
R-meas	0,08646 (0,8345)	0,07104 (0,7242)	0,08164 (0,5564)
R-pim	0,03289 (0,3205)	0,02705 (0,2846)	0,03098 (0,2116)
CC 1/2	0,997 (0,921)	0,998 (0,903)	0,998 (0,983)
R-work	0,2114 (0,3330)	0,1919 (0,2715)	0,2304 (0,4173)
R-free	0,2366 (0,3564)	0,2163 (0,3098)	0,2491 (0,4490)
RMS (ligação)	0,013	0,013	0,013
RMS (ângulos)	1,54	1,53	1,63
Ramachandran favorável (%)	97,88	98,22	98,04
Ramachandran permitido (%)	2,02	1,58	1,96
Ramachandran outliers (%)	0,10	0,20	0,00
B-fator médio	34,29	29,5	43,05

$$R = \frac{\sum_{hkl} |F_{obs} - F_{calc}|}{\sum_{hkl} F_{obs}}$$

$$R\text{-merge} = \frac{\sum_{hkl} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl}}$$

$$R\text{-pim} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl}}$$

	IT4	PB2	IT9
Comprimento de onda (Å)	0,97	0,97	0,97
Resolução (Å)	36,28 - 1,85 (1,916 - 1,85)	27,4 - 1,902 (1,97 - 1,902)	43,74 - 1,821 (1,886 - 1,821)
Grupo Espacial	P 1 2 1 1	C 1 2 1	P 1 2 1 1
Cela - Dimensão (Å)	82,221 73,628	109,653 74,488	82,348 74,497
Ângulos - cela α, β, γ (°)	88,709 90 100,489 90	82,286 90 126,97 90	89,147 90 101,095 90
Reflexões (total)	608545 (59037)	284321 (27155)	648752 (63215)
Reflexões (únicas)	88954 (8819)	41189 (4039)	94615 (9363)
Multiplicidade	6,8 (6,7)	6,9 (6,7)	6,9 (6,7)
Completeza (%)	99,40 (99,38)	98,32 (96,91)	99,48 (99,36)
I/sigma(I) médio	10,54 (2,36)	11,65 (3,28)	15,01 (2,63)
Wilson B-factor	23,93	32,71	25,21
R-merge	0,1194 (0,7921)	0,1539 (0,6123)	0,08 (0,7196)
R-meas	0,1294 (0,8597)	0,1667 (0,6645)	0,08663 (0,779)
R-pim	0,04921 (0,3307)	0,06351 (0,2562)	0,03293 (0,2958)
CC 1/2	0,998 (0,926)	0,985 (0,933)	0,997 (0,919)
R-work	0,2260 (0,3079)	0,2124 (0,3256)	0,2025 (0,2934)
R-free	0,2493 (0,3396)	0,2357 (0,3083)	0,2309 (0,3355)
RMS (ligação)	0,013	0,013	0,013
RMS (ângulos)	1,54	1,56	1,52
Ramachandran favorável (%)	97,76	97,82	98,06
Ramachandran permitido (%)	2,24	2,18	1,94
Ramachandran outliers (%)	0,00	0,00	0,00
B-fator médio	31,54	41,97	33,04

$$R = \frac{\sum_{hkl} |F_{obs} - F_{calc}|}{\sum_{hkl} F_{obs}}$$

$$R\text{-merge} = \frac{\sum_{hkl} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

$$R\text{-pim} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

	PT6	IT6	IT1
Comprimento de onda (Å)	0,97	0,97	0,97
Resolução (Å)	43,7 - 1,64 (1,699 - 1,64)	40,46 - 1,581 (1,638 - 1,581)	31,13 - 1,61 (1,668 - 1,61)
Grupo Espacial	P 1 2 1 1	P 1 2 1 1	P 1 2 1 1
Cela - Dimensão (Å)	82,511 74,198	82,41 74,837 89,029	82,312 74,621 89,09
Ângulos - cela α, β, γ (°)	88,954 90 100,724 90	90 100,945 90	90 101,067 90
Reflexões (total)	878766 (86532)	949872 (97880)	892336 (91589)
Reflexões (únicas)	129171 (12839)	138590 (14193)	132817 (13544)
Multiplividade	6,8 (6,7)	6,9 (6,9)	6,7 (6,7)
Completeza (%)	99,44 (99,33)	95,20 (98,14)	96,46 (99,43)
I/sigma(I) médio	13,61 (2,36)	15,82 (2,45)	13,94 (2,39)
Wilson B-factor	21,37	20,82	17,63
R-merge	0,08259 (0,6398)	0,06851 (0,9154)	0,1261 (0,7757)
R-meas	0,08939 (0,6934)	0,07421 (0,9896)	0,1367 (0,8403)
R-pim	0,03389 (0,2646)	0,02825 (0,3731)	0,05226 (0,3201)
CC 1/2	0,997 (0,934)	0,998 (0,896)	0,995 (0,913)
R-work	0,2016 (0,3213)	0,2071 (0,3171)	0,2060 (0,3146)
R-free	0,2337 (0,3286)	0,2296 (0,3274)	0,2324 (0,3304)
RMS (ligação)	0,013	0,013	0,013
RMS (ângulos)	1,51	1,55	1,52
Ramachandran favorável (%)	98,53	98,16	98,26
Ramachandran permitido (%)	1,47	1,84	1,74
Ramachandran outliers (%)	0,00	0,00	0,00
B-fator médio	30,87	29,42	25,66

$$R = \frac{\sum_{hkl} |F_{obs} - F_{calc}|}{\sum_{hkl} F_{obs}}$$

$$R\text{-merge} = \frac{\sum_{hkl} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

$$R\text{-pim} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

	D7_2	PB1/PT1	IT11
Comprimento de onda (Å)	0,97	0,97	0,97
Resolução (Å)	40,44 - 1,5 (1,554 - 1,5)	43,8 - 1,68 (1,74 - 1,68)	54,41 - 1,86 (1,927 - 1,86)
Grupo Espacial	P 1 21 1	P 1 21 1	C 1 2 1
Cela - Dimensão (Å)	82,354 74,24 89,126	82,503 74,852	109,14 74,257
Ângulos - cela α, β, γ (°)	90 100,848 90	89,138 90 100,635 90	82,191 90 127,051 90
Reflexões (total)	1148482 (112767)	802597 (78888)	302874 (30396)
Reflexões (únicas)	166702 (16340)	118230 (11883)	44094 (4357)
Multiplicidade	6,9 (6,9)	6,8 (6,6)	6,9 (6,9)
Completeza (%)	98,56 (97,31)	96,40 (97,56)	99,09 (98,66)
I/sigma(I) médio	16,38 (2,70)	17,17 (2,76)	11,28 (2,43)
Wilson B-factor	18,72	26,21	27,48
R-merge	0,06384 (0,78)	0,05391 (0,6945)	0,1427 (1,091)
R-meas	0,06915 (0,8429)	0,05856 (0,7533)	0,1542 (1,178)
R-pim	0,02631 (0,3173)	0,02257 (0,2895)	0,05795 (0,4429)
CC 1/2	0,998 (0,922)	0,998 (0,954)	0,992 (0,9)
R-work	0,1902 (0,3093)	0,2338 (0,4303)	0,2454 (0,4173)
R-free	0,2144 (0,3344)	0,2610 (0,4601)	0,2815 (0,4808)
RMS (ligação)	0,013	0,013	0,014
RMS (ângulos)	1,55	1,57	1,64
Ramachandran favorável (%)	98,27	98,52	97,66
Ramachandran permitido (%)	1,63	1,48	2,15
Ramachandran outliers (%)	0,10	0,00	0,20
B-fator médio	27,95	37,24	37,57

$$R = \frac{\sum_{hkl} |F_{obs} - F_{calc}|}{\sum_{hkl} F_{obs}}$$

$$R\text{-merge} = \frac{\sum_{hkl} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

$$R\text{-pim} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

	PT7	IT8	PT5
Comprimento de onda (Å)	0,97	0,97	0,97
Resolução (Å)	27,4 - 1,902 (1,97 - 1,902)	40,47 - 1,49 (1,543 - 1,49)	40,55 - 1,785 (1,848 - 1,785)
Grupo Espacial	C 1 2 1	P 1 2 1 1	P 1 2 1 1
Cela - Dimensão (Å)	109,653 74,488	82,454 74,896 89,13	82,658 75,8 89,841
Ângulos - cela α, β, γ (°)	82,286 90 126,97 90	90 100,982 90	90 101,152 90
Reflexões (total)	284321 (27155)	1180810 (116009)	707743 (69364)
Reflexões (únicas)	41189 (4039)	173590 (17219)	103535 (10259)
Multiplplicidade	6,9 (6,7)	6,8 (6,7)	6,8 (6,7)
Completeza (%)	98,32 (96,91)	99,54 (99,43)	99,41 (99,28)
I/sigma(I) médio	11,65 (3,28)	17,48 (2,68)	11,21 (2,42)
Wilson B-factor	32,71	20,52	25,93
R-merge	0,1539 (0,6123)	0,05881 (0,6945)	0,1113 (1,006)
R-meas	0,1667 (0,6645)	0,06376 (0,7541)	0,1206 (1,092)
R-pim	0,06351 (0,2562)	0,02437 (0,2902)	0,04589 (0,4201)
CC 1/2	0,985 (0,933)	0,998 (0,946)	0,996 (0,878)
R-work	0,2124 (0,3256)	0,2575 (0,3734)	0,2360 (0,3872)
R-free	0,2357 (0,3083)	0,2692 (0,3766)	0,2636 (0,4110)
RMS (ligação)	0,013	0,013	0,013
RMS (ângulos)	1,56	1,54	1,53
Ramachandran favorável (%)	97,82	98,37	98,26
Ramachandran permitido (%)	2,18	1,54	1,74
Ramachandran outliers (%)	0,00	0,10	0,00
B-fator médio	41,97	30,75	36,97

$$R = \frac{\sum_{hkl} |F_{obs} - F_{calc}|}{\sum_{hkl} F_{obs}}$$

$$R\text{-merge} = \frac{\sum_{hkl} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl}}$$

$$R\text{-pim} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl}}$$

	CC1-4/PT4	PT2	IT2
Comprimento de onda (Å)	0,97	0,97	0,97
Resolução (Å)	40,66 - 1,602 (1,659 - 1,602)	40,65 - 1,74 (1,802 - 1,74)	41,9 - 1,854 (1,92 - 1,854)
Grupo Espacial	C 1 2 1	C 1 2 1	P 1 2 1
Cela - Dimensão (Å)	115,128 72,661	109,998 74,177	82,38 74,664 89,157
Ângulos - cela α, β, γ (°)	82,119 90 126,926 90	82,685 90 127,141 90	90 100,601 90
Reflexões (total)	463260 (47853)	361118 (36316)	575538 (54186)
Reflexões (únicas)	66993 (6857)	53521 (5386)	87226 (8949)
Multiplicidade	6,9 (7,0)	6,7 (6,7)	6,6 (6,0)
Completeza (%)	94,00 (96,38)	97,29 (98,10)	96,13 (99,49)
I/sigma(I) médio	19,45 (2,16)	18,33 (2,91)	9,97 (2,64)
Wilson B-factor	15,02	28,03	24,24
R-merge	0,1274 (0,872)	0,06868 (0,5942)	0,151 (0,8724)
R-meas	0,1378 (0,9419)	0,0744 (0,6445)	0,1638 (0,9573)
R-pim	0,05204 (0,3536)	0,02832 (0,2471)	0,06284 (0,3887)
CC 1/2	0,996 (0,542)	0,998 (0,977)	0,991 (0,31)
R-work	0,1844 (0,2661)	0,2277 (0,4637)	0,2183 (0,3293)
R-free	0,2119 (0,2879)	0,2487 (0,4730)	0,2444 (0,3750)
RMS (ligação)	0,020	0,013	0,013
RMS (ângulos)	1,70	1,62	1,55
Ramachandran favorável (%)	97,69	96,72	97,77
Ramachandran permitido (%)	1,92	2,89	2,23
Ramachandran outliers (%)	0,38	0,39	0,00
B-fator médio	20,76	40,96	32,71

$$R = \frac{\sum_{hkl} |F_{obs} - F_{calc}|}{\sum_{hkl} F_{obs}}$$

$$R\text{-merge} = \frac{\sum_{hkl} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

$$R\text{-pim} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

	IT3	EDTA	IT15
Comprimento de onda (Å)	0,97	0,97	0,97
Resolução (Å)	27,24 - 1,95 (2,02 - 1,95)	41,69 - 1,78 (1,844 - 1,78)	27,26 - 1,903 (1,971 - 1,903)
Grupo Espacial	C 1 2 1	P 1 2 1 1	C 1 2 1
Cela - Dimensão (Å)	109,432 75,094	82,528 74,537	109,365 74,234
Ângulos - cela α, β, γ (°)	82,333 90 126,921 90	88,704 90 100,388 90	82,375 90 127,024 90
Reflexões (total)	254286 (25898)	687132 (67284)	281373 (27611)
Reflexões (únicas)	37147 (3812)	101424 (10035)	41343 (4075)
Multiplividade	6,8 (6,8)	6,8 (6,7)	6,8 (6,8)
Completeza (%)	94,69 (98,13)	98,68 (99,03)	98,83 (98,19)
I/sigma(I) médio	10,75 (2,66)	11,97 (2,93)	15,38 (2,86)
Wilson B-factor	33,16	21,33	35,79
R-merge	0,1736 (0,8784)	0,2726 (0,9084)	0,08153 (0,636)
R-meas	0,188 (0,9529)	0,2951 (0,9853)	0,08829 (0,6891)
R-pim	0,0715 (0,3654)	0,1122 (0,3778)	0,03359 (0,263)
CC 1/2	0,983 (0,918)	0,983 (0,898)	0,998 (0,97)
R-work	0,2407 (0,3661)	0,2260 (0,3177)	0,2673 (0,4493)
R-free	0,2802 (0,4436)	0,2543 (0,3276)	0,2916 (0,4459)
RMS (ligação)	0,013	0,013	0,013
RMS (ângulos)	1,64	1,60	1,68
Ramachandran favorável (%)	97,67	96,74	96,67
Ramachandran permitido (%)	2,33	2,49	3,33
Ramachandran outliers (%)	0,00	0,77	0,00
B-fator médio	43,28	28,98	50,33

$$R = \frac{\sum hkl |F_{obs} - F_{calc}|}{\sum hkl F_{obs}}$$

$$R\text{-merge} = \frac{\sum hkl \sum j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum hkl \sum j I_{hkl}}$$

$$R\text{-pim} = \frac{\sum hkl \sqrt{\frac{n}{n-1}} \sum j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum hkl \sum j I_{hkl}}$$

	IT14	IT19	IT20
Comprimento de onda (Å)	0,97	0,97	0,97
Resolução (Å)	43,73 - 1,84 (1,906 - 1,84)	43,64 - 1,782 (1,846 - 1,782)	25,41 - 2,192 (2,271 - 2,192)
Grupo Espacial	P 1 2 1 1	C 1 2 1	C 1 2 1
Cela - Dimensão (Å)	82,408 74,128	109,314 74,474	107,479 73,126
Ângulos - cela α, β, γ (°)	89,042 90 100,824 90	82,453 90 127,012 90	81,422 90 127,164 90
Reflexões (total)	525104 (55983)	339918 (33436)	171023 (16309)
Reflexões (únicas)	77867 (8433)	49941 (5026)	25891 (2540)
Multiplicidade	6,7 (6,6)	6,8 (6,6)	6,6 (6,4)
Completeza (%)	84,32 (92,29)	96,99 (96,87)	98,97 (98,79)
I/sigma(I) médio	13,44 (2,47)	17,15 (2,95)	7,91 (2,74)
Wilson B-factor	28,60	33,18	33,17
R-merge	0,07131 (0,7788)	0,05702 (0,4638)	0,2895 (0,9432)
R-meas	0,07736 (0,8435)	0,06181 (0,504)	0,3145 (1,027)
R-pim	0,0296 (0,3208)	0,02362 (0,195)	0,1215 (0,4017)
CC 1/2	0,999 (0,952)	0,999 (0,986)	0,978 (0,867)
R-work	0,2298 (0,3878)	0,2245 (0,4539)	0,2350 (0,3588)
R-free	0,2570 (0,3917)	0,2527 (0,4756)	0,2841 (0,3663)
RMS (ligação)	0,013	0,013	0,014
RMS (ângulos)	1,58	1,59	1,67
Ramachandran favorável (%)	98,04	98,43	98,04
Ramachandran permitido (%)	1,86	1,57	1,96
Ramachandran outliers (%)	0,10	0,00	0,00
B-fator médio	37,03	47,50	44,18

$$R = \frac{\sum_{hkl} |F_{obs} - F_{calc}|}{\sum_{hkl} F_{obs}}$$

$$R\text{-merge} = \frac{\sum_{hkl} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

$$R\text{-pim} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hklj}}$$

	IT18	IT17	PT8
Comprimento de onda (Å)	0,97	0,97	0,97
Resolução (Å)	26,01 - 1,843 (1,909 - 1,843)	43,64 - 1,94 (2,009 - 1,94)	36,53 - 1,84 (1,906 - 1,84)
Grupo Espacial	C 1 2 1	C 1 2 1	C 1 2 1
Cela - Dimensão (Å)	109,795 75,406	109,478 74,419	109,754 74,888
Ângulos - cela α, β, γ (°)	82,531 90 127,016 90	82,473 90 127,137 90	82,49 90 127,108 90
Reflexões (total)	316810 (31514)	266035 (25384)	290100 (28629)
Reflexões (únicas)	46472 (4606)	39154 (3869)	42671 (4336)
Multiplplicidade	6,8 (6,8)	6,8 (6,5)	6,8 (6,6)
Completeza (%)	99,53 (99,31)	98,52 (97,40)	91,31 (93,02)
I/sigma(I) médio	15,81 (2,72)	17,87 (3,22)	19,95 (2,87)
Wilson B-factor	33,36	34,21	36,14
R-merge	0,06676 (0,7339)	0,07406 (0,4522)	0,0484 (0,7183)
R-meas	0,07231 (0,7944)	0,08021 (0,4915)	0,05256 (0,7797)
R-pim	0,02755 (0,3017)	0,03053 (0,1904)	0,02025 (0,3)
CC 1/2	0,999 (0,962)	0,998 (0,983)	0,999 (0,961)
R-work	0,2062 (0,3350)	0,2732 (0,4433)	0,2098 (0,4076)
R-free	0,2281 (0,3656)	0,3161 (0,4267)	0,2357 (0,4750)
RMS (ligação)	0,013	0,013	0,013
RMS (ângulos)	1,63	1,72	1,60
Ramachandran favorável (%)	97,88	96,88	98,83
Ramachandran permitido (%)	1,73	3,12	1,17
Ramachandran outliers (%)	0,38	0,00	0,00
B-fator médio	44,49	46,78	50,23

$$R = \frac{\sum_{hkl} |F_{obs} - F_{calc}|}{\sum_{hkl} F_{obs}}$$

$$R\text{-merge} = \frac{\sum_{hkl} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl}}$$

$$R\text{-pim} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_j |I_{hklj} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl}}$$