

**FERNANDA CRISTINA RODRIGUES DE PAIVA**

**Estudos estruturais de enzimas envolvidas com ciclizações em vias de biossíntese de  
antibióticos**

Tese apresentada ao Programa de Pós-Graduação em Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do Título de Doutor em Ciências.

São Paulo  
2020

**FERNANDA CRISTINA RODRIGUES DE PAIVA**

**Estudos estruturais de enzimas envolvidas com ciclizações em vias de biossíntese de  
antibióticos**

Tese apresentada ao Programa de Pós-Graduação em Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do Título de Doutor em Ciências.

Área de concentração: Microbiologia

Orientador: Prof. Dr. Marcio Vinicius Bertacine Dias

Versão corrigida

São Paulo  
2020

CATALOGAÇÃO NA PUBLICAÇÃO (CIP)  
Serviço de Biblioteca e informação Biomédica  
do Instituto de Ciências Biomédicas da Universidade de São Paulo

Ficha Catalográfica elaborada pelo(a) autor(a)

Paiva, Fernanda Cristina Rodrigues de  
Estudos estruturais de enzimas envolvidas com  
ciclizações em vias de biossíntese de antibióticos /  
Fernanda Cristina Rodrigues de Paiva; orientador  
Marcio Vinicius Bertacine Dias. -- São Paulo, 2020.  
143 p.

Tese (Doutorado) -- Universidade de São Paulo,  
Instituto de Ciências Biomédicas.

1. Cristalografia de raios X. 2. Produtos  
naturais. 3. Antibióticos. 4. Ligase dependente de  
ATP. 5. Diels-alderases. I. Dias, Marcio Vinicius  
Bertacine, orientador. II. Título.



## ATA DE DEFESA

Aluno: 42132 - 8445025 - 1 / Página 1 de 1

Ata de defesa de Tese do(a) Senhor(a) Fernanda Cristina Rodrigues de Paiva no Programa: Ciências Biológicas (Microbiologia), do(a) Instituto de Ciências Biomédicas da Universidade de São Paulo.

Aos 12 dias do mês de agosto de 2020, no(a) via remota realizou-se a Defesa da Tese do(a) Senhor(a) Fernanda Cristina Rodrigues de Paiva, apresentada para a obtenção do título de Doutora intitulada:

"Estudos estruturais de enzimas envolvidas com ciclizações em vias de biossíntese de antibióticos"

Após declarada aberta a sessão, o(a) Sr(a) Presidente passa a palavra ao candidato para exposição e a seguir aos examinadores para as devidas arguições que se desenvolvem nos termos regimentais. Em seguida, a Comissão Julgadora proclama o resultado:


Nome dos Participantes da Banca	Função	Sigla da CPG	Resultado
Marcio Vinícius Bertacine Dias	Presidente	ICB - USP	Não Votante
Beatriz Gomes Guimarães	Titular	FIOCRUZ - Externo	<u>Aprovado</u>
Andrea Balan Fernandes	Titular	ICB - USP	<u>Aprovado</u>
Marli de Fatima Fiore	Titular	CENA - USP	<u>Aprovado</u>

Resultado Final: Aprovado

## Parecer da Comissão Julgadora \*

Eu, Lucianna Vicente da Silva \_\_\_\_\_, lavrei a presente ata, que assino juntamente com os(as) Senhores(as). São Paulo, aos 12 dias do mês de agosto de 2020.

  
Beatriz Gomes Guimarães

  
Andrea Balan Fernandes

  
Marli de Fatima Fiore

  
Marcio Vinícius Bertacine Dias  
Presidente da Comissão Julgadora

\* Obs: Se o candidato for reprovado por algum dos membros, o preenchimento do parecer é obrigatório.

A defesa foi homologada pela Comissão de Pós-Graduação em \_\_\_\_\_ e, portanto, o(a) aluno(a) \_\_\_\_\_ jus ao título de Doutora em Ciências obtido no Programa Ciências Biológicas (Microbiologia) - Área de concentração: Microbiologia.

\_\_\_\_\_  
Presidente da Comissão de Pós-Graduação





Cidade Universitária "Armando de Salles Oliveira", Butantã, São Paulo, SP - Av. Professor Lineu Prestes, 2415 - ICB III - 05508 000  
Comissão de Ética em Pesquisa - Telefone (11) 3091-7733 - e-mail: cep@icb.usp.br

## CERTIFICADO DE ISENÇÃO

Certificamos que o Protocolo CEP-ICB N° 752/15 referente ao projeto intitulado: "*Análise estrutural de enzimas relacionadas com a formação de éteres cíclicos em antibióticos ionóforos*" sob a responsabilidade de **Fernanda Cristina Rodrigues de Paiva**, e orientação do Prof. Dr. **Márcio Vinícius Bertacine Dias**, do Departamento de Microbiologia, foi analisado na presente data pela CEUA - COMISSÃO DE ÉTICA NO USO DE ANIMAIS e pela CEPESH- COMISSÃO DE ÉTICA EM PESQUISA COM SERES HUMANOS, tendo sido deliberado que o referido projeto não utilizará animais que estejam sob a égide da Lei n° 11.794, de 8 de outubro de 2008, nem envolverá procedimentos regulados pela Resolução CONEP n° 466 de 2012.

São Paulo, 31 de agosto de 2015.

PROF. DR. ANDERSON DE SÁ NUNES  
Coordenador da CEUA - ICB/USP

PROF. DR. PAOLO M. A. ZANOTTO  
Coordenador da CEPESH - ICB/USP



Cidade Universitária "Armando de Salles Oliveira", Butantã, São Paulo, SP - Av. Professor Lineu Prestes, 2415 - ICB III - 05508 000  
Comissão de Ética no Uso de Animais - Telefone (11) 3091-7733 - e-mail: cep@icb.usp.br

Decl. CEP.03.2020

## DECLARAÇÃO

Em adendo ao Certificado de Isenção CEP nº **752/2015**, aprovado em 31/08/2015, e por solicitação do Prof. Dr. **Marcio Vinicius Bertacine Dias**, do departamento de Microbiologia, informamos que o título do projeto da aluna **Fernanda Cristina Rodrigues de Paiva** foi alterado para: "**Estudos estruturais de enzimas envolvidas com ciclizações em vias de biossíntese de antibióticos**", visto que as proteínas alvo desse projeto tiveram de ser modificadas, e assim, melhor adequação do título ao estudo desenvolvido.

São Paulo, 05 de março de 2020.

Profa. Dra. **Luciane Valéria Sita**  
Coordenadora CEUA ICB/USP

Profa. Dra. **Camila Squarzoni Dale**  
Coordenadora CEPESH ICB/USP



Declaro, para os devidos fins, que

**Fernanda Cristina Rodrigues de Paiva**

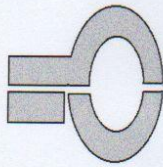
concluiu o Curso "Armazenamento, Manuseio e Descarte de Produtos Químicos",  
realizado no Instituto de Ciências Biomédicas da Universidade de São Paulo.

São Paulo, 29 fevereiro 2016  
(Declaração válida por 5 anos)

**Helayne Freitas**  
Presidente da Comissão de Resíduos Químicos

**Prof. Dr. Jackson Cioni Bittencourt**  
Diretor do ICB





# Certificado



Certificamos que **Fernanda Cristina Rodrigues de Paiva** participou do Treinamento em Biossegurança realizado no Departamento de Microbiologia/ICB-USP, no dia 27 de agosto de 2015, com carga horária total de 8 horas.

*Gabriel Padilla*

**Prof. Dr. Gabriel Padilla**  
Responsável pelo Treinamento

*Tatiana Alves dos Reis*

**Tatiana Alves dos Reis**  
Técnica Responsável pelo Treinamento

## AGRADECIMENTOS

A Deus e aos santos que sempre me abençoaram dando-me alento, força e coragem.

Aos meus pais que são minha vida, fonte de inspiração e porto seguro! Exemplos de caráter e amor incondicional! Obrigada por me enriquecerem com sabedoria e ensinamentos diários.

Ao meu Amor João pela amizade, suporte, paciência, carinho, incentivo e compreensão! Obrigada por deixar minha vida mais leve.

Aos meus familiares sem os quais não sou ninguém. Muito obrigada por serem luz e alegria em minha vida.

Ao meu orientador Prof. Dr. Marcio Dias pela oportunidade de realizar este projeto, por todos ensinamentos e também pelo carinho.

Aos funcionários do Departamento de Microbiologia da USP, em especial à Renata por toda ajuda e apoio.

Ao Prof. Dr. Siewert-Jan Marrink e aos Pesquisadores Paulo Cesar Telles de Souza e Eduardo T. de Oliveira por me receberem em Groningen e pela colaboração e discussão dos resultados.

Aos nossos colaboradores de Cambridge que contribuíram para o acontecimento deste projeto: Professor Peter Leadlay, Dra. Karen Chan e Dr. Rory Little.

À Profa. Dra. Andrea Balan pelas contribuições dentro e fora do ambiente acadêmico e ao Prof. Dr. Ariel Mariano Silber pelas colaborações científicas.

Aos colegas e amigos do LBEA pela ajuda, conselhos, momentos de descontração e troca de experiências.

À Universidade de São Paulo e à Universidade de Groningen.

Ao *Deutsches Elektronen-Synchrotron* (DESY) e ao Laboratório Nacional de Luz Síncrotron pelas infraestruturas oferecidas.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de doutorado – Processo: 141090/2016-2.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo financiamento do meu estágio sanduíche na *University of Groningen* - Processo: 88887.368441/2019-00.

Aos meus mestres e amigos dos laboratórios pelos quais passei (LTPNA-FCF-USP e Laboratório de Ficologia-UFMG) que me ensinaram a viver a biologia.

E a todos que contribuíram de alguma forma para a concretização desse sonho, meu eterno agradecimento!

Este projeto foi realizado no Laboratório de Biologia Estrutural Aplicada – LBEA, sendo financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo, processo 2018/00351-1.

## RESUMO

Paiva, F.C.R. De. Estudos estruturais de enzimas envolvidas com ciclizações em vias de biossíntese de antibióticos. [tese (Doutorado em Microbiologia)] - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2020.

Os produtos naturais são uma fonte de compostos bioativos utilizados pelo homem para o tratamento de diversas doenças. Apesar do campo da química biológica ter avançado consideravelmente no entendimento de como a natureza produz vários desses compostos, ainda há muitos eventos pouco compreendidos. Atualmente esse campo de pesquisa tem ganhado a atenção da comunidade científica devido ao crescente interesse no entendimento das enzimas relacionadas com a biossíntese de produtos naturais e, também, devido a sua aplicabilidade em programas de biologia sintética para a produção de novos compostos. A formação de anéis heterocíclicos em produtos naturais é um dos eventos em que os mecanismos moleculares não são claramente conhecidos. O objetivo deste trabalho é caracterizar estruturalmente diferentes enzimas responsáveis pela biossíntese de anéis heterocíclicos por cristalografia de raios-X. Um aspecto interessante e único da biossíntese de poliéteres ionóforos é a formação de anéis centrais aparentemente formados por um mecanismo utilizando Diels-alderases. Já para os anéis benzoxazóis, uma ligase dependente de ATP e uma amidohidrolase parecem estar envolvidas na formação do anel. A enzima AjiA1 que participa na formação de um anel benzoxazol e as proteínas ciclases Tmn9 e Tsn15 responsáveis pela formação de anéis em poliéteres ionóforos tetronatos são estudadas neste trabalho. As análises estruturais da enzima AjiA1 indicam que essa enzima possui um conjunto de movimentos completamente novo durante o ciclo catalítico nos membros das AFEs (*Adenylate-Forming Enzymes*). A alternância de domínio não deve ocorrer nessa enzima, já que ela apresenta um evento de troca de domínio envolvendo uma alça C-terminal longa (*loop swapping*) que restringe o domínio C-terminal e bloqueia a enzima em uma única conformação. A estrutura de Tsn15 em complexo com seu intermediário de reação e a estrutura da enzima Tmn9 em complexo com seu cofator FAD foram determinadas. Tsn15 e Tmn9 são estruturalmente homólogas às ciclases conhecidas como Diels-alderases. Apesar das reações Diels-Alder ocorrerem espontaneamente em água, nossos estudos evidenciam que a presença da enzima favorece a ocorrência da ciclização, permitindo que o substrato adquira uma posição favorável para a ocorrência da reação. A descrição dessas enzimas e as análises realizadas neste trabalho representam um avanço significativo no entendimento da biossíntese de diferentes produtos naturais.

**Palavras-chave:** Cristalografia de raios X. Produtos naturais. Antibióticos. Ligase dependente de ATP. Diels-alderases.

## ABSTRACT

Paiva, F.C.R. De. Structural studies of enzymes involved with cyclizations in antibiotic biosynthesis pathways. [thesis (Ph. D. thesis in Microbiology)] - Institute of Biomedical Sciences, University of São Paulo, São Paulo, 2020.

Natural products are a source of bioactive compounds used by men to treat various diseases. Although the field of biological chemistry has advanced considerably in the understanding of how nature produces several of these compounds, there are still many events not sufficiently understood. Currently, this research area has gained the attention of the scientific community due to the growing interest in understanding the enzymes involved in the biosynthesis of natural products and, also, due to its applicability in synthetic biology programs aiming at the production of new compound derivatives. The formation of heterocyclic rings in natural products is one of the events in which the molecular mechanisms are not clearly known. The objective of this work was to structurally characterize different enzymes involved in the biosynthesis of natural products heterocyclic rings through protein X-ray crystallography. An interesting and unique aspect of ionophore polyether biosynthesis is the formation of central rings which are catalyzed by Diels-alderases. For benzoxazole antibiotics, an ATP-dependent ligase and an amidohydrolase appear to be involved in the formation of the ring. The enzyme AjiA1 that participates in the formation of a benzoxazole ring and the cyclases Tmn9 and Tsn15 that are responsible for the formation of a pyran and a cyclohexane rings in tetronate containing ionophore polyethers are studied in this work. Structural analyzes of AjiA1 indicate that this enzyme has a completely new set of movements during the catalytic cycle, which is different from the other AFE family members (Adenylate-Forming Enzymes). Domain alternation should not occur in this enzyme, in contrast to the other members of this family, and instead, AjiA1 has a domain change event involving a long C-terminal loop (loop swapping) that restricts the C-terminal domain and blocks the enzyme in a single conformation. The structure of Tsn15 in complex with its reaction intermediate and the structure of the enzyme Tmn9 in complex with FAD coenzyme were determined. Both Tsn15 and Tmn9 are structurally homolog to cyclases known as Diels-alderases. Although the Diels-Alder reactions occur spontaneously in water, our studies propose that the presence of the enzyme favors the cyclization since the substrate can acquire a favorable position in the protein lumen for the reaction to occur. The description of these enzymes and the analyzes carried out in this work represent a significant advance in the understanding of the biosynthesis of different natural products.

**Key words:** X-ray crystallography. Natural products. Antibiotics. ATP-dependent ligase. Diels-alderases.

## LISTA DE FIGURAS

- Figura 1. Estruturas químicas da tetronomicina (a) e da tetronasina (b). .....26
- Figura 2. Estrutura da versipelostatina de *Streptomyces versipellis* 4083-SVS6.....27
- Figura 3. Formação de espirotetronato catalisada por enzima da biossíntese de versipelostatina. ....28
- Figura 4. Clusters gênicos da tetronomicina e da tetronasina. ....28
- Figura 5. a) Anel benzoxazol; b) Estrutura de AJI9561, um composto benzoxazol. ....29
- Figura 6. Reações catalisadas por AjiA1 (duas semi-reações) e pela amidohidrolase AjiH. A AjiA1 conduz uma ativação dependente de ATP e a dimerização do 3-HAA e a amidohidrolase AjiH realiza o fechamento do anel oxazol que produz o anel benzoxazol. ....43
- Figura 7. Análise de SDS-PAGE da purificação de AjiA1. As setas ajudam na identificação do tamanho aproximado da banda de interesse: AjiA1 (~48 kDa). M – marcador; AF – cromatografia de afinidade por íons metálicos imobilizados (IMAC); SE – cromatografia de exclusão molecular. ....45
- Figura 8. Cristais de AjiA1 obtidos pelo robô Oryx4 com o método de sitting drop (gota sentada) a 18 °C. Os cristais eram pequenos e possuíam um formato elíptico e sem faces bem definidas. Proteína: 13 mg.mL<sup>-1</sup> em 50 mM HEPES pH 8,0 e 300 mM de NaCl. (a) Condição: 37,7 % v/v PEG4000, 150mM AMPD/Tris pH 9,0 e 50mM sulfato de magnésio. (b) Condição: 30 % v/v Jeffamine M-600, 100mM MES pH 6,5 e 50mM cloreto de cézio. ....46
- Figura 9. Gráfico de Thermofluor para AjiA1. Curvas representativas do deslocamento térmico de AjiA1 em diferentes tampões. A curva de referência está representada pela linha vermelha.....47
- Figura 10. Cristais pequenos obtidos para AjiA1 pelo robô Oryx4 com o método de gota sentada a 18 °C após a troca do tampão 50 mM HEPES pH 8,0 e 300 mM de NaCl para 100 mM Bis-Tris propano pH 6,5 e 150 mM de KCl. Condição de cristalização: 40 % v/v PEG 300 e 100 mM fosfato citrato de potássio pH 4,2. Proteína: 13 mg.mL<sup>-1</sup>. ....47
- Figura 11. Cristais otimizados maiores e com faces definidas de AjiA1 obtidos manualmente com o método de gota pendurada a 18 °C. Proteína: 8 mg.mL<sup>-1</sup> em 100 mM Bis-Tris propano pH 6,5 e 150 mM de KCl. Condição de cristalização: 35 % v/v PEG 300 e 100 mM fosfato critrato de potássio pH 4,2.....48
- Figura 12. Estrutura geral de AjiA1 de *Streptomyces* sp. AJ9561. (a) Representação em cartoon da AjiA1 homodimérica (a cadeia A está colorida em amarelo e a cadeia B em azul). As esferas em cinza representam íons Mg<sup>2+</sup> e as esferas em verde representam íons Zn<sup>2+</sup> (b) Superfície de potencial eletrostático indica a região do sítio ativo (círculo). Azul indicando carga positiva e vermelho indicando carga negativa. ....50
- Figura 13. Estrutura e comparação entre AjiA1 reportada neste trabalho (amarelo) com os homólogos mais próximos PaaK1 (salmão) e NatL2:AMPPNP (cinza). ....50
- Figura 14. Estrutura do monômero da AjiA1 com o domínio N-terminal maior em verde, o domínio C-terminal em azul e o loop que conecta os domínios em amarelo. O sítio ativo está marcado com um círculo vermelho. ....51



- Figura 15. Dimerização de AjiA1. As interações íntimas de dimerização de AjiA1 criadas por folhas  $\beta$  compostas que abrangem os dois monômeros. A folha  $\beta$  de duas fitas que flanqueia a alça P (Ala82 roxo a Tyr99) é confinada pela folha  $\beta$  distorcida de cinco fitas (laranja) [Thr123 a Val126; Thr149 a Pro152; Leu176 a Trp178; Ala207 a Ala211; Val232 a Ser238] do segundo protômero. O domínio N-terminal maior é representado em verde no primeiro protômero e em lima no segundo protômero, o domínio C-terminal está em azul no primeiro protômero e azul claro no segundo protômero e o linker que conecta os domínios é mostrado em amarelo. ....52
- Figura 16. Interface de dimerização de AjiA1. O servidor online PDBsum (144) (EMBL-EBI) foi usado para gerar um diagrama esquemático mostrando as interações entre as duas cadeias de AjiA1. As cadeias de interação são unidas por linhas coloridas, cada uma representando um tipo diferente de interação, laranja = contatos não ligados, azul = ligações de hidrogênio e linhas vermelhas = pontes de sal. A área de cada círculo é proporcional à área da superfície da cadeia proteica correspondente. A extensão da região da interface em cada cadeia é representada pela fatia preta cujo tamanho significa a área da superfície da interface. ....53
- Figura 17. Homólogos estruturais de AjiA1. A cadeia Ca de AjiA1 foi alinhada com seus homólogos estruturais (a) 4R1M (ciano), (b) 3L2K (roxo), (c) 2D1R (laranja), (d) 5IE3 (vermelho) e (e) 6IJB (cinza ) O valor rmsd da superposição com AjiA1 e a identidade de aminoácidos (%) com AjiA1 são dados em cada análise. ....53
- Figura 18. Alinhamento de AjiA1 e PaaK1 com homólogos DALI. Os motivos A1-A10 conservados nas enzimas da superfamília ANL são indicados de acordo com o domínio de adenilação definido da peptídeo sintetase não-ribossomal (145). Os motivos principais A1-A10 são mostrados em quadros vermelhos. O alinhamento foi gerado pelo ClustalOmega (146) e importado para o Jalview Versão: 2.11.0 (147). Resíduos similares são coloridos de acordo: Azul, hidrofóbico; Verde, polar; Roxo, negativo; Vermelho, positivo; Amarelo, prolina; Laranja, glicina. A tabela exibe a identidade em porcentagem de AjiA1 para os homólogos alinhados. ....54
- Figura 19. Local de ligação do zinco na cadeia A de AjiA1 com 2Fo – Fc de  $Zn^{2+}$  (com contorno de 2,0  $\sigma$ ). Esta figura mostra as interações de coordenação.....55
- Figura 20. Representação esquemática do local de ligação ao  $Mg^{2+}$  na cadeia A de AjiA1. A figura foi preparada usando LigPlot + v.2.1 (152).....56
- Figura 21. AjiA1 loop swapping. O resíduo conservado Lis418 está representado por sticks.57
- Figura 22. Região de loop swapping. Os aminoácidos 422L, 423F e 424A parecem formar uma pequena fita beta que pode estar envolvida na formação de uma folha  $\beta$  com os aminoácidos 357A, 358K e 359A. A cadeia A está em amarelo e a cadeia B está em azul. ....58
- Figura 23. Representação de superfície do monômero AjiA1 mostrando a conservação de resíduos com outros membros da AFEs. A figura foi preparada usando o servidor ConSurf (155). Roxo e ciano indicam conservação de sequência alta e baixa, respectivamente.....59
- Figura 24. Comparação AjiA1 e estruturas em complexo com cosubstrato e substratos. (a) Sobreposição de AjiA1 e NatL2:AMPPNP (PDB ID: 6six). Os resíduos de AjiA1 estão representados em amarelo e o NatL2 em cinza. O principal resíduo catalítico Lys418 possui uma cor mais forte, pois pertence ao protômero adjacente. (b) Sobreposição de AjiA1 e NatL2:AMP:3-HAA (PDB ID: 6siy). Os resíduos AjiA1

estão representados em amarelo, em que K418 e Y435 estão em laranja e os resíduos da NatL2 estão em ciano, nos quais K418 e Y435 estão em azul. A cor diferente desses resíduos de aminoácidos representa que eles são do protômero adjacente. As águas do sítio ativo são representadas por esferas vermelhas. (c) Sobreposição de AjiA1 e NatL2 em complexo com AMPPNP (6SIX). Os protômeros estão representados em cartoon. AjiA1 está em amarelo e seus cinco loops diferentes estão em laranja. NatL2:AMPPNP está em cinza e seus loops estão em preto. (d) Sobreposição de AjiA1 e NatL2 em complexo com AMPPNP (6SIX). Os protômeros estão representados em ribbon. AjiA1 são representados em amarelo (cadeia A) e laranja (cadeia B) e NatL2 no complexo com AMPPNP (6SIX) são representados em cinza (cadeia A) e preto (cadeia B). Os principais resíduos catalíticos Lys418 de AjiA1 e NatL2:AMPPNP estão representados em palitos. O terminal C e, consequentemente, o loop do terminal C, também se move para uma posição mais próxima do local ativo no local ativo em cerca de 1,5 Å..... 60

- Figura 25. Representação de massa de fator B mostrando o núcleo rígido. (a) Os valores do fator B são ilustrados por cores, variando de baixo (azul) a alto (vermelho), e pela forma, um tubo mais largo indica regiões com fatores B mais altos, enquanto o tubo mais estreito indica regiões com fatores B mais baixos. b) Análise fatorial de temperatura. Gráficos do fator B médio versus resíduos para cada cadeia de AjiA1. .... 61
- Figura 26. Análise evolucionária pelo método de máxima verossimilhança de vários membros diferentes de AFEs que foram caracterizados estruturalmente, incluindo AjiA1. Os comprimentos das ramificações são mostrados. As proteínas e organismos utilizados são: DhbE de *Bacillus subtilis* (PDB ID: 1MDF); Acs fatty de *Thermus thermophilus* (PDB ID: 1ULT); Luciferase do vaga-lume de *Photinus pyralis* (PDB: 1LCI); AAE3 de *Arabidopsis thaliana* (PDB: 5IE3); PheA de *Brevibacillus brevis* (PDB: 1AMU); DltA de *Bacillus cereus* ATCC 14579 (PDB ID: 3DHV); Benzoílo-CL de *Paraburkholderia xenovorans* LB400 (PDB ID: 2V7B); Acs de *Salmonella enterica* (PDB: 1PG4); CBL de *Alcaligenes* sp. AL3007 (ID do PDB: 1T5S); A proteína deste estudo AjiA1 de *Streptomyces* sp. AJ9561 (ID do PDB: 6WUQ); NatL2 de *Streptomyces* sp. Tu 6176 (PDB: 6SIX); PaaK1 de *Burkholderia cenocepacia* (PDB: 2Y27) e BT\_0428 de *Bacteroides thetaiotaomicron* (PDB: 4R1M). As análises foram realizadas em MEGA X (160). .... 63
- Figura 27. As análises cinéticas de AjiA1 foram realizadas usando concentrações de 3HAA e 3HBA de 0 a 2500  $\mu$ M. Km e Vmax foram determinados a partir de regressão não linear para a equação de Michaelis-Menten a partir de três experimentos repetidos. .... 64
- Figura 28. Mecanismo proposto para a formação dos anéis ciclohexano e tetra-hidropirano na Tetronasina. Tsn11 catalisa uma reação hetero-Diels-Alder de demanda inversa de elétrons para formar um intermediário de oxidecalina. A oxidecalina é hidratada para formar um hemiacetal cíclico. Tsn15 então catalisa uma desidratação seguida de um rearranjo pericíclico para formar o anel de tetra-hidropirano e fragmentar a oxidecalina, produzindo Tetronasina. .... 67
- Figura 29. Análise de SDS-PAGE da purificação da enzima Tsn15. M – marcador; AF – cromatografia de afinidade por íons metálicos imobilizados (IMAC); SE – cromatografia de exclusão molecular. .... 68
- Figura 30. Cristais de Tsn15 obtidos por otimização manual. .... 68

- Figura 31. Representações da estrutura homodimérica de Tsn15 em complexo com seu intermediário e análises do sítio ativo.....70
- Figura 32. Interface de dimerização da Tsn15. O servidor online PDBsum (144) (EMBL-EBI) foi usado para gerar um diagrama esquemático mostrando as interações entre as duas cadeias de Tsn15. As cadeias estão unidas por linhas coloridas, cada uma representando um tipo diferente de interação, laranja = contatos não ligados, azul = ligações de hidrogênio e linhas vermelhas = pontes salinas. A área de cada círculo é proporcional à área da superfície da cadeia proteica correspondente. A extensão da região da interface em cada cadeia é representada pela fatia preta cujo tamanho significa a área da superfície de interface. ....71
- Figura 33. Homólogos estruturais de Tsn15. A cadeia C $\alpha$  de Tsn15 foi alinhada com seus homólogos estruturais AbyU, PyrI4, AOC2 e PodA. O RMSD e a identidade sequencial (%) em comparação com Tsn15 são mostrados próximos a cada estrutura. ....72
- Figura 34. Alinhamento de Tsn15 com outras ciclases [4+2]. O alinhamento foi gerado pelo ClustalOmega (146) e importado para o Jalview Versão: 2.11.0 (147). Resíduos similares são coloridos de acordo: Azul, hidrofóbico; Verde, polar; Roxo, negativo; Vermelho, positivo; Amarelo, prolina; Laranja, glicina. A tabela exibe a identidade em porcentagem de Tsn15 com os homólogos alinhados. Os resíduos que interagem com o ligante estão destacados com um asterisco. ....72
- Figura 35. Sobreposição das estruturas de Tsn15 em suas formas apo e em complexo com o ligante.....73
- Figura 36. Volume do sítio ativo da Tsn15. A ferramenta online CASTp 3.047 (Universidade de Illinois, EUA) (167) foi usada para calcular o volume da cavidade do barril  $\beta$  de Tsn15 de 388 Å<sup>3</sup>. ....74
- Figura 37. Sítio ativo da Tsn15 em complexo com o intermediário. Contatos de ligação detalhados entre o intermediário e os resíduos do barril  $\beta$  interno estão representados. As linhas cinza pontilhadas representam ligações de hidrogênio. ....75
- Figura 38. Representação da superfície do sítio ativo. Os resíduos hidrofóbicos ou polares (com carga positiva ou negativa e sem carga) são indicados em cores, sendo amarelo para resíduos hidrofóbicos, vermelho para resíduos carregados negativamente e azul para resíduos carregados positivamente.....75
- Figura 39. Via Biossintética Proposta. Um duplo exocíclico no anel tetronato sofre a cicloadição [4 + 2] com o dieno terminal do policetídeo linear. A Tmn9 catalisa a formação do anel transdecalina. ....78
- Figura 40. Análise de SDS-PAGE da purificação da enzima Tmn9. M - marcador; AF – cromatografia de afinidade por íons metálicos imobilizados (IMAC); SE – cromatografia de exclusão molecular. ....79
- Figura 41. Cristal de Tmn9 obtido pelo robolab do LNBIO. ....79
- Figura 42. Representações da estrutura homodimérica de Tmn9 com a utilização do Pymol (94) a partir dos dados obtidos. Está evidenciada a presença do FAD em cada um dos monômeros.....80
- Figura 43. Protômero da Tmn9. Estão evidenciados os três domínios: domínio de ligação ao FAD (verde), domínio médio (azul) e domínio C-terminal do tipo tiorredoxinana (rosa). A molécula do FAD ligada está mostrada no modelo em bastão. ....81

- Figura 44. Sítio de ligação ao FAD. As ligações de hidrogênio envolvidas estão mostradas como linhas pontilhadas em cinza. O mapa de densidade eletrônica de 2Fo-Fc é contornado em  $2,0 \sigma$ ..... 82
- Figura 45. Análises estruturais do Bolsão de Ligação ao Substrato de Tmn9. (A) A representação do potencial de carga de superfície evidencia o bolsão de ligação ao FAD positivamente carregado (azul indicando carga positiva e vermelho carga negativa). (B) Representação da superfície em que os resíduos que são hidrofóbicos estão em amarelo, os positivamente carregados em azul e os negativamente carregados em vermelho. (C) Representação de superfície mostrando que o sítio de ligação ao substrato em flavociclases é variável. Neste desenho, a superfície de Tmn9 é colorida de acordo com a conservação da sequência de aminoácidos baseada no alinhamento da sequência de diferentes flavociclases usando o servidor ConSurf (roxo e ciano indicam alta e baixa conservação de sequência, respectivamente).... 82
- Figura 46. Volume do sítio ativo de Tmn9. A ferramenta online CASTp 3.0 (169) (Universidade de Illinois, EUA) foi usada para calcular o volume de  $1716.720 \text{ \AA}^3$ ..... 83
- Figura 47. Homólogos estruturais de Tmn9 (verde). A cadeia de C $\alpha$  da Tmn9 foi alinhada com seus homólogos estruturais (a) 5xgv (ciano), (b) 6j0z (magenta), (c) 4k2x (amarelo), (d) 5kox (rosa), (e) 4x4j (cinza) e (f) 6brd (azul). Ao lado de cada estrutura está o valor de rmsd e a identidade de aminoácidos (%) com a Tmn9. .... 84
- Figura 48. Alinhamento de Tmn9 e PyrE3 com homólogos do DALI. O alinhamento foi gerado pelo ClustalOmega (146) e importado para o Jalview Versão: 2.11.0 (147). Resíduos similares são coloridos de acordo: Azul, hidrofóbico; Verde, polar; Roxo, negativo; Vermelho, positivo; Amarelo, prolina; Laranja, glicina. Os resíduos do sítio ativo proposto são destacados por estrelas..... 85
- Figura 49. Comparação entre Tmn9 e PyrE3. Loop 379V a 409K em Tmn9 e loop 368R a 381V em PyrE3..... 85
- Figura 50. Comparação estrutural da conformação do FAD em Tmn9 (verde), em AlpK com conformação “open” (salmão, PDB: 6J0Z) (175), em PgaE com conformação “in” (ciano, PDB: 2QA1) (176) e em WT pHBH com conformação “out” (magenta, PDB: 1DOD) (177). O alinhamento foi realizado com as proteínas. .... 86
- Figura 51. Análise evolucionária pelo método de máxima verossimilhança de proteínas com similaridade estrutural com a Tmn9. Os comprimentos das ramificações são mostrados. As proteínas e organismos utilizados são: A proteína deste estudo Tmn9 de *Streptomyces* sp. NRRL 11266 (PDB: 6UI5), PyrE3 de *Streptomyces rugosporus* (PDB: 5XGV), Rifampicina Monooxigenase complexada com rifampicina de *Nocardia farcinica* (PDB: RIFMO - 5KOX), Rifampicina Monooxigenase de *Streptomyces venezuelae* complexada com rifampina e FAD (PDB: Rox - 6BRD), Anidrotetraciclina Hidroxilase de *Streptomyces rimosus* (PDB: OxyS - 4K2X), Monooxigenase AlpK de *Streptomyces ambofaciens* (PDB: 6J0Z) e BexE de *Amocolatopsis orientalis* subsp. *Vinearia* (PDB: 4X4J) (a ser publicado). As análises foram realizadas em MEGA X (160)..... 87
- Figura 52. Melhores poses obtidas pela docagem e sítio ativo da homóloga Rox (PDB: 6BRDB). Comparativamente com o ligante da homóloga Rox depositado no PDB, foram escolhidas as melhores poses da docagem. É possível perceber que o anel decalina está próximo à isoaloxazina do cofator FAD em todas as posições. .... 88

Figura 53. Distâncias entre os átomos envolvidos na formação do anel decalina durante a simulação em água (preto) e a simulação na proteína (vermelho). Os pontos se referem às distâncias entre as ligações C18 e O8 e entre C25 e C20, átomos envolvidos na formação do anel decalina. Os dados representam as médias das triplicatas das dinâmicas moleculares de 500 ns realizadas..... 89

Figura 54. Comparação da flexibilidade dos protômeros de Tmn9 sem (a) e com ligante (b) obtida por meio da variação conformacional da enzima durante as simulações de dinâmica molecular. É possível verificar o cap em conformação mais aberta na presença do ligante em comparação com o protômero sem o ligante. A rigidez da região do núcleo central da proteína é evidenciada pela cor azul. .... 90

## LISTA DE TABELAS

Tabela 1. Dados gerais da proteína trabalhada. ....	45
Tabela 2. Dados coletados e estatísticas de refinamento de AjiA1 em sua forma apo. As estatísticas para a resolução mais alta são mostradas entre parênteses. ....	49
Tabela 3. Dados gerais da proteína trabalhada. ....	68
Tabela 4. Condições de cristalização da enzima Tsn15. ....	69
Tabela 5. Dados coletados e Estatísticas de Refinamento para Tsn15 em sua forma holo. As estatísticas para a resolução mais alta são mostradas entre parênteses. ....	69
Tabela 6. Dados gerais da proteína trabalhada. ....	78
Tabela 7. Condições de cristalização da enzima Tsn15. ....	79
Tabela 8. Dados coletados e Estatísticas de Refinamento para Tmn9 em sua forma apo. As estatísticas para a resolução mais alta são mostradas entre parênteses. ....	80

## LISTA DE ABREVIATURAS

Å: Angström

$\lambda$ : Comprimento de onda

KDa: (Kilo)Dalton

°C: Grau Celsius

3-HAA: Ácido 3-Hidroxiantranílico

3-HBA: Ácido 3-Hidroxibenzóico

6-MAS: Ácido 6-Metilsalicílico

Abs: Absorbância

Ala: Alanina

AMP: Adenosina Monofosfato

Arg: Arginina

Asn: Asparagina

Asp: Ácido Aspártico

ATP: Adenosina Trifosfato

BLAST: *Basic Local Alignment Search Tool*

CHARMM: *Chemistry at Harvard Molecular Mechanics*

CNPEM: Centro Nacional de Pesquisa em Energia e Materiais

Coa: Coenzima A

Cys: Cisteína

DESY: *Deutsche Elektronen-Synchrotron*

DMSO: Dimetilsulfóxido

DNA: Ácido Desoxirribonucleico

DNase: Deoxyribonuclease I

DO<sub>600</sub>: Densidade óptica a 600 nm

FAD: Flavina Adenina Dinucleotídeo

Gln: Glutamina

Glu: Ácido Glutâmico

Gly: Glicina

HEPES: Ácido 4-(2-Hidroxiethyl)-1-Piperazina Etanosulfônico)

His: Histidina

Ile: Isoleucina

IMAC: cromatografia de afinidade por íons metálicos imobilizados

IPTG: Isopropil B-D-1-Tiogalactopiranosida  
Kcat: Constante catalítica  
KCl: Cloreto De Potássio  
Km: Constante de Michaelis-Menten  
Leu: Leucina  
LNBIO: Laboratório Nacional de Biociências  
LNLS: Laboratório Nacional de Luz Síncrotron  
Lys: Lisina  
MES: Ácido 2-(N-Morfolino)Etanosulfônico  
Met: Metionina  
NaCl: Cloreto De Sódio  
NADH: Nicotinamida Adenina Dinucleotídeo, forma reduzida  
NCBI: *Nacional Center For Biotechnology Information*  
PDB: *Protein Data Bank*  
PEG: Polietilenoglicol  
Phe: Fenilalanina  
PKS: Policetídeo Sintase  
Pi: Pirofosfato Inorgânico  
Pro: Prolina  
RMSD: *Root Mean Square Deviation* (Raiz do desvio quadrático médio)  
RNA: Ácido Ribonucleico  
Rpm: Rotações Por Minuto  
SDS: Dodecil Sulfato De Sódio  
SDS-PAGE: Dodecil Sulfato De Sódio - Eletroforese em Gel de Poliacrilamida  
Ser: Serina  
TEMED: Tetrametiletlenodiamina  
Thr: Treonina  
T<sub>m</sub>: *Annealing Temperature*  
Tmn: Tetronomicina  
Trp: Triptofano  
Tsn: Tetronasina  
Tyr: Tirosina  
UV: Ultravioleta  
Val: Valina



## LISTA DE SÍMBOLOS

Ácido aspártico D

Ácido glutâmico E

Arginina R

Asparagina N

Glutamina Q

Histidina H

Lisina K

Serina S

Tirosina Y

Treonina T

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>23</b>
<b>1.1 Descoberta e desenvolvimento de fármacos.....</b>	<b>23</b>
<b>1.2 Formação de anéis heterocíclicos em produtos naturais .....</b>	<b>24</b>
<b>1.3 Poliéteres Ionóforos .....</b>	<b>25</b>
<b>1.3.1 Diels-alderases em poliéteres ionóforos tetronatos.....</b>	<b>26</b>
<b>1.4 Benzoxazóis .....</b>	<b>28</b>
<b>1.5 Métodos para o estudo da estrutura de uma proteína .....</b>	<b>30</b>
<b>1.6 Justificativa .....</b>	<b>33</b>
<b>2 OBJETIVOS .....</b>	<b>34</b>
<b>3 MATERIAL E MÉTODOS .....</b>	<b>35</b>
<b>3.1 Clonagem.....</b>	<b>35</b>
<b>3.2 Produção de proteínas recombinantes.....</b>	<b>35</b>
<b>3.3 Purificação.....</b>	<b>35</b>
<b>3.4 Eletroforese em gel de poliacrilamida-SDS - Análise da expressão e purificação das proteínas .....</b>	<b>36</b>
<b>3.5 Determinação da concentração proteica.....</b>	<b>36</b>
<b>3.6 Ensaio de cristalização das proteínas .....</b>	<b>37</b>
<b>3.7 Ensaio de <i>ThermoFluor</i> .....</b>	<b>37</b>
<b>3.8 Coleta e processamento de dados de difração de raios-x.....</b>	<b>38</b>
<b>3.9 Resolução e refinamento cristalográfico.....</b>	<b>38</b>
<b>3.10 Simulações de Dinâmica Molecular .....</b>	<b>38</b>
<b>3.10.1 Configuração do sistema e parâmetros do campo de força .....</b>	<b>39</b>
<b>3.11 Docagem ou Ancoramento Molecular .....</b>	<b>39</b>
<b>3.12 Ensaio enzimático .....</b>	<b>40</b>
<b>4 CAPÍTULO I - FORMAÇÃO DE ANÉIS BENZOXAZÓIS .....</b>	<b>42</b>

<b>5</b>	<b>CAPÍTULO II – DIELS-ALDERASES EM POLIÉTERES IONÓFOROS TETRONATOS: TSN15 .....</b>	<b>66</b>
<b>6</b>	<b>CAPÍTULO III - DIELS-ALDERASES EM POLIÉTERES IONÓFOROS TETRONATOS: TMN9.....</b>	<b>77</b>
<b>7</b>	<b>CONCLUSÕES E PERSPECTIVAS.....</b>	<b>92</b>
	<b>REFERÊNCIAS.....</b>	<b>94</b>
	<b>APÊNDICES A – Artigo: Crystal structure of AjiA1 shows a novel structural motion mechanism into the adenylate forming enzyme family .....</b>	<b>107</b>
	<b>APÊNDICES B – Artigo: Unexpected enzyme-catalysed [4+2] cycloaddition and rearrangement in polyether antibiotic biosynthesis .....</b>	<b>134</b>

## 1 INTRODUÇÃO

### 1.1 Descoberta e desenvolvimento de fármacos

Os compostos naturais são geralmente derivados do metabolismo secundário de microrganismos e plantas e evoluíram na natureza em resposta às pressões seletivas e aos desafios do ambiente natural. Dessa forma, a natureza tem continuamente executado sua própria versão de química combinatória por mais de três bilhões de anos, quando se acredita que ocorreu o surgimento das bactérias na Terra. Os metabólitos gerados pela síntese combinatória da natureza são representados por uma grande variedade de estruturas exóticas ricas em estereoquímica e anéis concatenados (1).

Historicamente, os produtos naturais têm sido fonte de muitas substâncias de uso medicinal. Produtos derivados de plantas dominaram a farmacopeia humana por milhares de anos. Em 1928, a penicilina foi descoberta por Alexander Fleming, acrescentando os microrganismos como importantes fontes de novos medicamentos. O papel dos produtos naturais na descoberta de novos fármacos foi diminuído com a descoberta de novas fontes de compostos bioativos, como: a síntese orgânica guiada pela estrutura-atividade, a síntese combinatória e pela química computacional/modelagem molecular (*in silico*). Os avanços da química, farmacologia e biologia molecular a partir de meados do século XX diminuíram sensivelmente o interesse em produtos naturais entre os anos de 1984 e 2003, o que se traduziu em um menor volume de investimentos da indústria farmacêutica nesse setor. Gradativamente os produtos sintéticos ganharam posição de destaque no arsenal terapêutico. Entretanto, diminuiu também, nos últimos anos, o número de novos fármacos lançados no mercado, principalmente aqueles que são totalmente sintéticos e, assim, pôde-se observar um ressurgimento no interesse por produtos naturais como fonte de novas substâncias bioativas. Os avanços nas técnicas de separação, purificação e identificação de misturas complexas de produtos e a diversidade estrutural presente na natureza são os fatores responsáveis por esse renovado interesse (2-4).

Dentre os microrganismos, os actinomicetos representam um grupo diversificado de bactérias filamentosas gram-positivas caracterizadas pelo alto conteúdo de GC (guanina e citosina) em seus genomas e por um ciclo de vida complexo envolvendo diferenciação morfológica. Os gêneros *Mycobacterium*, *Actinomyces*, *Nocardia*, *Streptomyces* e *Micromonospora* pertencem a esse grupo (5). Em particular, *Streptomyces spp.* produz uma ampla gama de produtos naturais bioativos, aproximadamente 60 % dos antibióticos conhecidos. Devido à complexidade dos produtos naturais, múltiplos genes são frequentemente

necessários para a sua biossíntese e esses genes são tipicamente reunidos em um agrupamento gênico biossintético (BGC – do inglês *Biosynthetic gene cluster*) (6-8).

Os produtos naturais provenientes de microrganismos podem ser biossintetizados por megaenzimas denominadas poliketídeos sintases (PKSs) (9), peptídeos sintases não ribossomais (NRPSs) (10) ou outros mecanismos como: peptídeos ribossomais (RIPPs) (11), vias biossintéticas que modificam precursores da via glicolítica (12) e vias que modificam o ácido corísmico produzido pela via do ácido chiquímico (13). Alguns produtos naturais são também sintetizados pela combinação das vias PKSs e NRPSs (14).

## 1.2 Formação de anéis heterocíclicos em produtos naturais

Os anéis heterocíclicos são componentes importantes de muitos produtos naturais bioativos, sendo frequentemente necessários para a bioatividade. Além disso, a presença desses anéis aumenta a complexidade estrutural desses compostos e confere rigidez conformacional e estabilidade para a ligação ao alvo (15, 16). A distribuição eletrônica e as propriedades moleculares dos anéis heterocíclicos contribuem ainda mais à sua versatilidade química, uma vez que esses anéis podem interagir com eletrófilos ou nucleófilos, podem se comportar como um ácido ou uma base e muitas vezes atuam como ligantes para coordenar complexos metálicos (17, 18). A natureza evoluiu para fornecer uma ampla variedade de anéis heterocíclicos por meio de diversas vias biossintéticas e novas atividades enzimáticas continuam a ser descobertas (19, 20). Os produtos naturais podem ter vários passos em suas biossínteses e suas estruturas podem ser modificadas por diversas classes de enzimas.

Em poliketídeos e peptídeos há enzimas que introduzem, retiram, ou modificam grupos químicos em suas cadeias (21), como a adição de grupos aminos, amidos, acetil, glicosídeos (22), a deleção de grupos hidroxilas, reduções, oxidações (23) ou ainda a formação de anéis em suas estruturas (24). As enzimas que catalisam essas reações, as denominadas enzimas alfaiates do inglês *tailoring enzymes*, são de grande interesse da indústria farmacêutica haja vista suas aplicabilidades em biologia sintética e biossíntese combinatorial. Elas podem auxiliar na produção de novos derivados de antibióticos que apresentem novos mecanismos de ação, aumento da eficiência biológica ou contornem mecanismos de resistência (21).

Os produtos naturais derivados do ácido corísmico ou de intermediários da via glicolítica também apresentam várias enzimas que modificam seus precursores, levando a produção de moléculas incomuns. Precursores são fusionados entre si formando estruturas complexas como o observado em antibióticos aminoglicosídeos, derivados de glicose, (12) ou no cloranfenicol, derivado do ácido corísmico (25).

As enzimas alfaiates e as enzimas que atuam na formação de diferentes tipos de anéis pela fusão de grupos ou cadeias laterais de aminoácidos têm cada vez mais despertado o interesse de pesquisadores devido às suas peculiaridades e aos seus mecanismos não usuais de catálise (26-28).

### 1.3 Poliéteres Ionóforos

Uma das famílias mais comuns entre os produtos naturais derivados de microrganismos é a dos policetídeos, que são diversos estruturalmente e apresentam um amplo repertório de atividades biológicas. Exemplos de policetídeos são antibióticos como a eritromicina, a penicilina, o imunossupressor tacrolimus (FK506), o antitumoral epotilona e os antibióticos de uso veterinário evermectina, monensina e lasalocida (29). Esses dois últimos compostos, a monensina e a lasalocida, são ainda pertencentes a uma classe de produtos naturais denominada poliéteres ionóforos.

Embora não se tenha no mercado nenhum poliéter utilizado para o tratamento humano, seu uso ocorre como aditivo em rações para ruminantes em alguns países (30). Porém, esses compostos já se mostraram ativos contra várias enfermidades humanas apresentando propriedades bioativas: antitumorais, antituberculosas e antimaláricas, o que remete a um crescente interesse por essa classe de moléculas (31-33).

Os poliéteres ionóforos policíclicos são geralmente ramificados, polioxigenados e seus éteres cíclicos têm afinidade por cátions metálicos monovalentes ou divalentes. A ligação desses cátions metálicos dentro de uma matriz hidrofóbica causa o seu sequestro e evita assim sua difusão através da membrana. Dessa forma, os poliéteres ionóforos exercem sua atividade biológica pela destruição fisiológica do gradiente de íons na membrana plasmática das células. Entre os antibióticos dessa classe estão inclusos a lasalocida, a monensina, a tetronomicina, a tetronasina, a nigericina e a salinomicina (34). Todos esses compostos têm suas rotas biossintéticas determinadas (35, 36).

A biossíntese dos poliéteres ionóforos policíclicos se inicia pela condensação de moléculas derivadas de acetato, propionato e butirato por uma PKS do tipo I e modificações posteriores são adicionadas por glicosiltransferases, *O*-metiltransferases, hidroxilases e óxido-redutases (34).

Os antibióticos tetronomicina e tetronasina são poliéteres ionóforos com estruturas extremamente similares, porém as configurações de seus centros quirais são opostas (Figura 1).

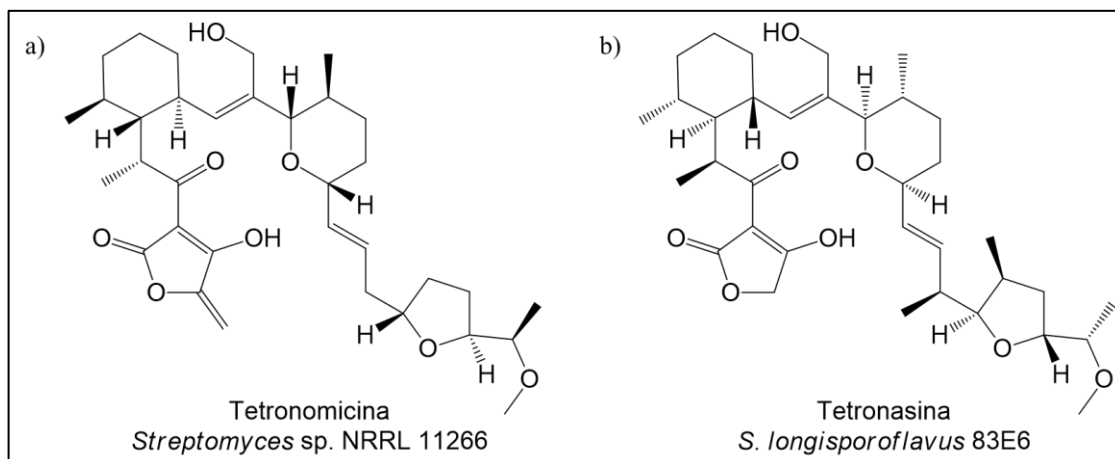


Figura 1. Estruturas químicas da tetronomicina (a) e da tetronasina (b).

A natureza apresenta um evento evolutivo extremamente elegante para a biossíntese desses dois compostos em organismos diferentes, sendo eles *S. longisporoflavus* e *Streptomyces* sp. NRRL 11266 para tetronasina e tetronomicina, respectivamente. Assim, a biossíntese desses compostos tem reações que envolvem reconhecimento por estereoespecificidade (37).

### 1.3.1 Diels-alderases em poliéteres ionóforos tetronatos

O ácido tetrônico (um anel de 5 membros) pode estar presente em diversos produtos naturais como alcaloides, terpenos, macrolídeos e taninos. Os produtos naturais que apresentam este grupo funcional são chamados de tetronatos (38). O ácido tetrônico pertence à classe dos 4-hidroxi-2(5H)-furanos, que são caracterizados por um anel 4-hidroxi-2(5H)-furanos e pode ser encontrado em poliéteres como a tetronomicina e a tetronasina (39). Os tetronatos são uma classe diversa e interessante de produtos naturais com potencial clínico (39, 40), apresentando uma grande diversidade de atividades biológicas, como a inibição da replicação do HIV (41) e a inibição do crescimento de células tumorais (39).

Os tetronatos podem ser divididos em dois grupos: tetronatos lineares e espirotetronatos. Os tetronatos lineares apresentam uma unidade de ácido tetrônico ligado a um policetídeo linear ou a um esqueleto de ácido graxo. Em contraste, os espirotetronatos são caracterizados por ter a fração de ácido tetrônico em ligação espiro com um anel ciclo-hexeno. Os espirotetronatos têm atraído a atenção de pesquisadores devido à formação da unidade espirotetronato que é provavelmente formada através de uma cicloadição [4 + 2] intramolecular ou reação Diels-Alder. A reação de Diels-Alder é uma reação química entre um dieno conjugado e um alceno substituído, comumente denominado dienófilo, que sob ação de calor forma um derivado de ciclohexeno substituído (42). Nos espirotetronatos, um duplo exocíclico presente no anel

tetronato sofre a cicloadição [4 + 2] com o dieno terminal do policetídeo linear. A presença de uma reação Diels-Alder biossinteticamente imprescindível sugere que uma enzima possa ser necessária para realizá-la, uma proposta amplamente discutida (39, 43-46).

Apesar de seu uso generalizado na química orgânica sintética, não é claro se enzimas evoluíram para catalisar as reações de Diels-Alder. Várias enzimas Diels-alderases foram descritas, SpnF da via biossintética da spinosina (26), LovB da via de biossíntese da lovastatina (47) e SPS da via biossintética de solanopirona (48), por exemplo. Porém, essas reações também podem ocorrer espontaneamente e as enzimas que as catalisam frequentemente têm outras funções distintas, tornando pouco claro se a ciclização é a função evolutiva dessas enzimas ou uma atividade secundária.

A primeira via para a qual a formação de espirotetronatos foi demonstrada *in vitro* com uma enzima que catalisa uma cicloadição [4 + 2] foi a via de biossíntese de versipelostatina (Figura 2).

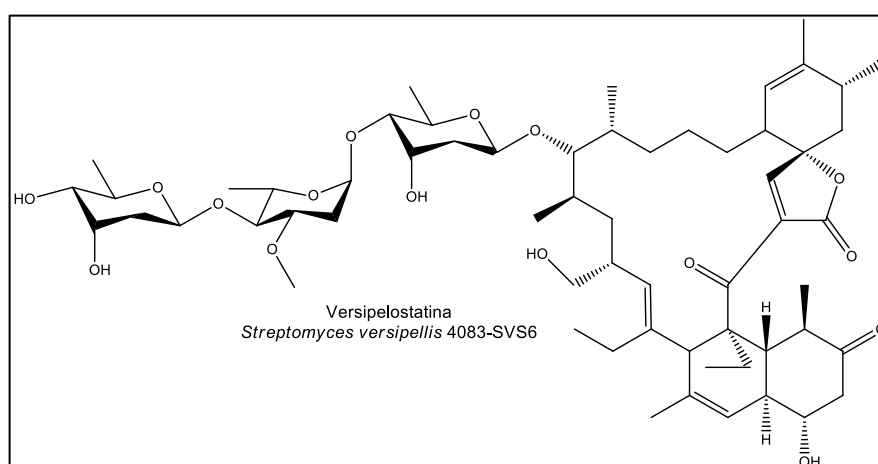


Figura 2. Estrutura da versipelostatina de *Streptomyces versipellis* 4083-SVS6.

A versipelostatina é um espirotetronato produzido por uma PKS modular de tipo I em *Streptomyces versipellis* 4083-SVS6. Ela regula a transcrição do gene que codifica a proteína Grp78, uma chaperona retículo endoplasmática envolvida em doenças como Alzheimer e Parkinson (49). Dentro do BGC para versipelostatina, uma proteína pequena de 142 aminoácidos, a VstJ, foi identificada como responsável por catalisar a reação de Diels-Alder entre a ligação dupla exocíclica do tetronato e o terminal 1,3-dieno do policetídeo intermediário (50) (Figura 3).



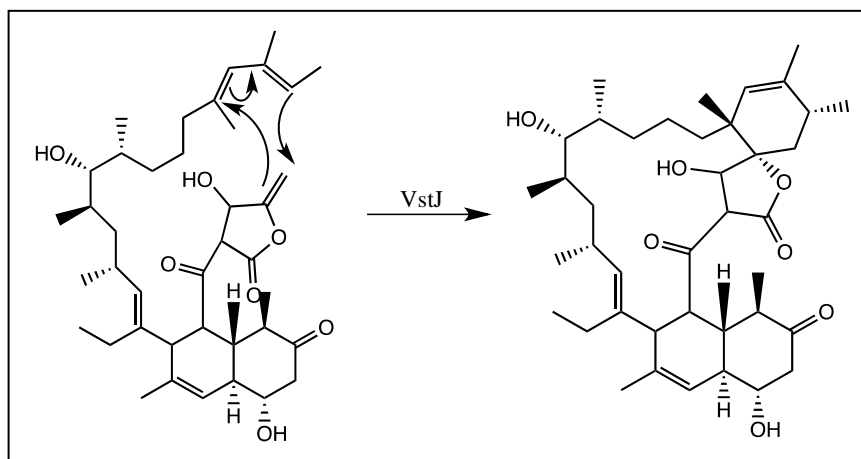


Figura 3. Formação de espirotetronato catalisada por enzima da biossíntese de versipelostatina.

A identificação de VstJ como uma Diels-alderase formadora de espirotetronatos permitiu que os genomas de outras Actinobactérias produtoras de espirotetronatos fossem pesquisados para identificar proteínas homólogas. Um resultado interessante dessa pesquisa por homologia foi a descoberta de que Tmn8 e Tsn15, enzimas de função desconhecida dentro dos *clusters* gênicos da biossíntese de tetronomicina (Tmn) e tetronasina (Tsn), respectivamente, são homólogas a VstJ (50). Tmn e Tsn são antibióticos poliéteres ionóforos tetronatos produzidos por enzimas PKS de Tipo I (37, 51).

A presença de duas ciclases presentes nos conjuntos de genes de tetronomicina, Tmn8 e Tmn9, e de tetronasina, Tsn11 e Tsn15, sugere que a formação desses anéis é catalisada por esses homólogos de ciclase (Figura 4). Entender como isso ocorre nos níveis mecanísticos e estrutural se faz de suma importância dentro da bioquímica de produtos naturais microbianos.

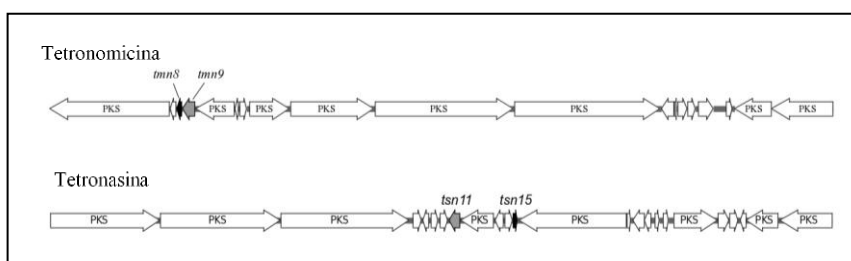


Figura 4. *Clusters* gênicos da tetronomicina e da tetronasina.

#### 1.4 Benzoxazóis

A natureza ainda apresenta outras estratégias para a formação de outros anéis, como por exemplo, a biossíntese de compostos com anéis benzoxazóis. A família de compostos benzoxazóis inclui produtos naturais com atividade antileishmania como o A33853 (52), antitumoral como o Nataxazol (53), atividade ionofórica como a Calcimicina (54),

antimicrobiana como a Caboxamicina (55) e citotóxica como UK-1 (56) e AJI9561 (57) produzidos por *Streptomyces* sp. A fração característica dessa família de compostos é um anel oxazol heterocíclico fundido com um anel benzeno que forma o motivo benzoxazol.

O composto A33853, isolado do *Streptomyces* sp. NRRL 12068, mostra uma excelente bioatividade contra leishmaniose e é um benzoxazol formado por duas moléculas de ácido 3-hidroxi-antranílico (3-HAA) e uma unidade de ácido 3-hidroxi-picolínico (3-HPA) derivado do corismato (58). A determinação do *cluster* gênico para esse composto bioativo mostra que na sua biossíntese há a participação de uma série de enzimas pouco usuais (52). Uma coenzima-A ligase (BomJ) dependente de ATP e uma amidohidrolase (BomN) parecem estar envolvidas na montagem do anel benzoxazol. BomJ catalisa uma adenilação dependente de ATP de um substrato de carboxilato na sua forma tiolada correspondente. A ativação do grupo carboxílico do precursor para a formação subsequente da ligação amida ocorre através de um ataque nucleofílico direto do grupo amino de um 3-HAA para o grupo acila de outro 3-HAA ligado a AMP. BomJ parece pertencer a superfamília das AFE (*Adenylate-Forming Enzymes*) por conter um motivo de ligação de AMP conservado (59). Uma análise preliminar utilizando a sequência da BomJ revela que ela apresenta uma identidade de aproximadamente 27 % com CoAs ligases. BomN catalisa a reação de heterociclicização do anel de benzoxazol por um mecanismo catalítico ainda não conhecido.

O composto AJI9561 (Figura 5) é estabelecido como um derivado do agente antitumoral UK-1, que exibe atividade citotóxica (57).

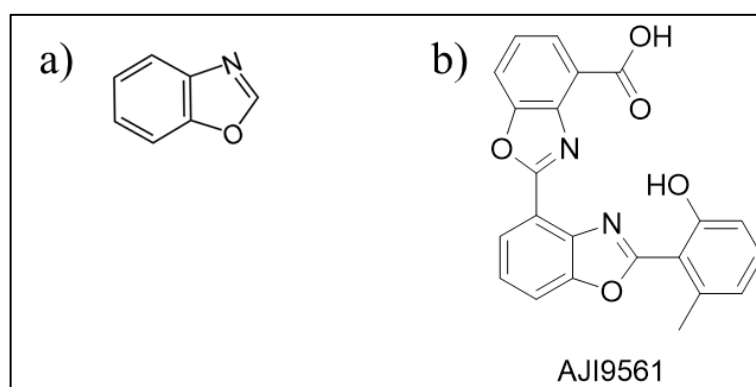


Figura 5. a) Anel benzoxazol; b) Estrutura de AJI9561, um composto benzoxazol.

AJI9561 contém um esqueleto composto por duas moléculas de ácido 3-hidroxi-antranílico (3-HAA) e uma unidade de ácido 6-metilsalicílico (6-MSA) (60). Embora o BGC desse composto ainda não tenha sido publicado, este foi identificado pelo grupo do Prof. Peter Leadlay, Universidade de Cambridge-UK, e a sequência gênica para *ajiA1* é altamente similar ao gene *bomJ*.

Devido à ampla gama de atividades biológicas, os benzoxazóis constituem uma importante classe de compostos terapêuticos e biotecnológicos e chamam a atenção dos químicos para a síntese de novos derivados. No entanto, a biossíntese de compostos benzoxazóis ainda não foi explorada em detalhes (60), bem como sua produção por actinomicetos.

### 1.5 Métodos para o estudo da estrutura de uma proteína

O estudo dos genes e enzimas responsáveis pela biossíntese de produtos naturais foi bastante facilitada pelos recentes avanços nos métodos para a determinação da estrutura tridimensional de uma proteína. A informação estrutural é essencial para entender como as moléculas biológicas funcionam e fornece conhecimento para a obtenção de novos medicamentos.

A cristalografia de raios-X é uma ciência dedicada ao estudo da estrutura molecular e cristalina, bem como das relações entre essa estrutura e suas propriedades. Embora a estrutura 3D de macromoléculas pode ser obtida através de diversos métodos experimentais, tais como a ressonância magnética nuclear (RMN) e a criomicroscopia eletrônica, a cristalografia de raios-X ocupa papel de destaque nesse processo, visto que 90 % das macromoléculas depositadas no PDB (*Protein Data Bank*) têm sua estrutura 3D determinada pelo método de cristalografia de raios-X (61).

Na cristalografia, o espaçamento dos átomos em uma estrutura cristalina pode ser determinado pela medida das fases e intensidades de pontos produzidos por um feixe de raios X de certo comprimento de onda, depois que o feixe tenha sido difratado pelos elétrons dos átomos de proteínas presentes no cristal. Um cristal arranja um número grande de moléculas na mesma orientação, dessa maneira é esperado que as ondas espalhadas possam ser adicionadas em fase aumentando o sinal a um nível que possa ser medido. A combinação dos pontos difração de diferentes imagens, mais a obtenção da informação de sua fase (obtida por métodos experimentais ou computacionais) utilizando uma transformada de Fourier permite a construção de um mapa de densidade eletrônica que está relacionado com a posição dos átomos na rede cristalina (62).

A produção de proteínas heterólogas em *Escherichia coli* tem facilitado a obtenção de proteínas para caracterização bioquímica, produção de anticorpos e determinação da estrutura tridimensional (63). Uma etapa fundamental para se alcançar sucesso na cristalização é a obtenção do alvo molecular em quantidade e pureza suficiente para os estudos cristalográficos. A concentração média utilizada para determinação das estruturas de proteínas depositadas no PDB é de  $14 \text{ mg.mL}^{-1}$ . No entanto, há exemplos de estruturas cristalizadas entre 2 e  $100 \text{ mg.mL}^{-1}$ .

<sup>1</sup>. A produção heteróloga da proteína tem como principais vantagens maior rendimento de proteína, uso de organismos de baixo nível de segurança com meios de cultura bem estabelecidos, maior facilidade na purificação e possibilidade de produzir diferentes construções da proteína (61, 64).

A expressão de proteínas com cauda de histidina (*His-tag*) tem possibilitado a obtenção de alto grau de pureza para ensaios cristalográficos, necessários para a compreensão da estrutura e função da proteína alvo (65). A coluna utilizada HisTrap HP consiste em uma matriz sólida de agarose contendo grupos quelantes imobilizados carregados com íons  $\text{Ni}^{2+}$ , os quais se ligam às histidinas expostas na superfície das proteínas. O imidazol é um composto semelhante ao radical da histidina e atua como competidor nas interações com o níquel, favorecendo o desligamento das proteínas da coluna. Já a cromatografia por exclusão molecular separa as proteínas de acordo com o tamanho. A coluna é constituída por uma matriz de pequenas esferas porosas empacotadas. Ao fazer passar a solução de proteínas pela matriz da coluna, as moléculas pequenas entram nos poros das esferas demorando a atravessá-los, enquanto que as moléculas grandes passam entre as esferas sendo separadas primeiro (62).

A cristalização de macromoléculas baseia-se em um conjunto diverso de princípios e não há uma teoria abrangente ou mesmo uma base de dados fundamentais para orientar os esforços experimentais. Como consequência, o crescimento do cristal macromolecular é em grande parte de natureza empírica e exige paciência, perseverança e intuição (66).

A obtenção de cristais adequados para os experimentos de difração de raios-X é fundamental para a determinação da estrutura tridimensional de macromoléculas. O fenômeno de cristalização ocorre quando a molécula em estudo precipita de modo lento e ordenado, formando cristais. O processo ocorre em condições controladas, incluindo uma solução supersaturada da proteína de interesse, agentes precipitantes, condições controladas de temperatura, força iônica e em pequenos intervalos de variação de pH. As propriedades da proteína, como distribuição de cargas na superfície, presença de regiões flexíveis e distribuição de conformações têm impacto significativo no fenômeno de cristalização. A técnica de *Thermal Shift (Thermofluor)* para realizar a troca do tampão da proteína para aquele que concede maior estabilidade à molécula pode também ajudar no sucesso da cristalografia (67, 68).

Para o maior entendimento da função de uma proteína, é preciso conhecer sua flexibilidade e movimentos ao longo do tempo, agregando informação às imagens estáticas de uma biomolécula obtidas por exemplo por cristalografia. Nas décadas de 70 e 80, a atividade biológica das proteínas começou a ser relacionada à sua dinâmica molecular e não apenas à sua estrutura (69). A dinâmica molecular (DM) é definida pela IUPAC (*International Union of Pure*

*and Applied Chemistry*) como um procedimento de simulação que consiste na computação do movimento dos átomos em uma molécula ou de átomos individuais ou moléculas em sólidos, líquidos e gases, de acordo com as leis de movimento de Newton. Essa técnica descreve, portanto, a variação do comportamento molecular como função do tempo. A DM faz parte dos métodos denominados Clássicos também chamados de métodos de mecânica molecular, em oposição aos métodos baseados na mecânica quântica. A maioria dos fenômenos biológicos, como por exemplo a associação de substratos a enzimas, está associada à flexibilidade de biomoléculas. Considerando esses movimentos, a DM nos possibilita obter modelos de moléculas muito mais próximos da realidade biológica (70).

Devido à complexidade dos sistemas biológicos, uma abordagem colaborativa e multidisciplinar que se vale da complementariedade de técnicas se faz essencial para uma compreensão global desses sistemas. Simulações de DM podem fornecer simultaneamente informações com alta resolução atômica e temporal de uma proteína em solução aquosa, auxiliando na descoberta das relações entre estrutura, dinâmica e função. As técnicas de simulações computacionais por DM são de uso comum nas ciências, sendo amplamente discutidas na literatura (71-76). A técnica se baseia na determinação do movimento dos átomos de um sistema de interesse a partir do conhecimento das configurações iniciais (posições), dos potenciais de interação entre eles e das equações que regem o movimento. Além disso, a correta representação do sistema em um determinado estado termodinâmico exige uma série de aproximações, tal como o uso de condições periódicas de contorno, somas de Ewald, termostatos, entre outros (70, 75-79). As interações interatômicas e intermoleculares são representadas por potenciais efetivos clássicos (campos de força) entre átomos individuais (*all-atoms models*) ou entre grupos específicos de átomos (*united-atoms models* ou *Coarse-Grain models*) (71, 79). As interações entre pares atômicos não diretamente ligados na molécula, bem como entre átomos em moléculas distintas são descritas pela soma de potenciais de Lennard-Jones e de Coulomb. Existem diversos campos de força desenvolvidos especificamente para proteínas, sendo OPLS (80, 81), GROMOS96 (82), CHARMM (83) e AMBER (84) os mais utilizados. A partir das configurações iniciais de um modelo, o método de DM requer velocidades iniciais para todos os átomos da caixa de simulação. Uma maneira de obter essas velocidades é atribuí-las de forma aleatória a todos os átomos, por meio da distribuição de Maxwell-Boltzmann em que a energia cinética do sistema é determinada de maneira aproximada pela temperatura.

Ao final de uma simulação de DM são obtidos conjuntos de coordenadas e velocidades que representam a evolução temporal do sistema nas condições termodinâmicas selecionadas.

A partir da média temporal de uma grandeza calculada ao longo de uma trajetória é possível obter propriedades microscópicas e macroscópicas do sistema representativas de comportamentos biológicos medidos experimentalmente (70).

## 1.6 Justificativa

Entre 1940 e 1970, quase todas as principais classes de antibióticos foram descobertas, incluindo as tetraciclina, os macrolídeos e os aminoglicosídeos (85, 86). No mesmo período, o número de cepas resistentes a esses novos medicamentos também cresceu, devido à pressão de seleção dos antibióticos sobre os patógenos. Os genes que conferem resistência a antibióticos (genes de resistência) podem inclusive serem transferidos de uma cepa resistente para uma cepa não resistente por um processo chamado transferência horizontal de genes (87). Inúmeros mecanismos de resistência a antibióticos em bactérias têm sido caracterizados (88). A pesquisa por novas moléculas bioativas é de suma importância, uma vez que várias doenças infecciosas não são ainda controladas adequadamente pelos antibióticos existentes e o desenvolvimento de patógenos resistentes exige a contínua descoberta de novos fármacos.

A formação de anéis heterocíclicos em compostos naturais é um evento bastante comum, porém pouco compreendido e somente recentemente as bases bioquímicas e moleculares de sua formação começaram a ser reveladas. A maioria desses anéis são formados por reações enzimáticas pouco usuais ou por enzimas que apresentam novas funções. Cada vez mais há interesse em entender os mecanismos enzimáticos e/ou a estrutura das enzimas envolvidas na biossíntese das moléculas que possuem esses anéis, a fim de se manipular os passos na tentativa de obtenção de compostos derivados, como demonstrado por Luhavaya et al., 2015 (28). Esses estudos têm ganhado um forte interesse da comunidade científica da área de química biológica e grupos de pesquisa que estudam esses mecanismos têm produzido trabalhos de altíssima excelência na área (26-28, 41, 89). Do ponto de vista estrutural, essas enzimas apresentam baixa similaridade com qualquer outra proteína que tenha sua estrutura resolvida, o que pode levar a descoberta de novos tipos de enovelamento.

Assim sendo, realizar estudos estruturais a respeito dessas classes de moléculas é de grande valor científico devido à possibilidade de se gerar resultados com alto impacto científico. Este trabalho não só poderá contribuir com aplicações em biologia sintética na manipulação das vias biossintéticas de produtos naturais, como também levou a determinação de estruturas cristalográficas completamente novas que poderão ser publicadas em revistas de alto impacto científico.

## 2 OBJETIVOS

### Objetivo Geral:

Caracterizar estruturalmente diferentes enzimas responsáveis pela biossíntese de anéis heterocíclicos em poliéteres ionóforos policíclicos e benzoxazóis. Com este estudo, espera-se obter informações que contribuam para o entendimento de reações incomuns realizadas pela natureza e que podem fornecer importantes *insights* para as áreas de biologia sintética e biossíntese combinatorial.

### Objetivos Específicos:

- Elucidar a estrutura tridimensional das enzimas: Tmn9 e Tsn15 da biossíntese de tetronomicina e tetronasina, respectivamente (responsáveis pela formação do anel central desses poliéteres ionóforos tetronatos) e da enzima AjiA1 da biossíntese do composto Aji9561 (responsável pela formação do anel benzoxazol desse composto) que foi recentemente identificada pelo nosso colaborador, o Prof. Peter Leadlay, Universidade de Cambridge-UK;
- Realizar análises de comparação estrutural com homólogos e propor mecanismos de catálise das enzimas: Tmn9, Tsn15 e AjiA1;
- Estabelecer a atividade enzimática de AjiA1 por meio de uma abordagem *in vitro*;
- Estudar a estrutura e a dinâmica do sítio de ligação da enzima Tmn9 computacionalmente por meio de modelagem molecular multiescala (ancoragem e Dinâmica Molecular).

### 3 MATERIAL E MÉTODOS

#### 3.1 Clonagem

Os genes das enzimas AjiA1, Tmn9 e Tsn15 foram clonados em vetor de expressão do tipo pET28a com cauda de histidina no N-terminal pelo grupo de Pesquisa do Prof. Peter Leadlay da *University of Cambridge-UK* para expressão recombinante em *E. coli*.

#### 3.2 Produção de proteínas recombinantes

Transformações foram realizadas em células competentes por choque térmico, inicialmente em *E. coli* DH5 $\alpha$  para amplificação do plasmídeo e posteriormente em cepas de *E. coli* BL21(DE3) para a expressão da proteína de interesse.

As células da cultura foram crescidas a 37 °C sob agitação de 200 rpm, sendo seu crescimento monitorado em espectrofotômetro até alcançar uma leitura de densidade óptica (DO<sub>600</sub>) de 0,6. Nesse momento, foi adicionado o indutor de expressão IPTG (Isopropil  $\beta$ -D-1-tiogalactopiranosídeo) para uma concentração final de 1 mM. Após a adição do indutor, os cultivos foram mantidos à 18 °C por 18 horas. Ao final desse processo, os meios foram centrifugados a 4500 rpm por 30 minutos para armazenamento do *pellet* a -20 °C. Foram expressos 8 litros de meio LB de cada enzima.

#### 3.3 Purificação

Para a purificação, o *pellet* foi descongelado e as células ressuspendidas em Tampão A (descrito a seguir), a fim de se homogeneizar a amostra. Em seguida foi adicionado 400  $\mu$ L do inibidor de protease PMSF (fluoreto de fenilmetilsulfonila) 10 mM para evitar a degradação proteolítica por proteases bacterianas, 400  $\mu$ L de DNase 1 mg.mL<sup>-1</sup> e 200  $\mu$ L de lisozima 10 mg.mL<sup>-1</sup>. As células foram rompidas com auxílio de sonificador ultrassônico (BransonSonifier) em 30 % de amplitude, com pulsação de 1 em 1 segundo, por 6 minutos e a solução celular foi centrifugada por 60 minutos à 12000 rpm para a separação das fases solúvel e insolúvel.

As proteínas expressas com os vetores utilizados nesse trabalho possuem em seu amino-terminal uma cauda de histidina que auxilia na purificação da proteína. Dessa forma, as proteínas solúveis foram purificadas inicialmente através de cromatografia de afinidade ao níquel em coluna HisTrap HP de 5 mL (GE Healthcare) e posteriormente por meio de cromatografia de exclusão molecular, ambas utilizando aparelho ÄKTA Purifier ou ÄKTA start (GE Healthcare).



Na cromatografia de afinidade por íons metálicos imobilizados (IMAC), a coluna foi previamente equilibrada com o mesmo tampão utilizado na lise das células (tampão A – 500 mM Hepes pH 8,0, 300 mM NaCl, 10 % glicerol). Após a aplicação da fração solúvel do lisado, a coluna foi lavada com 10 volumes de coluna de tampão A, ou até que nenhuma leitura de absorvância a 280 nm fosse observada. A eluição das proteínas que se ligaram à coluna foi feita utilizando um gradiente de concentração crescente com o tampão de lise contendo imidazol 500 mM (tampão B).

As frações de proteína coletadas foram centrifugadas até um volume de 1-2 mL e submetidas a uma nova etapa de purificação por cromatografia de exclusão molecular com a utilização das colunas HiLoad 16/60 Superdex 200 ou HiLoad 16/60 Superdex 75 (GE Healthcare Life Sciences). Foi utilizado para essa purificação o tampão 50 mM Hepes pH 8,0 e 300 mM de NaCl. A coluna foi equilibrada com 1,2 volume de coluna de tampão e eluída com 1 volume de coluna de tampão após a amostra ter sido injetada. Frações que continham as proteínas de interesse foram unidas, concentradas e tiveram suas concentrações dosadas para, se possível, serem submetidas a ensaios de cristalização.

Amostras recolhidas ao longo de todo o processo foram submetidas à análise por SDS-PAGE.

### **3.4 Eletroforese em gel de poliacrilamida-SDS - Análise da expressão e purificação das proteínas**

A expressão das proteínas e o resultado da purificação foram analisados em géis de SDS-PAGE 15 % corados com Coomassie Blue brilhante. O marcador de peso molecular utilizado foi o Marcador molecular Precision Plus – 250 kD (BioRad).

### **3.5 Determinação da concentração proteica**

As frações com a proteína de interesse foram concentradas em filtros de ultracentrifugação AmiconUltra-15 (Millipore) e a dosagem das proteínas foi realizada em Nanodrop® ou BioDrop  $\mu$ LITE (Analítica ou Biochrom). Os coeficientes de extinção para as proteínas fusionadas ou não com a cauda de histidina foram obtidos através do programa ProtParam (90), hospedado no servidor proteômico ExPASy - Expert Protein Analysis System (91). As proteínas purificadas foram armazenadas em alíquotas de 50  $\mu$ L em temperatura de - 80 °C.

### 3.6 Ensaio de cristalização das proteínas

O sistema de *sitting drop* (gota sentada) foi utilizado para os testes de cristalização com os *screens* padronizados e comercializados pela Jena Bioscience e Hampton Research. Nessa etapa foi utilizado um robô de cristalização Oryx 4 (Douglas Instruments), presente em nosso laboratório (LBEA-Laboratório de Biologia Estrutural Aplicada), para a realização de *screenings* utilizando vários kits de condições de cristalização. Foram utilizadas placas de 96 poços MRC2 (Swissci) onde foram depositados 30  $\mu\text{L}$  de condição em cada poço e na gota 0,3  $\mu\text{L}$  de condição mais 0,3  $\mu\text{L}$  de proteína a uma concentração em torno de 10  $\text{mg}\cdot\text{mL}^{-1}$  dependendo da proteína em questão.

Quando da obtenção de cristais, esses foram otimizados variando-se parâmetros como pH, concentração de proteínas ou precipitantes. A otimização foi realizada manualmente pelo método de gota pendurada (*hanging drop*) em placas VDX de 24 poços (Hampton Research). As placas foram armazenadas à 18 °C.

A proteína Tmn9 foi enviada ao robolab (plataforma de cristalização automatizada) presente nas dependências do Laboratório Nacional de Biociências (LNBIO-CNPEN) para tentativa de cristalização, uma vez que *hits* não surgiram nos testes de cristalização feitos em nosso laboratório.

Os cristais foram colhidos da gota de cristalização e transferidos para uma solução crioprotetora contendo a solução de cristalização e 30 % de glicerol (Tsn15 e Tmn9) ou 15 % de etilenoglicol (AjiA1). Os cristais foram então congelados rapidamente em nitrogênio líquido usando laços de náilon para posterior coleta de dados.

Quando hits não foram obtidos ou visando melhorar a qualidade dos cristais, foram feitas alterações no tampão no qual a proteína se encontrava ou ainda alterações na concentração da proteína.

### 3.7 Ensaio de *Thermofluor*

Com o propósito de buscar um tampão que estabilizasse as proteínas para aumentar a probabilidade de cristalização, realizou-se ensaios de *Thermofluor*. Os ensaios foram feitos no aparelho CFX Connect™ Real-Time PCR Detection System (Bio-Rad) em placa de PCR de 96 poços (Bio-Rad). Foram adicionados em cada poço 20  $\mu\text{L}$  de tampão (100 mM) contendo ou não 150 mM de sal (NaCl ou KCl) ou 5 % glicerol, 5  $\mu\text{L}$  de SYPRO Orange (5x), a proteína à concentração final de 10  $\mu\text{M}$  e água para completar o volume de 25  $\mu\text{L}$ . A emissão de

fluorescência foi monitorada durante a variação da temperatura do sistema. As placas foram aquecidas de 25 °C a 95 °C a uma taxa de calor de 0,5 °C.min<sup>-1</sup>. A relação de tampões utilizados foi retirada do artigo publicado por Reinhard et al, 2013 (92) e ligeiramente modificada pela disponibilidade de reagentes de nosso laboratório.

### **3.8 Coleta e processamento de dados de difração de raios-x**

A coleta de dados de difração de raios-X foi realizada remotamente na linha Petra III do laboratório DESY (*Deutsche Elektronen-Synchrotron*), Alemanha. O processamento dos dados de difração de raios-X foi realizado usando o programa XDS (93). A preparação de figuras foi realizada utilizando o programa PyMol Molecular Graphics System, Version 1.8 Schrödinger, LLC (94) ou COOT (95).

Os cristais foram selecionados de acordo com características que são esperadas para obtenção de um alto poder de difração, como tamanho e presença de monocristais. O processamento e a aferição da qualidade geral dos dados de difração de raios-X foram realizados usando os programas XDS (93), AIMLESS e SCALA (96) presentes no pacote de programas do CCP4 (97).

### **3.9 Resolução e refinamento cristalográfico**

As estruturas foram resolvidas por substituição molecular utilizando o programa Molrep presente no pacote de programas do CCP4 (97) e o programa Phaser do pacote de programas Phenix (98). O refinamento cristalográfico foi realizado utilizando o programa REFMAC (99, 100) e o Phenix-refine do pacote de programas Phenix (98). A inspeção e a construção dos modelos foram efetuadas por meio do programa Coot (95) utilizando mapas de densidade eletrônica 2Fo-Fc, Fo-Fc e omit map.

### **3.10 Simulações de Dinâmica Molecular**

As simulações de Tmn9 foram baseadas na estrutura cristalina de 2,4 Å descrita aqui (PDB 6UI5). As coordenadas foram preparadas removendo primeiro a cadeia B. Os resíduos de histidina foram simulados como o tautômero neutro de Nε, exceto His34 que foi considerada como carregada. Outros resíduos tituláveis foram simulados em seu estado de protonação dominante em pH 7,4, exceto Asp284 que foi considerado como não carregado. A protonação foi verificada calculando-se o pKa dos grupos ionizáveis pelos servidores H ++ (101), PROPKA (102) e Playmolecule (103). As 38 moléculas de água contidas na estrutura

cristalográfica (águas estruturais) foram mantidas. Os átomos de hidrogênio foram adicionados pelo protocolo CHARMM-GUI (104, 105).

### 3.10.1 Configuração do sistema e parâmetros do campo de força

O sistema foi preparado usando o “*Solution builder*” da interface gráfica para geração de configurações de sistema de simulação molecular CHARMM-GUI (104, 105). As simulações foram realizadas usando Gromacs (106), utilizando o campo de força CHARMM36m (107) para o modelo de proteína.

A Tmn9, incluindo o cofator FAD, foi solvatada com um sistema hidratado e equilibrado usando a interface CHARMM-GUI (104, 105). Foram adicionados íons sódio e cloreto para neutralizar o sistema e manter uma concentração fisiológica de sal de 0,15 M. A caixa retangular de simulação de  $96 \times 96 \times 96 \text{ \AA}^3$  continha 99 íons sódio, 71 íons cloreto e cerca de 25000 moléculas de água, num total de aproximadamente 85000 átomos.

Os parâmetros do FAD foram derivados dos parâmetros disponíveis para CHARMM das moléculas FMN e ADP (108). Os parâmetros do substrato e do produto foram gerados com o “CHARMM General Force Field” e algumas correções foram adicionadas para garantir a planaridade das ligações duplas conjugadas (109, 110). O modelo TIP3P foi utilizado para moléculas de água (111, 112). Todas as simulações foram realizadas como um conjunto isotérmico-isobárico a temperatura e pressão constantes de 310,15 K e 1 bar. A pressão constante foi mantida usando o algoritmo Parrinello-Rahman, com flutuação do pistão pela dinâmica de Langevin (113, 114), e a temperatura constante foi mantida pelo algoritmo de Nose-Hoover. O método PME (115) do inglês *Particle Mesh Ewald*, método do cálculo do potencial eletrostático, foi utilizado para interações eletrostáticas de longo alcance. Para a integração das equações de movimento foi utilizado o algoritmo Velocity-Verlet. O sistema foi simulado sob condições de contorno periódicas e pequenas etapas iniciais de minimização foram realizadas para remover maus contatos. O sistema foi minimizado em 5000 etapas e equilibrado por 25 ps com restrições de posição. Após essa etapa, outra equilibração sem restrições foi conduzida por 10 ns e, em seguida, o sistema foi simulado sem restrições por 500 ns em triplicata.

### 3.11 Docagem ou Ancoramento Molecular

Para estudar o modo de ligação e a dinâmica do substrato e do produto, o Autodock Vina (116) foi utilizado para gerar poses de ligação usando parâmetros de exaustividade de pesquisa global de 40 e o número máximo de poses de saída = 9. O espaço de busca foi

confinado a uma caixa de tamanho 25 Å centralizada perto do átomo N5 do FAD ( $x = 20,94$ ,  $y = 7,09$ ,  $z = 229,20$ ). O receptor foi tratado como um corpo rígido e preparado usando o software Openbabel (117) para tipagem de átomos e atribuição de cargas. Foram utilizadas as cargas de Gasteiger-Marsili (118).

Para tirar proveito da amostragem conformacional gerada durante as simulações de 500ns de Dinâmica Molecular da Tmn9 mencionadas anteriormente, obtivemos dez conformações representativas dos *clusters* gerados durante a dinâmica. Os *clusters* foram gerados comparando o RMSD das coordenadas de carbono alpha e a abertura do *cap* (domínio intermediário), selecionando o representante de cada *cluster*. Portanto, aumentamos a amostragem de conformações, em contraste com uma estrutura única usada na maioria dos estudos de docagem. Também foram utilizadas estruturas de homólogos (6brdB, 5kox, 6c7s, 5xgv) e a cadeia B de Tmn9 para aumentar nossa amostragem. As estruturas das proteínas homólogas selecionadas foram co-cristalizadas com RFP ou RFH. Os estados de protonação dos receptores foram previstos usando o servidor Protoss (119).

As poses geradas pela docagem foram agrupadas através da criação de *clusters* e da seleção do representante de cada *cluster* para análise posterior. As estruturas resultantes foram agrupadas com base em sua orientação comparativamente com a orientação dos ligantes nas estruturas homólogas, se estavam ou não enterradas na cavidade do sítio ativo e na proximidade com o FAD. As poses representativas foram validadas com novas simulações de Dinâmica Molecular utilizando agora um sistema com a Tmn9, o cofator FAD e o ligante (substrato ou produto). Foram usados os mesmos parâmetros descritos na seção 3.10 para essas simulações adicionais.

Para validar o protocolo do docking, realizamos um re-docking da estrutura da homóloga Rox (PDB: 6BRD). A adequação foi considerada pois o encaixe do ligante no sítio ativo reproduzia a conformação da estrutura cristalográfica.

### 3.12 Ensaio enzimático

Ensaio enzimáticos foram realizados com a enzima AjiA1 por um método espectrofotométrico contínuo acoplado usando o Kit de Ensaio de Pirofosfato EnzChek™ (Molecular Probes). As misturas de reação iniciais continham 0,2 mM de MESG (2-amino-6-mercaptopurina ribonucleosídeo), 1 U de nucleosídeo purina fosforilase, 0,01 U de pirofosfatase inorgânica, 10 mM ATP, tampão de reação 20X e AjiA1 (8 µM) em um volume total de 100 µL. As reações foram incubadas a 25 ° C por 10 min e depois iniciadas pela adição do substrato em várias concentrações. O aumento da absorbância a 360 nm resultante da

liberação fosforolítica da base de purina (2-amino-6-mercapto-7-metilpurina) do MESG foi monitorado durante um período de 15 min. Os parâmetros cinéticos  $K_m$  e  $V_{m\acute{a}x}$  foram calculados ajustando as taxas iniciais a um modelo hiperbólico de regressão não linear sob a premissa de uma cinética de Michaelis-Menten, utilizando o software GraphPad Prism 5.01. Para todas as medições, foram realizados três experimentos independentes, cada um em triplicata.

## 4 CAPÍTULO I - FORMAÇÃO DE ANÉIS BENZOXAZÓIS

### **Objetivo geral:**

- Determinação e análise estrutural da proteína AjiA1 da biossíntese do composto Aji9561

Este capítulo gerou o artigo “Crystal structure of AjiA1 shows a novel structural motion mechanism into the adenylate forming enzyme family.” Submetido em junho de 2020 para publicação na revista *Acta Crystallographica Section D*.

Autores: Fernanda C. R. de Paiva, Karen Chan, Markiyan Samborsky, Ariel M. Silber, Peter F. Leadlay, Marcio V. B. Dias.

A biossíntese dos antibióticos benzoxazóis envolve duas reações catalisadas por duas enzimas distintas, sendo uma ligase dependente de ATP, um membro das enzimas formadoras de adenilato e uma amidohidrolase (Figura 6).

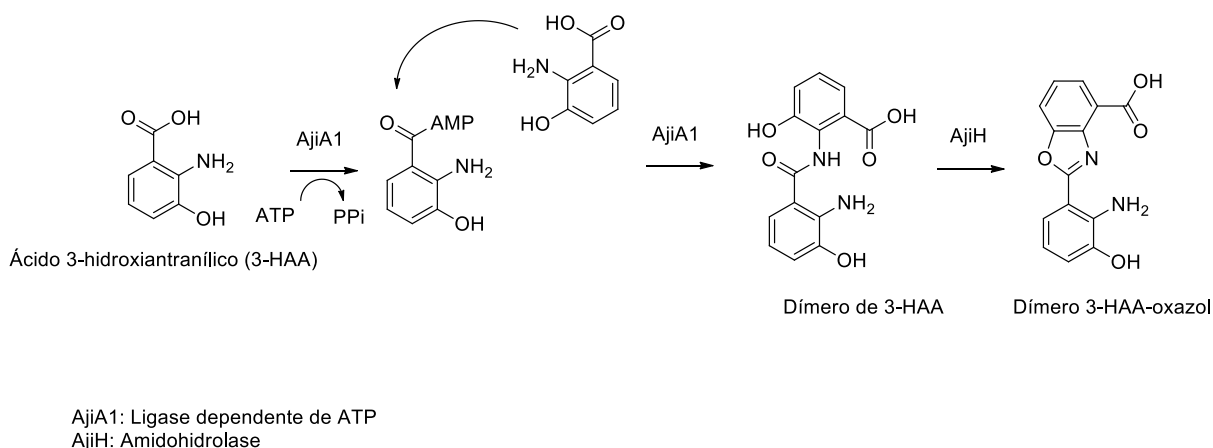


Figura 6. Reações catalisadas por AjiA1 (duas semirreações) e pela amidohidrolase AjiH. A AjiA1 conduz uma ativação dependente de ATP e a dimerização do 3-HAA e a amidohidrolase AjiH realiza o fechamento do anel oxazol que produz o anel benzoxazol.

A reação de adenilação catalisada pelas enzimas formadoras de adenilato (AFEs) é onipresente e desempenha papéis metabólicos ou catabólicos cruciais em todos os organismos, incluindo o metabolismo de ácidos graxos (120), detoxificações (121), bioluminescência (122-124), biossíntese de aminoácidos e proteínas (125) e biossíntese de produtos naturais (52, 126-129). AFEs são um grupo diverso de enzimas dividido em três classes principais: classe I, que inclui o domínio de adenilação de NRPS (Subclasse Ia), sintilases de Acila ou arila-CoA (subclasse Ib) e oxidoreductases ou luciferases de vaga-lume (subclasse Ic). Vários membros das AFEs da classe I, como PaaK1 e PaaK2, envolvidos na degradação do ácido fenilacético em *Burkholderia cenocepacia* (130) e 2-HIBA-CoA ligase (HCL) em *Aquicola tertiaricarbonis*, envolvendo a degradação do éter metil terc-butílico (MTBE) (131) não são adequados para esta classificação e são considerados membros de um novo subgrupo de AFEs da classe I; Classe II, que inclui um número de aminoacil-tRNA sintetases; e classe III, que inclui sintetases de sideróforos independentes de NRPS (NIS) (132, 133). No entanto, vários membros das AFEs não estão incluídos nessa superfamília e são descritos apenas como tendo uma reação de adenilação semelhante (124) como a pimeloil-CoA sintase BioW, que possui um enovelamento completamente diferente e deve pertencer a uma nova classe de enzimas (134).

Embora os membros das AFEs tenham divergências em características e funções estruturais, todos catalisam a mesma transformação bioquímica geral (133, 135). As AFEs geralmente catalisam sua reação em duas etapas: ocorre primeiramente a ativação dependente



de ATP de substratos de ácido carboxílico, como intermediários reativos de acila adenilato (acila-AMP), seguidos pela ligação da enzima a uma molécula aceitadora e a transferência consecutiva do adenilato de acila a um átomo de oxigênio, enxofre ou nitrogênio nucleofílico que gera produtos éster, tioéster e amida (133, 135, 136)

Nas AFEs da classe I, mecanicamente, a ativação dependente de ATP é via acila-CoA sintetases e a segunda reação parcial geralmente forma um tioéster. Estruturalmente, o sítio ativo é formado pela interface entre dois domínios, conectados por uma dobradiça flexível, um *linker* (132, 133). Para realizar a reação de adenilação, é relatado que os membros dessa classe de AFEs sofrem extensas alterações conformacionais durante a segunda metade da reação (135). Geralmente, essas enzimas usam uma rotação de domínio para apresentar faces opostas do domínio C-terminal dinâmico para as reações no sítio ativo (124). Portanto, duas distintas conformações são observadas nas AFEs, nas quais a orientação do domínio C-terminal difere uma da outra em cerca de 140° e cada conformação está envolvida em cada semirreação. Assim, apesar do papel catalítico, o domínio C-terminal está envolvido na formação das duas arquiteturas diferentes de sítio ativo para a ocorrência das duas etapas da reação, um fenômeno chamado alternância de domínio (124, 137, 138).

Embora as AFEs desempenhem um papel crucial na biossíntese de produtos naturais, particularmente peptídeos não-ribossômicos de NRPSs, que requerem um domínio de adenilação durante a incorporação e extensão da cadeia peptídica, recentemente, várias AFEs independentes de NRPS com baixa semelhança de sequência com o Fenilacetato CoA-Ligases, como PaaK1 e HCL, foram identificadas na biossíntese de vários produtos naturais, particularmente benzoxazóis, incluindo Nataxazol (60), Caboxamicina (139), A33853 (52) e AJI9561 (*cluster* gênico não publicado). Os *clusters* gênicos para produtos naturais contendo oxazol indicam a participação de uma série de enzimas incomuns e as rotas biossintéticas para essa classe estruturalmente única de produtos naturais eram pouco exploradas até recentemente. A condensação de moléculas e uma reação de acetalização têm sido propostas como essenciais para a geração dos anéis benzoxazol. Na calcimicina, demonstrou-se que o ácido 3-hidroxi-antranílico (3-HAA) é condensado com a cadeia policetídica ligada a CalA4 (140). No Nataxazol, sugeriu-se que a biossíntese da fração benzoxazol é realizada por meio de uma reação de condensação semelhante entre o ácido 6-metilsalicílico (6-MSA) e o ácido 3-HAA (60). Em A33853, uma coenzima-A ligase dependente de ATP (BomJ) catalisa uma adenilação dependente de ATP de um substrato de carboxilato à sua forma tiolada correspondente. A ativação do grupo carboxílico precursor para formação subsequente da ligação amida ocorre

através de um ataque nucleofílico direto do grupo amino de um 3-HAA ao grupo acila do 3-HAA ligado ao AMP e de uma amido-hidrolase (BomN) na montagem do anel benzoxazol(52).

Recentemente, Song e colegas relataram a estrutura da NatL2, ligase dependente de ATP complexada e, portanto, a estrutura apo ainda não foi elucidada. A existência de alternância de domínio nesse membro das enzimas formadoras de adenilato também não foi algo estudado (141).

A partir da clonagem da enzima AjiA1, ligase dependente de ATP, no vetor de expressão pET28a, foi possível produzir a proteína em sua forma solúvel com rendimento adequado para purificação e posteriores experimentos cristalográficos. Dados gerais da proteína trabalhada foram obtidos de acordo com análise feita pelo programa ProtParam (90) hospedado no servidor proteômico ExPASy (*Expert Protein Analysis System*) (91) e reunidos na Tabela 1.

Tabela 1. Dados gerais da proteína trabalhada.

Antibiótico	Organismo	Peso Molecular	Nº de resíduos de aminoácidos	pI teórico
AjiA1 AJI9561	<i>Streptomyces</i> sp. AJ9561	48,286 Da	436	5,32

Conforme mostrado na Figura 7, AjiA1 eluiu com alto grau de pureza pela técnica de separação por exclusão molecular. Os dados de purificação da proteína foram analisados por gel de SDS e a presença de uma única banda confirma a pureza da amostra, compatível com o esperado para uma proteína de massa molecular próximo a 50 kDa.

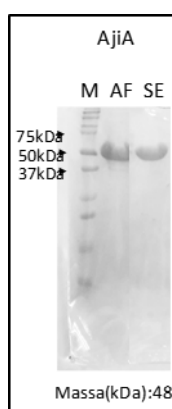


Figura 7. Análise de SDS-PAGE da purificação de AjiA1. As setas ajudam na identificação do tamanho aproximado da banda de interesse: AjiA1 (~48 kDa). M – marcador; AF – cromatografia de afinidade por íons metálicos imobilizados (IMAC); SE – cromatografia de exclusão molecular.

Após a purificação por cromatografia de exclusão molecular, as frações coletadas foram juntadas, concentradas e concentração verificada em um equipamento Biodrop (Biochrom). Ao final foram obtidos  $13 \text{ mg.mL}^{-1}$  de AjiA1. Testes de cristalização foram realizados e em duas condições obtidas a partir dos kits Basic 1-4 (Jena Bioscience) e Pi-Minimal (Jena Bioscience) surgiram pequenos cristais (Figura 8).

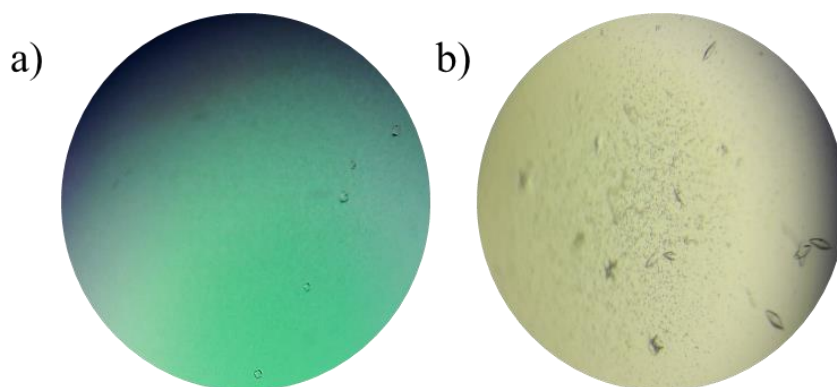


Figura 8. Cristais de AjiA1 obtidos pelo robô Oryx4 com o método de *sitting drop* (gota sentada) a  $18^\circ\text{C}$ . Os cristais eram pequenos e possuíam um formato elíptico e sem faces bem definidas. Proteína:  $13 \text{ mg.mL}^{-1}$  em  $50 \text{ mM}$  HEPES pH 8,0 e  $300 \text{ mM}$  de NaCl. (a) Condição:  $37,7 \%$  v/v PEG4000,  $150 \text{ mM}$  AMPD/Tris pH 9,0 e  $50 \text{ mM}$  sulfato de magnésio. (b) Condição:  $30 \%$  v/v Jeffamine M-600,  $100 \text{ mM}$  MES pH 6,5 e  $50 \text{ mM}$  cloreto de céσιο.

Foram realizadas tentativas de otimização dos pequenos cristais apresentados, porém não foi possível obter nenhuma outra estrutura cristalina. A técnica de *Thermofluor* pode ser realizada para se descobrir o tampão que concede maior estabilidade à proteína auxiliando na tentativa de sucesso da cristalografia (67, 68). A técnica se utiliza de fluoróforos que se ligam às regiões hidrofóbicas da proteína. Normalmente, quando a proteína se encontra devidamente enovelada não há muita exposição dessas regiões, conforme a temperatura aumenta há uma desestabilização da proteína que exhibe regiões de hidrofobicidade, onde os fluoróforos se ligam e emitem fluorescência. Para AjiA1 foi possível obter um resultado satisfatório que é apresentado no gráfico da Figura 9.

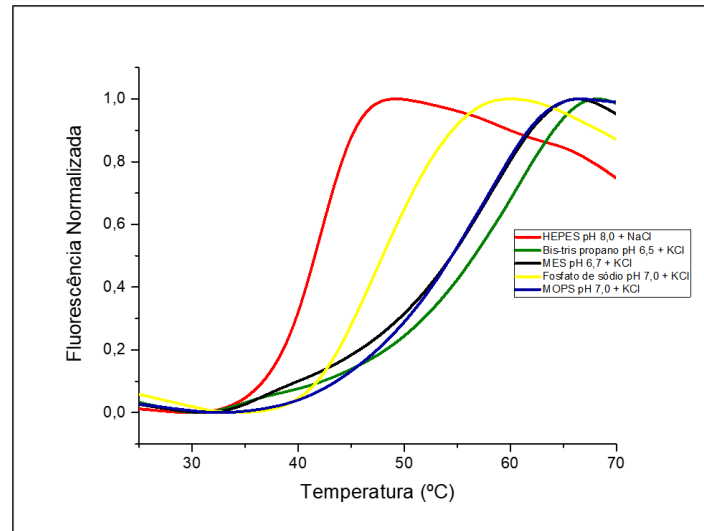


Figura 9. Gráfico de *Thermofluor* para AjiA1. Curvas representativas do deslocamento térmico de AjiA1 em diferentes tampões. A curva de referência está representada pela linha vermelha.

As curvas no gráfico representam emissão de fluorescência de acordo com o aumento da temperatura. O tampão representado pela curva verde seria o mais adequado à AjiA1, pois a mantém estável até aproximadamente 50 °C, temperatura em que há a exposição de regiões de hidrofobicidade da proteína onde os fluoróforos se ligam e há maior liberação de fluorescência. Com esses dados foi possível concluir que o tampão anteriormente utilizado para a purificação dessa enzima não está contribuindo para a estabilidade da mesma, possivelmente mantendo-a parcialmente desenovelada (92). Foi então realizada a troca de tampão para o tampão 100 mM Bis-Tris propano pH 6,5 e 150 mM de KCl para manter a proteína mais estável. Nessa condição, novas placas de cristalização foram submetidas ao robô e cristais foram conseguidos em uma condição do kit JCSG++ (Jena Biosciences) (Figura 10).

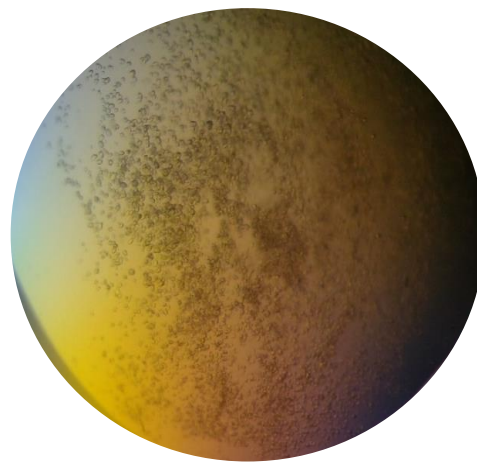


Figura 10. Cristais pequenos obtidos para AjiA1 pelo robô Oryx4 com o método de gota sentada a 18 °C após a troca do tampão 50 mM HEPES pH 8,0 e 300 mM de NaCl para 100 mM Bis-Tris propano pH 6,5 e 150 mM de KCl. Condição de cristalização: 40 % v/v PEG 300 e 100 mM fosfato citrato de potássio pH 4,2. Proteína: 13 mg.mL<sup>-1</sup>.

Após a obtenção de *hits* iniciais para AjiA1, as condições foram otimizadas alterando as concentrações de precipitante e proteína manualmente em placas no sistema de gota pendurada, a fim de se obter cristais de melhor qualidade e que aumentasse a chance de sucesso em nossos experimentos de difração. Esses cristais foram otimizados variando a concentração de PEG 300 (40 %, 37,5 %, 35 %, 32,5 %, 30 % e 27,5 %) e a concentração da proteína (10 e 8 mg.mL<sup>-1</sup>). Cristais maiores e com faces definidas foram conseguidos na condição de PEG 300 35 % e 8 mg.mL<sup>-1</sup> de proteína (Figura 11).

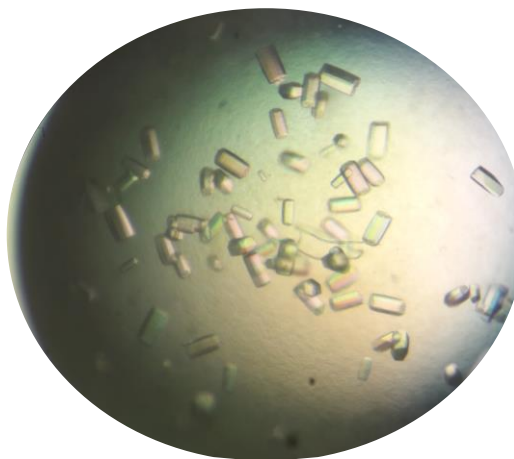


Figura 11. Cristais otimizados maiores e com faces definidas de AjiA1 obtidos manualmente com o método de gota pendurada a 18 °C. Proteína: 8 mg.mL<sup>-1</sup> em 100 mM Bis-Tris propano pH 6,5 e 150 mM de KCl. Condição de cristalização: 35 % v/v PEG 300 e 100 mM fosfato critrato de potássio pH 4,2.

Os cristais otimizados para a enzima AjiA1 foram enviados o DESY na Alemanha e uma solução crioprotetora de 15 % de etilenoglicol foi utilizada. A estrutura de AjiA1 na forma apo foi determinada com uma resolução de 2,0 Å. As tentativas de obter cristais da proteína em complexo com AMP, ADP ou ATP foram malsucedidas e as co-cristalizações com substrato e análogos de substrato não mostraram a densidade eletrônica correspondente no sítio ativo.

A estrutura foi resolvida pelo método de substituição molecular usando a estrutura de PaaK1 (PDB ID: 2Y27) como modelo. O cristal de AjiA1 pertence ao grupo espacial P 3<sub>1</sub> 2 1 com dois protômeros na unidade assimétrica (rmsd: 1,21) que devem corresponder à estrutura quaternária homodimérica funcional da proteína em solução indicada pelo PISA (142). Os dados processados e as estatísticas cristalográficas adicionais são mostrados na Tabela 2.

Tabela 2. Dados coletados e estatísticas de refinamento de AjiA1 em sua forma apo. As estatísticas para a resolução mais alta são mostradas entre parênteses.

<b>AjiA1</b>	
Código de acesso	6WUQ
Resolução (Å)	32,42 - 2,00 (2,07 - 2,00)
Grupo Espacial	P 31 2 1
Parâmetros de célula unitária (a, b, c e $\alpha$ , $\beta$ , $\gamma$ )	128,78 128,78 101,5 90 90 120
Reflexões totais	130802 (12838)
Reflexões únicas	65402 (6418)
Multiplicidade	2,0 (2,0)
Completeza (%)	1,00 (0,99)
Média I/sigma(I)	24,99 (1,80)
Wilson Fator-B	46,85
R-merge	0,014 (0,458)
R-meas	0,021 (0,647)
CC1/2	1 (0,754)
CC*	1 (0,927)
Reflexões usadas no refinamento	65388 (6417)
Reflexões usadas para R-free	3203 (277)
R-work	0,188 (0,325)
R-free	0,223 (0,376)
CC(work)	0,966 (0,841)
CC(free)	0,948 (0,705)
Número de átomos não-hidrogênio	6869
Macromoléculas	6713
Ligantes	4
Solvente	859
Resíduos de proteína	0,008
RMS (ligações)	0,93
RMS (ângulos)	96
Ramachandran regiões favoráveis (%)	4
Ramachandran regiões permitidas (%)	0,47
Ramachandran outliers (%)	1,2
Rotâmeros outliers (%)	5,33
Clashscore	63,61
Fator-B médio	63,74
Macromoléculas	53,38
Ligantes	58,01
Solvente	1

A análise da superfície do potencial eletrostático do homodímero indica uma cavidade positivamente carregada e hidrofóbica na interface entre o domínio C e o domínio N-terminal (Figura 12).

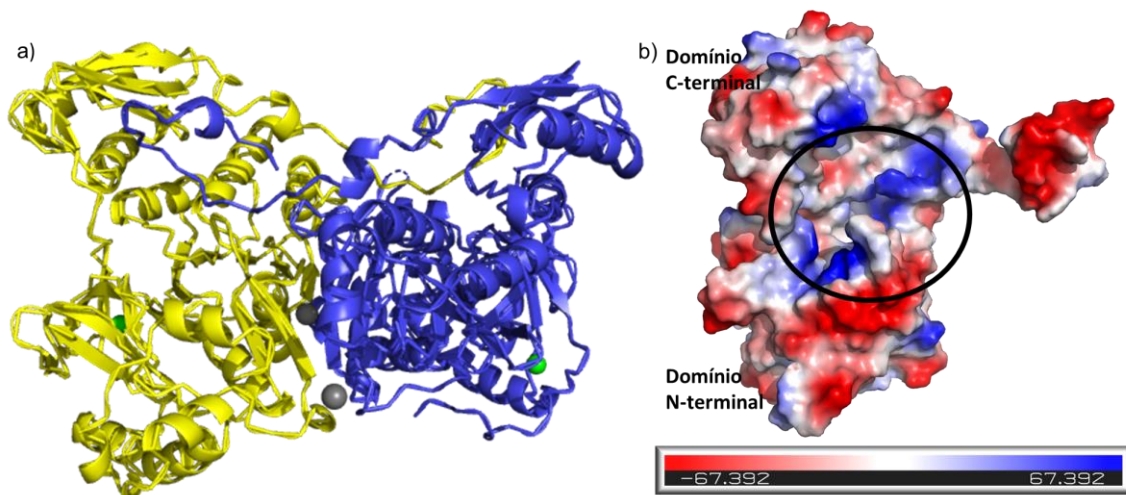


Figura 12. Estrutura geral de AjiA1 de *Streptomyces* sp. AJ9561. (a) Representação em *cartoon* da AjiA1 homodimérica (a cadeia A está colorida em amarelo e a cadeia B em azul). As esferas em cinza representam íons  $Mg^{2+}$  e as esferas em verde representam íons  $Zn^{2+}$  (b) Superfície de potencial eletrostático indica a região do sítio ativo (círculo). Azul indicando carga positiva e vermelho indicando carga negativa.

AjiA1 compartilha 30 % de identidade e 91 % de cobertura com PaaK1, uma ligase de fenilacetato-CoA envolvida na biossíntese de ácido fenilacético (PAA) (130, 137). A superposição do *backbone* de AjiA1 e PaaK1 apresentou um rmsd de 2,5 Å, indicando que essas estruturas são razoavelmente diferentes (Figura 13). Recentemente, a estrutura da ligase NatL2 dependente de ATP envolvida na biossíntese do Nataxazol foi resolvida na presença de cosubstrato e substrato e compartilha cerca de 91 % de identidade com AjiA1 (141) e a sobreposição das duas estruturas apresenta um rmsd de 1,14.

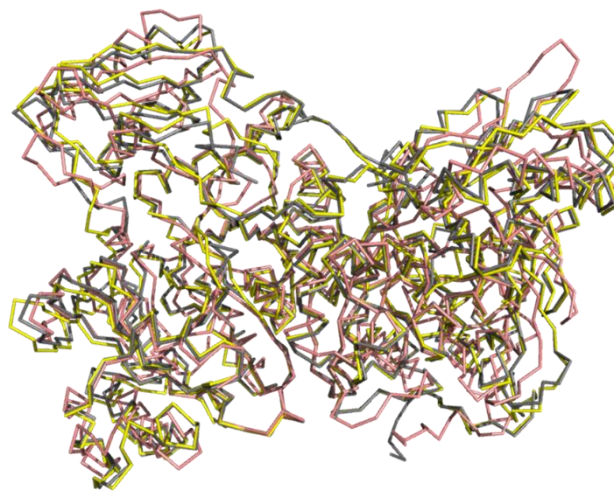


Figura 13. Estrutura e comparação entre AjiA1 reportada neste trabalho (amarelo) com os homólogos mais próximos PaaK1 (salmão) e NatL2:AMPPNP (cinza).

Semelhante à NatL2, AjiA1 pertence à família PFAM00501 das enzimas formadoras de adenilato (AFE - *Adenylate-Forming Enzyme*), que possui uma arquitetura geral constituída por dois domínios  $\alpha/\beta$  constituídos por um núcleo central de folhas  $\beta$  paralelas cercadas por  $\alpha$ -hélices com o sítio ativo formado na interface dos domínios N e C-terminal (Figura 14) (143). A AjiA1 também não possui o arranjo típico de  $\alpha/\beta$ -sanduíche no N-terminal que os membros da família possuem, mas possui um pequeno microdomínio de feixe helicoidal presente no N-terminal similarmente à NatL2 (130).



Figura 14. Estrutura do monômero da AjiA1 com o domínio N-terminal maior em verde, o domínio C-terminal em azul e o *loop* que conecta os domínios em amarelo. O sítio ativo está marcado com um círculo vermelho.

Cada monômero de AjiA1, enzima desse estudo que compartilha cerca de 91 % de identidade com NatL2, consiste em um grande domínio N-terminal (Ser4 a Gly321, o domínio N) e um pequeno domínio C-terminal (His327 a Ala432, o domínio C). O domínio N-terminal de AjiA1 consiste em 3 folhas  $\beta$  de 2 fitas antiparalelas entre nove  $\alpha$ -hélices, enquanto o domínio C-terminal é definido por duas  $\alpha$ -hélices, uma folha  $\beta$  curta antiparalela de duas fitas, uma folha trançada de três fitas e uma longa alça que se estende ao protômero adjacente e realiza contato extenso com esse protômero.

A formação do homodímero AjiA1 é estabilizada por uma extensa folha  $\beta$  envolvendo ambos os domínios. A montagem do dímero simétrico resulta da formação de duas folhas  $\beta$  de sete fitas de extensão intermoleculares, cada uma composta por uma folha  $\beta$  antiparalela de duas fitas de um monômero e uma folha  $\beta$  padrão de cinco fitas do segundo monômero (Figura 15) (130), que também foram observadas em outras enzimas da família das AFEs.



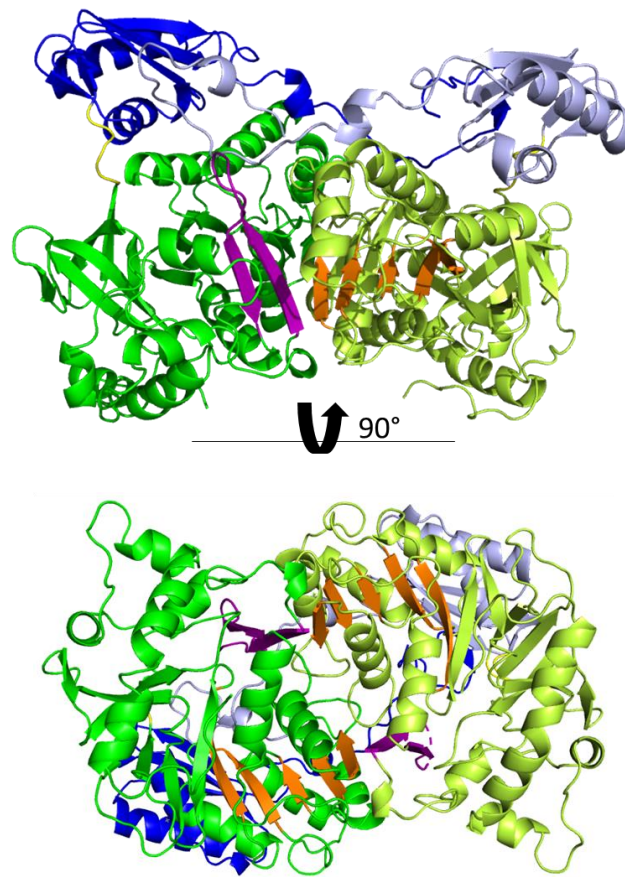


Figura 15. Dimerização de AjiA1. As interações íntimas de dimerização de AjiA1 criadas por folhas  $\beta$  compostas que abrangem os dois monômeros. A folha  $\beta$  de duas fitas que flanqueia a alça P (Ala82 roxo a Tyr99) é confinada pela folha  $\beta$  distorcida de cinco fitas (laranja) [Thr123 a Val126; Thr149 a Pro152; Leu176 a Trp178; Ala207 a Ala211; Val232 a Ser238] do segundo protômero. O domínio N-terminal maior é representado em verde no primeiro protômero e em lima no segundo protômero, o domínio C-terminal está em azul no primeiro protômero e azul claro no segundo protômero e o *linker* que conecta os domínios é mostrado em amarelo.

Na formação da interface do homodímero, cerca de  $3999 \text{ \AA}^2$  da área da superfície é estabilizada através de uma extensa complementaridade química. Essa interface é dominada por interações hidrofóbicas e é acompanhada por 46 ligações de hidrogênio e oito pontes salinas (Figura 16).

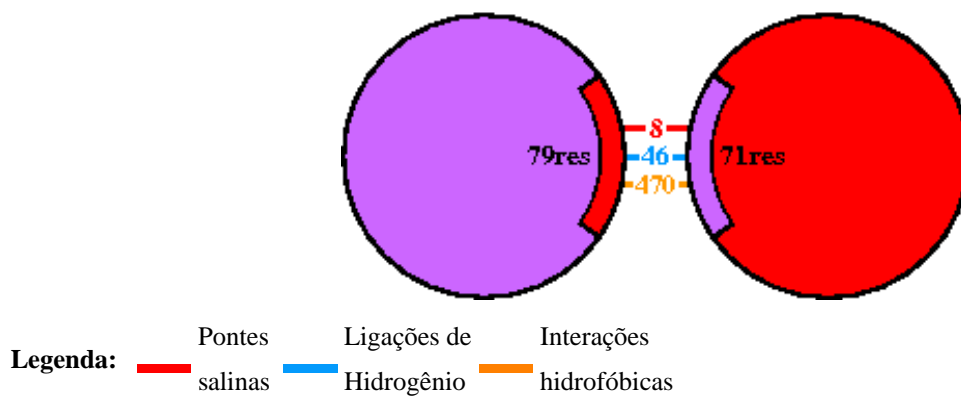


Figura 16. Interface de dimerização de AjiA1. O servidor online PDBsum (144) (EMBL-EBI) foi usado para gerar um diagrama esquemático mostrando as interações entre as duas cadeias de AjiA1. As cadeias de interação são unidas por linhas coloridas, cada uma representando um tipo diferente de interação, laranja = contatos não ligados, azul = ligações de hidrogênio e linhas vermelhas = pontes de sal. A área de cada círculo é proporcional à área da superfície da cadeia proteica correspondente. A extensão da região da interface em cada cadeia é representada pela fatia preta cujo tamanho significa a área da superfície da interface.

O grande domínio N-terminal se sobrepõe razoavelmente bem em AjiA1 e outras AFEs, enquanto outras regiões mostram diferenças significativas no sítio ativo, localizado entre os domínios C e N-terminal (Figura 17).

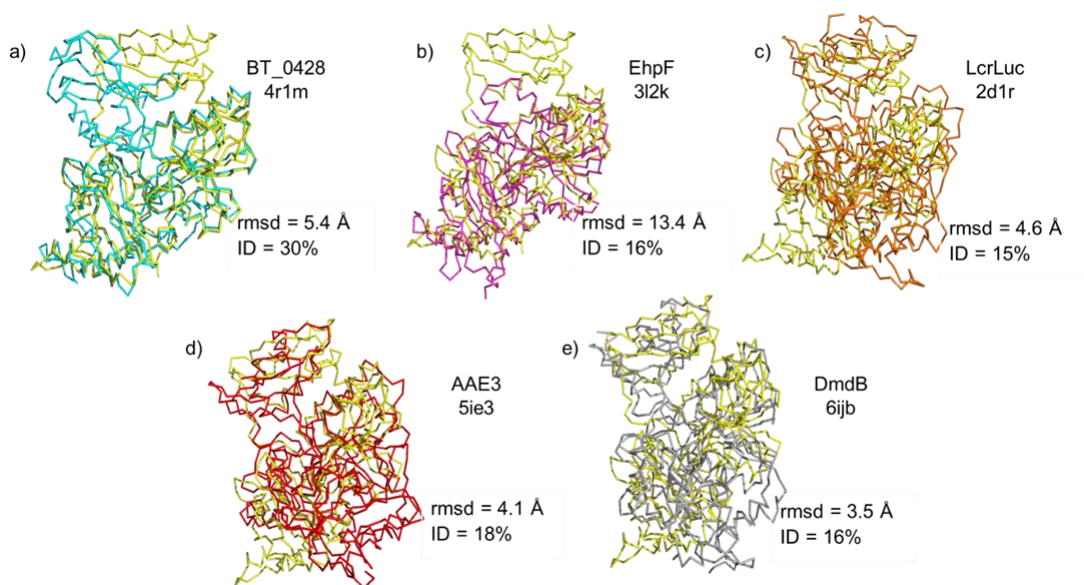
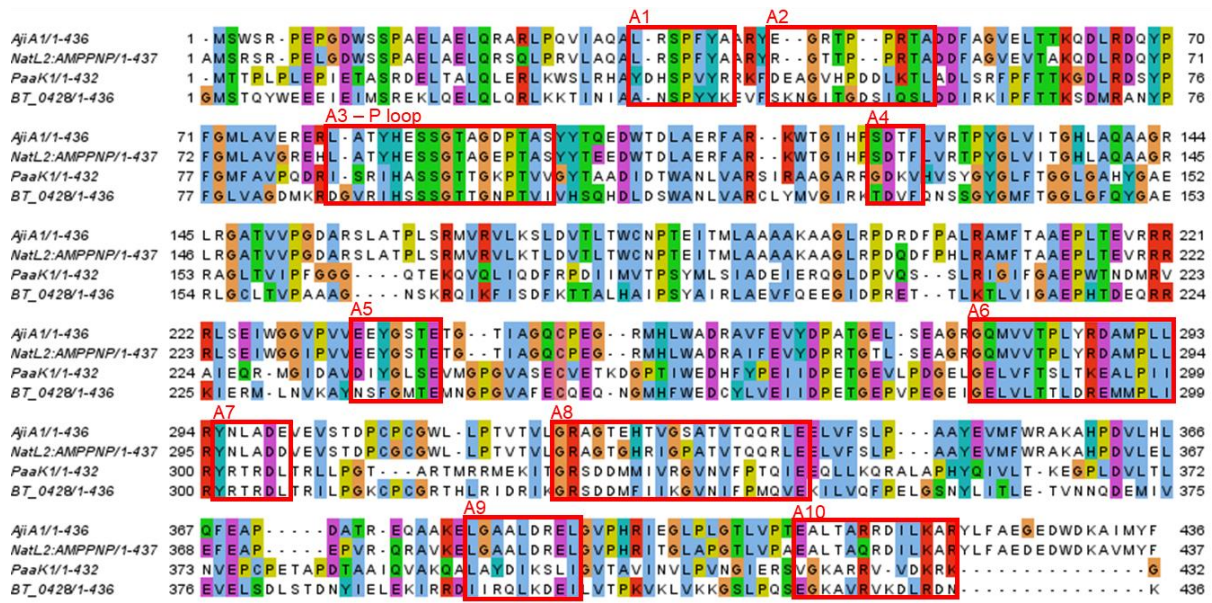


Figura 17. Homólogos estruturais de AjiA1. A cadeia Ca de AjiA1 foi alinhada com seus homólogos estruturais (a) 4R1M (ciano), (b) 3L2K (roxo), (c) 2D1R (laranja), (d) 5IE3 (vermelho) e (e) 6IJB (cinza). O valor rmsd da superposição com AjiA1 e a identidade de aminoácidos (%) com AjiA1 são dados em cada análise.

Além disso, todos os motivos de assinatura conservados (A1-A10), comumente observados em todos os domínios de adenilação, designados conforme a nomenclatura descrita por Marahiel et al. (145), são conservados em AjiA1, incluindo o P-loop e o motivo de adenina (A5), que contém resíduos críticos para a ligação da porção de adenina no ATP / AMP (Figura 18).



% Identidade com AjiA1	
NatL2:AMPPNP	91
PaaK1	30
BT_0428	30

Figura 18. Alinhamento de AjiA1 e PaaK1 com homólogos DALI. Os motivos A1-A10 conservados nas enzimas da superfamília ANL são indicados de acordo com o domínio de adenilação definido da peptídeo sintetase não-ribossomal (145). Os motivos principais A1-A10 são mostrados em quadros vermelhos. O alinhamento foi gerado pelo ClustalOmega (146) e importado para o Jalview Versão: 2.11.0 (147). Resíduos similares são coloridos de acordo: Azul, hidrofóbico; Verde, polar; Roxo, negativo; Vermelho, positivo; Amarelo, prolina; Laranja, glicina. A tabela exibe a identidade em porcentagem de AjiA1 para os homólogos alinhados.

Aproximadamente um terço de todas as proteínas requerem um cofator de metal para o funcionamento adequado. Em geral, os cofatores de metais têm papéis importantes no dobramento de proteínas e na estabilização estrutural (148). Em AjiA1, foi possível observar um átomo de zinco coordenado por Cys248, Cys308 e Cys310 e His254 (Figura 19). As enzimas podem conter locais catalíticos e estruturais, onde o zinco não é necessário para a reação enzimática, mas para estabilidade e dobramento correto da proteína. A coordenação do zinco também foi observada para reduzir o movimento da cadeia lateral e o espaço conformacional. Transferases e ligases contêm principalmente íons estruturais de zinco (148).

Em geral, os locais estruturais de zinco adotam um arranjo tetraédrico formado por grupos de cadeias laterais de cisteína, histidina e, ocasionalmente, resíduos de aspartato ou glutamato (149-151). Esse fato sugere que o íon zinco em AjiA1 desempenha um papel estrutural e, como esperado, em nossa estrutura possui coordenação tetraédrica, da mesma maneira que também foi observado em NatL2. Comparando com outros homólogos, o íon  $Zn^{2+}$  está presente apenas em AjiA1 e NatL2 e deve ser uma característica única desse grupo de proteínas. Em algumas famílias de proteínas, os ambientes aniônicos de zinco são estabilizados pelas cargas positivas de arginina e lisina. Na AjiA1, observamos a presença de Arg252 que pode desempenhar esse papel (150).

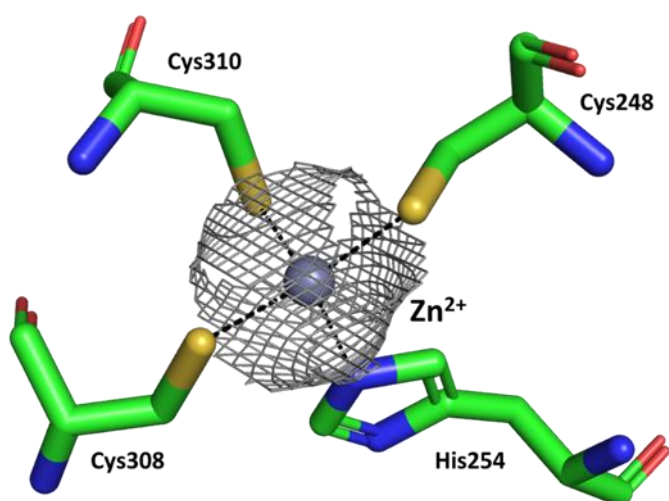


Figura 19. Local de ligação do zinco na cadeia A de AjiA1 com  $2Fo - Fc$  de  $Zn^{2+}$  (com contorno de  $2,0 \sigma$ ). Esta figura mostra as interações de coordenação.

Dentre as estruturas relatadas anteriormente da família AFE, algumas contêm um  $Mg^{2+}$  ligado no sítio ativo, atuando diretamente na catálise de ATP (137). AjiA1 possui um sítio de ligação ao  $Mg^{2+}$  em cada cadeia, porém distante do sítio ativo (Figura 20). A ausência do magnésio no sítio de ligação ao ATP pode ser devido ao fato de o ATP não estar presente na estrutura.



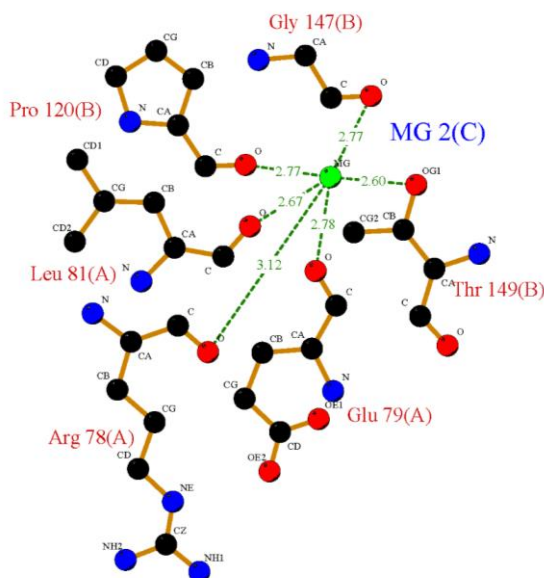


Figura 20. Representação esquemática do local de ligação ao  $Mg^{2+}$  na cadeia A de AjiA1. A figura foi preparada usando LigPlot + v.2.1 (152).

O sítio ativo de AjiA1 está posicionado na interface dos domínios N-terminal e C-terminal. Nessa região, temos a lisina 418 universal e cataliticamente essencial e uma arginina (Arg322) do *linker* presente da interface entre os domínios (Rx (D / K) x6G) (Figura 14) (124).

Em contraste com a maioria das AFEs resolvidas até agora, os resíduos de AjiA1 de ambos os protômeros de dímeros estão envolvidos na formação do sítio catalítico e a estrutura dimérica deve ser essencial para a atividade enzimática, que é uma característica nova dessa classe de enzimas. Apesar da Lys418 ocupar uma posição semelhante à Lys422 catalítica de PaaK1, a lisina catalítica de um protômero atua no protômero adjacente. Curiosamente, uma longa alça do C-terminal se estende ao sítio ativo do protômero adjacente e, conseqüentemente, resíduos de ambos os protômeros estão envolvidos na formação do sítio catalítico. Poderíamos descrever essa alça como um caso de *loop swapping*, que também é a segunda descrição desse recurso para essa classe de proteínas. O *loop swapping* observado pode refletir não apenas a dinâmica dos protômeros, mas também uma peculiaridade de AFEs envolvidas na formação de anéis benzoxazóis (153), já que esse evento também ocorre em NatL2 (Figura 21).

O *loop swapping* é um mecanismo não convencional de oligomerização, no qual os elementos estruturais de monômeros individuais são trocados entre parceiros idênticos, recrutando interações cruciais para estabilizar a proteína. O fenômeno foi proposto para desempenhar um papel na agregação, desdobração e regulação da função, mas suas implicações biológicas ainda não são claras. O custo de energia de quebrar e substituir as

interações da região de troca pode exigir a quebra de uma barreira de alta energia, e é surpreendente que qualquer proteína passe por uma transição de barreira tão grande (154).

Inicialmente, esperávamos que a ligação de cosubstrato pudesse desencadear movimentos nessa região para realizar a atividade catalítica. No entanto, a estrutura apo de AjiA1 indicou que esse *loop* não retorna ao seu protômero na ausência de ligante, sugerindo que essa característica estrutural não deve ser responsável pela regulação alostérica. Por outro lado, poderia desempenhar um papel na estabilização do dímero ou diminuir a flexibilidade do domínio C-terminal, diferente do que foi observado em outros membros das AFEs, incluindo PaaK1 (130) e HCL (131), que são relatadas como tendo mudanças conformacionais drásticas.

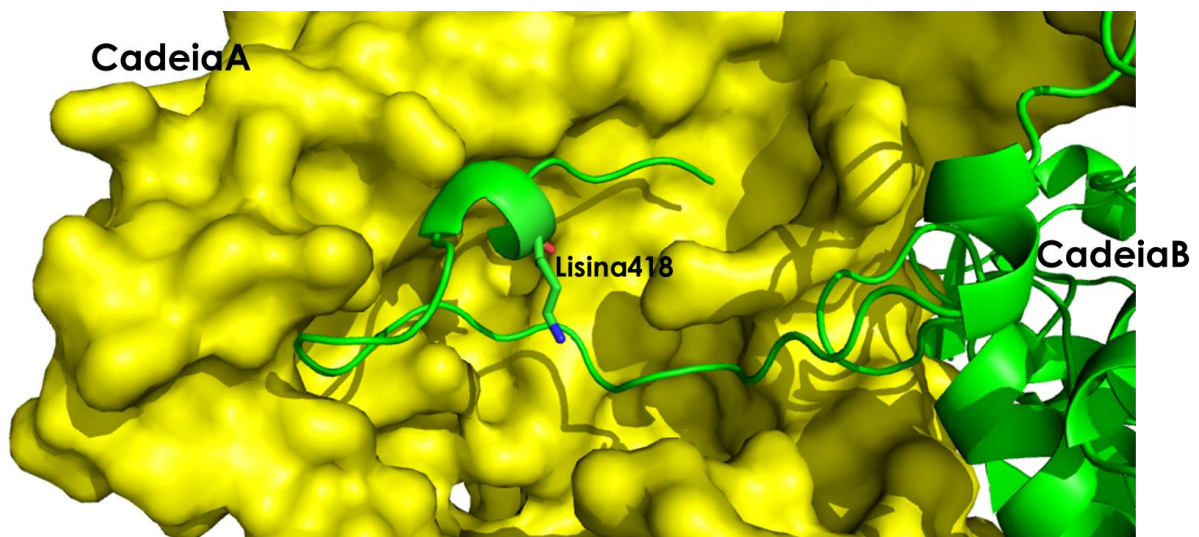


Figura 21. AjiA1 *loop swapping*. O resíduo conservado Lis418 está representado por *sticks*.

Foi também possível observar que os resíduos de aminoácidos de 422L a 424A da cadeia A formam uma folha  $\beta$  paralela de duas fitas com os aminoácidos 357A a 359A da cadeia B e vice-versa, na região de *loop swapping* (Figura 22).

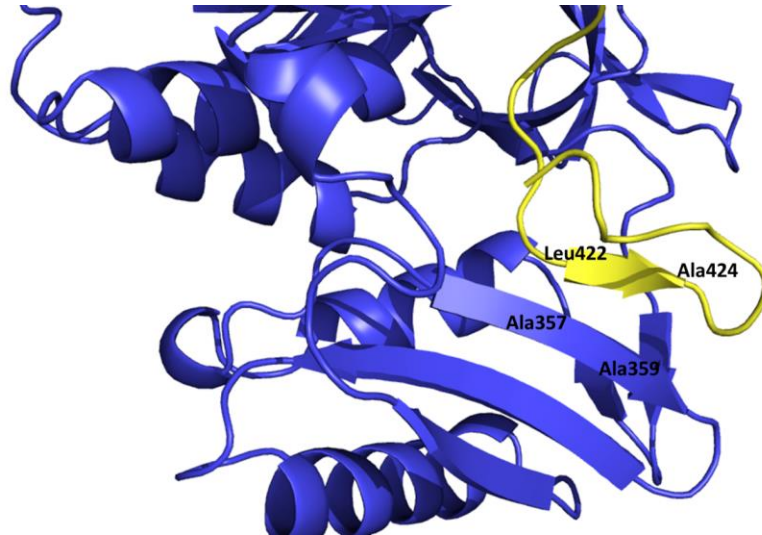


Figura 22. Região de *loop swapping*. Os aminoácidos 422L, 423F e 424A parecem formar uma pequena fita beta que pode estar envolvida na formação de uma folha  $\beta$  com os aminoácidos 357A, 358K e 359A. A cadeia A está em amarelo e a cadeia B está em azul.

Outro resíduo catalítico chave, Arg322, do motivo A8, que se propõe estar envolvido na estabilização do intermediário da reação, está localizado no *linker* curto exposto a solvente (124). O sítio de ligação ao ATP contém vários motivos altamente conservados que estão presentes em todos os membros da família e desempenham papéis semelhantes no posicionamento do nucleotídeo. O resíduo de ácido aspártico do motivo A7, Asp299 em AjiA1, é universalmente conservado e interage com as hidroxilas da ribose e Glu240, um resíduo do motivo A5, deve estar envolvido na ancoragem do intermediário adenilato. Um resíduo aromático do motivo A5, Tyr236, está empilhado contra a base de adenina. As principais diferenças observadas no sítio ativo na superposição de AjiA1 e NatL2:AMPPNP são: a aproximação de Ser88 (do *P-loop*) e Lys418 (a lisina A10) ao grupo fosfato do AMP, um reposicionamento de Ala212 para estar próxima ao anel de adenina, bem como a participação de seis moléculas de água (Figura 24a).

Análises estruturais adicionais revelaram, como esperado, que outras regiões funcionais dessa proteína são altamente conservadas (Figura 23).

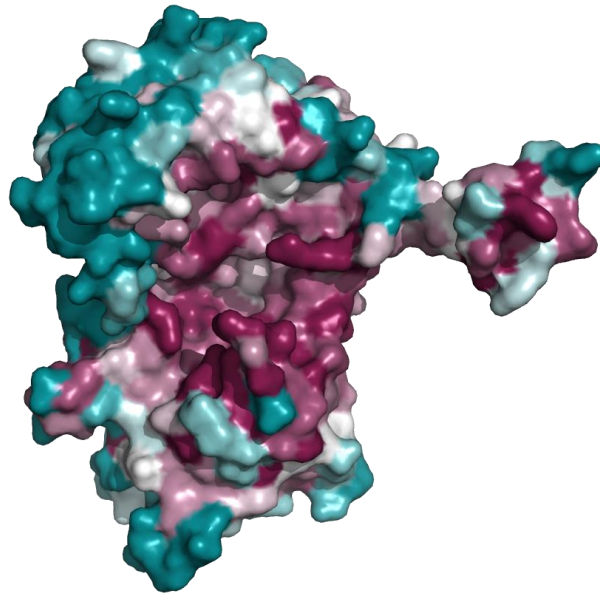


Figura 23. Representação de superfície do monômero AjiA1 mostrando a conservação de resíduos com outros membros da AFEs. A figura foi preparada usando o servidor ConSurf (155). Roxo e ciano indicam conservação de sequência alta e baixa, respectivamente.

Além disso, outros resíduos de aminoácidos, que estão no bolsão do sítio ativo, incluindo os motivos A3 ou *P-loop*, A5, A6 e A7, são substancialmente conservados (145, 155). No entanto, com base nos resíduos conservados do sítio ativo, AjiA1 possui resíduos catalíticos semelhantes aos relatados em PaaK1, HCL e NatL2. Consequentemente, o mesmo mecanismo catalítico deve atuar para essa enzima, que inclui a adenilação do ácido carboxílico do substrato, seguida pela reação de esterificação com uma segunda molécula de substrato, que consiste na condensação do anel do intermediário de AJI9561.



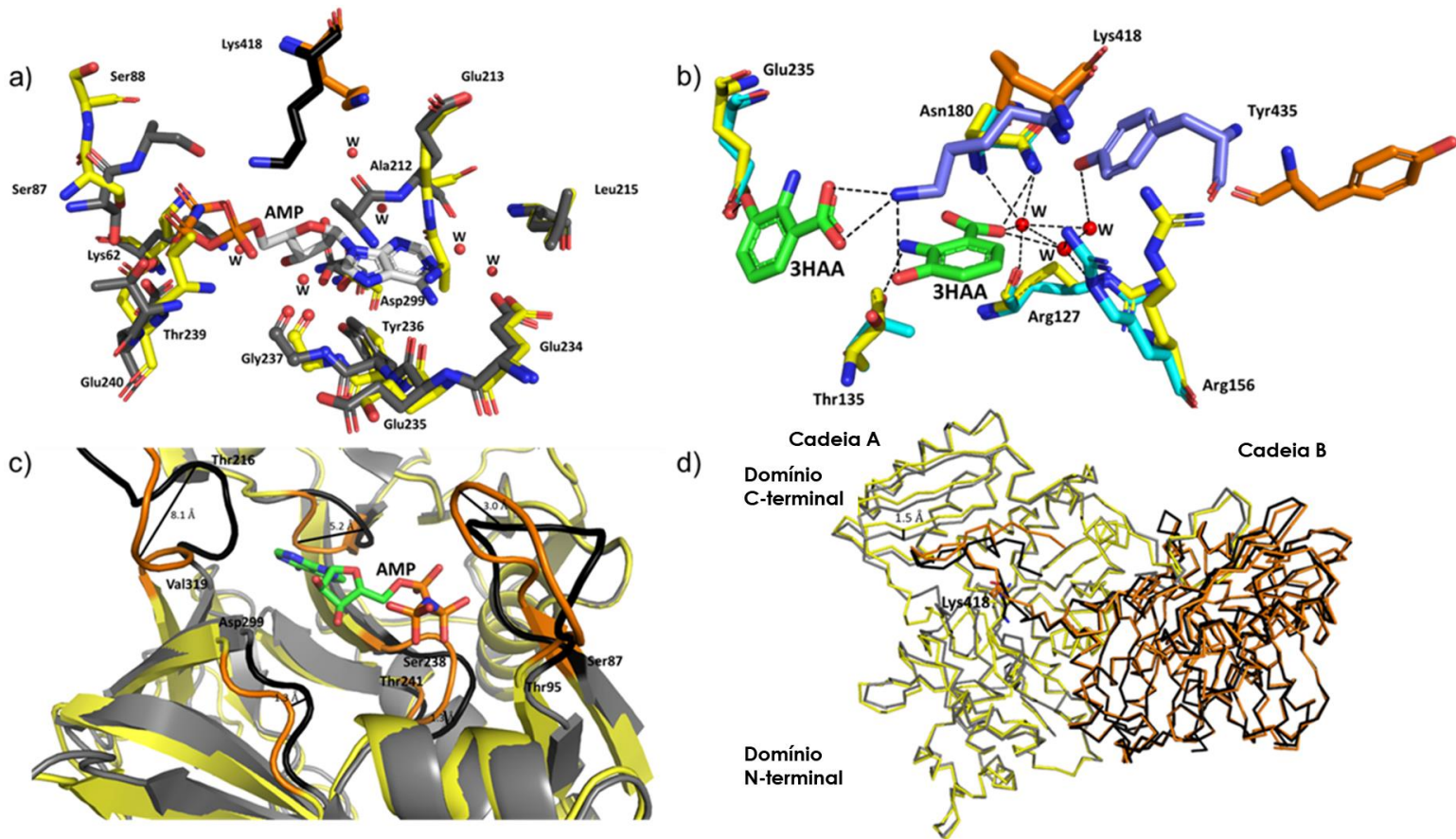


Figura 24. Comparação AjiA1 e estruturas em complexo com cosubstrato e substratos. (a) Sobreposição de AjiA1 e NatL2:AMPPNP (PDB ID: 6six). Os resíduos de AjiA1 estão representados em amarelo e o NatL2 em cinza. O principal resíduo catalítico Lys418 possui uma cor mais forte, pois pertence ao protômero adjacente. (b) Sobreposição de AjiA1 e NatL2:AMP:3-HAA (PDB ID: 6siy). Os resíduos AjiA1 estão representados em amarelo, em que K418 e Y435 estão em laranja e os resíduos da NatL2 estão em ciano, nos quais K418 e Y435 estão em azul. A cor diferente desses resíduos de aminoácidos representa que eles são do protômero adjacente. As águas do sítio ativo são representadas por esferas vermelhas. (c) Sobreposição de AjiA1 e NatL2 em complexo com AMPPNP (6SIX). Os protômeros estão representados em *cartoon*. AjiA1 está em amarelo e seus cinco *loops* diferentes estão em laranja. NatL2:AMPPNP está em cinza e seus *loops* estão em preto. (d) Sobreposição de AjiA1 e NatL2 em complexo com AMPPNP (6SIX). Os protômeros estão representados em *ribbon*. AjiA1 são representados em amarelo (cadeia A) e laranja (cadeia B) e NatL2 no complexo com AMPPNP (6SIX) são representados em cinza (cadeia A) e preto (cadeia B). Os principais resíduos catalíticos Lys418 de AjiA1 e NatL2:AMPPNP estão representados em palitos. O terminal C e, conseqüentemente, o *loop* do terminal C, também se move para uma posição mais próxima do local ativo no local ativo em cerca de 1,5 Å.

Lys418, a lisina do motivo A10, e Tyr435, resíduos do *loop swapping* que estão mais enterrados e mais próximos das moléculas de 3-HAA no sítio ativo da estrutura complexada de NatL2 mostram alterações conformacionais consideráveis quando comparados com os mesmos resíduos na estrutura de AjiA1. Arg156 também tem uma alteração conformacional relevante à

medida que se afasta do sítio ativo em AjiA1, enquanto Glu235, do motivo A5, parece ter um papel na posição de 3-HAA no sítio ativo, contribuindo com uma ligação de hidrogênio com o grupo hidroxila do substrato. Em NatL2, a Thr135 também sofre uma rotação em comparação com a de AjiA1 e sua cadeia lateral cria uma ligação de hidrogênio com uma das moléculas de 3-HAA (Figura 24b).

A comparação entre as estruturas de AjiA1, NatL2:AMPPNP (rmsd: 1,12) e NatL2:AMP:3-HAA (rmsd: 1,14) revelou movimentos importantes que devem ocorrer durante o ciclo catalítico e por causa da ligação do cosubstrato. Uma série de alças chave do sítio ativo sofrem alterações conformacionais quando NatL2 em complexo com AMP e AjiA1 em sua forma apo são sobrepostas. As maiores mudanças são observadas em seis alças próximas ao sítio ativo: 86E a 94P (motivo A3, *P-loop*), 210T a 213E, 238S a 242G (motivo A5), 296N a 298A (motivo A7), 305T a 313L e 320L a 328T (motivo A8) em AjiA1 (Figura 24c). Três dessas cinco alças têm o fator B mais alto (86E a 94P; 210T a 213E e 320L a 328T) em comparação com outras regiões de NatL2 e AjiA1, indicando também alta flexibilidade. Assim, o deslocamento médio sugere que é a parte flexível da proteína que deve atuar na função catalítica. O enovelamento da região central do domínio N-terminal é altamente conservada e rígida, como evidenciado por seu baixo fator B médio (Figura 25).

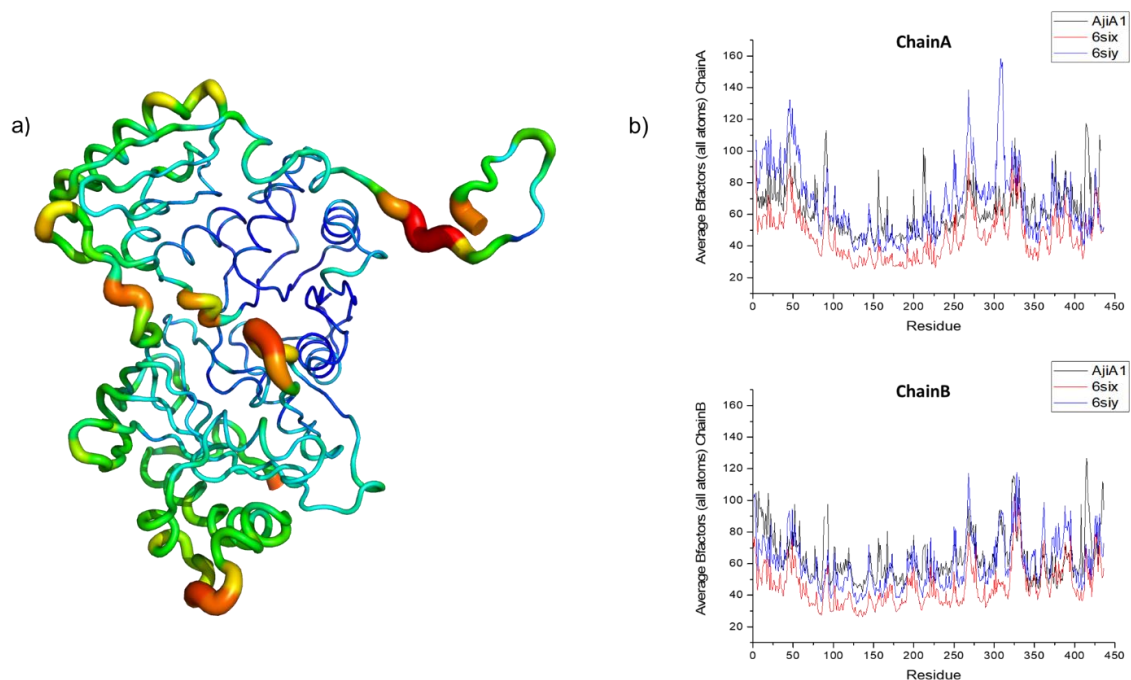


Figura 25. Representação de massa de fator B mostrando o núcleo rígido. (a) Os valores do fator B são ilustrados por cores, variando de baixo (azul) a alto (vermelho), e pela forma, um tubo mais largo indica regiões com fatores B mais altos, enquanto o tubo mais estreito indica regiões com fatores B mais baixos. b) Análise fatorial de temperatura. Gráficos do fator B médio versus resíduos para cada cadeia de AjiA1.

A alça P ou *P-loop* de AjiA1 (motivo 3), que inclui os resíduos Ser87 a Pro94 (Figura 18) é uma região altamente conservada em todos os membros de AFEs e possui uma sequência de aminoácidos de Gx(4)-GK-[TS]. Essa região é proposta para desempenhar um papel bifuncional na estrutura da PaaK1, que envolve a coordenação do ATP e a estabilização do domínio C-terminal. Na estrutura de PaaK1 em complexo com adenilato intermediário sem o grupo pirofosfato, essa região adota uma estrutura aberta e desordenada. Por outro lado, na ligação com ATP, essa região adota uma conformação fechada e ordenada, estabelecendo uma ligação do hidrogênio com o fosfato  $\beta$  e  $\gamma$ . Embora NatL2 tenha sido cristalizada na presença de AMP, a posição do *P-loop* difere da AjiA1. Na AjiA1, o *P-loop* adota uma conformação mais aberta com um deslocamento de cerca de 3 Å em comparação com NatL2 e aponta para o *loop* C-terminal, que também possui uma conformação diferente entre as duas estruturas. Além disso, a posição do domínio N-terminal também parece adotar uma conformação mais próxima do sítio ativo em cerca de 1,3 Å (Figura 24c). O domínio C-terminal e, conseqüentemente, a alça do C-terminal, também se move para uma posição mais próxima do sítio ativo em cerca de 1,5 Å, permitindo que a Lys418 adote uma posição catalítica e interaja com as porções fosfato e ribose do AMP (Figura 24d).

Na reação catalisada pelas enzimas dessa família, um substrato carboxilato é ativado usando ATP para formar um acil-adenilato (137). Estudos estruturais e funcionais indicaram que essas enzimas se alternam entre duas conformações durante a catálise (124, 137, 156, 157). A presença da lisina catalítica formando o *loop swapping* poderia sugerir um movimento intenso desse “braço”, funcionando como um “portão” para a estrutura aberta e fechada dessa família de enzimas. No entanto, nossa estrutura apo não mostrou evidências de uma grande mudança conformacional nessa região e o domínio C-terminal não devem adotar uma forma aberta e fechada, conforme descrito em outros membros de AFEs. A longa alça C-terminal deve ter um papel em manter a estrutura na forma fechada, como previsto por Song et al., 2020 (141). Curiosamente, como o *loop swapping* mantém as estruturas de AjiA1 e NatL2 na mesma conformação, no grupo de enzimas envolvidas na biossíntese dos anéis benzoxazol em produtos naturais, grandes movimentos conformacionais do domínio C-terminal são improváveis e, portanto, a alternância de domínio deve não ocorrer. Como grandes mudanças conformacionais envolvem uma barreira energética entálpica mais alta, essas enzimas podem ter evoluído usando uma estratégia diferente e elegante para realizar as duas meias-reações comumente encontradas em AFEs.

O *loop swapping* observado nessas estruturas poderia manter as duas semirreações em um único sítio ativo, evitando grandes rotações, mas com o custo de uma estrutura dimérica

ativa restritiva. No entanto, ainda é necessária uma investigação mais aprofundada para provar essa hipótese ou para entender se a *loop swapping* tem um papel alternativo para a alternância de domínio, como observado em outros membros da classe I de AFEs.

Para obter informações sobre divergência e evolução, realizamos uma análise filogenética usando o método de máxima verossimilhança (ML) (158, 159), que revela que AjiA1 e NatL2 são separadas em um ramo diferente (Figura 26), apoiando a ideia de que essas enzimas podem ter evoluído independentemente e devem ser classificadas distintamente como um novo subgrupo de ligases do tipo fenilacetato-CoA.

Portanto, propomos que AjiA1 e NatL2 são um novo subgrupo em AFEs, com base no fato de serem classificados em um clado diferente na análise filogenética (Figura 6) e que estruturalmente essas enzimas funcionam de maneira diferente que outras ligases do tipo fenilacetato-CoA, incluindo PaaK1, PaaK2 e HCL (130, 131).

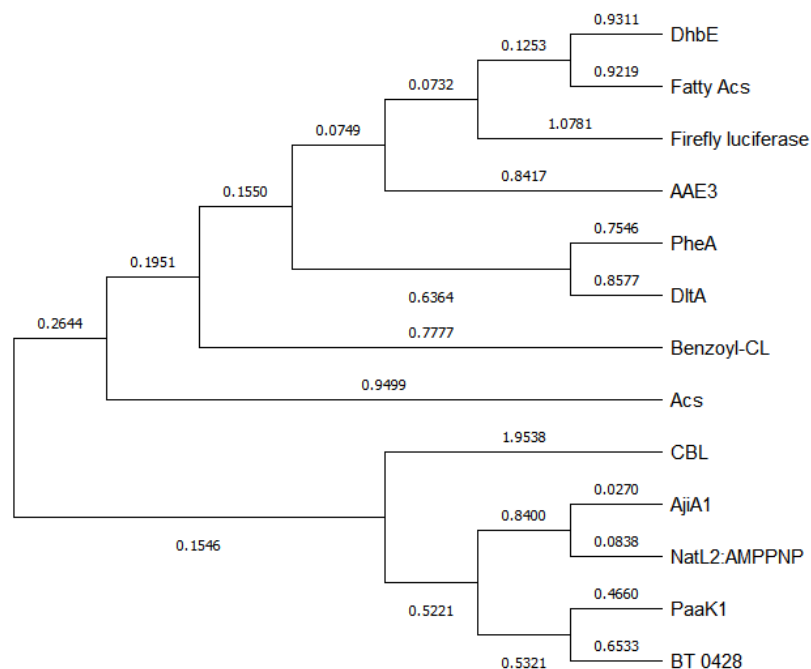


Figura 26. Análise evolucionária pelo método de máxima verossimilhança de vários membros diferentes de AFEs que foram caracterizados estruturalmente, incluindo AjiA1. Os comprimentos das ramificações são mostrados. As proteínas e organismos utilizados são: DhbE de *Bacillus subtilis* (PDB ID: 1MDF); Acs fatty de *Thermus thermophilus* (PDB ID: 1ULT); Luciferase do vaga-lume de *Photinus pyralis* (PDB: 1LCI); AAE3 de *Arabidopsis thaliana* (PDB: 5IE3); PheA de *Brevibacillus brevis* (PDB: 1AMU); DltA de *Bacillus cereus* ATCC 14579 (PDB ID: 3DHV); Benzoílo-CL de *Paraburkholderia xenovorans* LB400 (PDB ID: 2V7B); Acs de *Salmonella enterica* (PDB: 1PG4); CBL de *Alcaligenes sp.* AL3007 (ID do PDB: 1T5S); A proteína deste estudo AjiA1 de *Streptomyces sp.* AJ9561 (ID do PDB: 6WUQ); NatL2 de *Streptomyces sp.* Tu 6176 (PDB: 6SIX); PaaK1 de *Burkholderia cenocepacia* (PDB: 2Y27) e BT\_0428 de *Bacteroides thetaiotaomicron* (PDB: 4R1M). As análises foram realizadas em MEGA X (160).

Foi relatado que a alternância de domínio pode aumentar a promiscuidade das AFEs, pois a grande rotação do domínio C-terminal pode facilitar a ligação de substratos alternativos (124). Para obter informações sobre a promiscuidade de AjiA1, que possui uma estrutura muito mais restrita, realizamos um ensaio com um substrato alternativo ácido 3-hidroxibenzóico (3-HBA) e o substrato natural 3-HAA usando ATP como cosubstrato. Determinamos os parâmetros cinéticos que medem o aumento da absorbância a 360 nm usando um ensaio espectrofotométrico. Uma eficiência catalítica ligeiramente mais baixa foi observada para o 3HBA em comparação com o 3-HAA, o substrato natural. O resultado demonstra que AjiA1 é capaz de catalisar a formação de uma ligação éster na ausência do grupo amina. Em outras palavras, AjiA1 demonstrou catalisar a dimerização do 3-HBA (na ausência de um grupo amina 2-substituído) através de uma ligação éster.

Em contraste, Song et al. demonstram que NatL2 catalisa apenas a formação de éster e não amida, considerando-o como um produto derivado (141). No entanto, nosso resultado não implica necessariamente que um éster em vez de uma amida seja formado quando o substrato nativo 3-HAA é usado (Figura 27).

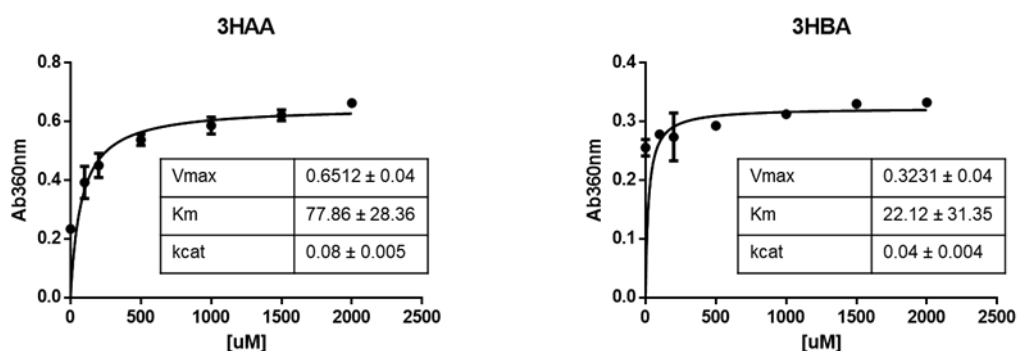


Figura 27. As análises cinéticas de AjiA1 foram realizadas usando concentrações de 3HAA e 3HBA de 0 a 2500  $\mu\text{M}$ .  $K_m$  e  $V_{\text{máx}}$  foram determinados a partir de regressão não linear para a equação de Michaelis-Menten a partir de três experimentos repetidos.

## Conclusão

Neste capítulo determinamos a estrutura de AjiA1 em sua forma apo e sua comparação com a enzima NatL2 em complexo com diferentes ligantes indica que esse grupo de enzimas possui um conjunto de movimentos completamente novo durante o ciclo catalítico nos membros das AFEs. A alternância de domínio não deve ocorrer nas enzimas envolvidas na biossíntese do benzoxazol. Em contraste, essas enzimas têm um evento de troca de domínio

envolvendo uma alça C-terminal longa. Essa alça restringe o domínio C-terminal e bloqueia a enzima em uma única conformação. Em conjunto com os novos movimentos que envolvem apenas alças do sítio ativo e uma ligeira rotação do protômero adjacente e as análises filogenéticas, propomos que essas enzimas sejam classificadas como um grupo distinto de ligases do tipo fenilacetato-CoA, mesmo com semelhanças de sequência entre homólogos. Além disso, mesmo com as restrições impostas pela ausência de alternância de domínio, AjiA1 é capaz de aceitar substratos alternativos indicando promiscuidade para essas enzimas. Consequentemente, essas descobertas trazem novos *insights* para a família crescente de enzimas formadoras de adenilato e a possibilidade de usá-las em abordagens biotecnológicas.

## 5 CAPÍTULO II – DIELS-ALDERASES EM POLIÉTERES IONÓFOROS TETRONATOS: TSN15

### **Objetivo geral:**

- Determinação e análise estrutural da proteína Tsn15 da biossíntese de tetronasina

Este capítulo gerou o artigo “Unexpected enzyme-catalysed [4+2] cycloaddition and rearrangement in polyether antibiotic biosynthesis.” Submetido em março de 2019 e aceito para publicação em agosto de 2019 na revista *Nature Catalysis*.

Autores: Rory Little, Fernanda C. R. Paiva, Robert Jenkins, Hui Hong, Yuhui Sun, Yuliya Demydchuk, Markiyan Samborskyy, Manuela Tosin, Finian J. Leeper, Marcio V. B. Dias, Peter F. Leadlay.

As enzimas são frequentemente classificadas pelas reações que realizam. Uma reação de cicloadição foi definida pela União Internacional de Química Pura e Aplicada (IUPAC) como “Uma reação na qual duas ou mais moléculas insaturadas (ou partes da mesma molécula) se combinam com a formação de um aduto cíclico no qual há uma redução líquida da multiplicidade de ligações ...” (161).

O trabalho descrito neste capítulo é uma contribuição significativa para a nossa compreensão da biossíntese de poliésteres tetronatos, uma vez que é relatada a estrutura da [4+2] ciclase Tsn15 em complexo com seu intermediário biossintético. A origem dos anéis ciclohexano central e tetra-hidropirano da tetronasina são formados em uma reação em cascata enzimática de duas etapas. As duas enzimas responsáveis, Tsn11 e Tsn15, são homólogas das enzimas [4+2] ciclases encontradas nas vias complexas de biossíntese de espirotetronatos/tetramatos, como a Tetronomicina. Tsn11 catalisa uma reação Diels-Alder para produzir um intermediário oxidecalina e Tsn15 catalisa um rearranjo que forma o anel tetra-hidropirano e desfaz a fração oxidecalina, produzindo a Tetronasina (Figura 28).

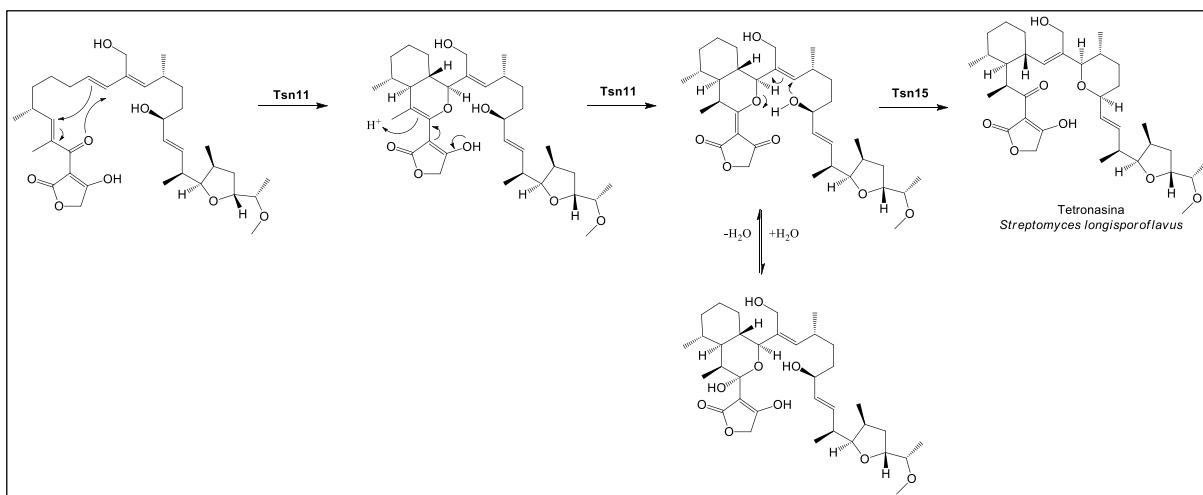


Figura 28. Mecanismo proposto para a formação dos anéis ciclohexano e tetra-hidropirano na Tetronasina. Tsn11 catalisa uma reação hetero-Diels-Alder de demanda inversa de elétrons para formar um intermediário de oxidecalina. A oxidecalina é hidratada para formar um hemiacetal cíclico. Tsn15 então catalisa uma desidratação seguida de um rearranjo pericíclico para formar o anel de tetra-hidropirano e fragmentar a oxidecalina, produzindo Tetronasina.

A partir da clonagem da enzima Tsn15, ciclase [4+2], no vetor de expressão pET28a, foi possível produzir a proteína em sua forma solúvel com rendimento adequado para purificação e posteriores experimentos cristalográficos. Dados gerais da proteína trabalhada foram obtidos de acordo com análise feita pelo programa ProtParam (90)



hospedado no servidor proteômico ExPASy (*Expert Protein Analysis System*) (91) e reunidos na Tabela 3.

Tabela 3. Dados gerais da proteína trabalhada.

Antibiótico	Organismo	Peso Molecular	Nº de resíduos de aminoácidos	pI teórico
Tsn15	<i>Streptomyces longisporoflavus</i>	22,447 Da	206	4,64

Conforme mostrado na Figura 29, Tsn15 eluiu com alto grau de pureza pela técnica de separação por exclusão molecular. Os dados de purificação da proteína foram analisados por gel de SDS e a presença de uma única banda confirma a pureza da amostra.

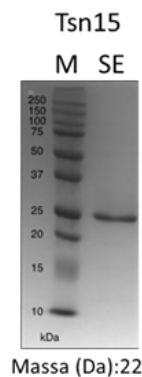


Figura 29. Análise de SDS-PAGE da purificação da enzima Tsn15. M – marcador; AF – cromatografia de afinidade por íons metálicos imobilizados (IMAC); SE – cromatografia de exclusão molecular.

Após a purificação por cromatografia de exclusão molecular, as frações coletadas foram juntadas, concentradas e analisadas em nanodrop para que fosse feita a medida da concentração das enzimas. Ao final foram obtidos 20 mg.mL<sup>-1</sup> de Tsn15. Testes de cristalização foram realizados e em uma condição surgiram cristais (Figura 30).

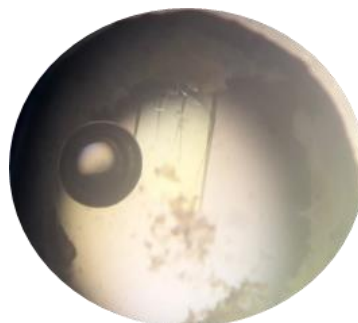


Figura 30. Cristais de Tsn15 obtidos por otimização manual.

A Tabela 4 apresenta as condições de cristalização para a enzima Tsn15 obtida a partir do kit Molecular Dimensions-1.

Tabela 4. Condições de cristalização da enzima Tsn15.

	<b>Precipitante</b>	<b>Precipitante 2</b>	<b>Tampão</b>	<b>Aditivo</b>
Tsn15	27 % PEG 1500		0,1 M PCPT pH 6	

A estrutura da proteína Tsn15 na sua forma apo já havia sido determinada por nossos colaboradores em Cambridge, mas neste projeto fomos capazes de determinar a estrutura de Tsn15 em complexo com seu intermediário. A estrutura de Tsn15 com o intermediário foi determinada com uma resolução de 1,7 Å. O modelo final possui 413 resíduos e 250 moléculas de água, com Rfactor e Rfree de 17 % e 20,4 %, respectivamente. A estatística do processamento de dados pode ser vista na Tabela 5.

Tabela 5. Dados coletados e Estatísticas de Refinamento para Tsn15 em sua forma holo. As estatísticas para a resolução mais alta são mostradas entre parênteses.

<b>Tsn15 em complexo com intermediário</b>	
Código de acesso	6NNW
Resolução (Å)	43,81 - 1,698 (1,759 - 1,698)
Grupo Espacial	C 1 2 1
Parâmetros de célula unitária (a, b, c e $\alpha$ , $\beta$ , $\gamma$ )	129,671 36,95 92,03 90 107,825 90
Reflexões totais	90988 (8254)
Reflexões únicas	45893 (4205)
Multiplicidade	2,0 (2,0)
Completeza (%)	98,81 (91,11)
Média I/sigma(I)	13,3 (2,0)
Wilson Fator-B	21,87
R-merge	0,030 (0,338)
R-meas	0,042 (0,479)
CC1/2	0,998 (0,788)
CC*	1 (0,939)
Reflexões usadas no refinamento	45883 (4202)
Reflexões usadas para R-free	2318 (195)
R-work	0,170 (0,299)
R-free	0,204 (0,355)
CC(work)	0,968 (0,867)
CC(free)	0,961 (0,812)
Número de átomos não-hidrogênio	3562
Macromoléculas	3211
Ligantes	82
Solvente	269
Resíduos de proteína	413
RMS (ligações)	0,007
RMS (ângulos)	0,9
Ramachandran regiões favoráveis (%)	99,51

Ramachandran regiões permitidas (%)	0,49
Ramachandran outliers (%)	0,00
Rotâmeros outliers (%)	0,00
Clashscore	4,16
Fator-B médio	28,97
Macromoléculas	27,84
Ligantes	39,71
Solvente	39,20

O cristal é pertencente ao grupo espacial C121 com duas moléculas na unidade assimétrica. Cada monômero possui um núcleo com enovelamento do tipo barril  $\beta$ . Duas  $\alpha$  hélices estão localizadas em cada protômero entre os núcleos de barril  $\beta$  e são importantes para a formação do homodímero (Figura 31). A interação dos protômeros dentro da unidade assimétrica cristalográfica é estabilizada por pontes salinas, ligações de hidrogênio e contatos não ligados (Figura 32).

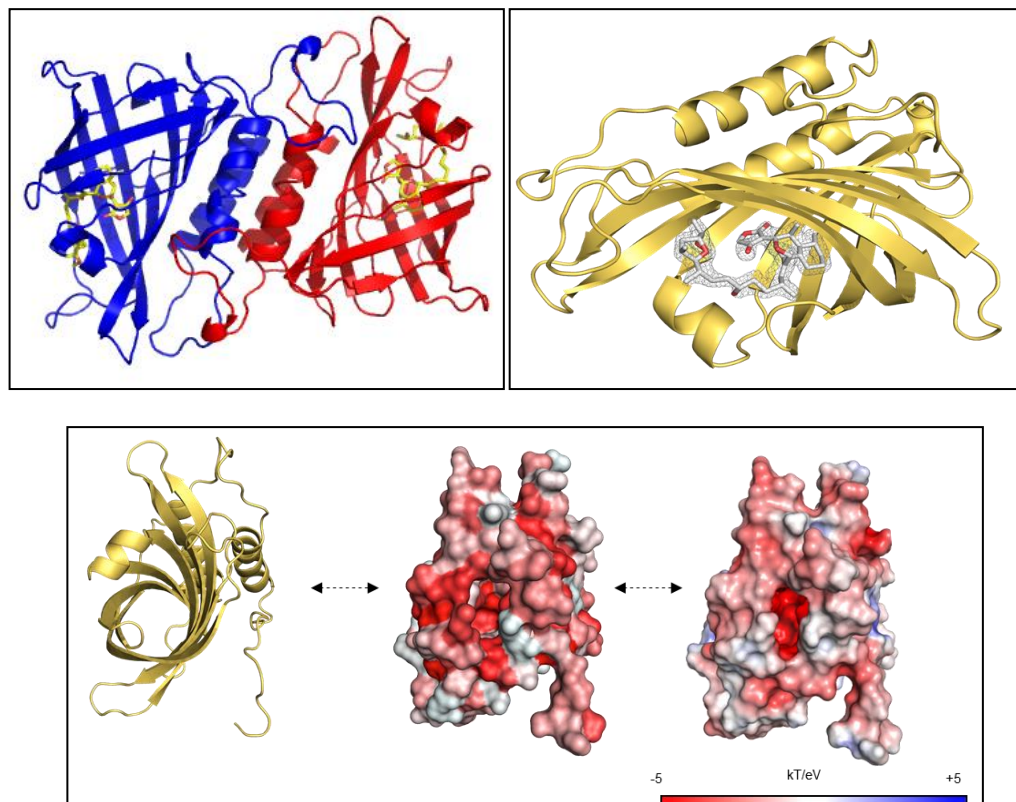


Figura 31. Representações da estrutura homodimérica de Tsn15 em complexo com seu intermediário e análises do sítio ativo.



Figura 32. Interface de dimerização da Tsn15. O servidor online PDBsum (144) (EMBL-EBI) foi usado para gerar um diagrama esquemático mostrando as interações entre as duas cadeias de Tsn15. As cadeias estão unidas por linhas coloridas, cada uma representando um tipo diferente de interação, laranja = contatos não ligados, azul = ligações de hidrogênio e linhas vermelhas = pontes salinas. A área de cada círculo é proporcional à área da superfície da cadeia proteica correspondente. A extensão da região da interface em cada cadeia é representada pela fatia preta cujo tamanho significa a área da superfície de interface.

Cada protômero de Tsn15 apresenta um enovelamento conservado que foi observado anteriormente nas ciclases [4 + 2] da biossíntese de espirotetronato e espirotetramato (PyrI4 e AbyU) (162, 163), apesar da baixa identidade de sequência (aproximadamente 20 %). Além dessas [4 + 2] ciclases homodiméricas, Tsn15 também compartilha um enovelamento com as enzimas ciclases homotriméricas de óxido de aleno, que catalisam uma reação de eletrociclicização  $4\pi$  (pericíclica) na biossíntese de ácido jasmônico (164, 165), e PodA uma desmetilase de pirocianina incomum de *Mycobacterium fortuitum* (166). As diferentes atividades dessas enzimas homólogas demonstram notável capacidade desse enovelamento de proteínas para catalisar uma gama diversificada de reações (Figura 33).

Cada barril  $\beta$  de oito fitas antiparalelas da Tsn15 possui uma  $\alpha$  hélice N-terminal, semelhante à estrutura de PyrI4 (163). Essa  $\alpha$ -hélice é descrita como uma tampa funcional e essencial em PyrE4 (163). Na estrutura da AOC (164), por exemplo, a hélice C-terminal não aparece.

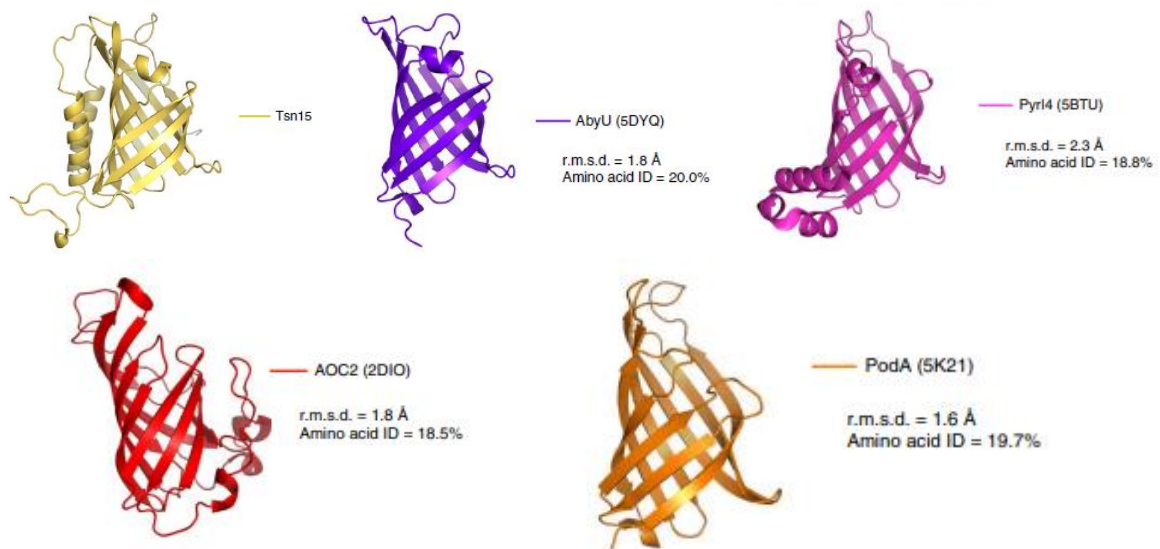


Figura 33. Homólogos estruturais de Tsn15. A cadeia  $C\alpha$  de Tsn15 foi alinhada com seus homólogos estruturais AbyU, Pyri4, AOC2 e PodA. O RMSD e a identidade sequencial (%) em comparação com Tsn15 são mostrados próximos a cada estrutura.

A Figura 34 apresenta o alinhamento de sequência da Tsn15 com outras ciclases [4+2]. A identidade sequencial dessas enzimas é baixa, apesar do enovelamento similar descrito anteriormente. É possível verificar que os resíduos que interagem com o ligante na estrutura complexada da proteína não são conservados. Com base nisso, possivelmente a Tsn15 possui um mecanismo catalítico exclusivo que não foi descrito em outras ciclases de tetronato.

```

Tsn 15/1-206   1 MTTSIDPTTPLTYNEVIDALVGSWRQIIDADYSADDTLRPLDLAVLARSTARAVAAVPRPLAEISA 66
Pyri4/1-184   1 MTT PQID ..... ERAMEAGAAALQETIVDPGGLD-VTALAVAAAALAGLH--SA 46
PodA/1-141    1 MDGRGGRSTT...EPVT..... 14
AbyU/1-160    1 MAHHHH.....HSSGLE.....VLFQGPMT... 21
AOC2/1-188    1 MRGSHHH.....HHHRSPSK..... 15

Tsn 15/1-206   67 PDAPDERGELVLLLEKVIQE-VADREYTPLSPEG.....PSVGDLLVLTKEIYNSD-REEI 119
Pyri4/1-184   47 ADDPAAALDKCIVLDELTE-FAEKLVVHDRP.....GGIGTTVEVVEYEDAS-GVRL 97
PodA/1-141    15 .....MTLDVKNQVAKH--DFGKPG.....MDVGDMDIFSDIL-SVD-GKQV 53
AbyU/1-160    22 ER-LETRPQALLIKVPE-IVVKVDDVDVAA.....PAVGQVGFDDDELIDEA-GAQI 72
AOC2/1-188    16 .....MQE-LSVYEINELDRHSPKILKNAFSLMFGGLGLVPFINKLYTGDLLKRV 64

Tsn 15/1-206   120 GADTGRRLRIIRKDPETGHHFTVSLVISTVQGNKLFAGGYTEMEAQLAGGRTTIQVACWDGWPWAGMS 185
Pyri4/1-184   98 GTATGNNAVVLKMEPHMW--QFHQSVSELADGSFEAVGVIDCTAMLRRTMQVLRVTRSGRYAGKS 160
PodA/1-141    54 GYDGGACFFTNVTPDNP-MTYCELIHLDAGEIFARSLTPH---TLAPFTMAITGGTGEYANSK 113
AbyU/1-160    73 GTSSGNFRIEYVRPTDGGMLTYQEDI TLSDGVIHAEGWADFNDVRTSKWVFYPATGVSGRYLGLT 138
AOC2/1-188    65 GITAGLCVVIIEHVPEKK--GERFEATYSFYFGDYGHLVQGG--PYLTYEDSFLAITGGAGIFEGAY 126

Tsn 15/1-206   186 GTLSW.....VINS.....MTAAESRYELR.....R... 206
Pyri4/1-184   161 GFMTL.....AISD.....PNQRPPHYSV--QVV-LC... 184
PodA/1-141    114 GELT.....VSG.....VATPDEKYEELKTKAENLYFQ... 141
AbyU/1-160    139 GFRQW.....RMTG.....VRKSAEARIL--MGE... 160
AOC2/1-188    127 GQVKLQQLVYPTKLFYTFYKGLANDLPLELGTTRVPPSKDI EPAPEAKAL---EPSSGVISNYTN 188
  
```

% Identidade com Tsn15	
Pyri4	18,8
PodA	19,7
AbyU	20
AOC2	18,5

Figura 34. Alinhamento de Tsn15 com outras ciclases [4+2]. O alinhamento foi gerado pelo ClustalOmega (146) e importado para o Jalview Versão: 2.11.0 (147). Resíduos similares são



coloridos de acordo: Azul, hidrofóbico; Verde, polar; Roxo, negativo; Vermelho, positivo; Amarelo, prolina; Laranja, glicina. A tabela exibe a identidade em porcentagem de Tsn15 com os homólogos alinhados. Os resíduos que interagem com o ligante estão destacados com um asterisco.

O RMSD C $\alpha$  entre Tsn15 apo e Tsn15 é de 0,5 Å (Figura 35). Nenhuma mudança estrutural significativa da enzima é necessária para acomodar o intermediário. Enquanto para PyrI4, os autores mencionaram que a enzima se comporta de maneira altamente dinâmica, com uma armadilha em forma de tampa para alternar a reação e a ciclização resultando de uma restrição conformacional sinérgica no barril  $\beta$ . Em AbyU, a comparação de simulações de DM das enzimas apo e ligada ao intermediário mostra que pouca alteração estrutural da enzima é necessária para acomodar o intermediário.

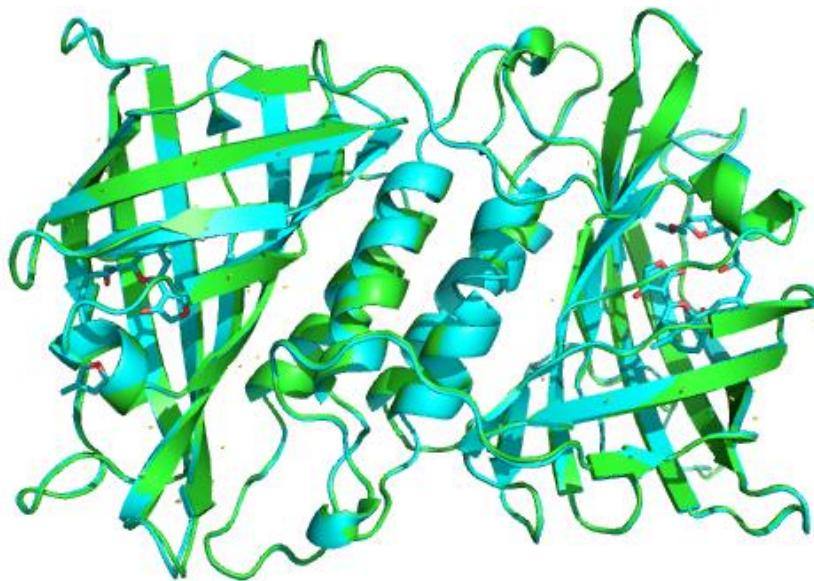


Figura 35. Sobreposição das estruturas de Tsn15 em suas formas apo e em complexo com o ligante.

O canal central do barril  $\beta$  forma uma cavidade em grande parte hidrofóbica e, usando o programa CASTp (167), foi possível calcular os volumes do sítio ativo de Tsn15, AbyU, PyrI4, PodA e AOC2 que são 388 Å<sup>3</sup>, 720 Å<sup>3</sup>, 840 Å<sup>3</sup>, 168 Å<sup>3</sup> e 77 Å<sup>3</sup>, respectivamente (Figura 36). Esse resultado indica que Tsn15 possui uma cavidade menor que PyrI4 e AbyU, mas possui uma cavidade maior que PodA e AOC2. O fato de o volume dos sítios ativos dessas enzimas ser diferente, muito provavelmente faz com que essas enzimas possam aceitar substratos diferentes.

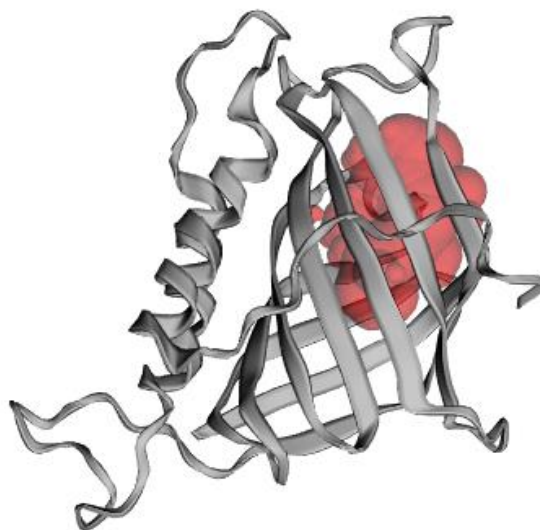


Figura 36. Volume do sítio ativo da Tsn15. A ferramenta online CASTp 3.047 (Universidade de Illinois, EUA) (167) foi usada para calcular o volume da cavidade do barril  $\beta$  de Tsn15 de 388 Å.

O anel decalina está em posição de *staking* com o triptofano 190 da estrutura, um tipo de interação não covalente entre anéis aromáticos. Esses anéis estão em uma configuração paralela e o triptofano está funcionando como uma espécie de “cama”, posicionando e estabilizando o intermediário no local correto no sítio de ativo. Além disso, podemos ver ao menos duas interações polares que parecem ser importantes na catálise entre o intermediário e os aminoácidos serina 142 e glutamina 164 (Figura 37). O anel tetronato está totalmente inserido na cavidade  $\beta$ -barril de Tsn15, com a fração de tetra-hidrofurano projetando-se no solvente. O ligante 1-OH-PHZ também parece se ligar por meio de uma interação *staking* p-p com F70 em PodA (166) e em AOC2 (164), o estado de transição com uma carboxilação do pentadienil e um oxiânion é estabilizado por uma molécula de água fortemente ligada e interações p-p favoráveis com resíduos aromáticos na cavidade. Além disso, o sítio ativo da Tsn15 possui quatro moléculas de água que podem ser importantes para o mecanismo catalítico (Figura 37). Essas águas fazem ligações de hidrogênio com as cadeias laterais de Ser142, Thr140, Gln164 e o intermediário em nossa estrutura. PodA, por exemplo, também requer água e substrato para a atividade (166).

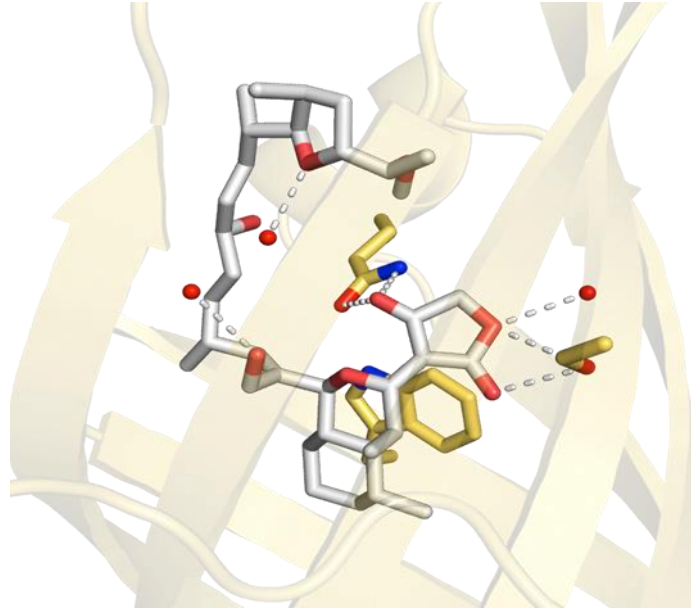


Figura 37. Sítio ativo da Tsn15 em complexo com o intermediário. Contatos de ligação detalhados entre o intermediário e os resíduos do barril  $\beta$  interno estão representados. As linhas cinza pontilhadas representam ligações de hidrogênio.

Cada barril  $\beta$  contém um sítio ativo hidrofóbico com uma região de potencial eletrostático negativo próximo ao anel decalina, semelhante a PyrI4, AOC2 e AbyU (Figura 38) (162-164). Diferentemente, no sítio ativo do PodA (166), existem vários resíduos de aminoácidos polares carregados e polares.

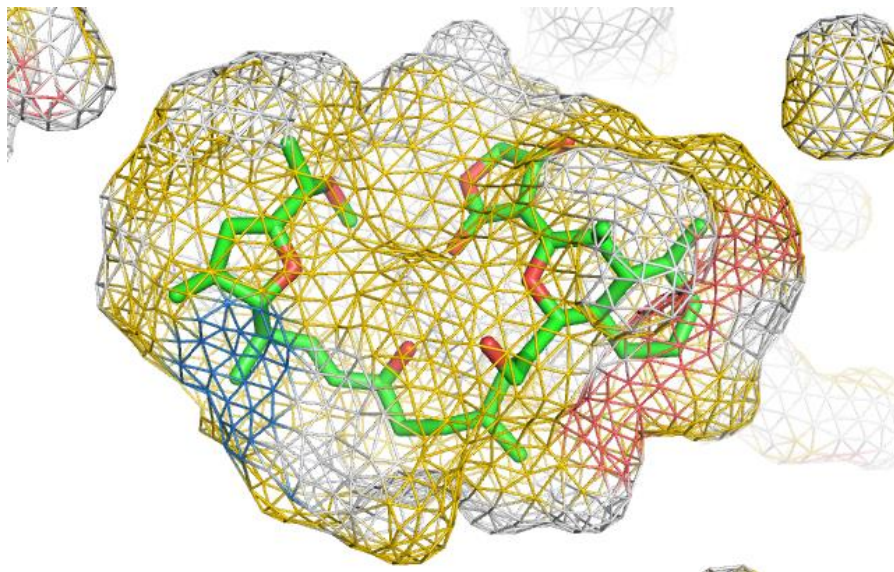


Figura 38. Representação da superfície do sítio ativo. Os resíduos hidrofóbicos ou polares (com carga positiva ou negativa e sem carga) são indicados em cores, sendo amarelo para resíduos hidrofóbicos, vermelho para resíduos carregados negativamente e azul para resíduos carregados positivamente.



## Conclusão

Neste capítulo determinamos a estrutura de Tsn15 em complexo com seu intermediário de reação. Tsn15, que é estruturalmente homóloga às ciclases conhecidas como Diels-alderases, evoluiu para catalisar uma transformação diferente (mas possivelmente também pericíclica). Do ponto de vista da biologia sintética, essa enzima pode ser útil para a síntese estereoespecífica das porções ciclohexano e tetra-hidropirano em análogos sintéticos de poliéteres ionóforos. Esses achados sugerem que novos exemplos de reações pericíclicas enzimáticas ainda precisam ser descobertos nas vias biossintéticas de produtos naturais. A elucidação da estrutura, do modo de ligação de substratos e também do mecanismo catalítico resolvem um problema instalado por anos dentro da biossíntese dessa classe de produtos naturais, na qual envolve a determinação da formação dos anéis centrais dos compostos tetronasina e tetronomicina. A presença de um intermediário de biossíntese latente com a presença de um anel decalina até então não era esperada para esse sistema e se acreditava que cada enzima era responsável pela formação de anéis distintos. Dessa maneira, este trabalho representa um avanço significativo no entendimento da biossíntese de policetídeos ionóforos contendo tetronatos.

## **6 CAPÍTULO III - DIELS-ALDERASES EM POLIÉTERES IONÓFOROS TETRONATOS: TMN9**

### **Objetivo geral:**

- Determinação e análise estrutural da proteína Tmn9 da biossíntese de tetronomicina

Este capítulo gerará o artigo “Tmn9: An essential [4+2] cyclase involved in the biosynthesis of tetranomycin.” que está em preparação.

Autores: Fernanda C. R. de Paiva, Rory Little, Paulo C. T. de Souza, Eduardo T. de Oliveira, Finian J. Leeper, Dick B. Janssen, Siewert-Jan Marrink, Peter F. Leadlay, Marcio V. B. Dias.

A via da Tetronomicina foi proposta com base na via da Tetronasina. Assim como para Tsn15, a origem dos anéis ciclo-hexano central e tetra-hidropirano da Tetronomicina é proposta como sendo uma reação em cascata enzimática de duas etapas. As duas enzimas responsáveis, Tmn8 e Tmn9 são ortólogas das enzimas [4+2] ciclases Tsn11 e Tsn15. Tmn9 é a penúltima enzima e Tmn8 é a última enzima na produção do antibiótico (Figura 39). Apesar de inúmeras tentativas, não conseguimos cristalizar Tmn9 em complexo com substrato ou produto. Diante desses resultados, foram realizadas experiências em múltiplas escalas.

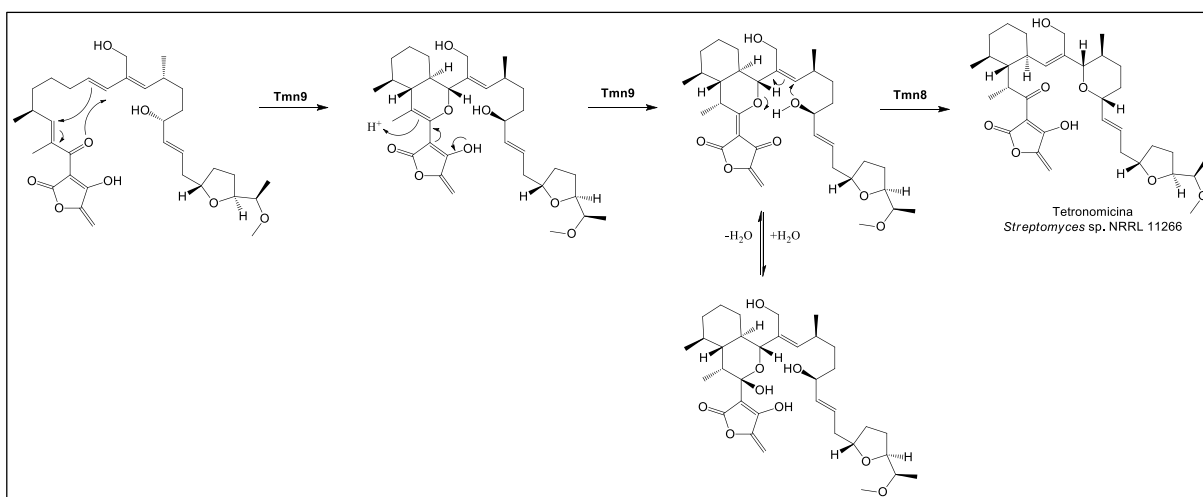


Figura 39. Via Biossintética Proposta. Um duplo exocíclico no anel tetronato sofre a cicloadição [4 + 2] com o dieno terminal do policetídeo linear. A Tmn9 catalisa a formação do anel transdecalina.

A partir da clonagem da enzima Tmn9, ciclase [4+2], no vetor de expressão pET28a, foi possível produzir a proteína em sua forma solúvel com rendimento adequado para purificação e posteriores experimentos cristalográficos. Dados gerais da proteína trabalhada foram obtidos de acordo com análise feita pelo programa ProtParam (90) hospedado no servidor proteômico ExPASy (*Expert Protein Analysis System*) (91) e reunidos na Tabela 6.

Tabela 6. Dados gerais da proteína trabalhada.

Antibiótico	Organismo	Peso Molecular	Nº de resíduos de aminoácidos	pI teórico
Tmn9	<i>Streptomyces</i> sp. NRRL 11266	54,277 Da	500	4,81

Conforme mostrado na Figura 40, Tmn9 eluiu com alto grau de pureza pela técnica de separação por exclusão molecular. Os dados de purificação da proteína foram analisados por gel de SDS e a presença de uma única banda confirma a pureza da amostra.

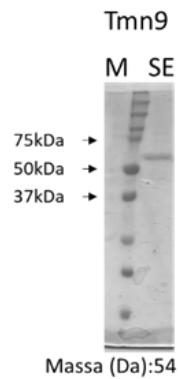


Figura 40. Análise de SDS-PAGE da purificação da enzima Tmn9. M - marcador; AF – cromatografia de afinidade por íons metálicos imobilizados (IMAC); SE – cromatografia de exclusão molecular.

Após a purificação por cromatografia de exclusão molecular, as frações coletadas foram juntadas, concentradas e analisadas no espectrofotômetro nanodrop (Biochrom) para que fosse feita a medida da concentração das enzimas. Ao final foram obtidos  $20 \text{ mg.mL}^{-1}$  de Tmn9. Testes de cristalização foram realizados e em uma condição surgiram cristais (Figura 41).



Figura 41. Cristal de Tmn9 obtido pelo robolab do LNBIO.

A Tabela 7 apresenta as condições de cristalização para a enzima Tmn9 obtida a partir do kit Crystal Screen (Hampton).

Tabela 7. Condições de cristalização da enzima Tsn15.

	<b>Precipitante</b>	<b>Precipitante 2</b>	<b>Tampão</b>	<b>Aditivo</b>
Tmn9	30 % w/v PEG 8000		100 mM cacodilato de sódio pH 6,5	200 mM acetato de sódio

A estrutura cristalina da proteína Tmn9 foi determinada com uma resolução de 2,4 Å. O cristal é pertencente ao grupo espacial  $P2_12_12_1$  com duas moléculas na unidade assimétrica (Figura 42). A proteína completa purificada (resíduos 1-497) era amarela brilhante, indicando a presença do cofator ligado, FAD.

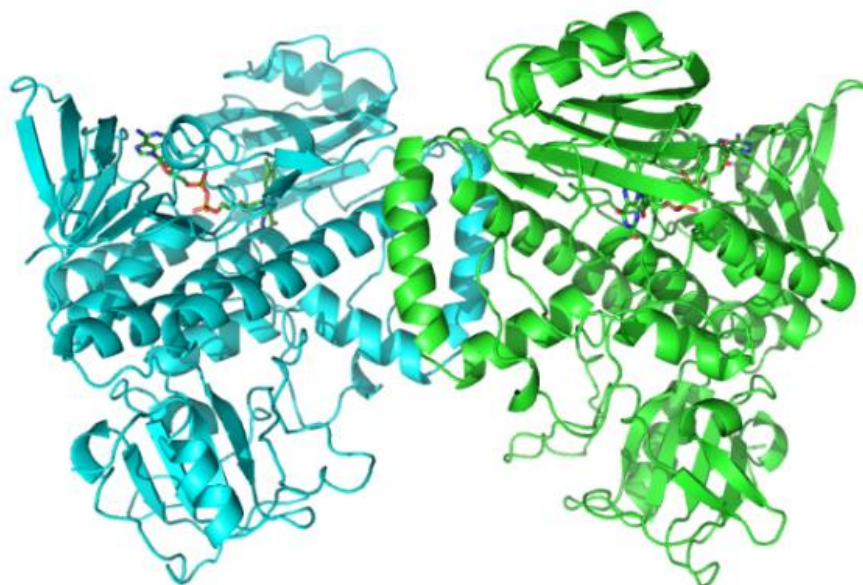


Figura 42. Representações da estrutura homodimérica de Tmn9 com a utilização do Pymol (94) a partir dos dados obtidos. Está evidenciada a presença do FAD em cada um dos monômeros.

A estatística do processamento de dados pode ser vista na Tabela 8.

Tabela 8. Dados coletados e Estatísticas de Refinamento para Tmn9 em sua forma apo. As estatísticas para a resolução mais alta são mostradas entre parênteses.

<b>Tmn9</b>	
Código de acesso	6UI5
Resolução (Å)	48,42 - 2,4 (2,486 - 2,4)
Grupo Espacial	$P 2_1 2_1 2_1$
Parâmetros de célula unitária (a, b, c e $\alpha$ , $\beta$ , $\gamma$ )	61,205 85,605 207,683 90 90 90
Reflexões totais	276007 (28306)
Reflexões únicas	43537 (4277)
Multiplicidade	6,3 (6,6)
Completeza (%)	99,69 (99,60)
Média I/sigma(I)	15,57 (2,10)
Wilson Fator-B	58,06
R-merge	0,06832 (0,8071)
R-meas	0,07481 (0,876)
CC1/2	0,996 (0,882)
CC*	0,999 (0,968)
Reflexões usadas no refinamento	43484 (4268)
Reflexões usadas para R-free	2156 (205)
R-work	0,2027 (0,3277)

R-free	0,2487 (0,3913)
CC(work)	0,966 (0,839)
CC(free)	0,986 (0,557)
Número de átomos não-hidrogênio	7728
Macromoléculas	7584
Ligantes	106
Solvente	38
Resíduos de proteína	994
RMS (ligações)	0,013
RMS (ângulos)	1,35
Ramachandran regiões favoráveis (%)	90,30
Ramachandran regiões permitidas (%)	7,88
Ramachandran outliers (%)	1,82
Rotâmeros outliers (%)	0,39
Clashscore	16,26
Fator-B médio	83,92
Macromoléculas	83,92
Ligantes	94,10
Solvente	56,41

O problema de fase foi resolvido pelo método de substituição molecular usando a estrutura da proteína PyrE3 (PDB ID: 5XGV) como modelo. Assim como PyrE3, Tmn9 é uma proteína com enovelamento semelhante das *para*-hidroxibenzoato hidroxilases. Verificou-se que os dois protômeros de Tmn9 ocupam a unidade assimétrica do cristal, na qual eles se agrupam para formar um dímero simétrico antiparalelo. Cada protômero é composto por 14  $\alpha$ -hélices e 25 fitas  $\beta$  e possui três domínios: um domínio de ligação ao FAD, um domínio médio e um C-terminal do tipo tiorredoxina  $\alpha/\beta$  (Embora faltem dois resíduos de cisteína essenciais para a catálise da tiorredoxina). O domínio de ligação ao FAD apresenta um enovelamento do tipo Rossmann. Somente o domínio de ligação ao FAD participa da dimerização dos dois protômeros (Figura 43) (168).

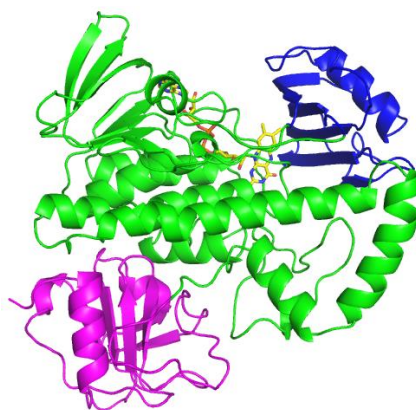


Figura 43. Protômero da Tmn9. Estão evidenciados os três domínios: domínio de ligação ao FAD (verde), domínio médio (azul) e domínio C-terminal do tipo tiorredoxinana (rosa). A molécula do FAD ligada está mostrada no modelo em bastão.

A interação entre Tmn9 e FAD é predominantemente mediada por ligações de hidrogênio envolvendo os resíduos A13, D32, H34, R38 e o resíduo conservado D284 e as principais cadeias dos resíduos D284 e L121 (Figura 44).

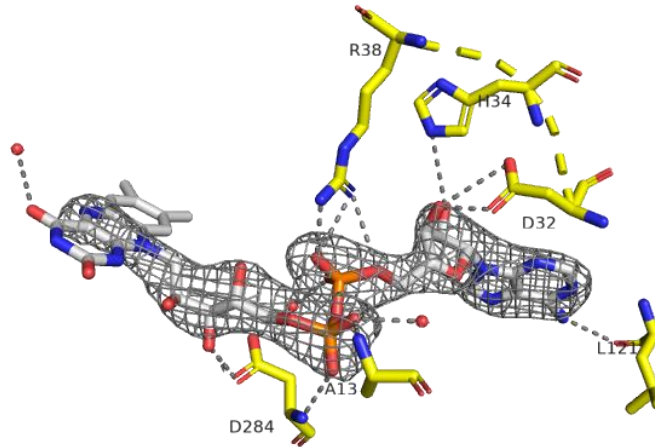


Figura 44. Sítio de ligação ao FAD. As ligações de hidrogênio envolvidas estão mostradas como linhas pontilhadas em cinza. O mapa de densidade eletrônica de  $2F_o - F_c$  é contornado em  $2,0 \sigma$ .

Apesar de não termos sido capazes de cristalizar Tmn9 em complexo com seu substrato, a análise estrutural de Tmn9 revelou uma cavidade positivamente carregada e hidrofóbica na interface entre o domínio de ligação FAD e o domínio médio. Verificamos que as cavidades de ligação ao substrato que é a mesma de ligação ao FAD é grande e abriga um espaço capaz de acomodar o substrato ( $1716,720 \text{ \AA}^3$ ) (Figura 45 e Figura 46).

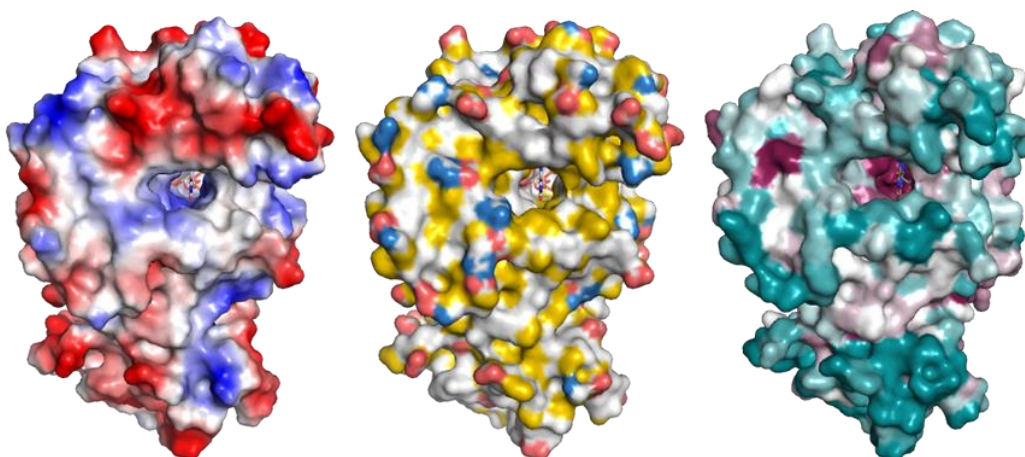


Figura 45. Análises estruturais do Bolsão de Ligação ao Substrato de Tmn9. (A) A representação do potencial de carga de superfície evidencia o bolsão de ligação ao FAD positivamente carregado (azul indicando carga positiva e vermelho carga negativa). (B) Representação da superfície em que os resíduos que são hidrofóbicos estão em amarelo, os positivamente carregados em azul e os negativamente carregados em vermelho. (C) Representação de superfície mostrando que o sítio de ligação ao substrato em flavociclases é

variável. Neste desenho, a superfície de Tmn9 é colorida de acordo com a conservação da sequência de aminoácidos baseada no alinhamento da sequência de diferentes flavociclases usando o servidor ConSurf (roxo e ciano indicam alta e baixa conservação de sequência, respectivamente).

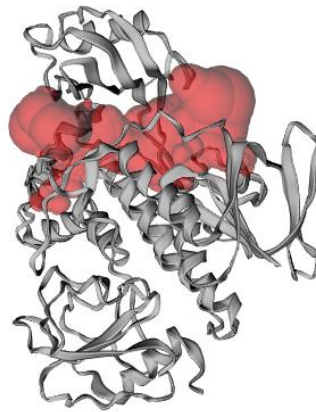


Figura 46. Volume do sítio ativo de Tmn9. A ferramenta online CASTp 3.0 (169) (Universidade de Illinois, EUA) foi usada para calcular o volume de  $1716.720 \text{ \AA}^3$ .

A pesquisa de similaridade estrutural com o DALI (170) revelou uma identidade de sequência inferior a 41 % com outras estruturas de homólogos depositados no Protein Data Bank (5xgv, 5kox, 6brd, 4k2x, 6j0z e 4x4j). A superposição de PyrE3 de *Streptomyces rugosporus* (5XGV) (171), rifampicina monooxigenase complexada com rifampicina de *Nocardia farcinica* (RIFMO - 5KOX) (172), rifampicina monooxigenase de *Streptomyces venezuelae* complexada com rifampina e FAD (Rox - 6BRD) (173), anidrotetraciclina hidroxilase de *Streptomyces rimosus* (OxyS - 4K2X) (174), monooxigenase AlpK de *Streptomyces ambofaciens* (6J0Z) (175) e BexE de *Amycolatopsis orientalis* subsp. *Vinearia* (4X4J) (a ser publicado) pode ser vista na Figura 47. Essas estruturas apresentam um alto grau de diferenças estruturais com um RMSD em torno de  $2,5 \text{ \AA}$ . As principais mudanças em todos estão nas alças do domínio tipo tiorredoxina.



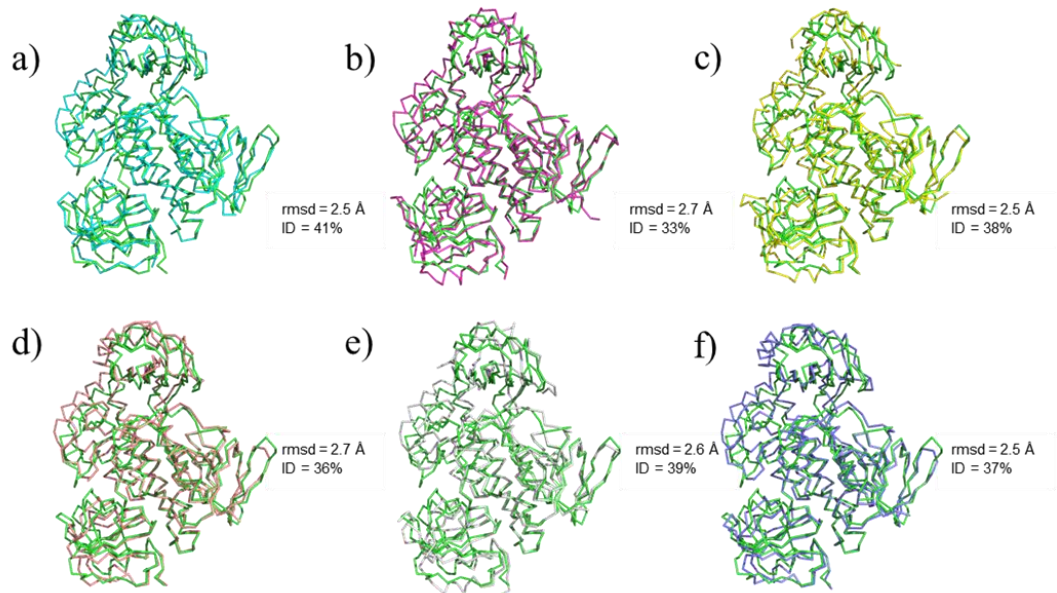


Figura 47. Homólogos estruturais de Tmn9 (verde). A cadeia de  $C\alpha$  da Tmn9 foi alinhada com seus homólogos estruturais (a) 5xgv (ciano), (b) 6j0z (magenta), (c) 4k2x (amarelo), (d) 5kox (rosa), (e) 4x4j (cinza) e (f) 6brd (azul). Ao lado de cada estrutura está o valor de rmsd e a identidade de aminoácidos (%) com a Tmn9.

As comparações estruturais revelaram uma considerável similaridade de sequência de Tmn9 com as monooxigenases da família das *para*-hidroxibenzoato-hidroxilase dependente de FAD (pHBH) (Figura 48). O alinhamento mostra os resíduos altamente conservados nos ortólogos do DALI (170). Um ponto interessante é a ausência de resíduos conservados no local de ligação do substrato, é possível notar apenas semelhanças entre os enovelamentos. Não foram identificados aminoácidos catalíticos essenciais no sítio ativo de Tmn9 (44M, 45A, 47N, 68V, 70F, 71P, 72E, 73V, 74R, 76A, 92T, 94M, 196Y, 197G, 198L, 206Q, 208L, 210A, 211P, 212L, 220F, 290Y, 291P, 292Y, 293N e 294G), sugerindo que, como PyrE3, o papel do sítio ativo é facilitar uma conformação próxima de ataque (NAC - *Near Attack Conformation*) (171). O sítio ativo possui apenas um aminoácido básico (74R) e um resíduo ácido (72E) que não estão posicionados de maneira a promover a catálise ácido-base. Apenas as prolina 211 e 291 são conservadas. A Pro291 está próxima do anel de isoaloxazina do FAD e possivelmente estabiliza a conformação desse cofator por meio de interação hidrofóbica, conforme proposto para AlpK (175). Resíduos não polares fazem as interações predominantes entre o ligante e a proteína.

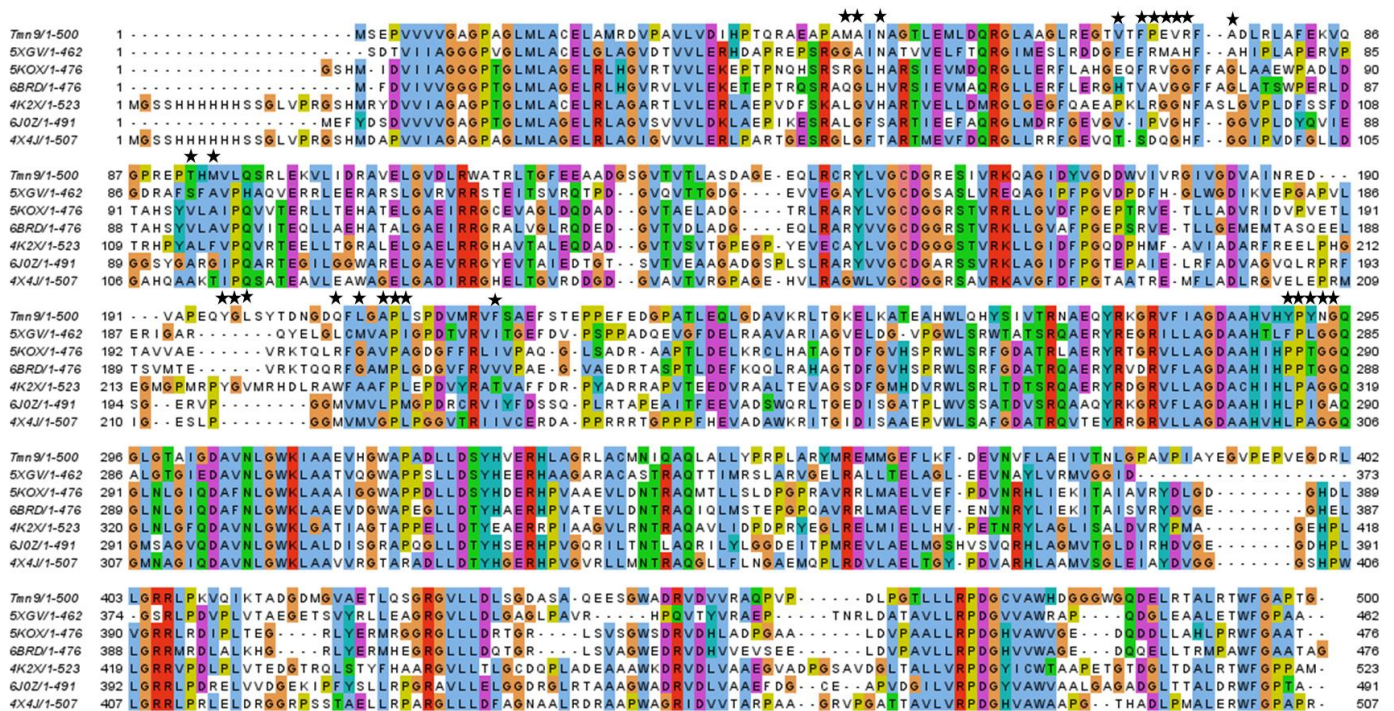


Figura 48. Alinhamento de Tmn9 e PyrE3 com homólogos do DALI. O alinhamento foi gerado pelo ClustalOmega (146) e importado para o Jalview Versão: 2.11.0 (147). Resíduos similares são coloridos de acordo: Azul, hidrofóbico; Verde, polar; Roxo, negativo; Vermelho, positivo; Amarelo, prolina; Laranja, glicina. Os resíduos do sítio ativo proposto são destacados por estrelas.

A comparação entre as estruturas também revelou uma alça intrigante de tamanho diferente nas estruturas. A alça da Tmn9 é muito grande quando comparado ao *loop* da PyrE3 na mesma posição, por exemplo (Figura 49).

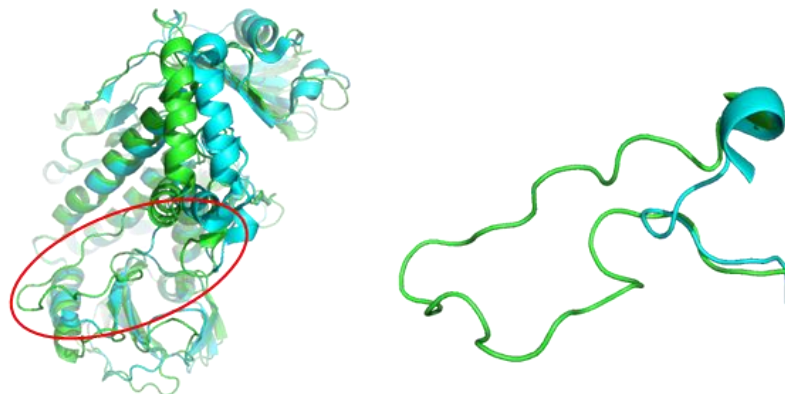


Figura 49. Comparação entre Tmn9 e PyrE3. *Loop* 379V a 409K em Tmn9 e *loop* 368R a 381V em PyrE3.

O cofator FAD adota nas estruturas cristalográficas das proteínas da família pHBH pelo menos três conformações: “in” (176), “open” (175) e “out” (177), dependendo das condições de cristalização (178). A conformação “open” é uma etapa crucial sobre o ambiente de reação no qual o FAD é exposto ao solvente e a proteína está apto para se ligar ao ligante (175, 178). A estrutura de AlpK, uma homóloga de Tmn9, também foi determinada na forma apo (apenas em complexo com o FAD), como Tmn9 neste estudo. Considerando isso, presume-se que as estruturas de Tmn9 e AlpK adotem uma conformação “open” (175). No entanto, podemos ver que o FAD está em posições diferentes nas duas estruturas. Pode ser devido ao fato de que AlpK é literalmente uma monooxigenase essencial que também se liga ao NADH e requer uma reação redox com o FAD. Aparentemente, o FAD na Tmn9 está na configuração “in”, como outro membro da família pHBH, a PgaE (176) (Figura 50). Em teoria, a conformação do FAD será alterada para “in” após a ligação do substrato (175, 178).

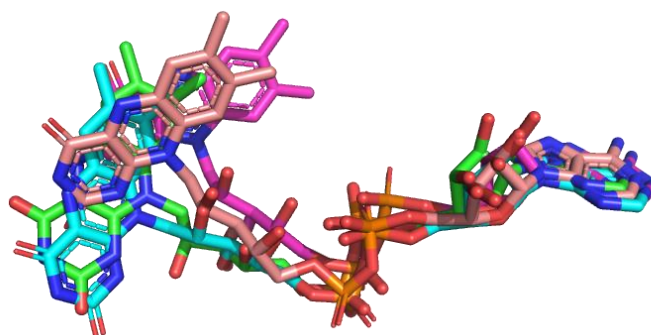


Figura 50. Comparação estrutural da conformação do FAD em Tmn9 (verde), em AlpK com conformação “open” (salmão, PDB: 6J0Z) (175), em PgaE com conformação “in” (ciano, PDB: 2QA1) (176) e em WT pHBH com conformação “out” (magenta, PDB: 1DOD) (177). O alinhamento foi realizado com as proteínas.

Sugerimos que o FAD seja apenas estrutural e atue mantendo a geometria da proteína e também posicionando o ligante para interações específicas com base em análises filogenéticas (Figura 51). Essas análises indicam que essas proteínas estão mais relacionadas às [4 + 2] ciclases do que às monooxigenases e carecem dos resíduos associados à ligação ao NADPH (179), sugerindo que o FAD tem uma função não-redox (180).

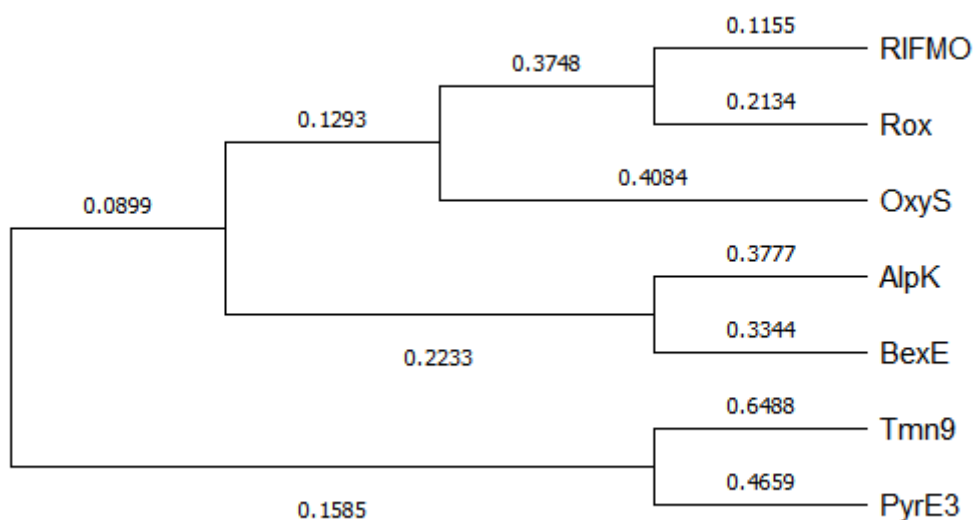


Figura 51. Análise evolucionária pelo método de máxima verossimilhança de proteínas com similaridade estrutural com a Tmn9. Os comprimentos das ramificações são mostrados. As proteínas e organismos utilizados são: A proteína deste estudo Tmn9 de *Streptomyces* sp. NRRL 11266 (PDB: 6UI5), PyrE3 de *Streptomyces rugosporus* (PDB: 5XGV), Rifampicina Monooxigenase complexada com rifampicina de *Nocardia farcinica* (PDB: RIFMO - 5KOX), Rifampicina Monooxigenase de *Streptomyces venezuelae* complexada com rifampina e FAD (PDB: Rox - 6BRD), Anidrotetraciclina Hidroxilase de *Streptomyces rimosus* (PDB: OxyS - 4K2X), Monooxigenase AlpK de *Streptomyces ambofaciens* (PDB: 6J0Z) e BexE de *Amycolatopsis orientalis* subsp. *Vinaria* (PDB: 4X4J) (a ser publicado). As análises foram realizadas em MEGA X (160).

Para apoiar a previsão da sugerida cavidade ser o bolsão de ligação ao substrato de Tmn9, o produto dessa enzima, que é relativamente rígido em comparação com o substrato linear altamente flexível, foi docado na estrutura cristalina de Tmn9 com o programa Autodock Vina (116). Uma busca pela melhor forma complementar comparativamente aos homólogos de Tmn9 levou a três melhores poses, que estão bem encaixas e enterradas na cavidade e que formam interações favoráveis com os resíduos circundantes, visto que a partir dessas poses foram produzidas dinâmicas moleculares dos ligantes na proteína e eles se mostraram estáveis em 500 ns. O FAD estrutural parece ser importante no posicionamento do anel decalina estabilizando o ligante no complexo. As melhores poses obtidas pela docagem, tanto para o substrato quanto para o produto, possuem o anel tetronato/anel decalina próximo à isoaloxazina do FAD, assim como todos os homólogos com ligantes usados como comparação, sendo eles: Rifampicina Monooxigenase complexada com rifampina e FAD (PDB: Rox - 6BRD) (Figura 52), Rifampicina Monooxigenase complexada com rifampicina (PDB: RIFMO - 5KOX), figura da docagem realizada com PyrE3 (PDB: 5XGV) e Rifampicina Monooxigenase complexada com produto (PDB: 6C7S).



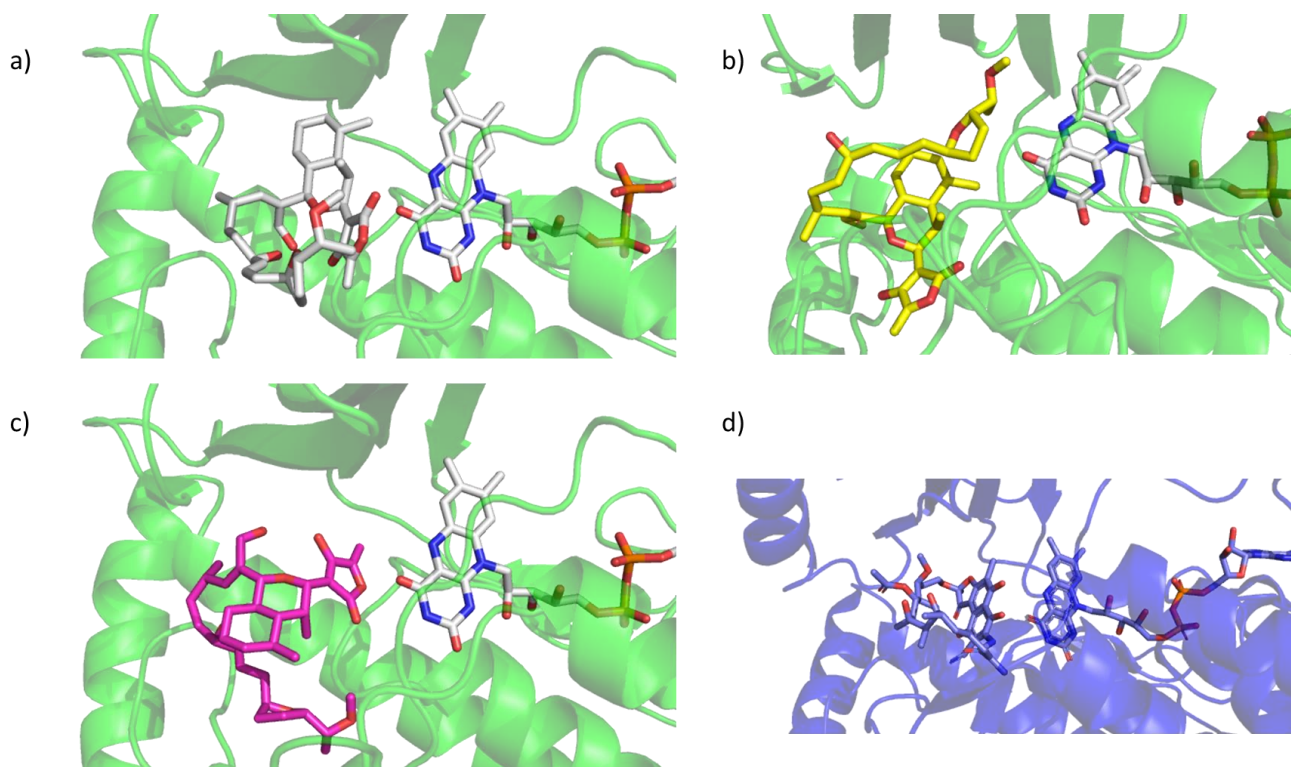


Figura 52. Melhores poses obtidas pela docagem e sítio ativo da homóloga Rox (PDB: 6BRDB). Comparativamente com o ligante da homóloga Rox depositado no PDB, foram escolhidas as melhores poses da docagem. É possível perceber que o anel decalina está próximo à isoaloxazina do cofator FAD em todas as posições.

O ambiente físico dentro de um cristal não é idêntico àquele da solução ou em uma célula viva. O cristal impõe um espaço e um tempo médios sobre a estrutura deduzida a partir das análises e os estudos de difração de raios X fornecem pouca informação sobre a movimentação molecular dentro da proteína (62).

Com base nos estudos computacionais realizados, propomos que o mecanismo catalítico ocorra possivelmente devido aos efeitos de proximidade e orientação. Os mecanismos catalíticos da reação enzimática são semelhantes aos das reações do modelo na água e, como afirmado anteriormente, reações Diels-Alder podem ocorrer espontaneamente. No entanto, a comparação das dinâmicas moleculares feitas com o substrato na água e com o substrato na proteína mostram que as posições mais favoráveis em que a molécula estava na posição de ataque foram encontradas com mais frequência na dinâmica do substrato na proteína do que na dinâmica do substrato na água (Figura 53). Portanto, pode-se concluir que as reações na presença da enzima são cataliticamente mais eficientes. Acreditamos que mesmo que a Tmn9 não apresente resíduos essenciais, assim como outras proteínas dessa família, ela é responsável por “catalisar” essa reação, não com base em ataques nucleofílicos, mas simplesmente colocando o substrato em uma conformação favorável para o ataque espontâneo dos grupos

químicos envolvidos. Dessa maneira, é dispensada a presença de resíduos cataliticamente relevantes, algo até então muito pouco observado na natureza nessa classe de proteínas.

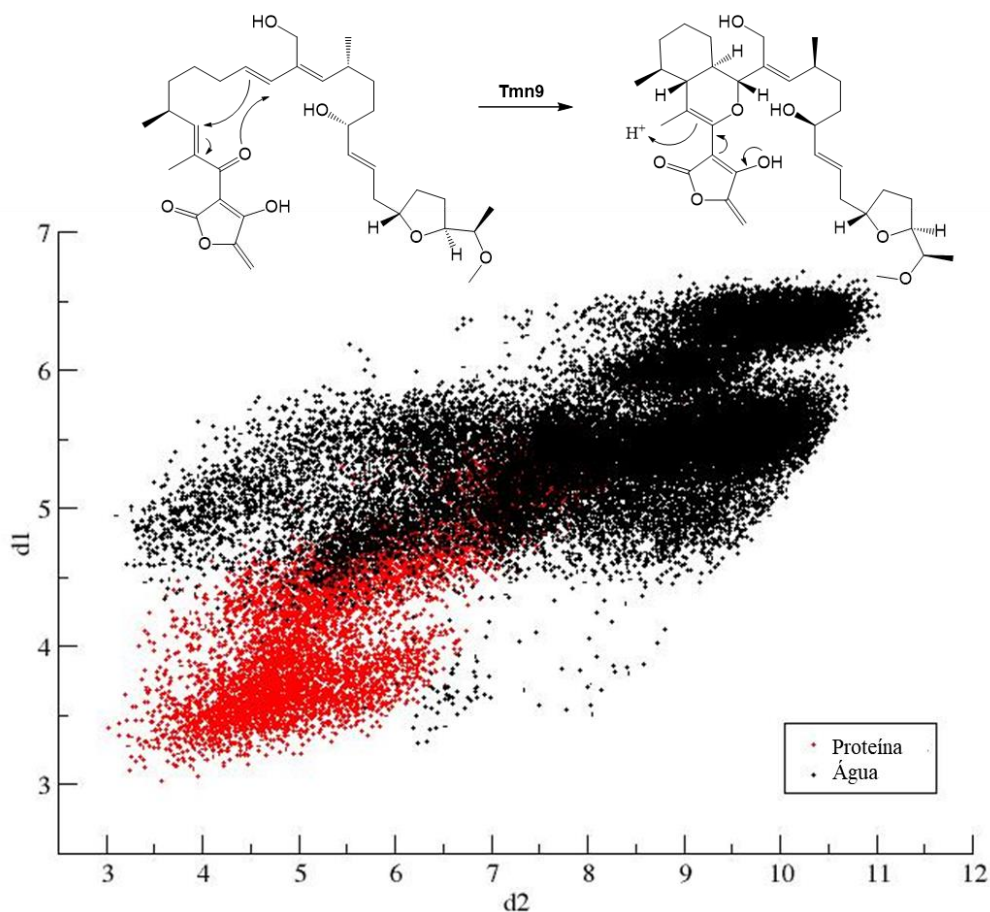


Figura 53. Distâncias entre os átomos envolvidos na formação do anel decalina durante a simulação em água (preto) e a simulação na proteína (vermelho). Os pontos se referem às distâncias entre as ligações C18 e O8 e entre C25 e C20, átomos envolvidos na formação do anel decalina. Os dados representam as médias das triplicatas das dinâmicas moleculares de 500 ns realizadas.

A dinâmica molecular revelou movimentos importantes que devem ocorrer durante o ciclo catalítico. Os dados das simulações de DM nos permitiram explorar as interações entre as poses do ligante geradas pela docagem e a proteína em escala atômica. Em particular, as simulações de DMs permitiram a comparação dos rearranjos estruturais e dinâmicos entre os estados holo e apo de Tmn9. Por meio de uma análise da média estrutural das conformações nos sistemas com e sem ligante, é possível verificar a variação conformacional da enzima e também do sítio ativo. O domínio médio atua como um *cap*, fechando o sítio catalítico como foi possível observar na dinâmica molecular. A mudança de conformação do *cap* também pode ser verificada na comparação entre os ortólogos. Esse domínio apresenta alta flexibilidade em comparação com outras áreas da proteína, como por exemplo o núcleo central, e isso foi descrito para outras enzimas semelhantes, evidenciando a flexibilidade do domínio (173, 176). Portanto,

foi possível verificar que as regiões mais móveis são basicamente o *cap* e boa parte da face envolvida na dimerização. Comparativamente, a estrutura com o ligante estabiliza o *cap* aberto, enquanto dados preliminares de simulações com o dímero indicam estabilização da interface entre os protômeros (Figura 54).

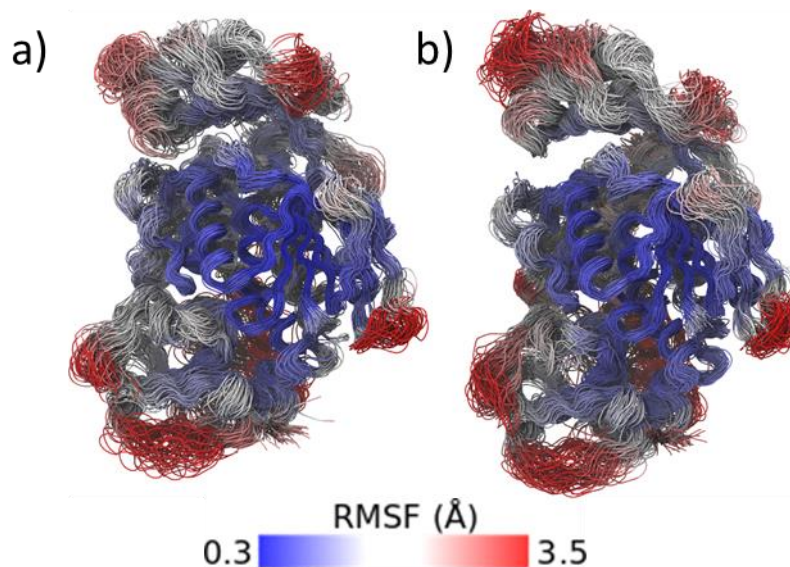


Figura 54. Comparação da flexibilidade dos protômeros de Tmn9 sem (a) e com ligante (b) obtida por meio da variação conformacional da enzima durante as simulações de dinâmica molecular. É possível verificar o *cap* em conformação mais aberta na presença do ligante em comparação com o protômero sem o ligante. A rigidez da região do núcleo central da proteína é evidenciada pela cor azul.

### Conclusão

Neste capítulo fornecemos *insights* estruturais e mecanicistas de Tmn9, uma ciclase [4+2] dependente de flavina. Os dados de Dinâmica Molecular e de docagem complementam os resultados estruturais, gerando informações em relação à variável movimento ao longo do tempo. O FAD desempenha um papel estrutural e o sítio ativo parece acomodar o substrato de forma semelhante às homólogas para as quais há estruturas cristalográficas em complexo depositadas. Propomos que a enzima atue na formação do anel transdecalina provavelmente por meio de um estado de transição [4 + 2] endo-seletivo. Apesar das reações Diels-Alder ocorrerem espontaneamente em água, nossos estudos evidenciam que a presença da enzima favorece a ocorrência da ciclização, permitindo que o substrato adquira uma posição favorável para a ocorrência da reação. A observação da baixa conservação sequencial e da ausência de resíduos relevantes corroboram a ideia de que embora essa enzima seja capaz de catalisar a formação do anel decalina, não há o envolvimento direto dela nesse mecanismo. Sendo assim, o papel dos resíduos do sítio ativo é o de somente dirigir a conformação do substrato para o

fechamento espontâneo do anel decalina por meio de uma reação de Diels-Alder. Estudos de QM/MM ainda estão sendo realizados para a confirmação da nossa hipótese que, se confirmada, será uma revelação importante dentro dos mecanismos enzimáticos presentes na natureza.



## 7 CONCLUSÕES E PERSPECTIVAS

Este trabalho permitiu descrever estruturalmente três enzimas envolvidas com ciclizações em vias de biossíntese de antibióticos: AjiA1, uma ligase dependente de ATP e Tsn15 e Tmn9, duas ciclases [4+2].

AjiA1 foi determinada em sua forma apo e as análises de sua comparação com a enzima NatL2 (forma holo) permitiram observar movimentos completamente novos durante o ciclo catalítico de AFEs. A análise estrutural de AjiA1 também permitiu identificar a presença de um *loop swapping*, um evento de troca de domínio envolvendo uma alça C-terminal longa. A presença desse elemento estrutural e a ausência de alternância de domínios fez com que propuséssemos que AjiA1 e NatL2 devam ser classificadas como um novo subgrupo de ligases do tipo fenilacetato-CoA. Além disso, foi verificado que AjiA1 é capaz de aceitar substratos alternativos, indicando promiscuidade para essa enzima. Para trabalhos adicionais, estudos cristalográficos da amidohidrolase AjiH podem esclarecer detalhes da enzimologia e da biossíntese do composto AJI9561. As enzimas Aji podem ser potencialmente usadas para a geração de novos antibióticos benzoxazóis por meio do uso de derivados do ácido benzóico como precursores.

A estrutura da Tsn15 foi determinada em complexo com um intermediário de reação no sítio ativo. Essa enzima apresentou dois protômeros na unidade assimétrica sendo cada um deles com enovelamento do tipo  $\beta$  barril. Quando comparada com outras ciclases [4+2], a identidade sequencial é baixa, apesar do enovelamento ser conservado e os resíduos que interagem com o ligante na estrutura complexada da proteína não serem conservados. Com base nisso, possivelmente a Tsn15 possui um mecanismo catalítico exclusivo que não foi descrito em outras ciclases de tetronato. Essa enzima pode ser útil para a síntese estereoespecífica das porções ciclohexano e tetra-hidropirano em análogos sintéticos de poliéteres ionóforos. A descrição dessa enzima, bem como as análises realizadas neste trabalho representa um avanço significativo no entendimento da biossíntese de policetídeos ionóforos contendo tetronatos. Estudos adicionais com a enzima Tsn11 da mesma via biossintética podem esclarecer detalhes da biossíntese das enzimas Diel-alderases.

Tmn9, uma ciclase [4+2], teve sua estrutura determinada em complexo com o cofator FAD. Os dados de Dinâmica Molecular e de docagem complementam os resultados estruturais, gerando *insights* mecanicistas para essa classe de enzimas. Apesar das reações Diels-Alder ocorrerem espontaneamente em água, nossos estudos evidenciam que a presença da enzima favorece a ocorrência da ciclização. Propomos que o mecanismo catalítico ocorra possivelmente

devido aos efeitos de proximidade e orientação, em que a proteína permite que ocorram conformações favoráveis de NAC (*Near Attack Conformations*) para o substrato. Comparativamente com seus ortólogos, há baixa identidade sequencial entre as enzimas e há a ausência de resíduos catalíticos conservados no sítio ativo, corroborando o fato de possível não envolvimento direto da enzima no mecanismo catalítico. Estudos de QM/MM ainda estão sendo realizados para a confirmação de nossa hipótese. Da mesma forma que para a Tsn15, estudos adicionais com a enzima Tmn8 da mesma via biossintética podem esclarecer detalhes da biossíntese das enzimas Diel-alderases.

Os resultados obtidos nesta tese de doutorado permitem a formulação de novas questões para futuros estudos tanto experimentais como computacionais.

## REFERÊNCIAS

1. Zhang A, Demain AL. Natural Products. *Drug Discovery and Therapeutical Medicine*–2005–382 p. 2005.
2. Schmidt B, Ribnicky DM, Poulev A, Logendra S, Cefalu WT, Raskin I. A natural history of botanical therapeutics. *Metabolism*. 2008;57:S3-S9.
3. Rishton GM. Natural products as a robust source of new drugs and drug leads: past successes and present day issues. *The American journal of cardiology*. 2008;101(10):S43-S9.
4. Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *Journal of natural products*. 2012;75(3):311-35.
5. Waksman SA, Henrici AT. The nomenclature and classification of the actinomycetes. *Journal of bacteriology*. 1943;46(4):337.
6. Berdy J. Bioactive microbial metabolites. *The Journal of antibiotics*. 2005;58(1):1-26.
7. Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Klenk H-P, et al. Taxonomy, physiology, and natural products of Actinobacteria. *Microbiol Mol Biol Rev*. 2016;80(1):1-43.
8. Subramani R, Sipkema D. Marine rare actinomycetes: a promising source of structurally diverse and unique novel natural products. *Marine drugs*. 2019;17(5):249.
9. Smith JL, Skiniotis G, Sherman DH. Architecture of the polyketide synthase module: surprises from electron cryo-microscopy. *Current opinion in structural biology*. 2015;31:9-19.
10. Hur GH, Vickery CR, Burkart MD. Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Natural product reports*. 2012;29(10):1074-98.
11. Burkhart BJ, Hudson GA, Dunbar KL, Mitchell DA. A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nature chemical biology*. 2015;11(8):564-70.
12. Park SR, Park JW, Ban YH, Sohng JK, Yoon YJ. 2-Deoxystreptamine-containing aminoglycoside antibiotics: recent advances in the characterization and manipulation of their biosynthetic pathways. *Natural product reports*. 2013;30(1):11-20.
13. Walsh CT, Haynes SW, Ames BD. Aminobenzoates as building blocks for natural product assembly lines. *Natural product reports*. 2012;29(1):37-59.
14. Boettger D, Hertweck C. Molecular diversity sculpted by fungal PKS–NRPS hybrids. *ChemBioChem*. 2013;14(1):28-42.
15. Joule JA, Mills K. *Heterocyclic chemistry*: John Wiley & Sons; 2008.
16. Storm DR, Strominger JL. Complex formation between bacitracin peptides and isoprenyl pyrophosphates the specificity of lipid-peptide interactions. *Journal of Biological Chemistry*. 1973;248(11):3940-5.
17. Liu H-K, Sadler PJ. Metal complexes as DNA intercalators. *Accounts of chemical research*. 2011;44(5):349-59.

18. Pozharskii AF, Soldatenkov AT, Katritzky AR. Heterocycles in life and society: an introduction to heterocyclic chemistry, biochemistry and applications: John Wiley & Sons; 2011.
19. Walsh CT. Insights into the chemical logic and enzymatic machinery of NRPS assembly lines. *Natural product reports*. 2016;33(2):127-35.
20. Hemmerling F, Hahn F. Biosynthesis of oxygen and nitrogen-containing heterocycles in polyketides. *Beilstein journal of organic chemistry*. 2016;12(1):1512-50.
21. Tibrewal N, Tang Y. Biocatalysts for natural product biosynthesis. *Annual review of chemical and biomolecular engineering*. 2014;5:347-66.
22. Yim G, Kalan L, Koteva K, Thaker MN, Waglechner N, Tang I, et al. Harnessing the Synthetic Capabilities of Glycopeptide Antibiotic Tailoring Enzymes: Characterization of the UK-68,597 Biosynthetic Cluster. *ChemBioChem*. 2014;15(17):2613-23.
23. Nguyen H-C, Darbon E, Thai R, Pernodet J-L, Lautru S. Post-PKS tailoring steps of the spiramycin macrolactone ring in *Streptomyces ambofaciens*. *Antimicrobial agents and chemotherapy*. 2013;57(8):3836-42.
24. Ueberbacher BT, Hall M, Faber K. Electrophilic and nucleophilic enzymatic cascade reactions in biosynthesis. *Natural product reports*. 2012;29(3):337-50.
25. He J, Magarvey N, Pirae M, Vining L. The gene cluster for chloramphenicol biosynthesis in *Streptomyces venezuelae* ISP5230 includes novel shikimate pathway homologues and a monomodular non-ribosomal peptide synthetase gene. *Microbiology*. 2001;147(10):2817-29.
26. Fage CD, Isiorho EA, Liu Y, Wagner DT, Liu H-w, Keatinge-Clay AT. The structure of SpnF, a standalone enzyme that catalyzes [4+ 2] cycloaddition. *Nature chemical biology*. 2015;11(4):256-8.
27. Hotta K, Chen X, Paton RS, Minami A, Li H, Swaminathan K, et al. Enzymatic catalysis of anti-Baldwin ring closure in polyether biosynthesis. *Nature*. 2012;483(7389):355-8.
28. Luhavaya H, Dias MV, Williams SR, Hong H, de Oliveira LG, Leadlay PF. Enzymology of pyran ring A formation in salinomycin biosynthesis. *Angewandte Chemie*. 2015;127(46):13826-9.
29. Du L, Lou L. PKS and NRPS release mechanisms. *Natural product reports*. 2010;27(2):255-78.
30. Callaway T, Edrington T, Rychlik J, Genovese K, Poole T, Jung Y, et al. Ionophores: their use as ruminant growth promotants and impact on food safety. *Current issues in intestinal microbiology*. 2003;4(2):43-51.
31. Gumila C, Ancelin M-L, Delort A-M, Jeminet G, Vial HJ. Characterization of the potent in vitro and in vivo antimalarial activities of ionophore compounds. *Antimicrobial agents and chemotherapy*. 1997;41(3):523-9.

32. Nakamura M, Kunimoto S, Takahashi Y, Naganawa H, Sakaue M, Inoue S, et al. Inhibitory effects of polyethers on human immunodeficiency virus replication. *Antimicrobial agents and chemotherapy*. 1992;36(2):492-4.
33. Fuchs D, Heinold A, Opelz G, Daniel V, Naujokat C. Salinomycin induces apoptosis and overcomes apoptosis resistance in human cancer cells. *Biochemical and biophysical research communications*. 2009;390(3):743-9.
34. Liu T, Cane DE, Deng Z. The enzymology of polyether biosynthesis. *Methods in enzymology*. 2009;459:187-214.
35. Smith L, Hong H, Spencer JB, Leadlay PF. Analysis of specific mutants in the lasalocid gene cluster: evidence for enzymatic catalysis of a disfavoured polyether ring closure. *ChemBioChem*. 2008;9(18):2967-75.
36. Oliynyk M, Stark CB, Bhatt A, Jones MA, Hughes-Thomas ZA, Wilkinson C, et al. Analysis of the biosynthetic gene cluster for the polyether antibiotic monensin in *Streptomyces cinnamonensis* and evidence for the role of monB and monC genes in oxidative cyclization. *Molecular microbiology*. 2003;49(5):1179-90.
37. Demydchuk Y, Sun Y, Hong H, Staunton J, Spencer JB, Leadlay PF. Analysis of the tetronomycin gene cluster: insights into the biosynthesis of a polyether tetronate antibiotic. *ChemBioChem*. 2008;9(7):1136-45.
38. Janecki T. *Natural Lactones and Lactams: Synthesis, Occurrence and Biological Activity*: John Wiley & Sons; 2013.
39. Vieweg L, Reichau S, Schobert R, Leadlay PF, Süßmuth RD. Recent advances in the field of bioactive tetronates. *Natural product reports*. 2014;31(11):1554-84.
40. Lacoske MH, Theodorakis EA. Spirotetronate polyketides as leads in drug discovery. *Journal of natural products*. 2014;78(3):562-75.
41. Sun Y, Hahn F, Demydchuk Y, Chettle J, Tosin M, Osada H, et al. In vitro reconstruction of tetronate RK-682 biosynthesis. *Nature chemical biology*. 2010;6(2):99-101.
42. Diels O, Alder K. Synthesen in der hydroaromatischen Reihe. *Justus Liebigs Annalen der Chemie*. 1928;460(1):98-122.
43. Gottardi EM, Krawczyk JM, von Suchodoletz H, Schadt S, Mühlenweg A, Uguru GC, et al. Abyssomicin Biosynthesis: Formation of an Unusual Polyketide, Antibiotic-Feeding Studies and Genetic Analysis. *ChemBioChem*. 2011;12(9):1401-10.
44. He H-Y, Pan H-X, Wu L-F, Zhang B-B, Chai H-B, Liu W, et al. Quartromicin biosynthesis: two alternative polyketide chains produced by one polyketide synthase assembly line. *Chemistry & biology*. 2012;19(10):1313-23.
45. Jia X-Y, Tian Z-H, Shao L, Qu X-D, Zhao Q-F, Tang J, et al. Genetic characterization of the chlorothricin gene cluster as a model for spirotetronate antibiotic biosynthesis. *Chemistry & biology*. 2006;13(6):575-85.

46. Zhang H, White-Phillip JA, Melançon CE, Kwon H-j, Yu W-l, Liu H-w. Elucidation of the kijanimicin gene cluster: insights into the biosynthesis of spirotetronate antibiotics and nitrosugars. *Journal of the American Chemical Society*. 2007;129(47):14670-83.
47. Auclair K, Sutherland A, Kennedy J, Witter DJ, Van den Heever JP, Hutchinson CR, et al. Lovastatin nonaketide synthase catalyzes an intramolecular Diels–Alder reaction of a substrate analogue. *Journal of the American Chemical Society*. 2000;122(46):11519-20.
48. Oikawa H, Suzuki Y, Naya A, Katayama K, Ichihara A. First direct evidence in biological Diels-Alder reaction of incorporation of diene-dienophile precursors in the biosynthesis of solanapyrones. *Journal of the American Chemical Society*. 1994;116(8):3605-6.
49. Park H-R, Furihata K, Hayakawa Y, Shin-ya K. Versipelostatin, a novel GRP78/Bip molecular chaperone down-regulator of microbial origin. *Tetrahedron letters*. 2002;43(39):6941-5.
50. Hashimoto T, Hashimoto J, Teruya K, Hirano T, Shin-ya K, Ikeda H, et al. Biosynthesis of versipelostatin: identification of an enzyme-catalyzed [4+ 2]-cycloaddition required for macrocyclization of spirotetronate-containing polyketides. *Journal of the American Chemical Society*. 2015;137(2):572-5.
51. Keller-Juslen C, King HD, Kuhn M, Loosli H-R, Pache W, Petcher TJ, et al. Tetronomycin, a novel polyether of unusual structure. *The Journal of antibiotics*. 1982;35(2):142-50.
52. Lv M, Zhao J, Deng Z, Yu Y. Characterization of the biosynthetic gene cluster for benzoxazole antibiotics A33853 reveals unusual assembly logic. *Chemistry & biology*. 2015;22(10):1313-24.
53. Sommer PS, Almeida RC, Schneider K, Beil W, Süssmuth RD, Fiedler H-P. Nataxazole, a New Benzoxazole Derivative with Antitumor Activity Produced by *Streptomyces* sp. Tü 6176 [dagger]. *Journal of antibiotics*. 2008;61(11):683.
54. Reed PW, Lardy HA. A23187: a divalent cation ionophore. *Journal of Biological Chemistry*. 1972;247(21):6970-7.
55. Hohmann C, Schneider K, Bruntner C, Irran E, Nicholson G, Bull AT, et al. Caboxamycin, a new antibiotic of the benzoxazole family produced by the deep-sea strain *Streptomyces* sp. NTK 937. *The Journal of antibiotics*. 2009;62(2):99-104.
56. Shibata K, Kashiwada M, Ueki M, Taniguchi M. UK-1, a novel cytotoxic metabolite from *Streptomyces* sp. 517-02. II. Structural elucidation. *The Journal of antibiotics*. 1993;46(7):1095-100.
57. Sato S, Kajiura T, Noguchi M, Takehana K, Kobayashi T, Tsuji T. AJI9561, a new cytotoxic benzoxazole derivative produced by *Streptomyces* sp. *The Journal of antibiotics*. 2001;54(1):102-4.
58. Wu Q, Liang J, Lin S, Zhou X, Bai L, Deng Z, et al. Characterization of the biosynthesis gene cluster for the pyrrole polyether antibiotic calcimycin (A23187) in *Streptomyces chartreusis* NRRL 3882. *Antimicrobial agents and chemotherapy*. 2011;55(3):974-82.

59. Seibert CM, Raushel FM. Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry*. 2005;44(17):6383-91.
60. Cano-Prieto C, García-Salcedo R, Sánchez-Hidalgo M, Braña AF, Fiedler HP, Méndez C, et al. Genome Mining of *Streptomyces* sp. Tü 6176: characterization of the nataxazole biosynthesis pathway. *ChemBioChem*. 2015;16(10):1461-73.
61. Maluf FVM, J. R. C.; Oliva, G.; Guido, R. V. C. *Cristalografia de Proteínas*. In: SBBq SBdBeBM-, editor. *Bioinformática: da Biologia à Flexibilidade Molecular*. São Paulo2014.
62. Nelson DL, Cox MM. *Lehninger princípios de bioquímica*. 4 ed2006.
63. Rücker E, Schneider G, Steinhäuser K, Löwer R, Hauber J, Stauber RH. Rapid evaluation and optimization of recombinant protein production using GFP tagging. *Protein Expression and Purification*. 2001;21(1):220-3.
64. Sahdev S, Khattar SK, Saini KS. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and cellular biochemistry*. 2008;307(1-2):249-64.
65. Sambrook J, Russell DW. *Molecular cloning: a laboratory manual 3rd edition*. Coldspring-Harbour Laboratory Press, UK. 2001.
66. McPherson A, Gavira JA. Introduction to protein crystallization. *Acta Crystallographica Section F: Structural Biology Communications*. 2014;70(1):2-20.
67. Deller MC, Kong L, Rupp B. Protein stability: a crystallographer's perspective. *Acta Crystallographica Section F: Structural Biology Communications*. 2016;72(2):72-95.
68. Forneris F, Orru R, Bonivento D, Chiarelli LR, Mattevi A. ThermoFAD, a ThermoFluor®-adapted flavin ad hoc detection system for protein folding and ligand binding. *FEBS journal*. 2009;276(10):2833-40.
69. McCammon J. Protein dynamics. *Reports on Progress in Physics*. 1984;47(1):1.
70. Verli H. Dinâmica Molecular. In: SBBq SBdBeBM-, editor. *Bioinformática: da Biologia à Flexibilidade Molecular*. São Paulo2014.
71. Morgon NH, Coutinho KR. *Métodos de química teórica e modelagem molecular*. 2007.
72. Levitt M, Warshel A. Computer simulation of protein folding. *Nature*. 1975;253(5494):694-8.
73. McCammon J, Karplus M. Internal motions of antibody molecules. *Nature*. 1977;268(5622):765-6.
74. Case DA, Karplus M. Dynamics of ligand binding to heme proteins. *Journal of molecular biology*. 1979;132(3):343-68.
75. Rapaport DC, Rapaport DCR. *The art of molecular dynamics simulation*: Cambridge university press; 2004.

76. Frenkel D, Smit B. Understanding molecular simulation: from algorithms to applications: Elsevier; 2001.
77. Tildesley D, Allen M. Computer simulation of liquids. Clarendon. Oxford; 1987.
78. Hünenberger P, van Gunsteren W. Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications. Kluwer Academic Publisher: Dordrecht, the Netherlands; 1997.
79. Leach AR. Molecular modelling: principles and applications: Pearson education; 2001.
80. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *The Journal of Physical Chemistry B*. 2001;105(28):6474-87.
81. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*. 1996;118(45):11225-36.
82. Schuler LD, Daura X, Van Gunsteren WF. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *Journal of Computational Chemistry*. 2001;22(11):1205-18.
83. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan Sa, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of computational chemistry*. 1983;4(2):187-217.
84. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham III TE, DeBolt S, et al. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Computer Physics Communications*. 1995;91(1-3):1-41.
85. Lewis K. Recover the lost art of drug discovery. *Nature*. 2012;485(7399):439-40.
86. Lyddiard D, Jones GL, Greatrex BW. Keeping it simple: lessons from the golden era of antibiotic discovery. *FEMS microbiology letters*. 2016;363(8).
87. Zaman SB, Hussain MA, Nye R, Mehta V, Mamun KT, Hossain N. A review on antibiotic resistance: alarm bells are ringing. *Cureus*. 2017;9(6).
88. Davies J, Davies D. Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev*. 2010;74(3):417-33.
89. Dunbar KL, Chekan JR, Cox CL, Burkhart BJ, Nair SK, Mitchell DA. Discovery of a new ATP-binding motif involved in peptidic azoline biosynthesis. *Nature chemical biology*. 2014;10(10):823-9.
90. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExpASy server: Springer; 2005.



91. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research*. 2003;31(13):3784-8.
92. Reinhard L, Mayerhofer H, Geerlof A, Mueller-Dieckmann J, Weiss MS. Optimization of protein buffer cocktails using Thermofluor. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*. 2013;69(2):209-14.
93. Kabsch W. Xds. *Acta Crystallographica Section D: Biological Crystallography*. 2010;66(2):125-32.
94. DeLano WL. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography*. 2002;40:82-92.
95. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*. 2004;60(12):2126-32.
96. Evans P. Scaling and assessment of data quality. *Acta Crystallographica Section D: Biological Crystallography*. 2006;62(1):72-82.
97. Collaborative CP. The CCP4 suite: programs for protein crystallography. *Acta crystallographica Section D, Biological crystallography*. 1994;50(Pt 5):760.
98. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*. 2010;66(2):213-21.
99. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*. 2011;67(4):355-67.
100. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D: Biological Crystallography*. 1997;53(3):240-55.
101. Gordon JC, Myers JB, Folta T, Shoja V, Heath LS, Onufriev A. H++: a server for estimating p K as and adding missing hydrogens to macromolecules. *Nucleic acids research*. 2005;33(suppl\_2):W368-W71.
102. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic acids research*. 2004;32(suppl\_2):W665-W7.
103. Martínez-Rosell G, Giorgino T, De Fabritiis G. PlayMolecule ProteinPrepare: a web application for protein preparation for molecular dynamics simulations. *Journal of chemical information and modeling*. 2017;57(7):1511-6.
104. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry*. 2008;29(11):1859-65.
105. Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul JA, et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM

simulations using the CHARMM36 additive force field. *Journal of chemical theory and computation*. 2015;12(1):405-13.

106. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1:19-25.

107. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods*. 2017;14(1):71.

108. Freddolino PL, Gardner KH, Schulten K. Signaling mechanisms of LOV domains: new insights from molecular dynamics studies. *Photochemical & Photobiological Sciences*. 2013;12(7):1158-70.

109. Vanommeslaeghe K, Raman EP, MacKerell Jr AD. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *Journal of chemical information and modeling*. 2012;52(12):3155-68.

110. Vanommeslaeghe K, MacKerell Jr AD. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *Journal of chemical information and modeling*. 2012;52(12):3144-54.

111. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*. 1983;79(2):926-35.

112. MacKerell Jr AD, Bashford D, Bellott M, Dunbrack Jr RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B*. 1998;102(18):3586-616.

113. Feller SE, Zhang Y, Pastor RW, Brooks BR. Constant pressure molecular dynamics simulation: the Langevin piston method. *The Journal of chemical physics*. 1995;103(11):4613-21.

114. Martyna GJ, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. *The Journal of chemical physics*. 1994;101(5):4177-89.

115. Darden T, York D, Pedersen L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *The Journal of chemical physics*. 1993;98(12):10089-92.

116. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*. 2010;31(2):455-61.

117. O'Boyle NM, Morley C, Hutchison GR. Pybel: a Python wrapper for the OpenBabel cheminformatics toolkit. *Chemistry Central Journal*. 2008;2(1):5.

118. Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron*. 1980;36(22):3219-28.

119. Bietz S, Urbaczek S, Schulz B, Rarey M. Protoss: a holistic approach to predict tautomers and protonation states in protein-ligand complexes. *Journal of cheminformatics*. 2014;6(1):12.
120. Hisanaga Y, Ago H, Nakagawa N, Hamada K, Ida K, Yamamoto M, et al. Structural basis of the substrate specific two-step catalysis of long chain fatty acyl-CoA synthetase dimer. *Journal of Biological Chemistry*. 2004.
121. Chang K-H, Xiang H, Dunaway-Mariano D. Acyl-adenylate motif of the acyl-adenylate/thioester-forming enzyme superfamily: a site-directed mutagenesis study with the *Pseudomonas* sp. strain CBS3 4-chlorobenzoate: coenzyme A ligase. *Biochemistry*. 1997;36(50):15650-9.
122. Conti E, Franks NP, Brick P. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure*. 1996;4(3):287-98.
123. Day JC, Goodall TI, Bailey MJ. The evolution of the adenylate-forming protein family in beetles: multiple luciferase gene paralogues in fireflies and glow-worms. *Molecular phylogenetics and evolution*. 2009;50(1):93-101.
124. Gulick AM. Conformational Dynamics in the Acyl-CoA Synthetases, Adenylation Domains of Non-ribosomal Peptide Synthetases, and Firefly Luciferase. *ACS Chemical Biology*. 2009;4(10):811-27.
125. Cusack S. Aminoacyl-tRNA synthetases. *Current opinion in structural biology*. 1997;7(6):881-9.
126. Stanišić A, Kries H. Adenylation Domains in Nonribosomal Peptide Engineering. *ChemBioChem*. 2019;20(11):1347-56.
127. Kudo F, Miyanaga A, Eguchi T. Structural basis of the nonribosomal codes for nonproteinogenic amino acid selective adenylation enzymes in the biosynthesis of natural products. *Journal of industrial microbiology & biotechnology*. 2019;46(3-4):515-36.
128. Labby KJ, Watsula SG, Garneau-Tsodikova S. Interrupted adenylation domains: unique bifunctional enzymes involved in nonribosomal peptide biosynthesis. *Natural product reports*. 2015;32(5):641-53.
129. Huang W, Xu H, Li Y, Zhang F, Chen X-Y, He Q-L, et al. Characterization of yatakemycin gene cluster revealing a radical S-adenosylmethionine dependent methyltransferase and highlighting spirocyclopropane biosynthesis. *Journal of the American Chemical Society*. 2012;134(21):8831-40.
130. Law A, Boulanger MJ. Defining a structural and kinetic rationale for paralogous copies of phenylacetate-CoA ligases from the cystic fibrosis pathogen *Burkholderia cenocepacia* J2315. *Journal of Biological Chemistry*. 2011;286(17):15577-85.
131. Zahn M, Kurteva-Yaneva N, Schuster J, Krug U, Georgi T, Müller RH, et al. Structures of 2-hydroxyisobutyric acid-CoA ligase reveal determinants of substrate specificity and describe a multi-conformational catalytic cycle. *Journal of molecular biology*. 2019;431(15):2747-61.

132. Schmelz S, Naismith JH. Adenylate-forming enzymes. *Current opinion in structural biology*. 2009;19(6):666-71.
133. D'Ambrosio HK, Derbyshire ER. Investigating the Role of Class I Adenylate-Forming Enzymes in Natural Product Biosynthesis. *ACS Chemical Biology*. 2019.
134. Estrada P, Manandhar M, Dong S-H, Deveryshetty J, Agarwal V, Cronan JE, et al. The pimeloyl-CoA synthetase BioW defines a new fold for adenylate-forming enzymes. *Nature chemical biology*. 2017;13(6):668.
135. Lux MC, Standke LC, Tan DS. Targeting adenylate-forming enzymes with designed sulfonfyladenosine inhibitors. *The Journal of antibiotics*. 2019;72(6):325-49.
136. P Duckworth B, M Nelson K, C Aldrich C. Adenylating enzymes in *Mycobacterium tuberculosis* as drug targets. *Current topics in medicinal chemistry*. 2012;12(7):766-96.
137. Reger AS, Wu R, Dunaway-Mariano D, Gulick AM. Structural characterization of a 140 domain movement in the two-step reaction catalyzed by 4-chlorobenzoate: CoA ligase. *Biochemistry*. 2008;47(31):8016-25.
138. Bandarian V, Patridge KA, Lennon BW, Huddler DP, Matthews RG, Ludwig ML. Domain alternation switches B 12-dependent methionine synthase to the activation conformation. *Nature structural biology*. 2002;9(1):53-6.
139. Losada AA, Cano-Prieto C, García-Salcedo R, Braña AF, Méndez C, Salas JA, et al. Caboxamycin biosynthesis pathway and identification of novel benzoxazoles produced by cross-talk in *Streptomyces* sp. NTK 937. *Microbial biotechnology*. 2017;10(4):873-85.
140. Wu Q, Liang J, Lin S, Zhou X, Bai L, Deng Z, et al. Characterization of the Biosynthesis Gene Cluster for the Pyrrole Polyether Antibiotic Calcimycin (A23187) in *Streptomyces chartreusis* NRRL 3882. *Antimicrobial Agents and Chemotherapy*. 2011;55(3):974-82.
141. Song H, Rao C, Deng Z, Yi Y, Naismith JH. The biosynthetic pathway of the benzoxazole in nataxazole proceeds via an unstable ester and has synthetic utility. *Angewandte Chemie International Edition*. 2020.
142. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology*. 2007;372(3):774-97.
143. Branden CI, Tooze J. *Introduction to Protein Structure*: CRC Press; 2012.
144. Laskowski RA, Jabłońska J, Pravda L, Vařeková RS, Thornton JM. PDBsum: Structural summaries of PDB entries. *Protein Science*. 2018;27(1):129-34.
145. Marahiel MA, Stachelhaus T, Mootz HD. Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chemical reviews*. 1997;97(7):2651-74.
146. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*. 2011;7:539.

147. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* (Oxford, England). 2009;25(9):1189-91.
148. Laitaoja M, Valjakka J, Jänis J. Zinc coordination spheres in protein structures. *Inorganic chemistry*. 2013;52(19):10983-91.
149. Alberts IL, Nadassy K, Wodak SJ. Analysis of zinc binding sites in protein crystal structures. *Protein Science*. 1998;7(8):1700-16.
150. Touw WG, van Beusekom B, Evers JM, Vriend G, Joosten RP. Validation and correction of Zn–CysxHisy complexes. *Acta Crystallographica Section D: Structural Biology*. 2016;72(10):1110-8.
151. Pace N, Weerapana E. Zinc-binding cysteines: diverse functions and structural motifs. *Biomolecules*. 2014;4(2):419-34.
152. Laskowski RA, Swindells MB. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. ACS Publications; 2011.
153. Schiering N, Casale E, Caccia P, Giordano P, Battistini C. Dimer formation through domain swapping in the crystal structure of the Grb2-SH2– Ac-pYVNV complex. *Biochemistry*. 2000;39(44):13376-82.
154. Yang S, Cho SS, Levy Y, Cheung MS, Levine H, Wolynes PG, et al. Domain swapping is a consequence of minimal frustration. *Proceedings of the National Academy of Sciences*. 2004;101(38):13786-91.
155. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic acids research*. 2016;44(W1):W344-W50.
156. Shao X, Cao H-Y, Zhao F, Peng M, Wang P, Li C-Y, et al. Mechanistic insight into 3-methylmercaptopropionate metabolism and kinetical regulation of demethylation pathway in marine dimethylsulfoniopropionate-catabolizing bacteria. *Molecular Microbiology*. 2019;111(4):1057-73.
157. Nakatsu T, Ichiyama S, Hiratake J, Saldanha A, Kobashi N, Sakata K, et al. Structural basis for the spectral difference in luciferase bioluminescence. *Nature*. 2006;440(7082):372-6.
158. Ziemert N, Jensen PR. Phylogenetic approaches to natural product structure prediction. *Methods in enzymology*. 517: Elsevier; 2012. p. 161-82.
159. Benevides LdJ, Carvalho DSd, Andrade RFS, Bomfim GC, Fernandes FMdC. Evolutionary analysis of apolipoprotein E by Maximum Likelihood and complex network methods. *Genetics and molecular biology*. 2016;39(4):665-73.
160. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*. 2018;35(6):1547-9.

161. Jeon B-s, Wang S-A, Rusczycky MW, Liu H-w. Natural [4+ 2]-Cyclases. *Chemical reviews*. 2017;117(8):5367-88.
162. Byrne MJ, Lees NR, Han L-C, van der Kamp MW, Mulholland AJ, Stach JE, et al. The catalytic mechanism of a natural Diels–Alderase revealed in molecular detail. *Journal of the American Chemical Society*. 2016;138(19):6095-8.
163. Zheng Q, Guo Y, Yang L, Zhao Z, Wu Z, Zhang H, et al. Enzyme-dependent [4+ 2] cycloaddition depends on lid-like interaction of the N-terminal sequence with the catalytic core in PyrI4. *Cell chemical biology*. 2016;23(3):352-60.
164. Hofmann E, Zerbe P, Schaller F. The crystal structure of *Arabidopsis thaliana* allene oxide cyclase: insights into the oxylipin cyclization reaction. *The Plant Cell*. 2006;18(11):3201-17.
165. Yoeun S, Cho K, Han O. Structural evidence for the substrate channeling of rice allene oxide cyclase in biologically analogous Nazarov reaction. *Frontiers in chemistry*. 2018;6:500.
166. Costa KC, Glasser NR, Conway SJ, Newman DK. Pyocyanin degradation by a tautomerizing demethylase inhibits *Pseudomonas aeruginosa* biofilms. *Science*. 2017;355(6321):170-3.
167. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic acids research*. 2018.
168. Zheng Q, Gong Y, Guo Y, Zhao Z, Wu Z, Zhou Z, et al. Structural Insights into a Flavin-Dependent [4+ 2] Cyclase that Catalyzes trans-Decalin Formation in Pyrroindomycin Biosynthesis. *Cell chemical biology*. 2018.
169. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic acids research*. 2018;46(W1):W363-W7.
170. Holm L, Sander C. Searching protein structure databases has come of age. *Proteins: Structure, Function, and Bioinformatics*. 1994;19(3):165-73.
171. Zheng Q, Gong Y, Guo Y, Zhao Z, Wu Z, Zhou Z, et al. Structural insights into a flavin-dependent [4+ 2] cyclase that catalyzes trans-decalin formation in pyrroindomycin biosynthesis. *Cell chemical biology*. 2018;25(6):718-27. e3.
172. Liu L-K, Abdelwahab H, Del Campo JSM, Mehra-Chaudhary R, Sobrado P, Tanner JJ. The structure of the antibiotic deactivating, N-hydroxylating rifampicin monooxygenase. *Journal of Biological Chemistry*. 2016;291(41):21553-62.
173. Koteva K, Cox G, Kelso JK, Surette MD, Zubyk HL, Ejim L, et al. Rox, a rifamycin resistance enzyme with an unprecedented mechanism of action. *Cell chemical biology*. 2018;25(4):403-12. e5.
174. Wang P, Bashiri G, Gao X, Sawaya MR, Tang Y. Uncovering the enzymes that catalyze the final steps in oxytetracycline biosynthesis. *Journal of the American Chemical Society*. 2013;135(19):7138-41.

175. Wang W, Li J, Li H, Fan K, Liu Y. Crystal structure of AlpK: An essential monooxygenase involved in the biosynthesis of kinamycin. *Biochemical and biophysical research communications*. 2019;510(4):601-5.
176. Koskiniemi H, Metsä-Ketelä M, Dobritsch D, Kallio P, Korhonen H, Mäntsälä P, et al. Crystal structures of two aromatic hydroxylases involved in the early tailoring steps of angucycline biosynthesis. *Journal of molecular biology*. 2007;372(3):633-48.
177. Gatti DL, Palfey BA, Lah MS, Entsch B, Massey V, Ballou DP, et al. The mobile flavin of 4-OH benzoate hydroxylase. *Science*. 1994;266(5182):110-4.
178. Brender JR, Dertouzos J, Ballou DP, Massey V, Palfey BA, Entsch B, et al. Conformational dynamics of the isoalloxazine in substrate-free p-hydroxybenzoate hydroxylase: single-molecule studies. *Journal of the American Chemical Society*. 2005;127(51):18171-8.
179. Bosserman MA, Downey T, Noinaj N, Buchanan SK, Rohr Jr. Molecular insight into substrate recognition and catalysis of Baeyer–Villiger monooxygenase MtmOIV, the key frame-modifying enzyme in the biosynthesis of anticancer agent mithramycin. *ACS chemical biology*. 2013;8(11):2466-77.
180. Koehn FE, Carter GT. The evolving role of natural products in drug discovery. *Nature reviews Drug discovery*. 2005;4(3):206-20.

## APÊNDICES A – Artigo: Crystal structure of AjiA1 shows a novel structural motion mechanism into the adenylate forming enzyme family

002  
003  
004  
005  
006  
007  
008  
009  
010  
011  
012  
013  
014  
015  
016  
017  
018  
019  
020  
021  
022  
023  
024  
025  
026  
027  
028  
029  
030  
031  
032  
033  
034  
035  
036  
037  
038  
039  
040  
041  
042  
043  
044  
045  
046  
047  
048  
049  
050  
051  
052  
053  
054  
055  
056  
057  
058  
059  
060  
061  
062  
063  
064  
065  
066  
067  
068  
069  
070  
071  
072  
073  
074  
075  
076

Page 1

IUCrJ

BIOLOGY | MEDICINE research papers

### Crystal structure of AjiA1 shows a novel structural motion mechanism into the adenylate forming enzyme family

Authors

Fernanda C. R. de Paiva<sup>a</sup>, Karen Chan<sup>b</sup>, Markiyana Samborsky<sup>b</sup>, Ariel M. Silber<sup>c</sup>, Peter F. Leadlay<sup>b</sup> and Marcio V. B. Dias<sup>ad\*</sup>

<sup>a</sup>Department of Microbiology, Institute of Biomedical Sciences II, University of São Paulo, São Paulo, Brazil

<sup>b</sup>Department of Biochemistry, University of Cambridge, Cambridge, UK

<sup>c</sup>LaBTryps – Laboratory of Biochemistry of Tryps. Department of Parasitology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil

<sup>d</sup>Department of Chemistry, University of Warwick, Coventry, UK

Correspondence email: mvbdias@usp.br

**Funding information** Fundação de Amparo à Pesquisa do Estado de São Paulo (grant No. 2018/00351-1 ; grant No. 2016/06034-2); Conselho Nacional de Desenvolvimento Científico e Tecnológico (grant No. 308351/2013-4; grant No. 404769/2018-7 ; grant No. 302848/2017-7 ; studentship No. 141090/2016-2).

**Synopsis** AjiA1 is an Adenylate-forming enzyme (AFE) member that catalyzes the condensation of two molecules of 3-hydroxyanthranilic acid (3-HAA) using ATP as co-substrate. The structure of AjiA1 in the apo form was solved and revealed key conformational changes, including an unusual loop swapping, reason that we suggest that it should be classified into a new subgroup in AFE members.

**Abstract** Adenylate-forming enzymes (AFEs) are a mechanistic superfamily of proteins that are involved in a large number of cellular roles. In the biosynthesis of benzoxazole antibiotics, an AFE member has been reported to have a key role in the condensation of cyclic molecules. In the gene cluster for AJI9561 benzoxazole, AjiA1 catalyzes the condensation of two molecules of 3-hydroxyanthranilic acid (3-HAA) using ATP as co-substrate. Herein, we report the enzymatic activity and structural analysis of AjiA1 in its apo form. The structure of AjiA1 was solved at 2.0 Å resolution and shows a conserved fold with other AFEs family members. AjiA1 exhibits activity in the presence of the 3-HAA ( $K_m = 77.86 \pm 28.36$  and  $K_{cat} = 0.04 \pm 0.004$ ) and also for the alternative substrate 3-hydroxybenzoic acid (3-HBA) ( $K_m = 22.12 \pm 31.35$  and  $K_{cat} = 0.08 \pm 0.005$ ). The structure of AjiA1 in the apo form also reveals key conformational changes that occur during the catalytic cycle of this

IMPORTANT: this document contains embedded data - to preserve data integrity, please ensure where possible that the IUCr Word tools (available from <http://journals.iucr.org/services/docxtemplate/>) are installed when editing this document.



002  
003  
004  
005  
006  
007  
008  
009  
010  
011  
012  
013  
014  
015  
016  
017  
018  
019  
020  
021  
022  
023  
024  
025  
026  
027  
028  
029  
030  
031  
032  
033  
034  
035  
036  
037  
038  
039  
040  
041  
042  
043  
044  
045  
046  
047  
048  
049  
050  
051  
052  
053  
054  
055  
056  
057  
058  
059  
060  
061  
062  
063  
064  
065  
066  
067  
068  
069  
070  
071  
072  
073  
074  
075  
076

enzyme, which has not been described for other AFE members and consequently the results showed here provide insights into the knowledge of this protein family and propose a new subgroup for those enzymes involved in the benzoxazole ring formation.

**Keywords:** ATP-dependent CoA ligase; Benzoxazole antibiotic; Loop swapping.

## 1. Introduction

The adenylation reaction catalyzed by the Adenylate-Forming Enzymes (AFEs) is ubiquitous and play crucial metabolic or catabolic roles in all organisms, including fatty acid metabolism (Hisanaga *et al.*, 2004), detoxification (Chang *et al.*, 1997), bioluminescence (Conti *et al.*, 1996, Day *et al.*, 2009, Gulick, 2009), biosynthesis of amino acids and proteins (Cusack, 1997) and natural products (Stanišić & Kries, 2019, Kudo *et al.*, 2019, Labby *et al.*, 2015, Huang *et al.*, 2012, Lv *et al.*, 2015). AFEs members are classified into three major groups: class I or AFN family, which includes the NRPS (Non-Ribosomal Peptide Synthetases) adenylation domains (Subclass Ia), Acyl or aryl CoA synthetases (subclass Ib) and oxidoreductases or firefly luciferases (subclass Ic). Several members of class I AFEs such as PaaK1 and PaaK2, involved in the degradation of phenylacetic acid in *Burkholderia cenocepacia* (Law & Boulanger, 2011) and 2-HIBA(Hydroxyisobutyric Acid)-CoA ligase (HCL) from *Aquicola tertiarycarbonis*, involved the degradation of oxygenate methyl *tert*-butyl ether (MTBE) (Zahn *et al.*, 2019) are not suitable for this classification and are considered members of an alternative subgroup of class I AFEs; Class II, which includes a number of aminoacyl-tRNA synthetases; and class III, which includes NRPS independent siderophore (NIS) synthetases (Schmelz & Naismith, 2009, D'Ambrosio & Derbyshire, 2019). However, a number of AFEs members are not included in this superfamily and they are only described to have a similar adenylation reaction (Gulick, 2009) such as the pimeloyl-CoA synthase BioW, which has a completely different fold and should belong to a new class of enzymes (Estrada *et al.*, 2017).

Although AFEs members have divergences in structural features and functions, they all catalyze the same general biochemical transformation (Lux *et al.*, 2019, D'Ambrosio & Derbyshire, 2019). AFEs usually catalyze its reaction in two steps: firstly occurs the ATP-dependent activation of a carboxylic acid substrates as a reactive acyl adenylate (acyl-AMP) intermediate followed by the binding of the enzyme with an acceptor molecule and the consecutive transfer of the acyl adenylate to a nucleophilic oxygen, sulfur or nitrogen atom generating an ester, thioester, and amide products (Lux *et al.*, 2019, D'Ambrosio & Derbyshire, 2019, P Duckworth *et al.*, 2012)

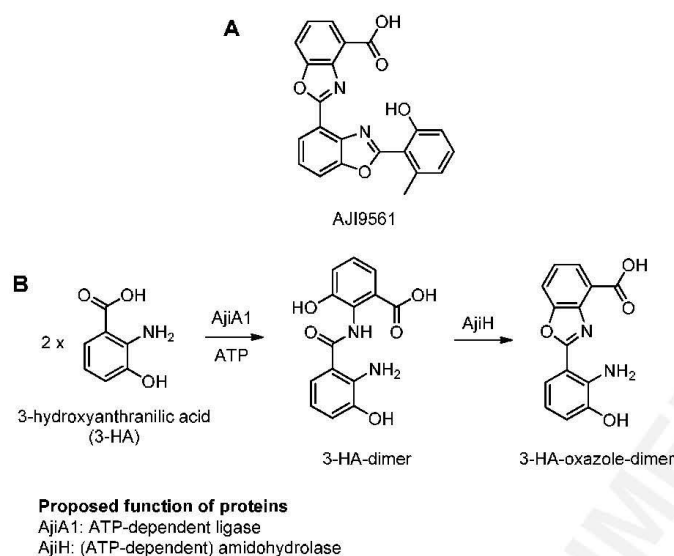
In class I AFEs, mechanistically, the ATP-dependent activation is via acyl-CoA synthetases and the second partial reaction commonly forms a thioester. Structurally, the active site is formed in the interface between two domains, which are connected by a flexible hinge (Schmelz & Naismith, 2009,

D'Ambrosio & Derbyshire, 2019). To perform the adenylation reaction, members of this class of AFEs undergo extensive conformational changes during the second half of the reaction (Lux *et al.*, 2019). Generally, these enzymes use a domain rotation to introduce opposing faces of the dynamic C terminal domain for the active site reactions (Gulick, 2009). Hence, two distinctly conformations are observed in AFEs, in which the C terminal domain orientation differs from each other in about 140°, and each conformation is involved in each half-reaction. Thus, despite the catalytic role, the C terminal domain is involved in the formation of the two distinct active site architectures, a phenomenon which is known as domain alternation (Reger *et al.*, 2008, Gulick, 2009, Bandarian *et al.*, 2002).

Although the AFEs from class I play a crucial role in the biosynthesis of natural products, particularly, nonribosomal peptides from NRPSs, which require an adenylation domain during incorporation and extension of the peptide chain, recently, a number of NRPS independent AFEs with modest sequence similarity to the phenylacetate-CoA ligases such as PaaK1 and HCL were identified in the biosynthesis of several natural products such as benzoxazoles, including nataxazole (Cano-Prieto *et al.*, 2015), caboxamycin (Losada *et al.*, 2017), compound A33853 (Lv *et al.*, 2015), calcimycin (Wu *et al.*, 2011) and compound AJI9561 (gene cluster not published). Particularly, compound AJI9561 isolated from *Streptomyces* sp. AJ9561 is a benzoxazole structurally similar to UK-1 and nataxazole and contains a backbone composed of two molecules of 3-HAA and a 6-methylsalicylic acid unit (6-MSA) (Sato *et al.*, 2001, Cano-Prieto *et al.*, 2015) (Fig. 1).

The gene clusters for oxazole-containing natural products indicate the participation of a series of unusual enzymes and the biosynthetic routes to this structurally unique class of natural products were underexplored until recently. The condensation of molecules and an acetalization reaction have been proposed to be essential for the generation of the benzoxazole rings and two enzymes have been proved to participate in this process, including an ATP dependent coenzyme A ligase and an amidohydrolase (Lv *et al.*, 2015).

In the biosynthesis of the compound A33853, the ATP-dependent coenzyme-A ligase BomJ was proved to catalyze the adenylation of a carboxylate substrate to its corresponding thiolated form. The activation of the precursor carboxylic group for the subsequent formation of an amide bond occurs through a nucleophilic attack of a 3-HAA amino group to the acyl group of an AMP-linked 3-HAA to produce a 3-HAA-dimer. This is the further substrate for the amidohydrolase BomN, which is involved in the assembly of the benzoxazole ring through the closure of the oxazole moiety. Although the gene cluster for AJI9561 has not been yet published (to be published somewhere else), an AFE member, herein denominated as AjiA1 was identified and it should play a similar role as BomJ in the formation of 3-HAA-dimer and a further amidohydrolase could complete the assembling of the benzoxazole ring (Fig. 1a) (Lv *et al.*, 2015).



**Figure 1** Chemical structure of Aji9561 and proposed reaction catalyzed by AjiA1 and the hypothetical amidohydrolase AjiH. A) Structure of Aji9561, which contains two benzoxazole rings and a unit derived from 6-MSA. B) AjiA1 conducts an ATP-dependent activation and the dimerization of 3-HAA and the hypothetical aminohydrolase AjiH performs the closure of the oxazole ring producing the benzoxazole ring.

We have investigated the structure of AjiA1 in its apo form and performed a structural analysis of this enzyme that shows key movements in the active site loops in comparison to the previously solved ortholog structures in complex with cofactor and substrate. Additionally, we have also used an in vitro approach to test the precursor 3-HAA and the alternative substrate 3-hydroxybenzoic acid (3-HBA).

## 2. METHODS

### 2.1. Heterologous expression in *E. coli* and purification of AjiA1

The *ajiA1* gene was cloned into pET28a in-frame with the N-terminal polyhistidine and this construct was used to transform chemically competent *E. coli* BL21 (DE3) cells (Thermo Fisher, USA). A single colony from each *E. coli* BL21 (DE3) transformant was cultured at 37 °C, 200 rpm in lysogeny-broth (LB) containing 50 µg/mL kanamycin. Once the culture had reached an  $A_{600} = 0.6$ , protein expression was induced by adding isopropyl β-D-1- thiogalactopyranoside (IPTG) to a final

concentration of 1 mM. The culture was further incubated for 18 h at 18 °C, 200 rpm. The cells were harvested by centrifugation (7000 g for 30 minutes at 4°C) and resuspended in 30 mL of 50 mM HEPES pH 8 buffer, 300 mM NaCl, 10% glycerol. Cells were lysed using an ultrasonic sonicator (BransonSonifier) at 30% amplitude, with a pulse every 1 second for 6 minutes. The insoluble fraction was removed by centrifugation (21700 g for 60 minutes at 4°C) and the soluble fraction was passed through a Ni-affinity column HisTrap HP 5 mL (GE Healthcare) to bind the recombinant protein. Lysis buffer containing increasing imidazole concentrations (25 – 500 mM) was used to remove contaminants and subsequently to elute the pure recombinant protein (monitored by SDS-PAGE). Fractions containing the AjiA1 was concentrated using an Amicon Ultra centrifugal filter units (Merck Millipore, USA). The protein concentration was determined by BioDrop  $\mu$ lite (Analítica). The protein was further purified using a preparative gel-filtration Hiload 16/60 column packed with Superdex 75 resin (GE Healthcare Lifescience, USA). For this purification, a buffer constituted by 100 mM Bis-Tris Propane pH 6.5 and 150 mM KCl was used. Fractions protein aliquots were pooled, concentrate, flash-frozen in liquid nitrogen and stored at -80 °C.

## 2.2. Protein crystallization

AjiA1 at 8 mg. mL<sup>-1</sup> was crystallized using a hanging-drop, vapor-diffusion method at 18 °C in a condition composed of 100 mM potassium phosphate citrate pH4.2 buffer and 35% PEG 300. Crystals were harvested from the crystallization drop and transferred into a cryoprotecting solution containing the crystallization solution and 15% ethylene glycol. Crystals were then flash-frozen in liquid nitrogen using nylon loops. X-ray diffraction data were collected at PETRA III-DESY, beamline 13, Hamburg, Germany. The X-ray data processing was performed using XDS (Kabsch, 2010) and scaled by Aimless (Evans & Murshudov, 2013) from the CCP4 suite (Winn *et al.*, 2011). The phase was determined by molecular replacement using the structure of PaaK1 (PDB: 2Y27) (Law & Boulanger, 2011) as the search model using the program Phaser (McCoy, 2007) from the Phenix suite (Adams *et al.*, 2011). The AjiA1 model was refined using Phenix.refine (Afonine *et al.*, 2012). Further visual inspection and real-space refinement were performed by COOT (Emsley & Cowtan, 2004). The stereochemical quality of the model was assessed using MolProbity (Chen *et al.*, 2010). The protein structure figures were prepared using PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

## 2.3. Enzymatic Assays

The enzymatic assays were performed by a coupled continuous spectrophotometric assay using the EnzChek™ Pyrophosphate Assay Kit (Molecular Probes). The initial reaction mixtures contained 0.2 mM MESG (2-amino-6-mercapto-7-methylpurine ribonucleoside), 1 U of purine nucleoside phosphorylase (PNP), 0.01 U of inorganic pyrophosphatase, 10 mM ATP, 20X reaction buffer, and AjiA1 (8  $\mu$ M) in a total volume of 100  $\mu$ L. The reaction mixtures were pre-incubated at 25 °C for 10

min and then initiated by the addition of the substrate in various concentrations. The increase of absorbance at 360 nm resulting from the phosphorolytic release of the purine base (2-amino-6-mercapto-7-methylpurine) from MESG was monitored over a 15 min period. The kinetic parameters  $K_m$  and  $V_{max}$  were calculated by fitting the initial rates to a nonlinear regression hyperbolic model under the assumption of a Michaelis Menten kinetics, using the GraphPad Prism 5.01 software. For all measurements, three independent experiments were performed, each one in triplicate.

#### 2.4. Alignment analysis and phylogeny

The protein sequences of AjiA1, and further 12 different AFEs members that have been structurally characterized were obtained in the Protein Data Bank (<https://www.rcsb.org/>) (Berman *et al.*, 2000). The alignments were performed using Clustal Omega (Sievers *et al.*, 2011) and imported to Jalview Version: 2.11.0 (Waterhouse *et al.*, 2009). The members of this enzyme family have been chosen based on the conserved sequence motifs in the ANL superfamily enzymes (Marahiel *et al.*, 1997, Gulick, 2009). The evolutionary history was inferred by using the Maximum Likelihood method and the JTT matrix-based model (Jones *et al.*, 1992). The tree with the highest log likelihood (-17500.85) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. Finally, the phylogenetic tree was constructed using the program MEGAX (Kumar *et al.*, 2018).

### 3. RESULTS and DISCUSSION

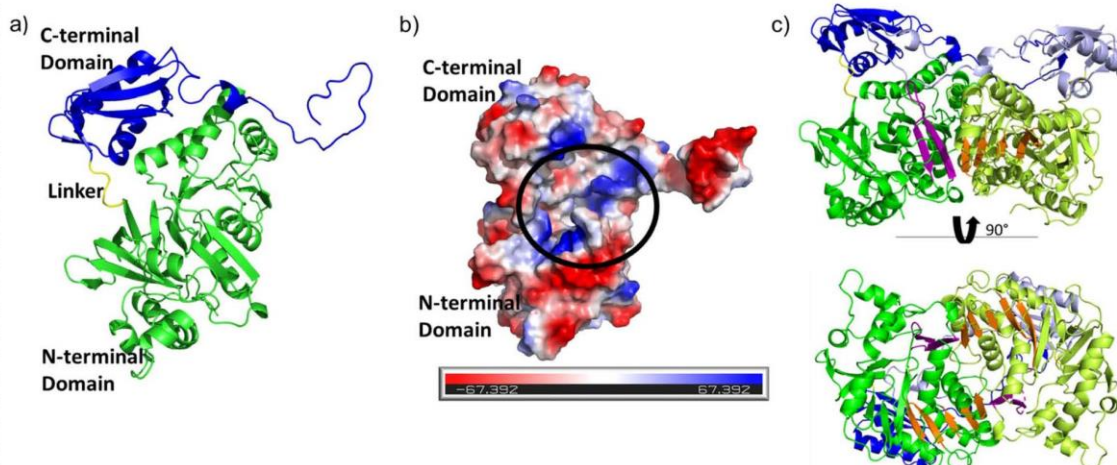
#### 3.1. Overall Structure of AjiA1

The structure of AjiA1 was determined at a resolution of 2.0 Å. Attempts to obtain crystals of the protein in complex with ATP were unsuccessful and the co-crystallization with substrate or substrate analogs did not show the corresponding electron density in the active site. The structure was solved by molecular replacement using the structure of PaaK1 (PDB ID: 2Y27) as a search model (Law & Boulanger, 2011). AjiA1 crystals belong to the space group  $P3_121$  with two protomers in the asymmetric unit (rmsd: 1.21 Å) which should correspond to the functional homodimeric quaternary structure of the protein in solution as predicted by PISA (Krissinel & Henrick, 2007) (Fig. 2). Further data processing and crystallographic statistics are shown in Table S1.

Similar to other AFEs members, AjiA1 has an overall architecture constituted by two  $\alpha/\beta$  domains with a central core of parallel  $\beta$ -sheets surrounded by an outer layer of  $\alpha$ -helices with the active site formed in the interface of the N and C terminal domains (Fig. 2c) (Branden & Tooze, 2012). AjiA1 lacks the typical  $\alpha/\beta$ -sandwich arrangement at the N terminus, which is common to other AFEs members and has a small helical bundle microdomain (Supplementary Figure 1) (Law & Boulanger,

2011). Each AjiA1 protomer, as expected, consists of a large N terminal domain (Ser4 to Gly321, the N domain) and a small C terminal domain (His327 to Ala432, the C domain) (Fig. 2a). The AjiA1 N domain consists of 3 antiparallel 2-stranded  $\beta$ -sheets sandwiched by nine  $\alpha$ -helices, whereas the C domain is defined by two helices, a short antiparallel two-strand  $\beta$ -sheet, a twisted three-stranded sheet, and a long loop that extends to the adjacent protomer. The analysis of the electrostatic potential surface of the dimer indicates a positively charged cavity in the interface between the N and C domain as observed previously in other AFEs members (Marahiel *et al.*, 1997, Law & Boulanger, 2011, Bera *et al.*, 2010, Nakatsu *et al.*, 2006, Fan *et al.*, 2016, Shao *et al.*, 2019) (Fig. 2b).

AjiA1 forms a dimer with 2-axis symmetry, which is stabilized by the formation of two large 7-stranded intersubunit  $\beta$ -sheet with contributions of two antiparallel strands from one protomer and five parallel strands from the second protomer (Fig. 2c). The dimer interface is about  $\sim 3999\text{\AA}^2$  and is stabilized through extensive shape and chemical complementarity (Supplementary Figure 2), including extensive hydrophobic interactions, by 46 hydrogen bonds and eight salt bridges (Supplementary Figure 3). However, similarly to other AFEs members, most of the C domain is not involved in the dimerization, except for the long C-terminal loop that extends to the adjacent protomer (Fig. 2c).

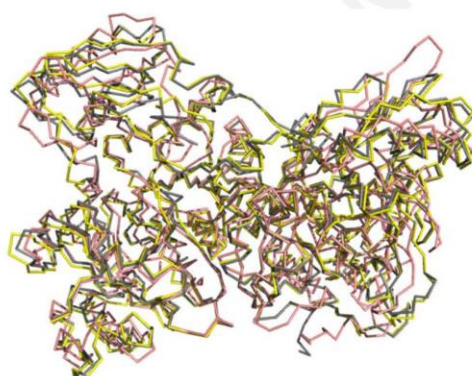


**Figure 2** The overall structure of AjiA1. (a) Monomer structure of AjiA1 with the larger N domain in green, the C domain in blue, and the linker that connects the domains in yellow. (b) The electrostatic potential surface indicates the groove of the active site (circle). In blue is indicated positive charge and red negative charge. The active center is shown in a circle. (c) Dimerization of AjiA1. Intimate dimerization interactions of AjiA1 create a composite  $\beta$ -sheets spanning both



monomers. The two-stranded sheet in purple (Ala82 to Tyr99) of the first protomer is abutted by the distorted, five-strand sheet in orange (Thr123 to Val126; Thr149 to Pro152; Leu176 to Trp178; Ala207 to Ala211; Val232 to Ser238) of the second protomer. The larger N domain is represented in green in the first protomer and lime in the second one, the C domain is in blue in the first protomer and light blue in the second protomer, and the linker that connects the two domains is shown in yellow.

AjiA1 shares 30% identity and 91% coverage with PaaK1, a phenylacetate-CoA ligase involved in phenylacetic acid biosynthesis (PAA) (Law & Boulanger, 2011, Holm & Laakso, 2016). The backbone superposition of AjiA1 and PaaK1 rendered an rmsd of 2.5 Å, indicating that these structures are considerably different (Fig. 3). Recently, the structure of the ATP dependent ligase NatL2 involved in the biosynthesis of nataxazole was also solved but in complex with cofactor and substrate and shares an identity of about 91% identity with AjiA1 (Song *et al.*, 2020). The superposition of two structures rendered an rmsd of about 1.14 Å (Fig. 3). The large N domain overlaps reasonably well with AjiA1 and other AFEs, while other regions show significant differences, particularly the active site region, which is located between the C and N terminal domains (Supplementary Figure 4). In addition, all the conserved signature motifs (A1–A10), commonly observed in all adenylation domains, are designated as per the nomenclature described by Marahiel *et al.* (Marahiel *et al.*, 1997) are conserved in AjiA1, including the P-loop, the linker motif (A8), and the adenine motif (A5), which contains critical residues for the binding of the ATP/AMP adenine moiety (Supplementary Figure 5).



**Figure 3** Superposition of AjiA1 (yellow) with the closest homologs PaaK1 (salmon) and NatL2 (grey).

### 3.2. Binding of metal ions

During the structure refinement, we observed a zinc ion coordinated by Cys248, Cys308 and Cys310, and His254 in AjiA1 (Supplementary Figure 6). As this ion is located remotely from the active site, it should only play a structural role, probably reducing side-chain movement and conformational changes as was also previously proposed for NatL2 structure analysis (Song *et al.*, 2020). Compared to other homologous, the Zn<sup>2+</sup> ion is only present in AjiA1 and NatL2 and should be a particular characteristic of this protein group.

Furthermore, in the previously reported structures of AFEs family, several of them contain a Mg<sup>2+</sup> coordinated at the active site, acting directly on ATP catalysis (Reger *et al.*, 2008). AjiA1 has a Mg<sup>2+</sup> binding site in each chain, however again remotely from the active site and should also play a structural role in reducing the structural flexibility of the region (Supplementary Figure 7). The structural analysis of AjiA1 did not indicate the presence of any metal in the active site, in contrast to the other AFEs structures in complex with ATP or ATP analogs and consequently, this ion should be associated with the co-factor.

### 3.3. Active site architecture, C terminal loop swapping and comparison with other AFEs members

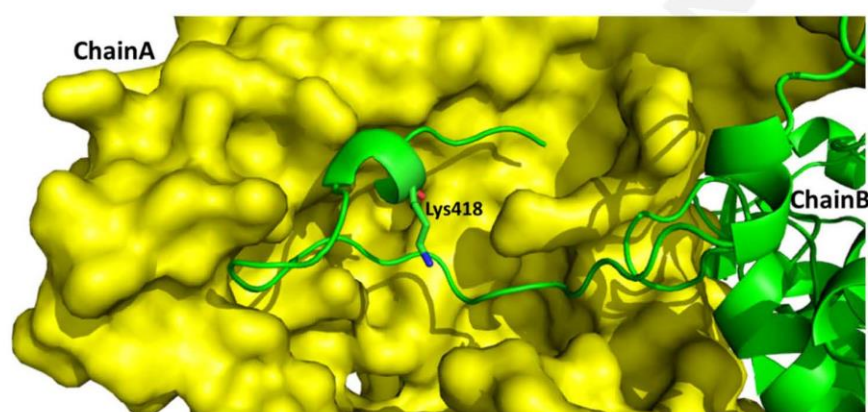
Although the structures of NatL2 were solved in complex with AMPPNP and with AMP in the presence of the substrate 3-HAA, the structure in apo form was not reported. In addition, as structures of AFEs members show different motions over the cofactor and substrate binding since these structures have a more mobile C domain, the analysis of the conformational changes during the catalytic cycle of enzymes involved in benzoxazole rings formation makes necessary.

The active site of AjiA1 is located in a groove in the interface between the N and the C domains. In contrast to most AFEs solved so far, AjiA1 residues of both protomers of the dimer are involved in the catalytic site formation and the dimeric structure should be a prerequisite for the enzyme activity, which is a novel feature of this class of enzymes. Although the key catalytic residue Lys418, which should have a role in the nucleophilic attack to the  $\alpha$ -phosphate of ATP, superposes well with PaaK1 catalytic Lys422, in AjiA1 occurs a loop swapping event. The long C-terminal loop from the adjacent protomer extends to the neighbor subunit active site. The observed loop swapping might reflect not only the dynamics of protomers but also a peculiarity of AFEs involving in the benzoxazoles ring formation since a similar event has also been observed in NatL2 (Schiering *et al.*, 2000, Song *et al.*, 2020) (Fig. 4).

Loops swapping is generally associated with an unconventional mechanism of oligomerization in which the structural elements of individual monomers are interchanged between identical partners by recruiting interactions that are crucial for stabilizing the protein. The phenomenon has been proposed



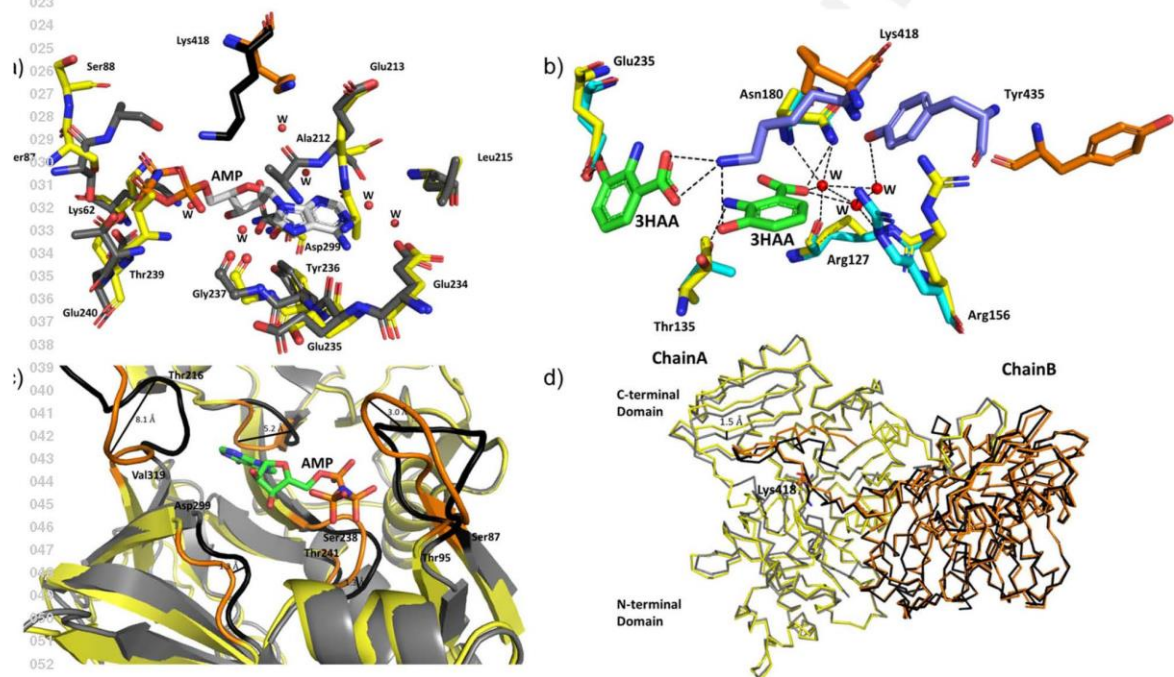
to play a role in aggregation, misfolding, and allosteric regulation, but all the biological implications are still unclear (Yang *et al.*, 2004). Initially, we expected that the cofactor binding could trigger movements in this region to perform the catalytic activity. However, the AjiA1 apo structure indicated that this loop does not return to its protomer in the absence of ligand, suggesting that this structural feature should not be responsible for allosteric regulation. Alternatively, it should play a role in the stabilization of the dimer or decreasing the flexibility of the C domain, in contrast to the C domains from other AFEs members, including PaaK1 (Law & Boulanger, 2011) and HCL (Zahn *et al.*, 2019), in which are reported to have drastic conformational changes. Nevertheless, amino acid residues from 422L to 424A of chain A form a two strains parallel  $\beta$ -sheet with amino acids 357A to 359A of chain B, and vice-versa, at the loop swapping region (Supplementary Figure 8).



**Figure 4** AjiA1 loop swapping. The conserved residue Lys418 is shown in the view of sticks.

Another key catalytic residue, Arg322 from the A8 motif, which is proposed to be involved in the stabilization of intermediate of reaction, is located at a short and solvent-exposed interdomain linker (Gulick, 2009). The ATP binding site contains several highly conserved motifs that are present in all family members and perform similar roles in positioning the nucleotide. The aspartic acid residue from the A7 motif, Asp299 in AjiA1, is universally conserved and interacts with both ribose hydroxyls and; Glu240, a residue from the A5 motif, should be involved in the anchoring the adenylate intermediate. An aromatic residue from the A5 motif, Tyr236, stacks against the adenine base. The major differences in the active site observed in the superposition of AjiA1 and NatL2:AMPPNP are: the approximation of Ser88 (from the P-loop) and Lys418 (the A10 lysine) to the AMP phosphate group, a repositioning of Ala212 to be close to the adenine ring, as well as the participation of six water molecules (Fig. 5a).

Further structural analysis revealed, as expected, that other several functional regions of this protein have their conformations highly conserved (Supplementary Figure 9). Besides, other amino acid residues, which are in the active site pocket, including the motifs A3 or P-loop, A5, A6 and 7 are substantially conserved (Ashkenazy *et al.*, 2016, Marahiel *et al.*, 1997). However, based on the conserved active site residues, AjiA1 has similar catalytic residues as reported to PaaK1, HCL and NatL2. Consequently, the same catalytic mechanism should act for this enzyme, which includes the adenylation of the carboxylic acid from the substrate followed by the esterification reaction with a second molecule of substrates, which consists in the ring condensation of AJI956 intermediate.



**Figure 5** Comparison AjiA1 and structures in complex with cofactor and substrates. (a) Superposition of AjiA1 and NatL2:AMPPNP (PDB ID: 6six). The AjiA1 residues are represented in yellow and those from NatL2 in grey. The key catalytic residue Lys418 is in a stronger color as it belongs to the adjacent protomer. (b) Superposition of AjiA1 and NatL2:AMP:3-HAA (PDB ID: 6siy). AjiA1 residues are represented in yellow, in which K418 and Y435 are in orange and the NatL2 residues are in cyan, in which K418 and Y435 are in blue. The different color of these amino acid residues is because they are from the adjacent protomer. The active site waters are represented by red

spheres. (c) Superimposition of AjiA1 and NatL2 in complex with AMPPNP (6SIX). The protomers are represented as a cartoon. AjiA1 is in yellow and its five different loops are in orange. NatL2:AMPPNP is in grey and its loops are in black. (d) Superimposition of AjiA1 and NatL2 in complex with AMPPNP (6SIX). The protomers are represented as a ribbon. AjiA1 are represented in yellow (chain A) and orange (chain B) and NatL2 in complex with AMPPNP (6SIX) are represented in grey (chain A) and black (chain B). The key catalytic residue Lys418 of AjiA1 and NatL2:AMPPNP is represented in sticks. The C domain, and consequently the C terminal loop, also moves to a closer position of the active site on the active site in about 1.5 Å.

Lys418 and Tyr435, from the C-terminal loop swapping, are more buried and closer to the 3-HAA molecules in the active site of NatL2 complex and they show considerable conformational changes when compared to the same residues in the structure of AjiA1. Arg156 also has a relevant conformational change as it moves away from the active site in AjiA1 while Glu235, from motif A5, seems to have a role in position 3-HAA on the active site contributing with a hydrogen bond with the substrate hydroxyl group. In NatL2 in complex with 3-HAA, Thr135 also undergoes a rotation in comparison to AjiA1 and its side chain makes a hydrogen bond with one of the 3-HAA molecules (Fig. 5b).

The comparison among the structures of AjiA1, NatL2:AMPPNP (rmsd: 1.12 Å) and NatL2:AMP:3-HAA (rmsd: 1.14 Å) revealed key motions that should occur during the catalytic cycle and also because of cofactor binding. A series of movements in key loops of the active sites have conformational changes when the NatL2 in complex with AMP and AjiA1 in apo form are superposed. The largest changes are observed in six loops near the active site: 86E to 94P (motif A3, P-loop), 210T to 213E, 238S to 242G (motif A5), 296N to 298A (motif A7), 305T to 313L and 320L to 328T (motif A8) in AjiA1 (Fig. 5c). Three of these five loops have the highest B-factors (86E to 94P; 210T to 213E and 320L to 328T) in comparison to other regions of NatL2 and AjiA1, indicating high flexibility. Thus, the average displacement suggests that the most flexible part of the protein that should act on the catalytic function. The core fold in the N domain is highly conserved and rigid, as evidenced by its low average B-factor (Supplementary Figure 10).

The AjiA1 P-loop (motif 3), which includes the residues Ser87 to Pro94 (Supplementary Figure 5), is a highly conserved region in all AFEs members and has an amino acid sequence of Gx(4)-GK-[TS]. This region is proposed to play a bifunctional role in the structure of PaaK1, which involves both the coordination of ATP and the stabilization of the C domain. In the PaaK1 structure in complex with adenylate intermediate without the pyrophosphate group, this region adopts an opened and disordered structure. In contrast, on the binding of the cofactor ATP, this region adopts a closed and ordered conformation establishing hydrogen bonding contact with cofactor  $\beta$  and  $\gamma$  phosphate. Although

NatL2 has been crystallized in the presence of AMP, the position of P-loop differs from AjiA1. In AjiA1, the P-loop adopts an opener conformation with a displacement of about 3 Å in comparison to NatLA2 and moves towards the C terminal loop. Additionally, the position of the N domain also seems to adopt a closer conformation on the active site in about 1.3 Å (Fig. 5c). The C domain, and consequently the C-terminal loop, also moves to a closer position to the active site in about 1.5 Å, allowing the Lys418 to adopt a catalytic position and interact with the phosphate and ribose moieties of AMP (Fig. 5d).

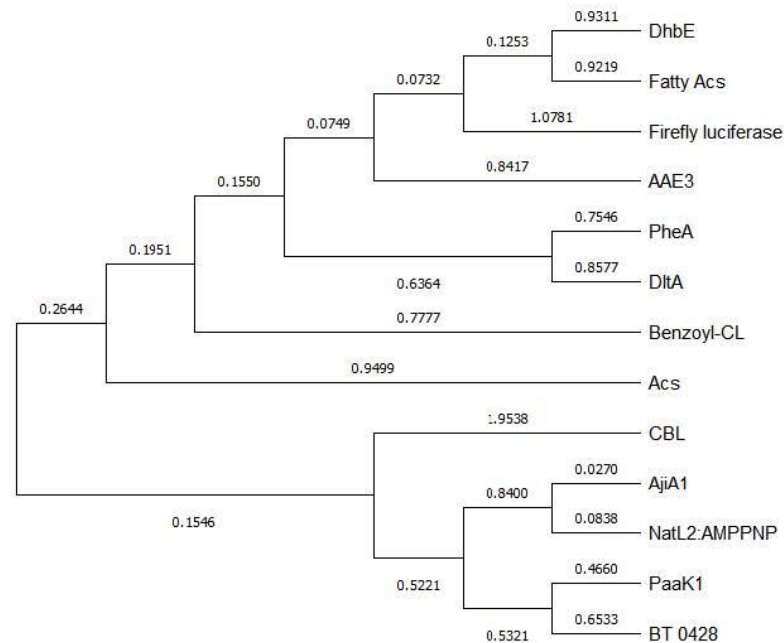
In the reaction catalyzed by the enzymes within this family, a substrate carboxylate is activated using ATP to form an acyl-adenylate (Reger *et al.*, 2008). Structural and functional studies indicated that these enzymes alternate between two conformations during catalysis (Reger *et al.*, 2008, Shao *et al.*, 2019, Nakatsu *et al.*, 2006, Gulick, 2009). The presence of catalytic lysine in the loop-swapping could suggest an intense movement of this “arm”, functioning as the “gate” for the open and closed structure on this family of enzymes, however, our apo structure did not show evidence of a large conformational change on this region and the C domain should not adopt a closed and open form as described in other members of AFEs. The long C terminal loop should have a role in the lock the structure and the closed-form as predicted by Song *et al.*, 2020 (Song *et al.*, 2020). Interestingly, as the loop swapping lock the structures of AjiA1 and NatL2 in the same conformation, large conformational motions on the C domain are unlikely and hence the domain alternation should not occur. Since large conformational changes involve a higher enthalpic energetic barrier, these enzymes could have evolved using a different and elegant strategy to perform the two half-reactions commonly for AFEs.

The loop swapping observed in these structures could maintain both half-reactions in a single active site avoid large rotations but costing a constrain functional dimeric structure. However, further investigation is needed to prove this hypothesis or to understand whether the loop swapping has an alternative role for the domain alternation as observed in other class I AFEs members.

### 3.4. Phylogenetic analysis

To gain insights concerning the divergence and evolution, we performed a phylogenetic analysis using the Maximum Likelihood (ML) method (Ziemert & Jensen, 2012, Benevides *et al.*, 2016), which reveals that AjiA1 and NatL2 are separated in a different branch (Fig. 6), supporting that these enzymes may have evolved independently and that they could be classified distinctly as a new subgroup of phenylacetate-CoA like ligases.





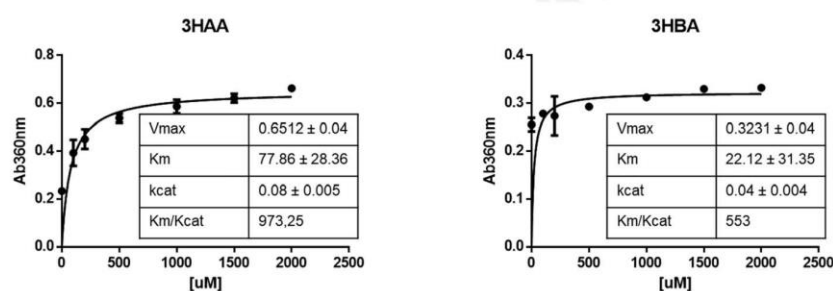
**Figure 6** Evolutionary analysis by Maximum Likelihood method of several different AFEs members that have been structurally characterized including AjiA1. The branch lengths are shown. The proteins and organisms used are: DhbE from *Bacillus subtilis* (PDB ID: 1MDF); Fatty Acs from *Thermus thermophilus* (PDB ID: 1ULT); Firefly Luciferase from *Photinus pyralis* (PDB ID: 1LCI); AAE3 from *Arabidopsis thaliana* (PDB ID: 5IE3); PheA from *Brevibacillus brevis* (PDB ID: 1AMU); DltA from *Bacillus cereus* ATCC 14579 (PDB ID: 3DHV); Benzoyl-CL from *Paraburkholderia xenovorans* LB400 (PDB ID: 2V7B); Acs from *Salmonella enterica* (PDB ID: 1PG4); CBL from *Alcaligenes* sp. AL3007 (PDB ID: 1T5S); The protein of this study AjiA1 from *Streptomyces* sp. AJ9561 (PDB ID: 6WUQ); NatL2 from *Streptomyces* sp. Tu 6176 (PDB ID: 6SIX); PaaK1 from *Burkholderia cenocepacia* (PDB ID: 2Y27) and BT\_0428 from *Bacteroides thetaiotaomicron* (PDB ID: 4R1M). The analyses were conducted in MEGA X (Kumar *et al.*, 2018).

### 3.5. Kinetics and activity of AjiA1

It has been reported that the domain alternation could increase the promiscuity of AFEs as the large rotation of the C domain could facilitate the binding of alternative substrates (Gulick, 2009). In order to gain insights into the promiscuity of AjiA1, which has a much more restrained structure, we performed an assay with an alternative substrate 3-hydroxybenzoic acid (3-HBA) and the natural

substrate 3-HAA using ATP as a cofactor. We determined the kinetic parameters for 3-HAA and 3-HBA by measuring the increase of absorbance at 360 nm using a spectrophotometric assay. As shown, the  $K_m$  and the catalytic efficiency ( $k_{cat}$ ) as well as the relationship  $K_m/K_{cat}$  diminished approximately by half when 3-HBA was used as a substrate instead of 3-HAA. This indicates that the enzyme has a higher affinity but a lower catalytic efficiency for 3-HBA than for 3-HAA, the natural substrate. The result demonstrates that AjiA1 is able to catalyze the formation of an ester bond in the absence of the amine group. In other words, AjiA1 was shown to catalyze the dimerization of 3-HBA (in the absence of a 2-substituted amine group) via an ester linkage.

In contrast, Song *et al.* demonstrated that NatL2 only catalyzes the formation of ester and not amide, considering this as a derivative product (Song *et al.*, 2020). Nevertheless, our result does not necessarily imply an ester instead of an amide is formed when the native substrate 3-HAA is used (Fig. 7).



**Figure 7** The kinetic analyses of AjiA1 were performed using 3HAA and 3HBA concentrations from 0 to 2500  $\mu\text{M}$ .  $K_m$  and  $V_{max}$  were determined from nonlinear regression to the Michaelis–Menten equation from three replicate experiments.

#### 4. Conclusion

In summary, we have determined the structure of AjiA1 in its uncomplexed form and its comparison with the enzyme NatL2 in complex with different ligands indicates that this group of enzymes has a completely new set of motion during the catalytic cycle into the AFEs members. The domain alternation should not occur in the enzymes involved in the benzoxazole biosynthesis. In contrast, these enzymes have an event of domain swapping involving a long C terminal loop. This loop constrains the C domain and locks the enzyme in a single conformation. Taken together, the new movements, which involve only loops of the active site and a slight rotation of the adjacent protomer, and phylogenetic analysis, we propose that these enzymes should be classified as a distinct group of phenylacetate-CoA like ligases even with the modest sequence similarity. In addition, even with the

constraints imposed by the absence of domain alternation, AjiA1 was still able to accept alternative substrates indicating promiscuity for these enzymes. Consequently, these findings bring new insights into the crescent AFEs family and indicate the possibility to use these enzymes in biotechnological approaches.

**Acknowledgements** The author declares that there are no conflicts of interest. MVBD and PFL designed the research; FCRP performed the protein production, purification, crystallization, structure determination and analysis; KC performed cloning and protein expression optimization; MS performed the DNA sequence analysis and AJ9561 gene cluster identification on on *Streptomyces* sp. AJ9561 whole-genome; FCRP, KC and AMS performed the protein kinetics and analysis. FCRP and MVBD wrote the manuscript with contributions and approve of all authors.

## References

- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L.-W., Jain, S., Kapral, G. J. & Kunstleve, R. W. G. (2011). *Methods* **55**, 94-106.
- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Crystallographica Section D: Biological Crystallography* **68**, 352-367.
- Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T. & Ben-Tal, N. (2016). *Nucleic acids research* **44**, W344-W350.
- Bandarian, V., Patridge, K. A., Lennon, B. W., Huddler, D. P., Matthews, R. G. & Ludwig, M. L. (2002). *Nature structural biology* **9**, 53-56.
- Benevides, L. d. J., Carvalho, D. S. d., Andrade, R. F. S., Bomfim, G. C. & Fernandes, F. M. d. C. (2016). *Genetics and molecular biology* **39**, 665-673.
- Bera, A. K., Atanasova, V., Gamage, S., Robinson, H. & Parsons, J. F. (2010). *Acta crystallographica. Section D, Biological crystallography* **66**, 664-672.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic acids research* **28**, 235-242.
- Branden, C. I. & Tooze, J. (2012). *Introduction to Protein Structure*. CRC Press.
- Cano-Prieto, C., García-Salcedo, R., Sánchez-Hidalgo, M., Braña, A. F., Fiedler, H. P., Méndez, C., Salas, J. A. & Olano, C. (2015). *ChemBioChem* **16**, 1461-1473.
- Chang, K.-H., Xiang, H. & Dunaway-Mariano, D. (1997). *Biochemistry* **36**, 15650-15659.

- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Crystallographica Section D: Biological Crystallography* **66**, 12-21.
- Conti, E., Franks, N. P. & Brick, P. (1996). *Structure* **4**, 287-298.
- Cusack, S. (1997). *Current opinion in structural biology* **7**, 881-889.
- D'Ambrosio, H. K. & Derbyshire, E. R. (2019). *ACS Chemical Biology*.
- Day, J. C., Goodall, T. I. & Bailey, M. J. (2009). *Molecular phylogenetics and evolution* **50**, 93-101.
- Emsley, P. & Cowtan, K. (2004). *Acta Crystallographica Section D: Biological Crystallography* **60**, 2126-2132.
- Estrada, P., Manandhar, M., Dong, S.-H., Deveryshetty, J., Agarwal, V., Cronan, J. E. & Nair, S. K. (2017). *Nature chemical biology* **13**, 668.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Crystallographica Section D: Biological Crystallography* **69**, 1204-1214.
- Fan, M., Xiao, Y., Li, M. & Chang, W. (2016). *Molecular Plant* **9**, 1349-1352.
- Gulick, A. M. (2009). *ACS Chemical Biology* **4**, 811-827.
- Hisanaga, Y., Ago, H., Nakagawa, N., Hamada, K., Ida, K., Yamamoto, M., Hori, T., Arii, Y., Sugahara, M. & Kuramitsu, S. (2004). *Journal of Biological Chemistry*.
- Holm, L. & Laakso, L. M. (2016). *Nucleic acids research* **44**, W351-W355.
- Huang, W., Xu, H., Li, Y., Zhang, F., Chen, X.-Y., He, Q.-L., Igarashi, Y. & Tang, G.-L. (2012). *Journal of the American Chemical Society* **134**, 8831-8840.
- Jones, D. T., Taylor, W. R. & Thornton, J. M. (1992). *Bioinformatics* **8**, 275-282.
- Kabsch, W. (2010). *Acta Crystallographica Section D: Biological Crystallography* **66**, 125-132.
- Krissinel, E. & Henrick, K. (2007). *Journal of molecular biology* **372**, 774-797.
- Kudo, F., Miyanaga, A. & Eguchi, T. (2019). *Journal of industrial microbiology & biotechnology* **46**, 515-536.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. (2018). *Molecular biology and evolution* **35**, 1547-1549.
- Labby, K. J., Watsula, S. G. & Garneau-Tsodikova, S. (2015). *Natural product reports* **32**, 641-653.
- Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S. & Thornton, J. M. (2018). *Protein Science* **27**, 129-134.
- Laskowski, R. A. & Swindells, M. B. (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery: ACS Publications.
- Law, A. & Boulanger, M. J. (2011). *Journal of Biological Chemistry* **286**, 15577-15585.
- Losada, A. A., Cano-Prieto, C., García-Salcedo, R., Braña, A. F., Méndez, C., Salas, J. A. & Olano, C. (2017). *Microbial biotechnology* **10**, 873-885.
- Lux, M. C., Standke, L. C. & Tan, D. S. (2019). *The Journal of antibiotics* **72**, 325-349.
- Lv, M., Zhao, J., Deng, Z. & Yu, Y. (2015). *Chemistry & biology* **22**, 1313-1324.



- 002  
003  
004  
005  
006  
007  
008  
009  
010  
011  
012  
013  
014  
015  
016  
017  
018  
019  
020  
021  
022  
023  
024  
025  
026  
027  
028  
029  
030  
031  
032  
033  
034  
035  
036  
037  
038  
039  
040  
041  
042  
043  
044  
045  
046  
047  
048  
049  
050  
051  
052  
053  
054  
055  
056  
057  
058  
059  
060  
061  
062  
063  
064  
065  
066  
067  
068  
069  
070  
071  
072  
073  
074  
075  
076
- Marahiel, M. A., Stachelhaus, T. & Mootz, H. D. (1997). *Chem Rev* **97**, 2651-2674.
- McCoy, A. J. (2007). *Acta Crystallographica Section D: Biological Crystallography* **63**, 32-41.
- Nakatsu, T., Ichiyama, S., Hiratake, J., Saldanha, A., Kobashi, N., Sakata, K. & Kato, H. (2006). *Nature* **440**, 372-376.
- P Duckworth, B., M Nelson, K. & C Aldrich, C. (2012). *Current topics in medicinal chemistry* **12**, 766-796.
- Reger, A. S., Wu, R., Dunaway-Mariano, D. & Gulick, A. M. (2008). *Biochemistry* **47**, 8016-8025.
- Sato, S., Kajiura, T., Noguchi, M., Takehana, K., Kobayashi, T. & Tsuji, T. (2001). *The Journal of antibiotics* **54**, 102-104.
- Schiering, N., Casale, E., Caccia, P., Giordano, P. & Battistini, C. (2000). *Biochemistry* **39**, 13376-13382.
- Schmelz, S. & Naismith, J. H. (2009). *Current opinion in structural biology* **19**, 666-671.
- Shao, X., Cao, H.-Y., Zhao, F., Peng, M., Wang, P., Li, C.-Y., Shi, W.-L., Wei, T.-D., Yuan, Z., Zhang, X.-H., Chen, X.-L., Todd, J. D. & Zhang, Y.-Z. (2019). *Molecular Microbiology* **111**, 1057-1073.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Rimmert, M., Soding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol Syst Biol* **7**, 539.
- Song, H., Rao, C., Deng, Z., Yi, Y. & Naismith, J. H. (2020). *Angewandte Chemie International Edition*.
- Stanišić, A. & Kries, H. (2019). *ChemBioChem* **20**, 1347-1356.
- Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. (2009). *Bioinformatics (Oxford, England)* **25**, 1189-1191.
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. & McCoy, A. (2011). *Acta Crystallographica Section D: Biological Crystallography* **67**, 235-242.
- Wu, Q., Liang, J., Lin, S., Zhou, X., Bai, L., Deng, Z. & Wang, Z. (2011). *Antimicrobial Agents and Chemotherapy* **55**, 974-982.
- Yang, S., Cho, S. S., Levy, Y., Cheung, M. S., Levine, H., Wolynes, P. G. & Onuchic, J. N. (2004). *Proceedings of the National Academy of Sciences* **101**, 13786-13791.
- Zahn, M., Kurteva-Yaneva, N., Schuster, J., Krug, U., Georgi, T., Müller, R. H., Rohwerder, T. & Sträter, N. (2019). *Journal of molecular biology* **431**, 2747-2761.
- Ziemert, N. & Jensen, P. R. (2012). *Methods in enzymology*, pp. 161-182: Elsevier.

---  
 002  
 003  
 004  
 005  
 006  
 007  
 008  
 009  
 010  
 011  
 012  
 013  
 014  
 015  
 016  
 017  
 018  
 019  
 020  
 021  
 022  
 023  
 024  
 025  
 026  
 027  
 028  
 029  
 030  
 031  
 032  
 033  
 034  
 035  
 036  
 037  
 038  
 039  
 040  
 041  
 042  
 043  
 044  
 045  
 046  
 047  
 048  
 049  
 050  
 051  
 052  
 053  
 054  
 055  
 056  
 057  
 058  
 059  
 060  
 061  
 062  
 063  
 064  
 065  
 066  
 067  
 068  
 069  
 070  
 071  
 072  
 073  
 074  
 075  
 076

IUCrJ

BIOLOGY | MEDICINE research papers

Page 19

## Supporting information

**Table S1** Data collection and processing

Values for the outer shell are given in parentheses.

Diffraction source	PETRA III, EMBL c/o DESY BEAMLIN P13 (MX1)
Wavelength (Å)	0.976195 Å
Temperature (K)	100
Detector	DECTRIS PILATUS 6M PIXEL
Crystal-detector distance (mm)	438.35
Rotation range per image (°)	0.1
Total rotation range (°)	360
Exposure time per image (s)	0.04
Space group	$P3_121$
$a, b, c$ (Å)	128.78, 128.78, 101.45
$\alpha, \beta, \gamma$ (°)	90, 90, 120
Resolution range (Å)	48.870–2.000 (2.050–2.000)
Total No. of reflections	1320343 (92631)
No. of unique reflections	65461 (4483)
Completeness (%)	99.800 (97.900)
Redundancy	20.200 (20.700)
$\langle I/\sigma(I) \rangle$	24.400 (1.8)
$R_{\text{r.i.m.}}$	0.069 (2.637)
$R_{\text{p.i.m.}}$	0.015
Overall $B$ factor from Wilson plot (Å <sup>2</sup> )	46.820

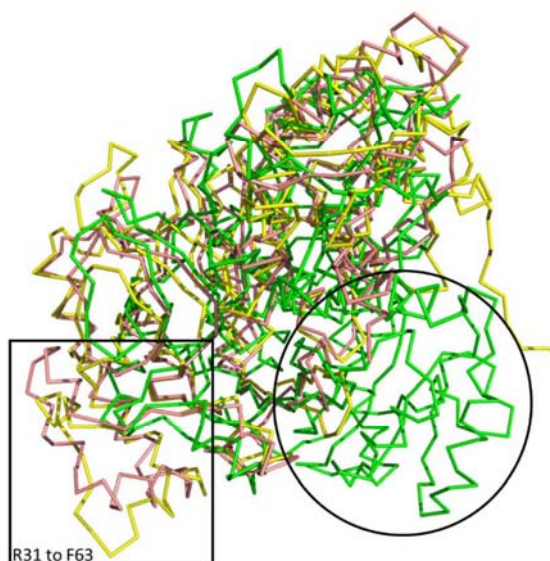
**Table S2** Structure solution and refinement

Values for the outer shell are given in parentheses.

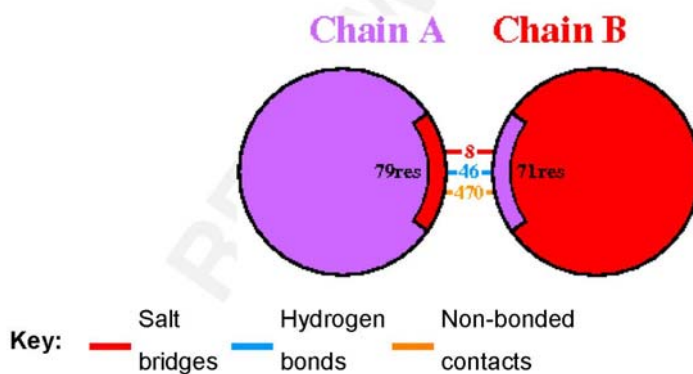
Resolution range (Å)	33.8160–2.0030 (2.0325–2.0026)
Completeness (%)	99.8
$\sigma$ cutoff	$F > 1.350\sigma(F)$
No. of reflections, working set	62196 (2564)
No. of reflections, test set	3205 (133)
Final $R_{\text{cryst}}$	0.188 (0.3501)
Final $R_{\text{free}}$	0.224 (0.3892)
No. of non-H atoms	6869
Protein	6713
Ligand	<u>0</u>
Solvent	152 <u>0</u>
Total	6869
R.m.s. deviations	
Bonds (Å)	0.008
Angles (°)	0.924
Average $B$ factors (Å <sup>2</sup> )	63.61
Protein	63.74
Ligand	<u>0.0</u>
Ramachandran plot	
Most favoured (%)	96
Allowed (%)	4

002  
003  
004  
005  
006  
007  
008  
009  
010  
011  
012  
013  
014  
015  
016  
017  
018  
019  
020  
021  
022  
023  
024  
025  
026  
027  
028  
029  
030  
031  
032  
033  
034  
035  
036  
037  
038  
039  
040  
041  
042  
043  
044  
045  
046  
047  
048  
049  
050  
051  
052  
053  
054  
055  
056  
057  
058  
059  
060  
061  
062  
063  
064  
065  
066  
067  
068  
069  
070  
071  
072  
073  
074  
075  
076

IUCrJ



**Figure S1** N terminal of AjiA1. AjiA1 (yellow) and PaaK1 (2y27) (salmon) contain a small helical bundle arrangement at the N-terminus (boxed) but lack the typical N-terminal arrangement exhibited by family members (circle). overlays of AjiA1, PaaK1 and 1mdb a homologous family member (Protein Data Bank code 1MDB, 2,4-dihydroxybenzoateAMPligase, green) demonstrate the typical  $\alpha/\beta$ -sandwich arrangement at the N terminus (circle).



**Figure S2** Dimerization interface of AjiA1. The online server PDBsum (Laskowski *et al.*, 2018) (EMBL-EBI) was used to generate a schematic diagram showing the interactions between the two chains of AjiA1. Interacting chains are joined by colored lines, each representing a different type of

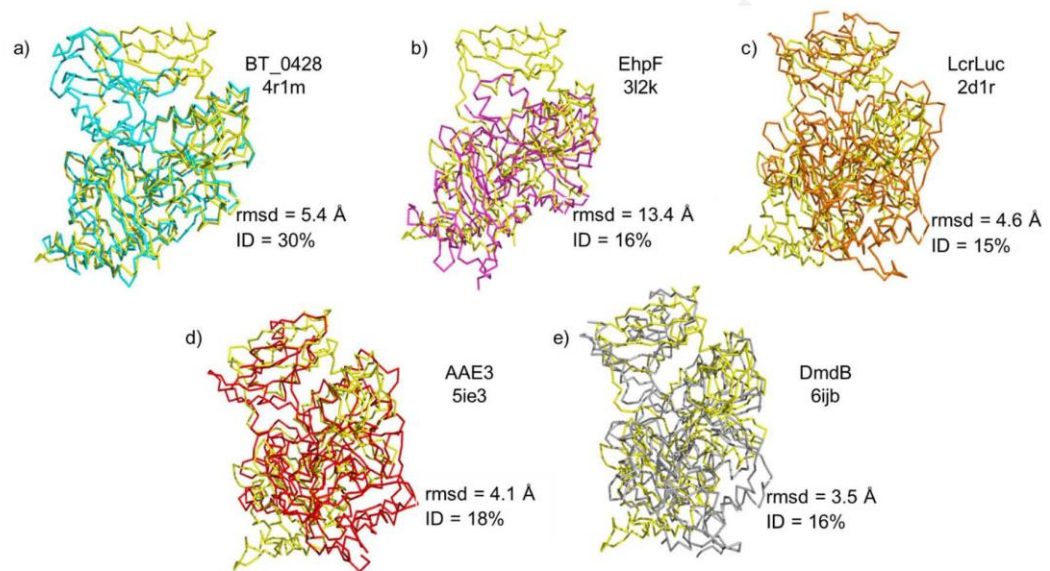
interaction, orange = non-bonded contacts, blue = hydrogen bonds, and red lines = salt bridges. The area of each circle is proportional to the surface area of the corresponding protein chain. The extent of the interface region on each chain is represented by the black wedge whose size signifies the interface surface area.

### Chain A Chain B



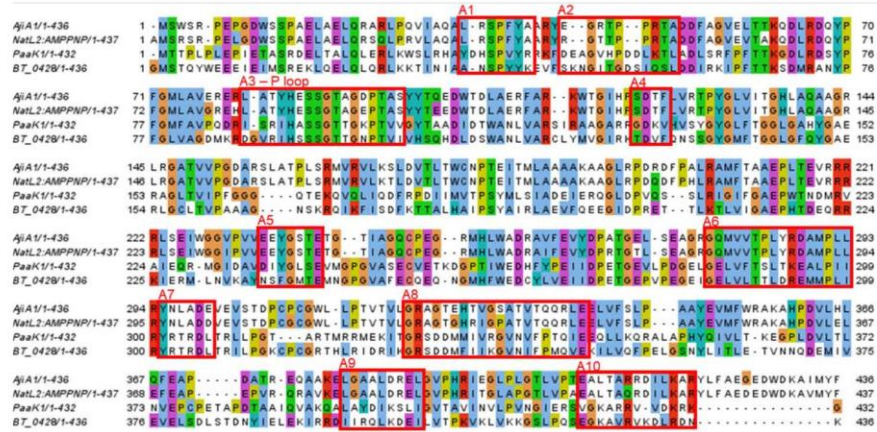
**Figure S3** Dimerization interface of AjiA1. The online server PDBsum (Laskowski et al., 2018) (EMBL-EBI) was used to generate a residue interactions net across the interface between the two chains of AjiA1 coloured by residue type. The number of H-bond lines between any two residues

indicates the number of potential hydrogen bonds between them. For nonbonded contacts, which can be plentiful, the width of the striped line is proportional to the number of atomic contacts. Interacting chains are joined by colored lines, each representing a different type of interaction, orange = non-bonded contacts, blue = hydrogen bonds, and red lines = salt bridges. The residue colours are: Positive blue (H,K,R); negative red (D,E); neutral green (S,T,N,Q); aliphatic gray (A,V,L,I,M); aromatic purple (F,Y,W); Proline and Glycine orange.



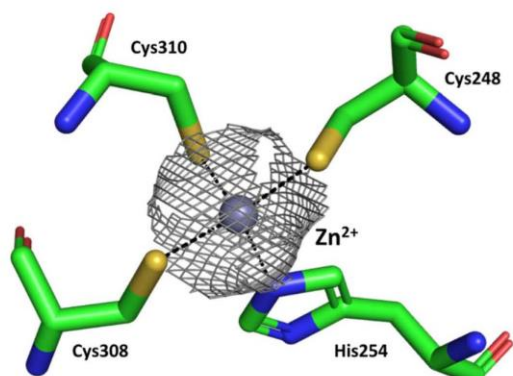
**Figure S4** Structural homologs of AjiA1. The C $\alpha$  chain of AjiA1 was aligned with its structural homologues (a) 4R1M (cyan), (b) 3L2K (purple), (c) 2D1R (orange), (d) 5IE3 (red), and (e) 6IJB (gray). The rmsd value for the superposition with AjiA1, and amino acid identity (%) to AjiA1 is given in each analysis.



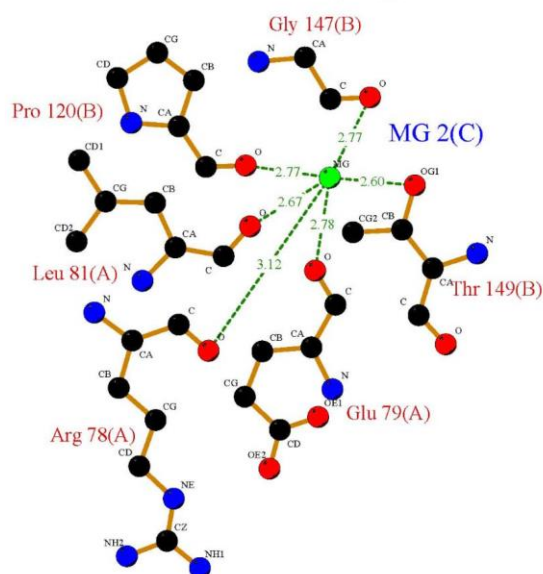


% Identity to AjiA1	
NatL2:AMPPNP	91
PaaK1	30
BT_0428	30

**Figure S5** Alignment of AjiA1 and PaaK1 with DALI homologs. The A1–A10 motifs conserved in the ANL superfamily enzymes are indicated according to the defined adenylation domain of nonribosomal peptide synthetase (Marahiel *et al.*, 1997). Core motifs A1-A10 are shown in red frames. The alignment was generated by ClustalOmega (Sievers *et al.*, 2011) and imported into Jalview Version: 2.11.0 (Waterhouse *et al.*, 2009). Similar residues are colored accordingly: blue, hydrophobic; green, polar; purple, negative; red, positive; yellow, proline; orange, glycine. The table displays the % identity of AjiA1 to the aligned homologs.

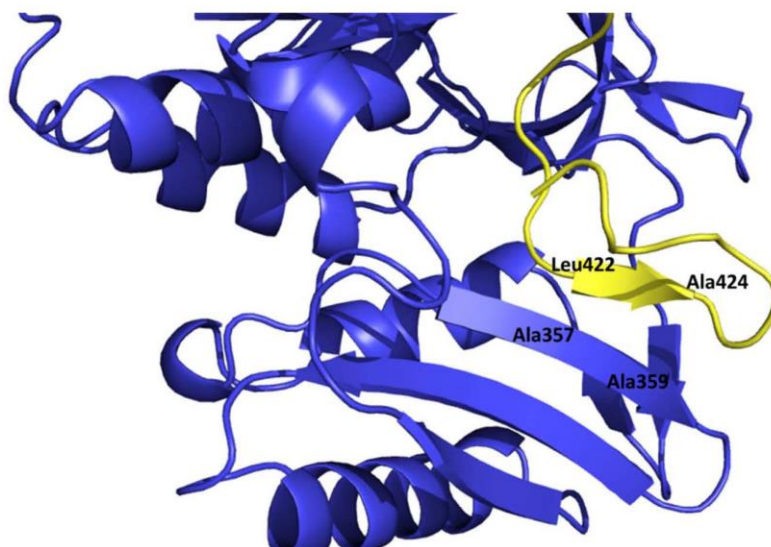


**Figure S6** Zinc binding site in the chain A of AjiA1 with 2Fo-Fc of  $Zn^{2+}$  (contoured to  $2.0 \sigma$ ). This figure shows the coordination interactions.

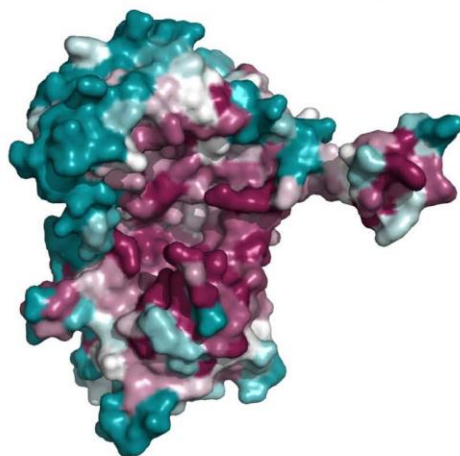


**Figure S7** Schematic representation of the  $Mg^{2+}$  binding site at chain A of AjiA1. The figure was prepared using LigPlot+ v.2.1 (Laskowski & Swindells, 2011)





**Figure S8** Region of the loop swapping in cartoon view. Amino acids 422L, 423F and 424A appear to form a small beta-strand that could be involved in the formation of a  $\beta$ -sheet with amino acids 357A, 358K and 359A. Chain A is in yellow and chain B is in blue.



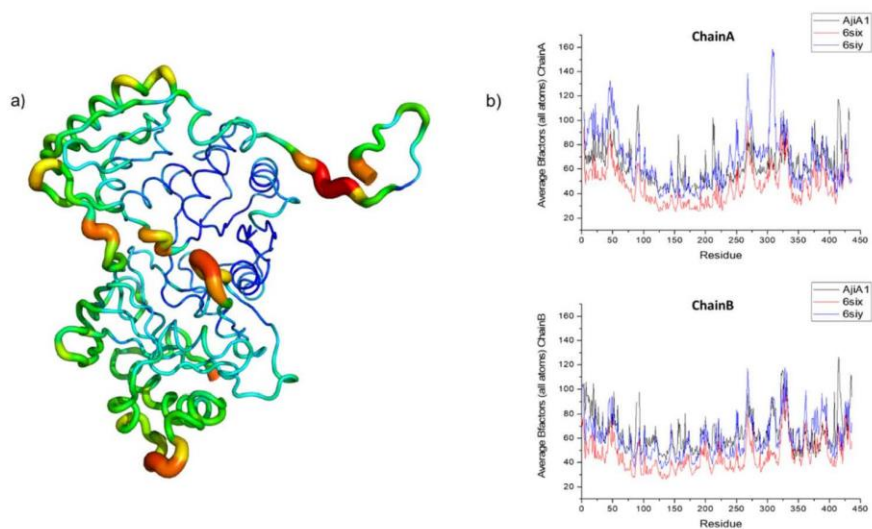
**Figure S9** Surface representation of the AjiA1 monomer showing the conservation of residues with other AFEs members. The figure was prepared using the ConSurf server (Ashkenazy et al., 2016). Purple and cyan indicate high and low sequence conservation, respectively.

---  
 002  
 003  
 004  
 005  
 006  
 007  
 008  
 009  
 010  
 011  
 012  
 013  
 014  
 015  
 016  
 017  
 018  
 019  
 020  
 021  
 022  
 023  
 024  
 025  
 026  
 027  
 028  
 029  
 030  
 031  
 032  
 033  
 034  
 035  
 036  
 037  
 038  
 039  
 040  
 041  
 042  
 043  
 044  
 045  
 046  
 047  
 048  
 049  
 050  
 051  
 052  
 053  
 054  
 055  
 056  
 057  
 058  
 059  
 060  
 061  
 062  
 063  
 064  
 065  
 066  
 067  
 068  
 069  
 070  
 071  
 072  
 073  
 074  
 075  
 076

IUCrJ

Page 27

BIOLOGY | MEDICINE research papers



**Figure S10** The B-factor putty representation showing the rigid core. (a) The B-factor values are illustrated by color, ranging from low (blue) to high (red), and by shape, a wider tube indicates regions with higher B-factors, whereas narrower tube indicates regions with lower B-factors. (b) Temperature factor analysis. Graphs of average B-factor versus residue for each chain in AjiA1.

## APÊNDICES B – Artigo: Unexpected enzyme-catalysed [4+2] cycloaddition and rearrangement in polyether antibiotic biosynthesis

nature  
catalysis

ARTICLES

<https://doi.org/10.1038/s41929-019-0351-2>

# Unexpected enzyme-catalysed [4+2] cycloaddition and rearrangement in polyether antibiotic biosynthesis

Rory Little<sup>1</sup>, Fernanda C. R. Paiva<sup>2</sup>, Robert Jenkins<sup>3</sup>, Hui Hong<sup>1</sup>, Yuhui Sun<sup>4</sup>, Yuliya Demydchuk<sup>1,6</sup>, Markiyan Samborsky<sup>1</sup>, Manuela Tosin<sup>3</sup>, Finian J. Leeper<sup>5</sup>, Marcio V. B. Dias<sup>2,3</sup> and Peter F. Leadlay<sup>1\*</sup>

**Enzymes that catalyse remarkable Diels–Alder-like [4+2] cyclizations have been previously implicated in the biosynthesis of spirotreronate and spirotreronate antibiotics. Biosynthesis of the polyether antibiotic tetronasin is not expected to require such steps, yet the tetronasin gene cluster encodes enzymes Tsn11 and Tsn15, which are homologous to authentic [4+2] cyclases. Here, we show that deletion of Tsn11 led to accumulation of a late-stage intermediate, in which the two central rings of tetronasin and four of its twelve asymmetric centres remain unformed. In vitro reconstitution showed that Tsn11 catalyses an apparent inverse-electron-demand hetero-Diels–Alder-like [4+2] cyclization of this species to form an unexpected oxadecalin compound that is then rearranged by Tsn15 to form tetronasin. To gain structural and mechanistic insight into the activity of Tsn15, the crystal structure of a Tsn15-substrate complex has been solved at 1.7 Å resolution.**

The Diels–Alder reaction<sup>1</sup>, in which a 1,3-diene and an alkene (dienophile) undergo a concerted [4+2] cycloaddition to form a cyclohexene ring, is of central importance to synthetic organic chemistry<sup>2</sup>. Transformations during the biosynthesis of numerous cyclic microbial metabolites have been speculated to be catalysed by naturally evolved [4+2] cyclase (‘Diels–Alderase’) enzymes<sup>3,4</sup>, and several enzymes have been shown to catalyse [4+2] cyclizations consistent with this mechanism<sup>5–12</sup>. Almost all examples studied so far involve the combination of an electron-rich 1,3-diene with an electron-poor dienophile to form cyclohexene-containing products. Important exceptions are the hetero-[4+2] cycloadditions that are catalysed by the pyridine synthases of thiopeptide antibiotic biosynthesis<sup>13</sup> and by LepI, which installs a dihydropyran during leporin B biosynthesis<sup>14</sup> (Fig. 1a). In the case of LepI, an electron-poor (oxygen-based) diene reacts with an electron-rich dienophile in an inverse-electron-demand hetero-Diels–Alder reaction<sup>14</sup>.

Tetronasin (1) (Fig. 1b) from *Streptomyces longisporoflavus*, which has found use as an antibiotic and antiparasitic agent, is an unusual polyether ionophore that contains an acyltetrone acid moiety<sup>15</sup>. It bears a near-mirror-image structural relationship to tetronomycin (*tmn*) from *Streptomyces* sp. NRRL 11266 (ref. 16), whose biosynthesis has been shown to involve assembly of the polyketide backbone on a modular polyketide synthase (PKS)<sup>16</sup>, with release of the chain by tetrone acid ring formation<sup>17</sup>. Formation of the tetrahydrofuran ring occurs by epoxidation followed by ring-opening and concomitant cyclization by an epoxide hydrolase<sup>16</sup>. The origins and the formation timing of the central cyclohexane and tetrahydropyran rings remain unknown, although a plausible metal-assisted cascade mechanism has been suggested in which both rings are formed in a single step following release of the polyketide chain from the PKS<sup>18,19</sup>.

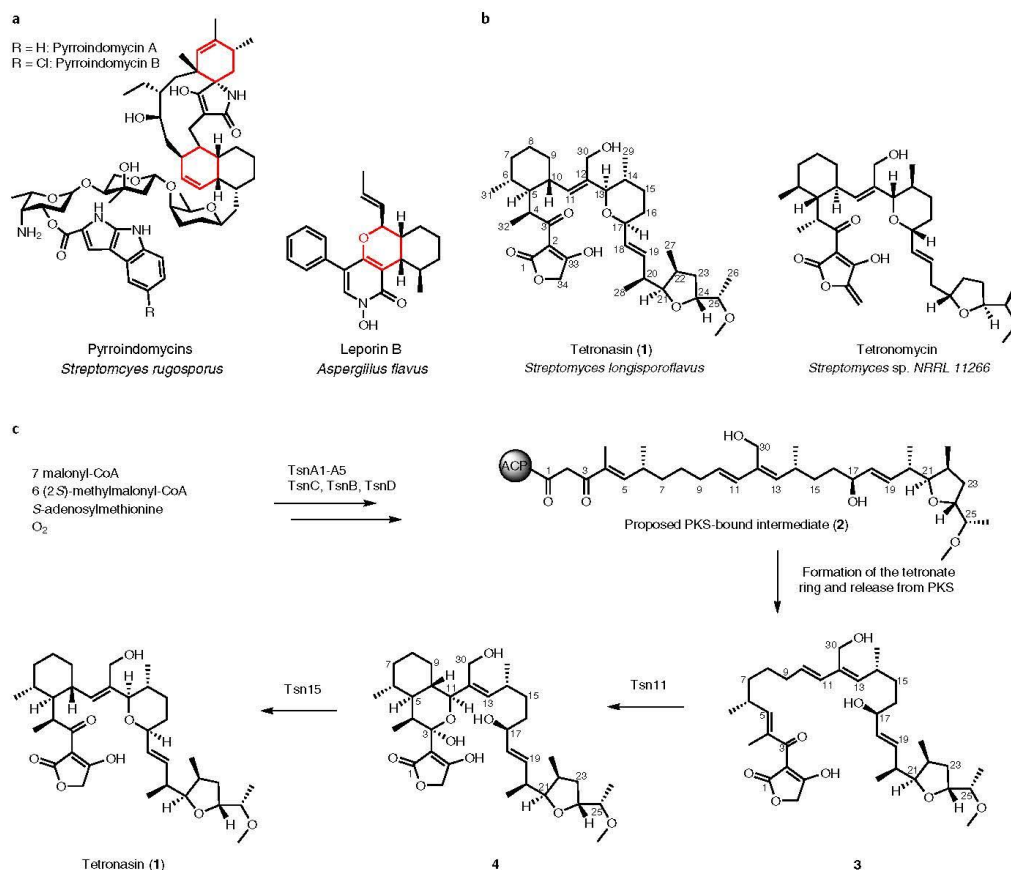
Intrigued by the biosynthetic mystery of the origin of the central cyclohexane and tetrahydropyran rings of tetronasin, herein we show that they are formed in a two-step enzymatic cascade reaction. The two enzymes responsible, Tsn11 and Tsn15, are homologues of [4+2] cyclase enzymes that are found in complex spirotreronate/tetramate pathways. Using a tetronasin precursor (3) isolated from a *S. longisporoflavus*  $\Delta$ tsn11 mutant, we show that Tsn11 catalyses an apparent inverse-electron-demand hetero-Diels–Alder reaction of this species to produce an unexpected oxadecalin intermediate (4); Tsn15 then catalyses a rearrangement that forms the tetrahydropyran ring and dismantles the oxadecalin moiety, producing tetronasin. To gain insight into the structure and mechanism of Tsn15, the crystal structure of a Tsn15-substrate complex was solved at 1.7 Å. Mutagenesis experiments then indicated that—akin to its homologues that catalyse [4+2] cycloadditions—Tsn15 also uses a pericyclic mechanism for ring formation.

### Results

**The PKS-bound linear intermediate of tetronasin.** We have reanalysed the previously sequenced tetronasin (*tsn*) biosynthetic gene cluster (GenBank: FJ462704; Supplementary Fig. 1) and used domain analysis of the PKSs to predict the structure of the hypothetical full-length polyketide (Supplementary Fig. 2). To investigate whether further enzymatic reactions tailor the structure before release from the PKS enzymes, we used a chemical probe (methyl 6-decanamido-2-fluoro-3-oxohexanoate (6), ref. 20) that is designed to intercept PKS-bound intermediates in vivo on the *tsn* PKS enzymes. We detected an incomplete undecaketide polyketide chain in which tetrahydrofuran ring formation, C25 O-methylation and hydroxylation at C30 had all apparently occurred, forming 2, the postulated PKS-bound intermediate (Fig. 1c, Supplementary Fig. 3). Although we cannot rule out that

<sup>1</sup>Department of Biochemistry, University of Cambridge, Cambridge, UK. <sup>2</sup>Department of Microbiology, Institute of Biomedical Sciences II, University of São Paulo, São Paulo, Brazil. <sup>3</sup>Department of Chemistry, University of Warwick, Coventry, UK. <sup>4</sup>Key Laboratory of Combinatorial Biosynthesis and Drug Discovery, Ministry of Education, School of Pharmaceutical Sciences, Wuhan University, Wuhan, China. <sup>5</sup>Department of Chemistry, University of Cambridge, Cambridge, UK. <sup>6</sup>Present address: Bicycle Therapeutics Limited, Babraham Research Campus, Cambridge, UK. \*e-mail: pfl10@cam.ac.uk





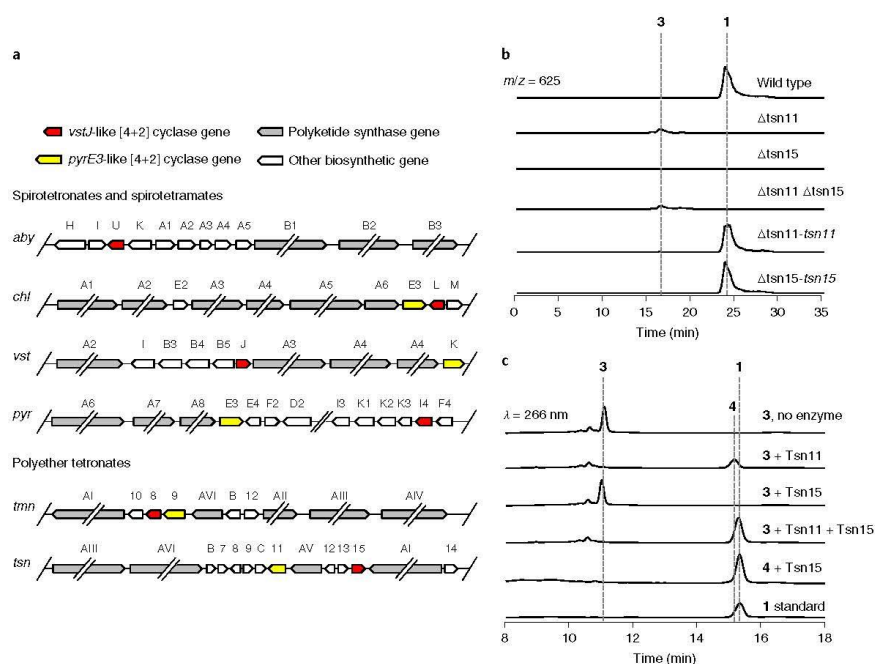
**Fig. 1** | [4+2] cyclases in polyether tetronate biosynthesis. **a**, The structures of the pyrroindomycins and leporin B, two polyketide natural products that require a [4+2] cyclase in their biosynthesis to create the rings highlighted in red. **b**, The structures of the polyether tetronate antibiotics tetronasin (1) and tetronomycin. **c**, The proposed role of the [4+2] cyclase homologues Tsn11 and Tsn15 in tetronasin biosynthesis.

these enzymatic modifications occurred after being offloaded by 6, the detection of 2 strongly suggests that they occur before polyketide chain release from the PKS enzymes, as has been previously established for other polyethers<sup>21</sup>.

**Homologues of [4+2] cyclase present in the tetronasin biosynthetic gene cluster.** The tetronasin cluster includes *tsn11*, a gene which is the homologue of *tmn9* (38% identity at the protein sequence level); the latter has previously shown to be essential for tetronomycin biosynthesis<sup>16</sup>. Both of these gene products show considerable sequence similarity to the monooxygenase-like [4+2] cyclases that catalyse dialkyldecalin formation in the biosynthetic pathways to the spirotetramate pyrroindomycin (PyrE3)<sup>11</sup> and the spirotetronates versipelostatin (VstK)<sup>3</sup> and chlorothricin (ChIE3)<sup>11</sup> (42–44% protein sequence identity) (Fig. 2a, Supplementary Fig. 4). Similar to PyrE3, Tsn11 contains mutations in several of the four positively charged amino acids (typically arginine) that are involved in interacting with NADPH (refs. 22,23), immediately suggesting that it is not a functional monooxygenase (Supplementary Fig. 4).

Co-located with *tsn11* is a second, previously unannotated gene, *tsn15*, which also has a counterpart in the *tmn* cluster (*tmn8*)<sup>16</sup>. The products of the latter two genes resemble a second family of Diels-Alderase-like cyclases (VstJ, PyrI4 and AbyU), which create the spiro moiety in spirotetronates and spirotetramates<sup>9–11</sup> (Fig. 2a, Supplementary Fig. 5).

**Tsn11 and Tsn15 are essential for tetronasin biosynthesis.** To analyse the potential roles of Tsn15 and Tsn11 in the biosynthesis of tetronasin (which contains neither a dialkyldecalin nor a spirotetronate moiety), we specifically deleted each gene in *S. longisporoflavus* (Supplementary Fig. 6). High performance liquid chromatography–mass spectrometry (HPLC–MS) analysis of fermentation extracts (Fig. 2b) showed that deletion of either gene abolished tetronasin production and that complementation in trans using the wild-type gene restored production in both cases. In the  $\Delta$ tsn11 mutant, a new metabolite (3) with the same molecular weight as 1 accumulated that was also produced by a  $\Delta$ tsn11  $\Delta$ tsn15 double mutant (Fig. 2b and Supplementary Fig. 6c), implying that Tsn11 acts before



**Fig. 2 | Functional characterization of the Diels-Alderase homologues Tsn11 and Tsn15 in tetronasin biosynthesis.** **a**, The biosynthetic gene clusters of tetronomycin (*tmn*) and tetronasin (*tsn*) encode *vstJ*- (red) and *pyrE3*-like (yellow) [4+2] cyclases that are characteristic of spirotetronate and spirotetramate biosynthesis pathways, including abyssomicin (*aby*), chlorothricin (*chl*), versipolostatin (*vst*) and pyrroindomycin (*pyr*). Only partial biosynthetic gene clusters are displayed and individual genes are not shown to scale. **b**, An HPLC-MS analysis of the production of tetronasin **1** ( $m/z = 625$ ) and intermediate **3** ( $m/z = 625$ ) from wild type *S. longisporoflavus*; *S. longisporoflavus*  $\Delta$ tsn11,  $\Delta$ tsn15 and  $\Delta$ tsn11  $\Delta$ tsn15 gene deletion mutants, and genetically complemented *S. longisporoflavus* deletion mutants. Data are representative of three independent experiments. **c**, An HPLC analysis of the in vitro conversion of intermediate **3** to tetronasin **1** using purified Tsn11 and Tsn15. Data are representative of three independent experiments.

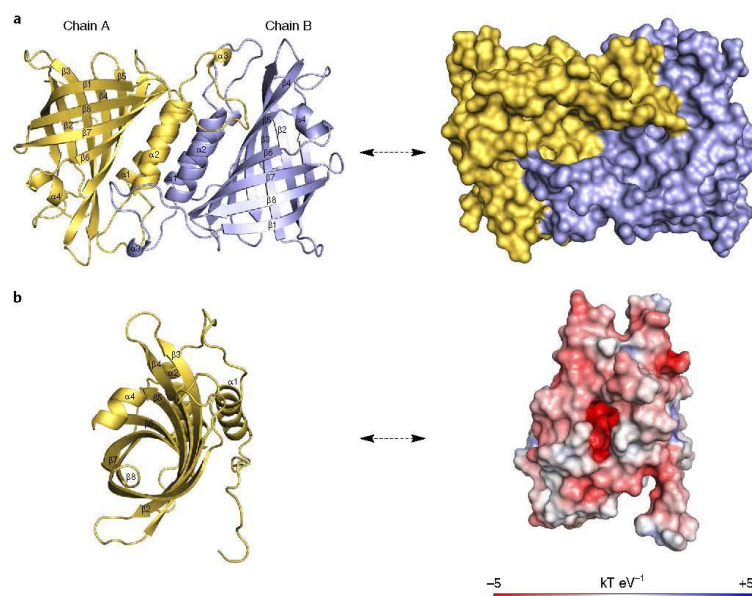
Tsn15. The new metabolite was isolated from the  $\Delta$ tsn11 mutant and its structure was determined to be **3** by MS and NMR (Fig. 1c, Supplementary Figs. 7a, 8, 9 and 24–29, Supplementary Table 1 and Supplementary Note 1). The new metabolite **3** differs from **1** in that it lacks both of the central six-membered rings of tetronasin, directly implicating Tsn11 and Tsn15 as catalysing ring formation in the final steps of the biosynthetic pathway.

**In vitro reconstitution of the cyclohexane and tetrahydropyran rings of tetronasin.** To attempt in vitro reconstitution of **1** from **3**, Tsn11 and Tsn15 were individually expressed and purified from *Escherichia coli* (Supplementary Fig. 10); Tsn11 was found to have flavin adenine dinucleotide as a tightly bound co-factor (Supplementary Fig. 10b). Incubation of **3** with Tsn11 in vitro led to its disappearance within 1 h (Fig. 2c) and HPLC-MS analysis showed the accumulation of a new compound **4**. Tsn11 could not oxidize NADPH or NADH and assaying the activity of Tsn11 pre-incubated with sodium dithionite (to reduce the bound flavin adenine dinucleotide) showed that changing its redox state did not inhibit the reaction of Tsn11 with **3** (Supplementary Fig. 11), which is consistent with previous reports<sup>1,22</sup> that bound flavin adenine dinucleotide is a passive spectator in the active site of flavocyclase PyrE3 in pyrroindomycin biosynthesis. The incubation of **3** with Tsn11 was scaled up to isolate larger amounts of **4**, and its structure was determined by MS and NMR analysis (Fig. 1c, Supplementary Figs. 7b, 12, 13

and 30–35, Supplementary Table 2 and Supplementary Note 2). Similar to **3**, compound **4** still lacks the tetrahydropyran ring but the cyclohexane ring has closed within an unexpected oxadecalin hemiacetal moiety. The formation of an oxadecalin intermediate by Tsn11 is strikingly reminiscent of the formation of carbocyclic decalin rings that are catalysed by its [4+2] cyclase homologues (PyrE3, ChlE3) in spirotetronate/spirotetramate biosynthesis<sup>11</sup>.

Conversion of **4** to **1** requires dehydration of the hemiacetal, fragmentation of the oxadecalin ring and a bond to form between C17-OH and C13 to establish the tetrahydropyran ring. Remarkably, incubation of **4** with Tsn15 alone—or of **3** with both Tsn11 and Tsn15—led to complete conversion to **1** in 10 min (Fig. 2c and Supplementary Figs. 13 and 14). Neither the Tsn11- nor the Tsn15-catalysed reaction proceeded in the absence of enzyme. Despite the evident similarity between the tetronasin and tetronomycin pathways, purified Tmn9 could not substitute for Tsn11 in the assay, nor could purified Tmn8 substitute for Tsn15 (Supplementary Fig. 16).

To probe the timing of cyclohexane and tetrahydropyran formation, the  $\Delta$ tsn11 and  $\Delta$ tsn15 mutant strains of *S. longisporoflavus* were treated in vivo with the chemical probe **6** (ref. 20). Exactly the same partially assembled polyketide species were observed as from the wild-type cells (Supplementary Fig. 3). These data are consistent with the view that tetronate and tetrahydrofuran ring formation, C30-hydroxylation and C25-O-methylation may precede tetrahydropyran and cyclohexane formation. Therefore, **3** and **4**



**Fig. 3 | The crystal structure of Tsn15.** The crystal structure of Tsn15 was solved at 1.8 Å resolution using single anomalous dispersion phasing. **a**, Ribbon and surface representations of the homodimeric Tsn15 are shown. **b**, The active-site cavity in the  $\beta$ -barrel of Tsn15 is shown. The surface charge representation is contoured at  $\pm 5$  kT eV $^{-1}$  (blue/red), as shown to the right.

are probably true biosynthetic intermediates *in vivo* and the central six-membered rings of tetronasin are formed after polyketide chain release as final steps in the pathway. Between them, these two enzymes establish the configuration at four asymmetric centres in tetronasin<sup>18</sup>.

Cycloadditions of 1-oxa-1,3-butadienes through inverse-electron-demand hetero-Diels–Alder reactions are important transformations in synthetic chemistry<sup>25,26</sup>; thus, an enzymatic counterpart such as Tsn11 is of considerable interest. The formation of *trans*-oxadecalin **4** seems consistent with such a mechanism but a stepwise process cannot be ruled out (Supplementary Fig. 17). Alternative mechanisms are also possible for the reaction catalysed by Tsn15 to convert **4** to **1** (Supplementary Fig. 17). A previous isotope-feeding study established that the ketone oxygen atom at C3 is derived from propionate rather than water<sup>15</sup>, indicating that any mechanism in which the C3 oxygen is derived from water is unlikely (Supplementary Fig. 17).

**The structure of Tsn15.** We propose that Tsn11—akin to its spiro-tetronate/spiro-tetramate homologues—catalyses a [4+2] cycloaddition even though the resulting ring is latent; that is, it is not observed in the final antibiotic structure. The structure and mechanism of Tsn11 is currently under investigation. To gain insight into the specificity and mechanism of the unusual transformation catalysed by Tsn15, we have solved the crystal structure of Tsn15 at 1.8 Å resolution using single anomalous dispersion phasing (Fig. 3), and also that of a Tsn15–substrate 1:1 complex (obtained by co-crystallization of Tsn15 and **4**) at 1.7 Å (Fig. 5).

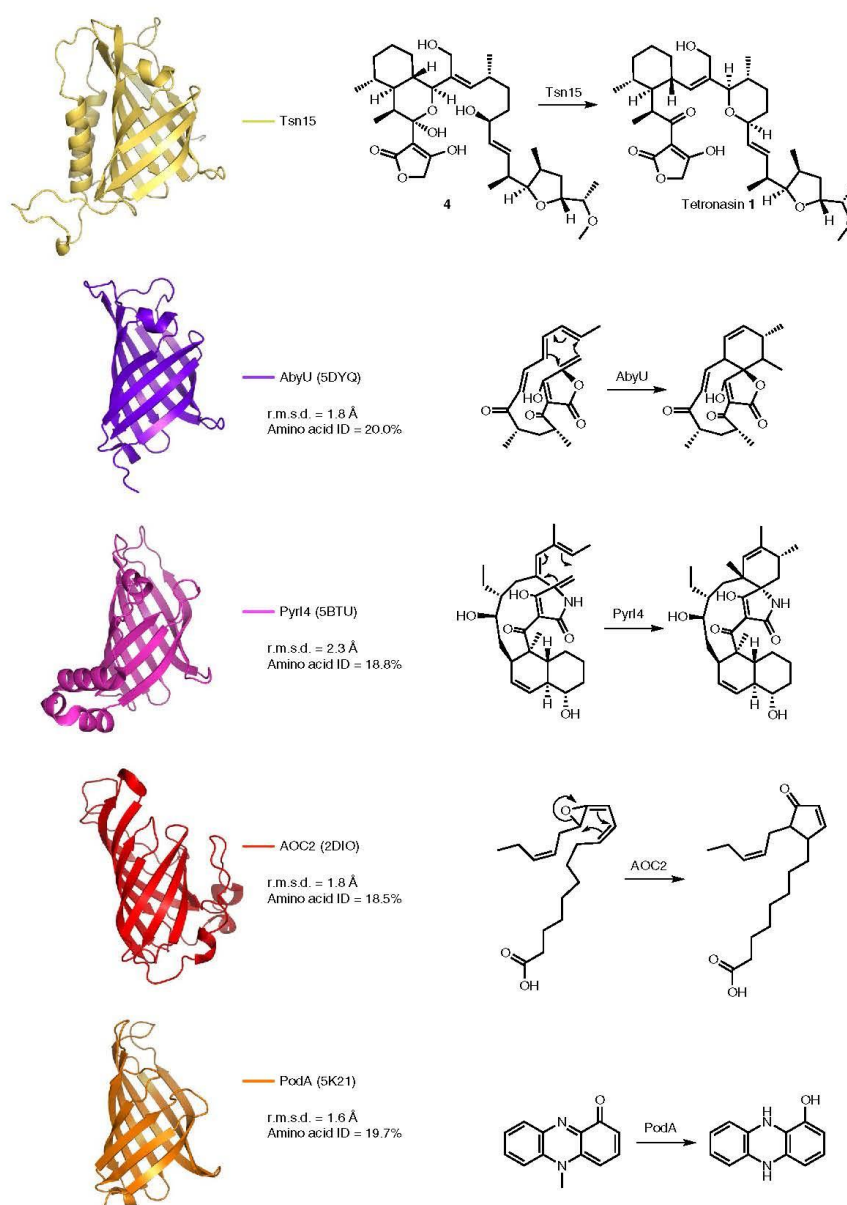
Consistent with its quaternary structure in solution (Supplementary Fig. 18), Tsn15 crystallized as a homodimer of two antiparallel-facing monomers (Fig. 3a). Each monomer chain comprises an *N*-terminal  $\alpha$ -helical dimerization region ( $\alpha 1$ – $\alpha 3$ ) that

is followed by an eight-strand  $\beta$ -barrel ( $\beta 1$ – $\beta 8$ ,  $\alpha 4$ ) enclosing a 388 Å<sup>3</sup> hydrophobic internal cavity (Fig. 3b, Supplementary Figs. 19 and 20). Each Tsn15 monomer shows the conserved domain fold that has previously been seen in the [4+2] cyclases of spiro-tetronate and spiro-tetramate biosynthesis (PyrI4 and AbyU)<sup>16,27</sup>, despite low sequence identity (approximately 20%). In addition to these homodimeric [4+2] cyclases, Tsn15 also shares its fold with the homotrimeric allene oxide cyclase enzymes (which catalyse a 4 $\pi$  electrocyclicization (pericyclic) reaction in jasmonic acid biosynthesis)<sup>28,29</sup> and PodA (an unusual pyocyanin demethylase from *Mycobacterium fortuitum*)<sup>30</sup>. The different activities of these homologous enzymes demonstrate remarkable ability of this protein fold to catalyse a diverse range of reactions (Fig. 4).

There are no substantial conformational differences between the  $\alpha$  chains of the two crystal structures (Supplementary Fig. 21). In the Tsn15–substrate co-crystal the hemiacetal C3 hydroxyl of **4** had undergone dehydration to produce **5** (Fig. 5a). The tetronate ring of **5** is fully inserted into the  $\beta$ -barrel cavity of Tsn15, with the tetrahydrofuran moiety protruding into the solvent (Fig. 5b). The side chains of amino acid residues S142 and Q164 form hydrogen bonds with the tetronate of **5** (Fig. 5c) and R89 is hydrogen-bonded to the hydroxyl group at C30 (chain B only) (Supplementary Fig. 22).

Our initial hypothesis was that Tsn15 might use acid/base catalysis to activate the C17 hydroxyl and promote nucleophilic ring closure of the tetrahydropyran ring. However, mutagenesis of individual acidic and basic amino acid residues in the  $\beta$ -barrel (R89, E109 and D122) showed that none of them are essential for Tsn15 activity (Supplementary Fig. 23). Further alanine mutagenesis highlighted W190 (essential) and Q164 (the mutant shows ~30% wild-type activity), both of which are probably important for substrate binding and selectivity (Supplementary Fig. 23); W190 seems to orient the oxadecalin moiety of **5** through  $\pi$ -stacking interactions (an equivalent

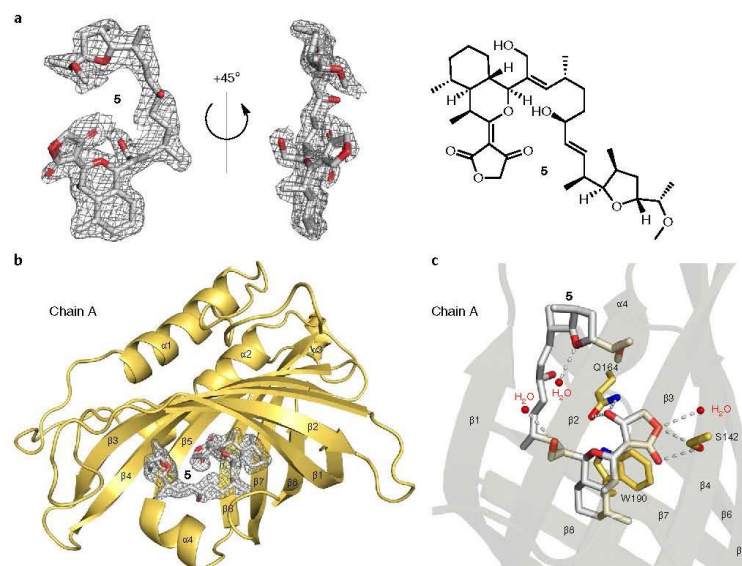




**Fig. 4 | Structural homologues of Tsn15 and their respective reactions.** The  $\alpha$  chain of Tsn15 was aligned with its structural homologues AbyU, Pyl14, AOC2 and PodA. The r.m.s.d. and the amino acid ID (%) in comparison with Tsn15, and the reaction that each homologue catalyses, are shown next to each structure.

tryptophan is conserved in Tmn8, suggesting a conserved role in binding oxadecalin polyether tetronate intermediates (Fig. 5c, Supplementary Fig. 5). The twelve *N*-terminal amino acids of Tsn15 do not form an essential substrate-enclosing lid, as removing these

residues did not abolish enzyme activity as in the case of Pyl14 (ref. 27; Supplementary Fig. 23). The inability to identify essential general acid/base catalytic residues suggests that the active site of Tsn15 is primarily a hydrophobic pocket that promotes a reactive conformation,



**Fig. 5 | Structure of Tsn15 and a Tsn15-substrate complex.** **a**, The  $F_o - F_c$  of **5** (left, contoured to  $2.0 \sigma$ ) and its chemical structure (right). **b**, Tsn15 in complex with intermediate **5**. The electron density difference map ( $F_o - F_c$ ) of **5** could be clearly assigned in the  $\beta$ -barrel cavity of Tsn15 chain A. The contouring is to  $2.0 \sigma$ . **c**, Detailed binding interactions between intermediate **5** and the internal  $\beta$ -barrel residues of Tsn15 (chain A) are shown. The dotted grey lines represent hydrogen bonds.

possibly producing the tetrahydropyran ring via a pericyclic rearrangement (Fig. 6) similar to the mechanisms proposed for its homologue AOC2 (ref. 28) and the [4+2] cyclases PyrI4 (ref. 27) and AbyU (ref. 19). One complication with this proposal is that the conformation of **5** in the active site seems to be inert for tetrahydropyran formation, with the atoms that must form a bond (C13 and the oxygen of the C17 hydroxyl) being 4.6 Å apart. A possible explanation for this could be that the conformational rigidity imposed on Tsn15 when it crystallized prevented **5** from adopting a near attack conformation. It is also notable that **5** is present in the active site of Tsn15 rather than **4** or tetronasin **1**. We had thought it more likely that **4** would be converted by Tsn15 and tetronasin **1** would therefore co-crystallize in its active site, analogous to the PyrI4 cocrystallization experiments<sup>27</sup>. One explanation for this could be that **5** is an inert spontaneous dehydration product of **4**, the true substrate. If **4** is the true substrate, however, then tetrahydropyran formation and dehydration at C3 must occur simultaneously (Supplementary Fig. 17b), which would probably require an acidic residue to donate a proton to the hemiacetal hydroxyl, thereby creating a better leaving group. However, as demonstrated, the acidic groups in the active site are not essential and no acidic amino acid is adjacent to C3, so it is not clear how this could occur. Although we cannot rule out **5** being inert, in our view, **5** is more likely to be the Tsn15 substrate as the dehydration of **4** to form **5** should be a favourable first step for tetronasin formation. Dehydration of the hemiacetal hydroxyl (whether spontaneous or Tsn15 catalysed) would stabilize the anion that forms at the C3 oxygen during tetrahydropyran formation through conjugation with the C1 and C33 carbonyl groups (Supplementary Fig. 17a,c,e). Although the exact mechanism used by Tsn15 to convert **4** to tetronasin remains unknown, our data are more consistent with Tsn15 catalysing a pericyclic rearrangement rather than a nucleophilic acid/base reaction (Fig. 6). Regardless, the crystal structure of Tsn15 with

**5** in its active site delineates the substrate-binding residues of Tsn15 and provides a robust foundation for conducting future mechanistic studies on this fascinating class of cyclase.

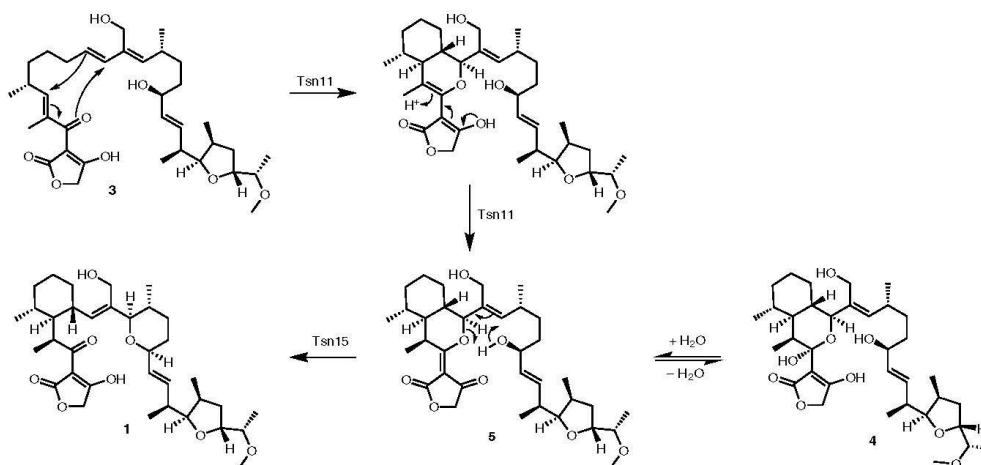
Our results show that the standalone enzyme Tsn11 catalyses a previously unsuspected intramolecular [4+2] cycloaddition that, in formal terms, represents an inverse-electron-demand hetero-Diels-Alder reaction<sup>26</sup> and constitutes a latent but compelling mechanistic link with spirotetronate and spirotetramate biosynthesis. Likewise, Tsn15, which is structurally homologous to known Diels-Alderase-like cyclases, has evolved to catalyse a different (but possibly also a pericyclic) transformation. From a synthetic biology perspective, these enzymes could be useful for stereospecifically synthesizing cyclohexane and tetrahydropyran moieties in synthetic polyether ionophore analogues. The tolerance of Tsn15 to mutations in its active site and the diverse range of reactions catalysed by its protein fold (Fig. 4) further suggests it may have some potential as a biotransformation catalyst. Meanwhile, these findings suggest that further examples of novel—and potentially latent—enzymatic pericyclic reactions remain to be discovered in natural product biosynthetic pathways.

## Methods

**Materials, bacterial strains and culture conditions.** All chemicals and biomaterials were purchased from Merck or Thermo Fisher unless otherwise stated. All of the enzymes were purchased from Thermo Fisher or New England Biolabs. All of the *E. coli* strains were grown in liquid or on solid (1.5% agar) lysogeny broth medium at 37 °C. For general maintenance of strain stocks, *S. longisporoflavus* was grown at 30 °C and 220 r.p.m. in trypticase soy broth with yeast extract (30 g l<sup>-1</sup> tryptic soy broth, 5 g l<sup>-1</sup> yeast extract, 103 g l<sup>-1</sup> sucrose) or at 30 °C on soya flour mannitol (SFM) agar (20 g l<sup>-1</sup> soya flour, 20 g l<sup>-1</sup> D-mannitol, 20 g l<sup>-1</sup> agar).

**General DNA manipulation techniques.** The PCR amplification of DNA was performed using Q5 DNA polymerase (New England Biolabs). DNA cloning was performed using standard restriction digestion or isothermal DNA assembly.





**Fig. 6 | The proposed mechanism for formation of the cyclohexane and tetrahydropyran rings of tetronasin.** Tsn11 catalyses an inverse-electron-demand hetero-Diels-Alder reaction to form an oxadecalin intermediate. The oxadecalin is hydrated to form the cyclic hemiacetal **4** that was purified from the Tsn11 + **3** assay and structurally characterized. Tsn15 then catalyses dehydration of **4** to form **5** (which crystallized in the active site of Tsn15), followed by a pericyclic rearrangement to form the tetrahydropyran ring and fragment the oxadecalin, producing tetronasin **1**.

Sanger sequencing was used to confirm the sequences of all cloned DNA. Genomic DNA from *S. longisporoflavus* was isolated by chloroform extraction followed by isopropanol precipitation and washing using 70% ethanol. Mutagenesis of *tsn15* was performed for *tsn15*-E109A, *tsn15*-D122A and *tsn15*-W190A using a Quikchange II mutagenesis kit (Agilent). The remaining *tsn15* mutant constructs were ordered from GenScript.

#### Deletion and complementation of *tsn11* and *tsn15* in *S. longisporoflavus*.

The deletions of *tsn11* and *tsn15* were made in the coding region to avoid introducing a frameshift that could have deleterious polar effects on downstream gene expression. Genomic DNA of *S. longisporoflavus* was purified and used as a template for all PCR amplifications necessary to create a *tsn11* gene-deletion plasmid construct. The *tsn11*\_Up\_Fw/*tsn11*\_Up\_Rv and *tsn11*\_Dn\_Fw/*tsn11*\_Dn\_Rv primers were used to amplify 2 kbp stretches of the genomic DNA upstream and downstream of *tsn11*, respectively. The primers contained additional nucleotides on their 5' ends that enabled the two fragments to be joined and ligated to the *Nde*I site of pYH7 using isothermal DNA assembly<sup>31</sup>, producing the recombinant construct pYH7-*tsn11* in which 1,356 nt of the *tsn11* coding frame was deleted. The pYH7-*tsn11* construct was transformed to *E. coli* ET12657/pUZ8002<sup>32</sup> cells that were mixed into a fresh liquid culture of *S. longisporoflavus*. After mixing, the cell mixture was plated on SFM agar that was supplemented with 20 mM MgCl<sub>2</sub> and incubated at 30 °C for 16 h. The SFM agar was then overlaid with 1 ml of sterilized water that contained nalidixic acid and apramycin (0.875 mg of each per 35 ml plate). After several days of growth at 30 °C, apramycin-resistant *S. longisporoflavus* exconjugants that featured a single homologous cross-over event with pYH7-*tsn11* were observed. These single colonies were grown on SFM agar that lacked the antibiotics to promote a second recombination event via plasmid loss. The *S. longisporoflavus* exconjugants that had undergone a double homologous cross-over event (those that had regained apramycin sensitivity) were screened using *tsn11*\_KO\_Fw/*tsn11*\_KO\_Rv primers to detect the 572 bp PCR product that is indicative of a successful genomic deletion of *tsn11*. The PCR product was sequenced to confirm the *S. longisporoflavus*  $\Delta$ *tsn11* genotype.

For the deletion of Tsn15, genomic DNA of *S. longisporoflavus* was purified and used as a template for all of the PCR amplifications necessary to create a *tsn15* gene-deletion plasmid construct. The *tsn15*\_Up\_Fw/*tsn15*\_Up\_Rv and *tsn15*\_Dn\_Fw/*tsn15*\_Dn\_Rv primers were used to amplify 2 kbp stretches of the genomic DNA upstream and downstream of *tsn15*, respectively. The primers contained additional nucleotides on their 5' ends that enabled the two fragments to be joined and ligated to the *Nde*I site of pYH7 by Gibson assembly<sup>31</sup>, producing the recombinant construct pYH7-*tsn15* in which the entire 621 bp *tsn15* coding frame was deleted. The protocol for deleting *tsn15* was identical to that previously described for creating *S. longisporoflavus*  $\Delta$ *tsn11*. Exconjugants that had undergone two homologous crossing-over events with pYH7-*tsn15* were screened using *tsn15*\_KO\_Fw/*tsn15*\_KO\_Rv primers to detect the 713 bp PCR product indicative

of a successful genomic deletion of *tsn15*. The PCR product was sequenced to confirm the *S. longisporoflavus*  $\Delta$ *tsn15* genotype.

The *S. longisporoflavus*  $\Delta$ *tsn11*  $\Delta$ *tsn15* double mutant was created by deleting *tsn15* in *S. longisporoflavus*  $\Delta$ *tsn11*. The same protocol for deleting *tsn15* was followed, except that *S. longisporoflavus*  $\Delta$ *tsn11*—rather than wild-type *S. longisporoflavus*—was conjugated with pYH7-*tsn15*.

To complement the *S. longisporoflavus*  $\Delta$ *tsn11* mutant, *tsn11* was amplified from the genomic DNA of *S. longisporoflavus* using *tsn11*-pIB139-Fw/*tsn11*-pIB139-Rv primers. The amplified *tsn11* was then cloned by isothermal DNA assembly at the *Xba*I site of pIB139 (downstream of the *ermE*<sup>+</sup> promoter)<sup>33</sup>, creating pIB139-*tsn11*; pIB139-*tsn11* was conjugated to *S. longisporoflavus*  $\Delta$ *tsn11* using *E. coli* ET12657/pUZ8002. Apramycin-resistant exconjugants were selected and grown for 7 d before being subjected to HPLC-MS analysis. To complement the *S. longisporoflavus*  $\Delta$ *tsn15* mutant, *tsn15* was amplified from the genomic DNA of *S. longisporoflavus* using *tsn15*-pIB139-Fw/*tsn15*-pIB139-Rv primers. The amplified *tsn15* was then cloned by isothermal DNA assembly at the *Xba*I site of pIB139 (downstream of the *ermE*<sup>+</sup> promoter), creating pIB139-*tsn15*; pIB139-*tsn15* was conjugated to *S. longisporoflavus*  $\Delta$ *tsn15* using *E. coli* ET12657/pUZ8002.

**Analysis of metabolites from *S. longisporoflavus*.** To test for the production of either tetronasin **1** or intermediate **3**, a colony of *S. longisporoflavus* (or a *S. longisporoflavus* deletion mutant/complementation strain) was inoculated into tsn medium A (30 g l<sup>-1</sup> tryptic soy broth, 3 g l<sup>-1</sup> CaCO<sub>3</sub>, trace elements: 4 mg l<sup>-1</sup> FeSO<sub>4</sub>, 4 mg l<sup>-1</sup> ZnSO<sub>4</sub>, 0.6 mg l<sup>-1</sup> CuSO<sub>4</sub>, 0.4 mg l<sup>-1</sup> MnSO<sub>4</sub>, 0.4 mg l<sup>-1</sup> KMoO<sub>4</sub>) and grown for 2 d at 30 °C and 200 r.p.m. The culture was then plated onto petri dishes that contained tsn medium B (30 g l<sup>-1</sup> tryptic soy broth, 3 g l<sup>-1</sup> CaCO<sub>3</sub>, 100 g l<sup>-1</sup> dextrin, 20 g l<sup>-1</sup> agar, trace elements: 4 mg l<sup>-1</sup> FeSO<sub>4</sub>, 4 mg l<sup>-1</sup> ZnSO<sub>4</sub>, 0.6 mg l<sup>-1</sup> CuSO<sub>4</sub>, 0.4 mg l<sup>-1</sup> MnSO<sub>4</sub>, 0.4 mg l<sup>-1</sup> KMoO<sub>4</sub>) and grown at 30 °C for 7 d. The agar and cells were then extracted using ethyl acetate. The ethyl acetate was evaporated under reduced pressure and the organic extract was dissolved in methanol for analysis through HPLC (Hewlett Packard, Agilent Technologies 1200 series) coupled to a mass spectrometer (Thermo Finnigan MAT LTQ). The HPLC-MS was fitted with a 250 mm × 4.6 mm 5  $\mu$ m C18 reverse-phase column (5  $\mu$  OSD3, 100 Å; Phenomenex). The mobile phase comprised 20 mM ammonium acetate and increasing methanol (v/v) at a flow rate of 0.7 ml min<sup>-1</sup>: 0–5 min, 5–75% methanol; 5–30 min, 75–95% methanol, 30–34 min, 95% methanol 35–36 min, 95–5% methanol. A normalized collision energy of 35% was used for the molecular fragmentation of tetronasin. High-resolution mass spectra were obtained using a Vion IMS QTOF (Waters) operated by the MS service of the University of Cambridge Department of Chemistry.

**Heterologous expression of cyclase genes in *E. coli*.** The *tsn15*, *tsn11*, *tsn9* and *tsn8* genes were each amplified by PCR. The PCR products were

individually digested using *Nde*I and *Xho*I and cloned between the *Nde*I and *Xho*I sites of pET28a(+), in-frame with the N-terminal polyhistidine tag (MGSSHHHHSSGLVPRGSH). Each expression plasmid was individually used to transform chemically competent *E. coli* BL21 (DE3) cells (Thermo Fisher). A single colony from each *E. coli* BL21 (DE3) transformant was cultured at 37 °C and 200 r.p.m. in lysogeny broth that contained 50 µg ml<sup>-1</sup> kanamycin. Once the culture had reached  $A_{600} = 0.5$ , protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside to a final concentration of 0.5 mM. The culture was then incubated for 16 h at 20 °C and 200 r.p.m. The cells were harvested by centrifugation and resuspended in 30 ml of 20 mM Tris-Cl pH 7.9 buffer, 500 mM NaCl. The cells were lysed using an Emulsiflex C5 (Avestin) according to the manufacturer's instructions. The insoluble fraction was removed by centrifugation and the soluble fraction was passed through nickel-affinity resin (Qiagen) to bind the recombinant protein. Lysis buffer with increasing imidazole concentrations (50–400 mM) was used to remove contaminants and subsequently elute pure recombinant protein (monitored by SDS-PAGE). Fractions that contained the purified recombinant proteins were concentrated using Amicon Ultra-15 centrifugal filter units (Merck Millipore) and the buffer was exchanged to 20 mM Tris-Cl, 100 mM NaCl, 10% glycerol (pH 7.9) using a PD-10 column (GE Healthcare) according to the manufacturer's instructions. The protein concentration was determined by the Bradford assay using a bovine serum albumin standard curve. The protein aliquots were then snap frozen in liquid nitrogen and stored at -80 °C.

#### Protein quaternary structure analysis by analytical ultracentrifugation.

Analytical ultracentrifugation was used to measure the molecular weight of Tsn11 and Tsn15 in solution. Tsn15 and Tsn11 were purified and individually concentrated to 0.75 mg ml<sup>-1</sup> in a final volume of 800 µl. When Tsn11 and Tsn15 were sedimented together, each was present in the same chamber at a concentration of 0.75 mg ml<sup>-1</sup>. Analytical ultracentrifugation experiments were performed using an Optima XL-I (Beckman Coulter) centrifuge fitted with an An60 Ti four-hole rotor. Absorbance and interference data were acquired in the continuous mode at time intervals of 170 s, at a rotor speed of 40,000 r.p.m. and 20 °C with systematic noise subtracted, but without averaging. The density and viscosity of the buffer (20 mM Tris-Cl pH 7.9, 100/150 mM NaCl) and the partial specific volume of the protein were both calculated using Sednterp<sup>34</sup>. The multi-component sedimentation coefficient distributions were obtained from 128 scans by direct boundary modelling of the Lamm equation using Sedfit v14.1<sup>35</sup>.

**Isolation of Intermediate 3 for NMR analysis.** Intermediate 3 was isolated from *S. longisporoflavus* Δtsn11. Well-grown *S. longisporoflavus* Δtsn11 colonies were inoculated into 5 ml of tsn medium A and grown for 2 d at 30 °C and 200 r.p.m. The cultures were then inoculated into flasks that contained 50 ml of tsn medium A and grown for 2 d at 30 °C and 200 r.p.m., or until a thick mycelial culture had grown. These cells were plated on 400 12 cm × 12 cm agar plates that each contained 50 ml of tsn medium B agar (30 g l<sup>-1</sup> tryptic soy broth, 3 g l<sup>-1</sup> CaCO<sub>3</sub>, 100 g l<sup>-1</sup> dextrin, 20 g l<sup>-1</sup> agar, trace elements: 4 mg l<sup>-1</sup> FeSO<sub>4</sub>, 4 mg l<sup>-1</sup> ZnSO<sub>4</sub>, 0.6 mg l<sup>-1</sup> CuSO<sub>4</sub>, 0.4 mg l<sup>-1</sup> MnSO<sub>4</sub>, 0.4 mg l<sup>-1</sup> KMoO<sub>4</sub>) and were grown for 7 d at 30 °C. The agar was then cut into small squares and combined in a large glass flask where they were extracted three times by being submerged in 2 l of ethyl acetate. The ethyl acetate was evaporated under reduced pressure to yield 7 g of a brown crude organic extract. The crude extract was dissolved in 8 ml of methanol and loaded onto eight 10 g per 70 ml C18 reverse-phase isolate cartridges (Biotage) according to the manufacturer's instructions. Fractions were eluted from the columns using a mixture of 20 mM ammonium acetate and increasing methanol (v/v). Fractions containing intermediate 3 were identified by HPLC-MS and pooled together (1.15 g in total) before a second round of solid phase extraction purification. The final purification step was performed using an Infinity II semipreparative HPLC (Agilent) fitted with a Phenomenex C18 Prodigy column (5 µm ODS-3 100 Å, 250 × 10 mm). Gradient elution of intermediate 3 was achieved using a mobile phase of 5 mM ammonium acetate and methanol with a flow rate of 3 ml min<sup>-1</sup>: 0–5 min, 5–75% methanol; 5–15 min, 75–85% methanol; 15–19 min, 85–100% methanol; 19–20 min, 100–5% methanol. Fractions containing intermediate 3 were identified by detecting its characteristic chromophore ( $\lambda_{max} = 236$  nm) and pooled together, was followed by solvent evaporation under reduced pressure. The dried extract was dissolved in ethyl acetate to remove the ammonium acetate, followed by freeze drying to yield purified 3 (12.5 mg). For structural determination, 3 was dissolved in deuterated chloroform and analysed using a 500 MHz DCH Cryoprobe Spectrometer (Bruker).

**Isolation of intermediate 4 for NMR analysis.** Intermediate 4 was isolated from an in vitro reaction of 3 and Tsn11. The assay mixture (60 ml in total) contained 20 mM Tris-Cl, 100 mM NaCl (pH 7.9), 50 µM Tsn11, 400 µM of 3 and 5% v/v methanol. After incubation at 30 °C for 1.5 h, the reaction mixture was extracted six times with 30 ml of ethyl acetate. The ethyl acetate was evaporated and the organic residue was dissolved in methanol. The extract was chromatographed on semipreparative HPLC as described above. Fractions that contained intermediate 4 were identified by detecting its characteristic chromophore ( $\lambda_{max} = 254$  nm) and pooled together for solvent evaporation. The dried residue was dissolved in ethyl acetate to remove the ammonium acetate, followed by freeze drying

to yield purified 4 (3.3 mg). For the structural determination, 4 was dissolved in deuterated methanol and analysed using a 500 MHz DCH NMR Cryoprobe Spectrometer (Bruker).

**In vitro assays.** In vitro activity assays of Tsn11 and Tsn15 were performed in 20 mM Tris-Cl buffer, 100 mM NaCl, 5% v/v methanol (pH 7.9). Typically, reaction volumes were 100 µl and contained 200 µM of 3 or 4; Tsn11 and/or Tsn15 were each added to a final concentration of 5 µM. The reactions were carried out at 30 °C for 1 h unless stated otherwise. The reactions were terminated by adding 400 µl of methanol before being completely dried under reduced pressure. The dried extract was redissolved in 100 µl methanol and analysed by HPLC fitted with a Phenomenex C18 Prodigy column (5 µm ODS-3 100 Å, 250 × 10 mm) using a gradient programme of 20 mM ammonium acetate and increasing methanol (v/v) at a flow rate of 3 ml min<sup>-1</sup>: 0–5 min, 5–75% methanol; 5–15 min, 75–85% methanol; 15–19 min, 85–100% methanol; 19–20 min, 100–5% methanol. For all of the in vitro assays, the identities of intermediates 3 and 4 were confirmed by their unique ultraviolet chromophores and MS<sup>n</sup> fragmentation patterns. For analysis of the Tsn11 sodium dithionite assays and the Tsn15 point mutation assays, a Poroshell 120, EC-C18, 27 µm, 46 × 100 mm (Agilent) column was eluted with 5 mM ammonium acetate and methanol at 1 ml min<sup>-1</sup>: 0–5 min, 5–75% methanol; 5–15 min 75–95% methanol; 15–19 min, 95% methanol; 19–20 min, 95–5% methanol. When the Tsn11 activity assay was conducted in the presence of sodium dithionite, 100 µM of 3 was used with 1 µM of Tsn11 and 1 mM of sodium dithionite. Parallel reactions were carried out and terminated at 0, 2, 4 and 8 min, respectively.

**Protein crystallization.** For protein crystallization, *tsn15* was expressed in *E. coli* BL21 (ΔDE3) and purified by nickel-affinity chromatography as previously described. The protein was further purified using a preparative gel-filtration ÅKTA (GE Healthcare Lifescience) connected to a Hiloal 16/60 column packed with Superdex 200 resin. Fractions that contained Tsn15 were pooled and concentrated; Tsn15 was concentrated to 15 mg ml<sup>-1</sup> (20 mM Tris-Cl pH 7.9 buffer, 100 mM NaCl) and crystallized in 0.1 M PCTP (pH 6.0) (0.1 M of sodium propionate, sodium cacodylate trihydrate and bis-Tris propane), 25% polyethylene glycol 1500 using the sitting-drop vapour-diffusion method at 19 °C. The Tsn15-substrate complex was crystallized using the hanging-drop vapour-diffusion method at 18 °C. To obtain the protein complex, a solution composed of Tsn15 protein (15 mg ml<sup>-1</sup> in 20 mM Tris-Cl pH 7.9 buffer, 0.5 M NaCl) and 10 mM of 4 was incubated for 10 min and then crystallized in 0.1 M PCTP pH 6.0 buffer 27% polyethylene glycol 1500. To solve the phases for the Tsn15 structure, a selenomethionine labelled version of Tsn15 (SeMet-Tsn15) was created and the structure was determined using the single anomalous dispersion method. To achieve this, *E. coli* BL21 (ΔDE3) cells transformed with pET28a(+)-*tsn15* were cultured in M9 medium. Once the culture had reached  $A_{600} = 0.5$ , 0.1 g l<sup>-1</sup> L-lysine, 0.1 g l<sup>-1</sup> L-threonine, 0.1 g l<sup>-1</sup> L-phenylalanine, 0.05 g l<sup>-1</sup> L-leucine, 0.05 g l<sup>-1</sup> L-isoleucine and 0.05 g l<sup>-1</sup> L-valine were added to the growing culture, followed by 0.06 g l<sup>-1</sup> of L-selenomethionine. SeMet-Tsn15 was purified and exchanged into a buffer that contained 20 mM Tris-Cl pH 7.9, 100 mM NaCl, 2 mM EDTA and 2 mM tris(2-carboxyethyl)phosphine, then crystallized as before. Crystals were harvested from the crystallization drop and transferred to a cryoprotecting solution that contained the crystallization solution and 25–27% ethylene glycol. The crystals were then flash frozen in liquid nitrogen using nylon loops. X-ray diffraction data for SeMet-Tsn15 were collected at the Diamond Light Source I04 beamline, whereas for the complex Tsn15-substrate data were collected at PETRA III, beamline 13. The X-ray data processing was performed using XDS<sup>36</sup> and scaled using Aimless<sup>37</sup> from CCP4 suite<sup>38</sup>. The phases for SeMet-Tsn15 was solved using AutoSol<sup>39</sup> from Phenix suite<sup>40</sup>. The structural model was initially built by AutoBuild<sup>41</sup> also from PHENIX suite<sup>40</sup>. The phases for the Tsn15-substrate complex was determined by molecular replacement using the program Phaser<sup>42</sup> from Phenix suite<sup>40</sup>. The SeMet-Tsn15 structure was used as the search model for molecular replacement. Both SeMet-Tsn15 and Tsn15-substrate complex models were refined using Phenix.refine<sup>43</sup> and further visual inspection and real space refinement was performed by COOT<sup>44</sup>. The stereochemical quality of the models was assessed using MolProbity<sup>45</sup>. The protein structure figures were prepared using PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC<sup>46</sup>.

**The revised sequence of Tsn11.** Closer examination of the *tsn11* gene compared to the deposited sequence revealed the start codon to be an additional 12 bp 5' of the previously annotated start codon. The new start codon was a methionine (ATG) rather than a leucine (TTG) and was 6 bp downstream from a likely Shine-Dalgarno sequence (GGAAGAA). The revised protein sequence of the updated Tsn11 is presented below.

Tsn11:  
MEIPLTGVTVIAGAGPVGLFLASELRLAGVEAVVLEERSPKANEHT  
VGGTLHARTADLFDQRGIMDTLRAGNPLPWRPLHFA5YWLDLAPH  
MEDEYSLLLPQQYTEEMLEAHATELGADIRRGHTCVSLTQDADGVT  
VGVRADSGDYELRGAYLVGCDGGDSTVRELAAPFVQESGPRWYGLA  
DVESIEGDWHPGNYPGGQFAVIRSPHEGGYRSMITLEFNETTQPQAD  
QPVSVEEVIATERITGRTPVVGEVQWLHRYTNTTREAENYRQGRVFA  
GDAHLHVAFAHGISTGLHDAANLWVKLAAVLDGRAPDSLLDTYDE  
ERRPVGHRAVCFTQSMALLTQGGQLDILRQLFTDLVVKLPEVNHLLIT



VTVRYALDGAEKEDTHPLLRPVNQLVKDADGQATAVAEALRAGR  
VLIDLTDGAAALPDTSGRRGHLDVSRGPADAVDATALLRPDGFVAWA  
ATADTNDGLEPALRRWFGDTA

**The revised sequence of Tmn9.** During cloning of *tmn9* it was noted that the sequence differed from that previously published<sup>6</sup> (as confirmed by sequencing independent PCR reactions). The revised Tmn9 protein sequence is as follows:

Tmn9:  
MSEPVVVVGGAGPAGMLACELAMRDVPAVLVDIHTPQRAEAPA  
MAINAGTLEMLDQRGLAAGLREGTVTPPEVRFADLRLEFAEKVQGPREP  
THMVLQRLEKVLIDRAVELGDLRWATRLTGFEAAADGSGVTVTLAS  
DAGEEQLRCRYIVGCDGRESIRVKQAGIDYVDDWVIRGVGDVAINR  
EDVAPQYGLSYTDNGDQFLGAPLSPDVMRVSAEFSTPEPEFDGPATLEQL  
GDVAKRLITGKELKATEAHWLQHYHSYTRNAEQYRKRVRVFIAGDAAHVHYPP  
NGQGLGTAIGDAVNLGWKIAAEVHWGAPADLLDSYHVERHLAARLACMNI  
QAQLALLYPRPLARYMREMMGEFLKFEDEVNVLFAEIVTNLGPVPIAYEGV  
PEPVEGDRLLGRLLPKVQIKTADGDMGVAETLQSGRGLVLLDLSGDASAQE  
ESGWA DRV DV RAVQVPDLPGLTLLRPDGCVAWHDDGGGWGQDELRTA  
LRTWFGAPTV

**Sequence of Tsn15.** Before the submission of this manuscript, a different open reading frame in the tetronasin gene cluster—that is, an ATP-binding cassette transporter—was named Tsn15. The Tsn15 described in this paper was not annotated. As such, the ATP-binding cassette transporter has been renamed Tsn15b and Tsn15 now refers to the [4+2] cyclase homologue. For clarity the DNA and protein sequence of Tsn15 are presented below:

tsn15:  
ATGACCACTTCCATCGATCCGACGACCCCGCTGACCTACAACCCC  
GTCACTCGACGCGCTCGTGGGCTCGTGGCGCCAGATCATCGACGCCGA  
CTACTCGGCGGACGACACCCGGCTGCCCGATCTCGCGGTGCTGGCC  
CGCTCCACCGCGCGGGCGGTGGCGGCTGCCGTACCCCGTCCCGCTCG  
CGGAGATCTCGGCGGACGCGCGGACGAGCGCGCGGAACTC  
GTGTGCTCGGAGAAGTGATCCAGGAAGTGGCGACCCGGAGTAC  
ACCCCGCTGAGCCCGAGGGCGGCGGTGGGGACCTCGTCTGG  
TGACGGAGAAGATCTACAACCTCCGACCGCGAGGAGATCGGCGCGGAC  
ACCGGGCGGCTCGGATCATCCGCAAGGACCCGGAGACCGGGCACC  
ACTTCACGGTCTCGCTCGTCCAGTCCAGGTGCGAGGCAACAAGC  
TGTTCCGCTTCGGCTACACCGAGATGGAGGCCAGCTCGCCGGGG  
GCCGACCCACCATCCAGTCCGCTGCTGGGACGGCGCCCTGGGC  
CGGATGAGCGGACCCCTGTCTGGGTGATCAACTCCATGACGG  
CCGCGGATCGCGGTACGAGCTGCGCCGCTGA

Tsn15:  
MTTSDIPTPLTYNPVIVALVGSWRQIDADYSADTRLPDLAVL  
ARSTARVAANAAPRLAEISAPDAPDERGELVLLKVIQEVADREYTP  
LSPGEPVSGDIVIVTEKYNDSREEIGADTGRLLRIIRKDPETGHHFTVS  
LVSTVQGNKLEAFGYTEMEAQLAGGRTTIQVACWDGPWAGMSGT  
LSWVINSMTAAESRYELRR

**Interception and detection of PKS-bound polyketide intermediates.** All *S. longisporoflavus* strains were grown in 10 ml of tsn medium A for 2 d at 30 °C and 200 r.p.m. Seed cultures (100 µl) were used to inoculate tsn medium A (10 ml of liquid culture, in duplicate, in 50 ml Erlenmeyer flasks with springs). They were then incubated at 30 °C for 5 d. After the first day of incubation, 6 (at a final concentration of 1 mM) was dissolved in 80 µl of MeOH and added portionwise over 4 d (20 µl each day, days 2–5). Control liquid cultures that omitted 6 were also prepared (in duplicate copy). After 5 d of fermentation, the liquid cultures were extracted with ethyl acetate (20 ml). The extracts were concentrated and the residues were redissolved in HPLC-grade methanol (1 ml) for MS analysis.

High performance liquid chromatography-high resolution-electrospray ionization-mass spectrometry analyses of *S. longisporoflavus* extracts were carried out on an LTQ-T Orbitrap Fusion instrument. Reverse-phase chromatography was used to separate the mixtures before MS analysis. Two columns were used: an Acclaim PepMap µ-precolumn cartridge (300 µm i.d. x 5 mm, 5 µm, 100 Å) and an Acclaim PepMap rapid-separation liquid chromatography column (75 µm x 15 cm 2 µm 100 Å) (Thermo Scientific). The columns were installed on an Ultimate 3000 RSLCnano system (Dionex). The mobile phase buffer A was 0.1% aqueous formic acid, whereas mobile phase B comprised 100% acetonitrile that contained 0.1% formic acid. Samples were loaded onto the µ-precolumn equilibrated in 50% aqueous acetonitrile that contained 0.1% trifluoroacetic acid for 8 min at 10 µl min<sup>-1</sup>. Compounds were then eluted onto the analytical column following a 75 min gradient for which the concentration of mobile phase B was increased from 50% to 80% B over 15 min, maintained at 80% B for 50 min and then decreased to 50% over 1 min, followed by a 9 min wash at 50% B. The species were analysed by electrospray ionization MS. Survey scans of precursors from 150 to 1500 *m/z* were performed at 60 K resolution (at 200 *m/z*) with a 4 × 10<sup>5</sup> ion count target. Tandem MS was performed by isolation at 1.6 Th with the quadrupole, higher energy collisional dissociation fragmentation with normalized collision energy of 32 and rapid-scan MS analysis in the ion trap. The MS/MS ion count target was set to

2 × 10<sup>5</sup> and the maximum injection time was 50 ms. A filter targeted inclusion mass list was used to select the precursor ions.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

#### Data availability

The tetronasin biosynthetic gene cluster sequence is available on GenBank (accession code: FJ462704). The crystal structure data is available on the PDB (accession codes: 6NOI (Tsn15) and 6NNW (Tsn15-substrate complex)). All other data that supports the findings of this study are available from the corresponding author on reasonable request.

Received: 21 March 2019; Accepted: 17 August 2019;

Published online: 14 October 2019

#### References

- Diels, O. & Alder, K. Synthesen in der hydroaromatischen Reihe. *Justus Liebigs Ann. der Chem.* **460**, 98–122 (1928).
- Corey, E. J. Catalytic enantioselective Diels–Alder reactions: methods, mechanistic fundamentals, pathways, and applications. *Angew. Chem. Int. Ed. Engl.* **41**, 1650–1667 (2002).
- Lichman, B. R., O'Connor, S. E. & Kries, H. Biocatalytic strategies towards [4+2] cycloadditions. *Chem. Eur. J.* **25**, 6864–6877 (2019).
- Jeon, B., Wang, S.-A., Ruzsyczky, M. W. & Liu, H. Natural [4+2]-cyclases. *Chem. Rev.* **117**, 5367–5388 (2017).
- Ma, S. M. et al. Complete reconstitution of a highly reducing iterative polyketide synthase. *Science* **326**, 589–592 (2009).
- Kasahara, K. et al. Solanapyrone synthase, a possible Diels–Alderase and iterative Type I polyketide synthase encoded in a biosynthetic gene cluster from *Alternaria solani*. *ChemBioChem* **11**, 1245–1252 (2010).
- Kim, H. J., Ruzsyczky, M. W., Choi, S., Liu, Y. & Liu, H. Enzyme-catalysed [4+2] cycloaddition is a key step in the biosynthesis of Spinosyn A. *Nature* **473**, 109–112 (2011).
- Page, C. D. et al. The structure of SpmF, a standalone enzyme that catalyzes [4+2] cycloaddition. *Nat. Chem. Biol.* **11**, 256–258 (2015).
- Hashimoto, T. et al. Biosynthesis of versipolstatin: identification of an enzyme-catalyzed [4+2]-cycloaddition required for macrocyclization of spirotronate-containing polyketides. *J. Am. Chem. Soc.* **137**, 572–575 (2015).
- Byrne, M. J. et al. The catalytic mechanism of a natural Diels–Alderase revealed in molecular detail. *J. Am. Chem. Soc.* **138**, 6095–6098 (2016).
- Tian, Z. et al. An enzymatic [4+2] cyclization cascade creates the pentacyclic core of pyrroindomycins. *Nat. Chem. Biol.* **11**, 259–265 (2015).
- Zhang, Z. et al. Enzyme-catalyzed inverse-electron demand Diels–Alder reaction in the biosynthesis of antifungal linciclin H. *J. Am. Chem. Soc.* **141**, 5659–5663 (2019).
- Cogan, D. P. et al. Structural insights into enzymatic [4+2] *aza*-cycloaddition in thiopeptide antibiotic biosynthesis. *Proc. Natl. Acad. Sci. USA* **114**, 12928–12933 (2017).
- Ohashi, M. et al. SAM-dependent enzyme-catalysed pericyclic reactions in natural product biosynthesis. *Nature* **549**, 502–506 (2017).
- Demetriadou, A. K. et al. Biosynthesis of the polyketide polyether Antibiotic ICI 139603 in *Streptomyces longisporoflavus* from <sup>13</sup>O-labelled acetate and propionate. *J. Chem. Soc., Chem. Commun.* **19**, 408–410 (1985).
- Denydzuk, Y. et al. Analysis of the tetronasin gene cluster: insights into the biosynthesis of a polyether tetronate antibiotic. *ChemBioChem* **9**, 1136–1145 (2008).
- Sun, Y., Hong, H., Gillies, F., Spencer, J. B. & Leadlay, P. F. Glycyl-S-Acyl carrier protein as an intermediate in the biosynthesis of tetronate antibiotics. *ChemBioChem* **9**, 150–156 (2008).
- Hailes, H. C., Jackson, C. M., Leadlay, P. F., Ley, S. V. & Staunton, J. Biosynthesis of tetronasin: part 1 introduction and investigation of the diketide and triketide intermediates bound to the polyketide synthase. *Tetrahedron Lett.* **35**, 307–310 (1994).
- Boons, G.-J. et al. Novel polyene cyclisation routes to the acyl tetronic acid ionophore tetronasin (ICI 139603). *Tetrahedron Lett.* **35**, 323–326 (1994).
- Riva, E. et al. Chemical probes for the functionalization of polyketide intermediates. *Angew. Chem. Int. Ed.* **53**, 11944–11949 (2014).
- Tosin, M., Smith, L. & Leadlay, P. F. Insights into lasalocid A ring formation by chemical chain termination in vivo. *Angew. Chem. Int. Ed. Engl.* **50**, 11930–11933 (2011).
- Zheng, Q. et al. Structural insights into a flavin-dependent [4+2] cyclase that catalyzes *trans*-decalin formation in pyrroindomycin biosynthesis. *Cell Chem. Biol.* **25**, 718–727 (2018).
- Bosserman, M. A., Downey, T., Noinaj, N., Buchanan, S. K. & Rohr, J. Molecular insight into substrate recognition and catalysis of Baeyer–Villiger monooxygenase MtmOIV, the key frame-modifying enzyme in the biosynthesis of anticancer agent mithramycin. *ACS Chem. Biol.* **8**, 2466–2477 (2013).

24. Vieweg, L., Reichau, S., Schobert, R., Leadlay, P. F. & Süßmuth, R. D. Recent advances in the field of bioactive tetronates. *Nat. Prod. Rep.* **31**, 1554–1584 (2014).
25. Jiang, X. & Wang, R. Recent developments in catalytic asymmetric inverse-electron-demand Diels–Alder reaction. *Chem. Rev.* **113**, 5515–5546 (2013).
26. Palasz, A. Recent advances in inverse-electron-demand hetero-Diels–Alder reactions of 1-oxa-1,3-butadienes. *Top. Curr. Chem.* **374**, 24 (2016).
27. Zheng, Q. et al. Enzyme-dependent [4+2] cycloaddition depends on lid-like interaction of the N-terminal sequence with the catalytic core in PyrI4. *Cell Chem. Biol.* **23**, 352–360 (2016).
28. Hofmann, E., Zerbe, P. & Schaller, F. The crystal structure of *Arabidopsis thaliana* allene oxide cyclase: insights into the oxylipin cyclization reaction. *Plant Cell* **18**, 3201–3217 (2006).
29. Yoon, S., Cho, K. & Han, O. Structural evidence for the substrate channeling of rice allene oxide cyclase in biologically analogous Nazarov reaction. *Front. Chem.* **6**, 500 (2018).
30. Costa, K. C., Glasser, N. R., Conway, S. J. & Newman, D. K. Pyocyanin degradation by a tautomertizing demethylase inhibits *Pseudomonas aeruginosa* biofilms. *Science* **355**, 170–173 (2017).
31. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
32. MacNeil, D. J. et al. Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**, 61–68 (1992).
33. Wilkinson, C. I. et al. Increasing the efficiency of heterologous promoters in actinomycetes. *J. Mol. Microbiol. Biotechnol.* **4**, 417–426 (2002).
34. Harding S. E. & Rowe S. E. *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Royal Society of Chemistry, 1992).
35. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).
36. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
37. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D* **69**, 1204–1214 (2013).
38. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
39. Terwilliger, T. C. et al. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D* **65**, 582–601 (2009).
40. Adams, P. D. et al. The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94–106 (2011).
41. Terwilliger, T. C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D* **64**, 61–69 (2008).
42. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D* **63**, 32–41 (2007).
43. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68**, 352–367 (2012).
44. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
45. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
46. DeLano, W. L. The PyMOL Molecular Graphics System (Delano Scientific, 2002).

### Acknowledgements

R.L. was supported by the Woolf Fisher Trust and the Cambridge Commonwealth European and International Trust. M.V.B.D. was supported by the São Paulo Research Foundation under grant nos 2015/09188-8, 2017/50140-4 and 2018/00351-1. F.C.R.P. was supported by a CNPq (National Council for Scientific and Technological Development) fellowship (141090/2016-2). M.T. and R.J. gratefully acknowledge EPSRC (DTA PhD studentship to R.J.); BBSRC (project grant no. BB/J007250/1 to M.T.) and the FAPESP–Warwick Joint Fund (for a SPRINT award to M.V.B.D. and M.T.).

### Author contributions

R.L., P.F.L., F.J.L., M.T. and M.V.B.D. developed the hypothesis and designed the study. Y.D., Y.S. and M.S. cloned, sequenced and analysed the gene clusters. R.L., Y.S. and H.H. performed gene deletions, whereas M.T. and R.J. performed and analysed the experiments with chain-terminating probes. R.L. performed protein expressions and purifications, in vitro experiments and compound isolations. R.L. and F.J.L. performed compound characterizations. F.C.R.P., R.L. and M.V.B.D. solved the crystal structures. All authors analysed and discussed the results. P.F.L., R.L. and F.C.R.P. prepared the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41929-019-0351-2>.

Correspondence and requests for materials should be addressed to P.F.L.

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019