



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA

RENATA SIGRIST

***GENOME MINING E ESTRATÉGIAS DE CLONAGEM PARA
CARACTERIZAÇÃO DE VIAS BIOSSINTÉTICAS DE PKS E NRPS EM
STREPTOMYCES SP. CBMAI 2042***

CAMPINAS
2019

RENATA SIGRIST

***GENOME MINING* E ESTRATÉGIAS DE CLONAGEM PARA
CARACTERIZAÇÃO DE VIAS BIOSSINTÉTICAS DE PKS E NRPS EM
STREPTOMYCES SP. CBMAI 2042**

**Tese de Doutorado apresentada ao Instituto de
Química da Universidade Estadual de Campinas como
parte dos requisitos exigidos para a obtenção do título
de Doutora em Ciências.**

Orientadora: Profa. Dra. Luciana Gonzaga de Oliveira

**O arquivo digital corresponde à versão final da Tese defendida pela aluna
Renata Sigrist e orientada pela Profa. Dra. Luciana Gonzaga de Oliveira.**

CAMPINAS

2019

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Química
Camila Barleta Fullin - CRB 8462

Si26g Sigrist, Renata, 1988-
Genome mining e estratégias de clonagem para caracterização de vias biossintéticas de PKS e NRPS em *Streptomyces* sp. CBMAI 2042 / Renata Sigrist. – Campinas, SP : [s.n.], 2019.

Orientador: Luciana Gonzaga de Oliveira.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Química.

1. *Streptomyces*. 2. Produtos naturais. 3. Mineração do genoma. I. Oliveira, Luciana Gonzaga de, 1975-. II. Universidade Estadual de Campinas. Instituto de Química. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Genome mining and cloning strategies for characterization of PKS and NRPS biosynthetic pathways from *Streptomyces* sp. CBMAI 2042

Palavras-chave em inglês:

Streptomyces

Natural products

Genome mining

Área de concentração: Química Orgânica

Titulação: Doutora em Ciências

Banca examinadora:

Luciana Gonzaga de Oliveira [Orientador]

Daniela Barretto Barbosa Trivella

Mônica Tallarico Pupo

Paulo Cezar Vieira

Anita Jocelyne Marsaioli

Data de defesa: 30-07-2019

Programa de Pós-Graduação: Química

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

- ORCID do autor: <https://orcid.org/0000-0002-9342-7470>

- Currículo Lattes do autor: <http://lattes.cnpq.br/4449688845822815>

BANCA EXAMINADORA

Profa. Dra. Luciana Gonzaga de Oliveira (Orientadora)

Dra. Daniela Barretto Barbosa Trivella (LNBio – CNPEM)

Profa. Dra. Mônica Tallarico Pupo (FCFRP – USP)

Prof. Dr. Paulo Cezar Vieira (FCFRP – USP)

Profa. Dra. Anita Jocelyne Marsaioli (IQ – UNICAMP)

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

Este exemplar corresponde à redação final da Tese de Doutorado defendida pela aluna **RENATA SIGRIST**, aprovada pela Comissão Julgadora em 30 de Julho de 2019.

“Nothing in life is to be feared, it is only to be understood.”

Marie Curie

“Let us strive for the impossible. The great achievements throughout history have been the conquest of what seemed the impossible.”

Charles Chaplin

“Um homem precisa viajar para lugares que não conhece para quebrar essa arrogância que nos faz ver o mundo como o imaginamos, e não simplesmente como é ou pode ser. Que nos faz professores e doutores do que não vimos, quando deveríamos ser alunos, e simplesmente ir ver.”

Amyr Klink

*Dedico esse trabalho à minha mãe, Ana Maria,
pela sua força e alegria... é imensurável a sua
importância em minha vida!*

AGRADECIMENTOS

Aos meus pais, Ana e José, e à minha irmã, Luciana, por sempre apoiarem minhas decisões, especialmente minha mãe, que sempre me acompanhou nos estudos e insistiu na sua importância para meu futuro. Vocês são essenciais!

À minha orientadora e amiga, Profa. Dra. Luciana Gonzaga de Oliveira, pela oportunidade de fazer parte do seu grupo de pesquisa. Agradeço a orientação, confiança, paciência e amizade. Por me abrir portas e apontar novos caminhos, me incentivando sempre na busca do novo e por todo o processo de crescimento profissional/ pessoal no decorrer desses looooongos anos. Obrigada por acreditar em mim!

Aos professores de ensino básico, fundamental, médio e técnico pelo conhecimento compartilhado, dedicação, incentivo e por me mostrarem o poder transformador da educação. Ao Prof. Marinho (FIEC-CEPIN), agradeço especialmente, por me aconselhar (insistentemente) a não desistir do curso técnico. A formação técnica garantiu meu emprego durante a graduação e me aproximou dos laboratórios de pesquisa da universidade. Aos professores da graduação e pós-graduação que me ensinaram uma profissão.

Aos amigos que a vida me presenteou. Talita, Thaís, Dani, Érica, Lili, Paula, Andreza, Gabriel, Ulisses, Pedro, Pira, Augusto, Palace II e Mansão&Eld. Foram boas conversas, conselhos, experiências compartilhadas, incentivos, inspirações, viagens, churrascos, festas, formaturas, casamentos e ressacas. Foram momentos especiais e essenciais para atravessar essa jornada. Cada um de vocês contribuiu, de alguma forma, para essa conquista. Àquelas que fizeram do meu intercâmbio nos EUA uma experiência ainda mais especial e uma grande aventura. Obrigada Dri, Lu e Bibi pela amizade e parceria.

Aos alunos que passaram pelos Labs D245 e D250, especialmente ao LaBiComb (Gaby, Ana Beatriz, Bruno, Douglas), Gi e Célio, pelos anos de doutorado, colaborações, conversas, trocas de conhecimento, cafés, cervejas e momentos de descontração. À Profa. Dra. Wanda Pereira Almeida, minha primeira orientadora, você foi muito importante para que eu chegasse até aqui. Agradeço por ter conhecido e convivido com tantas pessoas incríveis e inspiradoras nesses 11 anos.

Ao Prof. Dr. Bradley S. Moore, por me receber em seu grupo de pesquisas e ter me mostrado que a vida acadêmica pode ser mais leve e divertida, principalmente

quando da janela do lab temos uma bela vista do Pacífico. Ao MooreLab, pela experiência acadêmica tão enriquecedora. Especialmente, à Dra. Hanna Luhavaya (Anyá) por todo puxão de orelha, ajuda, conhecimento compartilhado e pela parceria nas aulas de kettlebell.

À Dra. Daniela B. B. Trivella e Dr. Rafael Felício (LNBio); Prof. Dr. Roberto G. S. Berlinck e Fabiana T. R. Martinelli (IQSC-USP); e ao LaCTAD, pelas análises em espectrometria de massas. À Prof. Dra. Anita Jocelyne Marsaioli pelo apoio à nossa pesquisa, empréstimos de reagentes e pela discussão científica.

À banca examinadora composta por Profa. Dra. Mônica Tallarico, Dra. Daniela Barretto Barbosa Trivella, Prof. Dr. Paulo Cezar Vieira e Profa. Dra. Anita Jocelyne Marsaioli, pela disponibilidade, pela rica discussão científica e pela contribuição para a versão final deste trabalho de tese.

Ao Instituto de Química pela estrutura e suporte à pesquisa realizada, e aos funcionários pelo trabalho e por garantirem o bom funcionamento e contribuírem com o nível de excelência da instituição.

À UNICAMP pela acolhida nesses longos 12 anos. Por ter sido minha segunda casa, ter me permitido conviver com o pluralismo e me ensinado tanto sobre diversidade racial, social, cultural e de gênero.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro às pesquisas (Processo nº 313492/2017-4).

Ao apoio financeiro concedido pelo Centro de Excelência para Pesquisa em Química Sustentável (CERSusChem) financiado pela FAPESP e GSK (Processo nº 2014/50249-8)

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelas bolsas concedidas (Processos nº 2013/12598-8, 2015/01013-4 e 2016/25735-1) e apoio financeiro (Processo nº 2014/12727-5).

RESUMO

Actinobacteria é um filo de bactérias Gram-positivas que habitam ambientes terrestres e aquáticos e são de grande relevância para a pesquisa acadêmica e a indústria farmacêutica devido ao grande potencial de biossintetizar diferentes classes de produtos naturais. Os policetídeos e os peptídeos não ribossômicos destacam-se por combinar alta complexidade estrutural a uma ampla gama de atividades terapêuticas, tais como antibiótica, antitumoral e imunossupressora. Os genes biossintéticos que codificam para enzimas produtoras desses compostos podem ser mapeados pelo sequenciamento e *genome mining* de diferentes linhagens de *Streptomyces*, de modo a auxiliar e agilizar a busca por novos e importantes alvos terapêuticos. O microrganismo endofítico *Streptomyces* sp. CBMAI 2042, isolado de *Citrus sinensis*, o genoma foi sequenciado e a análise *in silico* do genoma permitiu identificar um *cluster* de genes órfão composto por uma PKS do tipo 1, *trans*-AT PKS e NRPS. O estudo aqui apresentado visou a clonagem desse *cluster* de genes biossintético através de bibliotecas genômicas em vetores do tipo BAC e ESAC (empresa Bio S&T) e de técnicas de recombinação de DNA, como λ -RED, Gibson Assembly® e TAR-*cloning*, para posterior expressão em organismo heterólogo *S. coelicolor*. O *cluster* órfão obtido via TAR-*cloning* e expresso em *S. coelicolor* M1146 foi associado a produção de alpiniamida. A deleção do domínio de adenilação confirmou a correlação da via de biossíntese com a molécula isolada e caracterizada. Experimentos com precursores isotopicamente marcados embasaram a proposta de biossíntese dessa via mista *trans*-AT PKS/NRPS. Adicionalmente, o *cluster* de genes codificando para biossíntese da valinomicina, composto já descrito na literatura e produzido majoritariamente pela linhagem selvagem em diferentes cultivos, foi identificado no genoma e foi também selecionado para triagem de clones na biblioteca genômica. A clonagem e expressão heteróloga levou ao incremento de três vezes a produção do metabólito. Além disso, a associação da estratégia de *molecular networking*, levou a identificação de pelo menos seis análogos estruturais, entre eles a montanastatina, sendo pela primeira vez associados a mesma via de biossíntese.

ABSTRACT

Actinobacteria is a phylum of terrestrial and aquatic Gram-positive bacteria of significant importance for academic research and pharmaceutical industries due to their great potential to biosynthesize different types of natural products. Polyketides and nonribosomal peptides are produced by enzymes known as polyketide synthases and nonribosomal synthetase and stand out by combining high structural complexity with a wide range of therapeutic activities, such as antibiotics, antitumor and immunosuppressor. In order to accelerate identification of new therapeutic compounds, the biosynthetic genes encoding to these and other class of enzymes can be mapped by genome sequence and mining of different *Streptomyces* strains. The endophytic strain *Streptomyces* sp. CBMAI 2042 was isolated from *Citrus sinensis*, and the genome was completely sequenced. In silico analysis endorsed identification of the orphan biosynthetic genes comprising a type I PKS, *trans*-AT PKS and NRPS. This study aimed to obtain this biosynthetic gene *cluster* from genomic libraries constructed in BAC and ESAC vectors and by recombination methodologies λ -RED, Gibson Assembly® and TAR-cloning, followed by heterologous expression in *S. coelicolor*. The orphan *cluster* obtained by TAR-cloning and transferred to *S. coelicolor* M1146 was associated to alpiniamide biosynthesis. Deletion of adenylation domain confirmed the biosynthetic origin of isolated molecule and experiments with isotope-labeled precursors embased the proposed biosynthetic steps for alpiniamide assembly. Additionally, the gene cluster related to biosynthesis of the already described valinomycin was identified from genome annotation and also screened for clones in the genomic library, as it is the major metabolite produced by CBMAI 2042. Cloning and heterologous expression in *S. coelicolor* M1146 resulted in 3 times increase in production of valinomycin and association with molecular networking allowed the identification of at least six analogues, including montanastatin, related for the first time to the same biosynthetic pathway.

LISTA DE FIGURAS

Figura 1. Produtos naturais de grande contribuição para a humanidade isolados na era pré-antibiótica.....	31
Figura 2. Exemplos das principais classes de antibióticos descobertos durante a “Era de Ouro”.	34
Figura 3. Estratégias via <i>genome mining</i> para obtenção de novos produtos naturais bioativos. (Adaptado de Corre & Challis ⁶⁴).....	37
Figura 4. Biossíntese da Eritromicina A. O <i>cluster</i> de três genes codifica para as proteínas DEBS1-3 e cada proteína possui dois módulos extensores contendo três domínios obrigatórios (KS, AT e ACP) e domínios redutores opcionais (KR, DH e ER). DEBS1 contém o módulo iniciador de cadeia e DEBS3 contém uma tioesterase (TE), responsável pela ciclização e liberação da cadeia do complexo enzimático (adaptado de Menzella et al ⁸⁰).....	40
Figura 5. Sintase de policetídeo (PKS) do tipo I. A. Mecanismo de reação de incorporação de “n” unidades extensoras na cadeia policetídica via reação de condensação de Claisen catalisada por um domínio obrigatório de KS em uma PKS do tipo I. Ao final, o policetídeo é ciclizado no domínio TE (Adaptado de Kapur et al ⁸²). B. Domínios redutivos opcionais. KR: cetorredutase; DH: desidratase; ER: enoilredutase.	41
Figura 6. Exemplos de policetídeos biossintetizados por PKS do tipo I modular (salinomicina) e interativa (lovastatina).	42
Figura 7. Sintase de policetídeo (PKS) do tipo II. A. Domínios catalíticos envolvidos na biossíntese de policetídeos aromáticos e polienos (adaptado de Hertweck et al ⁸⁷). B. Estrutura do policetídeo doxorrubicina biossintetizado via uma PKS do tipo II. ...	43
Figura 8. Sintase de policetídeo (PKS) do tipo III. A. Domínio catalítico essencial composto apenas por KS B. Resveratrol (20), ⁹¹ exemplo de policetídeo biossintetizado por PKS do tipo III.....	43
Figura 9. Estruturas de policetídeos derivados de vias biossintéticas de <i>trans</i> -AT PKS.	44

Figura 10. Etapas de biossíntese de peptídeo não ribossomal catalisadas pelos domínios obrigatórios (C, A e PCP) de NRPS. (adaptado de Fischbach & Walsh¹⁰⁰)

..... 45

Figura 11. Reações catalisadas por domínios opcionais em NRPS. **A.** Domínio de epimerização. **B.** Domínio de oxidação de anéis tiazolina e oxazolina. **C.** Domínios redutivos e de transferência de metila. **D.** Domínio de ciclização intramolecular e formação de ligação amida. 46

Figura 12. Determinação da concentração de HindIII para digestão parcial de DNA genômico. Eletroforese de campo pulsado (6 V.cm⁻¹; P_{Inicial} 1 s; P_{final} 40 s; 18 h; 12 °C) em gel de agarose 0,8% TAE1X. M: PFGE Lambda *Ladder* (New England BioLabs). Colunas “a” ao “k”: *plugs* digeridos. a: 0 unidades.mL⁻¹; b: 0,5 unidades.mL⁻¹; c: 1 unidades.mL⁻¹; d: 1,5 unidades.mL⁻¹; e: 2 unidades.mL⁻¹; f: 3 unidades.mL⁻¹; g: 4 unidades.mL⁻¹; h: 6 unidades.mL⁻¹; i: 10 unidades.mL⁻¹; j: 20 unidades.mL⁻¹; k: 200 unidades.mL⁻¹. 74

Figura 13. Seleção I de fragmentos de DNA. *Plugs* de DNA parcialmente digeridos com enzima HindIII. Eletroforese de campo pulsado em agarose Seaken Gold (Lonza) 1% TAE 0.75X. Parâmetros: 1) 6 V.cm⁻¹; P_{Inicial} 1 s; P_{final} 40 s; 13 h; 2) 5 V.cm⁻¹; P_{Inicial} 4 s; P_{final} 5 s; 8 h; 10 °C para 1 e 2. M: PFGE Lambda *Ladder*. “PLUGS”: contem todos os *plugs* digeridos. A porção removida do gel corresponde aos fragmentos entre 100 e 350 kb. Antes do corte da porção contendo os fragmentos de interesse, apenas as partes do gel contendo os marcadores foram coradas em brometo de etídeo. 75

Figura 14. Seleção II de fragmentos de DNA. Eletroforese de campo pulsado em agarose Sealplaque LMP (Lonza) 1% TAE 1X. Parâmetros: 5 V.cm⁻¹; P_{Inicial} 3 s; P_{final} 4,5 s; 13 h; 12 °C. A, B e C: blocos obtidos da seleção I. M: PFGE Lambda *Ladder*. Antes dos cortes no gel contendo os fragmentos de interesse, apenas as partes dos marcadores foram coradas em brometo de etídeo. 75

Figura 15. Média de insertos da biblioteca genômica em vetor BAC. Eletroforese de campo pulsado em gel de agarose 0.8% TBE1X. Parâmetros: 6 V.cm⁻¹; 16 h; P_{inicial} 5 s e P_{final} 15 s). M: Marcador MidiRange PFG Marker (New England BioLabs). Demais colunas: BACs após digestão com enzima HindIII. A banda ~7,5 kb representa o vetor pIndigoBAC-5. A seta vermelha indica colônias “vazias”. 77

Figura 16. Produto de amplificação com enzima Phi29. Eletroforese em gel de agarose 1% TAE 1X. controle positivo (+) e negativo (-). Colunas 1 a 23: cada uma representa o *pool* de cada placa. 77

Figura 17. Triagem de clones em biblioteca BAC por qPCR. Eletroforese em gel de agarose 2% TAE 1X. **A.** Reação amplificando início do *cluster* (Primer_Início_F_1 e Primer_Início_R_1). **B.** Reação amplificando final do *cluster* (Contig51_Final_F3 e Contig51_Final_R3). A seta vermelha representa a placa (21) com produto de amplificação para ambas regiões. M: Marcador GeneRuler™ 1 kb Plus DNA Ladder (Thermo Scientific). (+): Controle positivo (gDNA). 78

Figura 18. Produtos de amplificação da região final do *cluster* obtidos por qPCR. Eletroforese em gel de agarose 2% TAE1X. Colunas marcadas representam clones positivos. M: Marcador GeneRuler™ 1 kb Plus DNA Ladder (Thermo Scientific). Segunda coluna (+): controle positivo. 78

Figura 19. Mapa do vetor pESAC13A. P1 rep: origem de replicação; *oriT*: origem de transferência; ϕ C31 *attP-int*: sistema para integração sítio-específica em *Streptomyces*; *aac(3)/IV*: resistência apramicina; BamHI: sítio de restrição e clonagem. 80

Figura 20. Biblioteca genômica em vetor ESAC de *Streptomyces* sp. CBMAI 2042. **A.** Média de insertos clonados. **B.** Triagem de clones. 5N: clone 5076 – placa9-5N; 16C: clone 5076 – placa11-16C; 4H: Clone 5076 – placa11-4H; 16L: clone 5076 – placa13-16L; CK: controle (gDNA). M: Marcador. Acima: clones t1PKS – *trans*-AT PKS – NRPS e abaixo clones da valinomicina. Os experimentos de triagem foram realizados por Bio S&T Inc. 80

Figura 21. Triagem de clones obtidos por conjugação triparental. Produtos de amplificação por PCR de colônia. Eletroforese em gel de agarose 1% TAE 1X. Linhas 1, 2, 3 e 4 correspondem aos clones 5076 – placa9-5N, 5076 – placa11-16C, 5076 – placa11-4H e 5076 – placa13-16L, respectivamente. 82

Figura 22. Experimentos de conjugação com *Streptomyces*. **A.** Placa de SFM-Nal/Thio com exconjugantes selecionados separados em quadrantes, para confirmação da integração do ESAC no cromossomo de *Streptomyces*. **B.** Confirmação do evento de conjugação por PCR com os *primers* correspondentes às extremidades do *cluster* biossintético. I = início do *cluster* e II = final do *cluster*;

exconjugante #11 e #21 - M1146/ESAC16L; exconjugante #12 e #16 - M1152/ESAC16L; exconjugante #5 e #17 - M1146/ESAC4H; exconjugante #2 e #16 - M1152/ESAC4H; exconjugante #12 e #16 - M1146/ESAC5N; exconjugante #3 e #18 - M1152/ESAC5N. (+): controle positivo (gDNA); M: Marcador de peso molecular 1 kb (Sinapse). 83

Figura 23. Representação *in silico* dos 9 fragmentos do *cluster* de genes t1PKS–*trans*-AT PKS–NRPS contendo sobreposição entre si. Imagem obtida pelo Programa Geneious Pro 5.6.7..... 84

Figura 24. Produtos de amplificação dos fragmentos correspondentes ao *cluster* de genes t1PKS – *trans*-AT PKS – NRPS. Eletroforese em gel de agarose 0,5%. M: GeneRuler™ 1 kb DNA Ladder. 84

Figura 25. Construção do plasmídeo de recombinação. **A.** Esquema para reações de PCR e construção do plasmídeo com regiões de homologies. Regiões de homologia do BGC: roxo e laranja; fragmento 1B: vermelho; fragmento 1A: verde; fragmento azul: pSET152cos linear. **B.** Produtos de amplificação 1A, 1B, R2 e R3 (plasmídeo para recombinação – pRS_recBGC03). Eletroforese em gel de agarose 0,7% TBE1X. pSET: pSET152cos digerido com BamHI; M: GeneRuler™ 1 kb DNA Ladder (Thermo Scientific). 85

Figura 26. Triagem de clones obtidos após recombinação. Eletroforese em gel de agarose 1% TAE1X dos produtos de amplificação por PCR de colônia. **A.** PCR de colônia com *primers* Contig51_5R e 6F. **B.** PCR de colônia com *primers* pSET_Contig51_início_F e Contig51_início_R; Contig51_Final_F e Contig51_pSET_Final_R. M: 1 kb Plus DNA Ladder (Invitrogen™). Controle (+): gDNA. 86

Figura 27. Vetores obtidos para experimentos de TAR-*cloning*. **A.** pCAP03-*acc*(3)/IV, derivado de pCAP01, contendo gene de resistência a apramicina entre os sítios de restrição XhoI e NdeI, sítio de clonagem.¹³⁹ **B.** pCAP03_p21, derivado de pCAP03 com promotor p21 *upstream* ao sítio de clonagem EcoRI e resistência a ampicilina entre o sítios EcoRI e NdeI, obtido por Dra. Hanna Luhavaya..... 88

Figura 28. Representação dos genes clonados por TAR-*cloning*. **A.** *Cluster* de genes biossintéticos de *Streptomyces* sp. CBMAI 2042 codificando para o complexo enzimático t1PKS-*trans*-AT PKS-NRPS (acima), representado pelos genes

biossintéticos essenciais (abaixo). **B.** Esquema representando a divisão do *cluster* biossintético em 7 fragmentos contendo ~150 pb de homologia entre si e ~45 nt de homologia com o vetor pCAP03_p21..... 89

Figura 29. Fragmentos obtidos por PCR para clonagem do *cluster* t1PKS – *trans*-AT PKS – NRPS. **A.** Eletroforese em gel de agarose 0,5% TAE 1X dos produtos de amplificação por PCR dos fragmentos 1 a 7 (3,5 – 5 kb). **B.** Confirmação dos fragmentos obtidos por análise de restrição..... 90

Figura 30. Triagem de colônias de levedura após recombinação. **A.** Colônias de *Saccharomyces cerevisiae* VL6-48N obtidas em meio ágar seletivo contendo 5-FOA após recombinação dos fragmentos, incubadas por 5 dias. **B.** Triagem de colônias recombinantes. Eletroforese em gel de agarose 1,5% de produtos de amplificação por PCR com o par de *primer* BGC03_117_176bp_R/F. (+): controle positivo (gDNA). Lane 1 a 6: colônias selecionadas. 91

Figura 31. Confirmação da construção dos plasmídeos obtidos por TAR-cloning. **A.** Construção do plasmídeo pCAP03_p21_BGC03 (40,3 kb) esperada após recombinação dos fragmentos. **B.** Análise de restrição dos plasmídeos 2C, 6B e 1A. Análise de restrição com enzimas EcoRI + EcoRV; KpnI + HindIII; HindIII + BamHI e BamHI..... 92

Figura 32. Confirmação da integração do plasmídeo pRSalp_6B ao cromossomo de *S. coelicolor* M1146. Eletroforese em gel de agarose 1%. Pares de *primers*: BGC03_117_176bp_F/R, BGC03_118_153bp_F/R; BGC03_119_168bp_F/R..... 93

Figura 33. Construção do plasmídeo de deleção. Representação dos fragmentos I e II flanqueando a deleção *in-frame* do domínio de adenilação (A) no gene *ctg1_119*. As regiões em laranja representam a região de homologia entre os fragmentos I e II; em azul são as regiões de homologia ao plasmídeo pYH7. 94

Figura 34. Padrão de digestão obtidos por restrição com NdeI e SmaI dos plasmídeos de deleção. Eletroforese em gel de agarose 0,5%. Foram observados três possíveis clones positivos quando digeridos com NdeI, dos quais dois foram confirmados com a enzima SmaI..... 95

Figura 35. Triagem de mutantes. Eletroforese em gel de agarose 0,7% TAE1X. **A.** PCR de colônia. Linhas 1 a 13: colônias isoladas; gDNA: controle negativo. **B.** PCR

com DNA genômico. Linhas 1 a 10: gDNA de cada colônia isolada; gDNA: controle negativo; P (plasmídeo de deleção): controle positivo. M: marcador 1 kb Sinapse.. 96

Figura 36. Cromatogramas de pico base do extrato metanólico dos cultivos em meio MP após análise por LC-ESI(+)-MS. *S. coelicolor* M1146 (verde) – 1; M1146_RSalp (vermelho) – 2; *Streptomyces* sp. CBMAI 2042 (azul) – 3. Em amarelo destaca-se o pico correspondente a alpiniamida e observado apenas nas amostras de 2 e 3..... 97

Figura 37. Espectro de massas correspondente ao pico com tempo de retenção em 13,442 min..... 98

Figura 38. Cromatogramas de pico base obtidos para todas as linhagens no estudo de identificação da alpiniamida após análise por LC-MS. As amostras são resultantes de extração por resina dos cultivos em meio MP. *S. coelicolor* M1146 (verde), M1146/ESAC5N (preto), M1146_RSalp (vermelho), RSalp ΔA_{domain} (laranja) e *Streptomyces* sp. CBMAI 2042 (azul). 99

Figura 39. Cromatogramas de íon extraído para m/z 326,19 obtidos para todas as linhagens no estudo de identificação da alpiniamida após análise por LC-MS. As amostras são resultantes da extração dos cultivos em meio MP com resina. *S. coelicolor* M1146 (verde), M1146/ESAC5N (preto), M1146_RSalp (vermelho), RSalp ΔA_{domain} (laranja) e *Streptomyces* sp. CBMAI 2042 (azul). 99

Figura 40. (+)-HRESIMS da alpiniamida isolada por HPLC-preparativo. m/z 366,1886 $[M+Na]^+$ (calcd 366,1887); m/z 326,1960 $[M'+H]$ ou $[M-H_2O+H]^+$ (calcd 326,1962); m/z 344,2074 $[M+H]^+$ (calcd 344,1994). 101

Figura 41. Espectro de RMN de 1H adquirido para alpiniamida (CD_3OD , 500 MHz). 104

Figura 42. Espectro de RMN do tipo COSY adquirido para alpiniamida (CD_3OD , 500 MHz). 105

Figura 43. Espectro de RMN do tipo HSQC adquirido para alpiniamida (CD_3OD , 500 MHz). 106

Figura 44. Espectro de RMN do tipo HMBC (500 MHz) adquirido para alpiniamida (CD_3OD , 500 MHz). 107

Figura 45. Atribuição de dados obtidos por espectroscopia de RMN para alpiniamida (25). Representação da correlações obtidas por COSY (negrito) e HMBC (setas em azul).....	108
Figura 46. Estereoquímica sugerida para alpiniamida (25) corresponde aquela descrita por Paulus et al. ¹⁷¹	109
Figura 47. Anotação do <i>cluster</i> t1PKS – <i>trans</i> -AT PKS – NRPS no genoma da linhagem <i>Streptomyces</i> sp. CBMAI 2042.....	110
Figura 48. Proposta de biossíntese da alpiniamida A envolvendo AlpA1, AlpA2 e AlpA3, apresentada por Paulus e colaboradores (adaptado de Paulus et al. ¹⁷¹). ...	111
Figura 49. Experimentos com metionina marcada. Espectros de massas de alta resolução obtidos a partir de extratos de cultivo da linhagem M1146_RSalp em meio MP A. sem adição de precursor marcado e B. com adição de L-metionina-(metil- ¹³ C).	112
Figura 50. Experimentos com propanoato marcado. Espectros de massas de alta resolução obtidos a partir de extratos de cultivo da linhagem M1146_RSalp em meio MP A. sem adição de precursor marcado e B. com adição de propanoato-1- ¹³ C de sódio.	113
Figura 51. Espectro MS/MS obtido para íon de m/z 326,1944 $[M-H_2O+H]^+$	114
Figura 52. Espectro MS/MS obtido para íon de m/z 326,1944 $[M-H_2O+H]^+$ ampliado para destacar os íons de fragmentos de m/z 137 e m/z 154.	115
Figura 53. Espectro MS/MS obtido para íon de m/z 326,1944 $[M-H_2O+H]^+$ ampliado para destacar os íons de fragmentos de m/z 137,05971 e m/z 137,09609 e o padrão isotópico.	115
Figura 54. Espectro MS/MS obtido para íon de m/z 326,1944 $[M-H_2O+H]^+$ ampliado para destacar os íons de fragmentos de m/z 154,08626 e m/z 154,12264 e o padrão isotópico.	116
Figura 55. Espectro MS/MS obtido para íon de m/z 326,1944 $[M-H_2O+H]^+$ ampliado para destacar o íon de fragmentos de m/z 109,10118.....	116

Figura 56. Alinhamento de sequências dos domínios de cetossintase (KS) anotados do genoma de <i>Streptomyces</i> sp. CBMAI 2042 (ClustalX). Setas amarelas indicam os resíduos conservados do sítio catalítico.....	117
Figura 57. Proposta de biossíntese da alpiniamida baseada nos resultados experimentais obtidos pelo presente estudo.	118
Figura 58. Valinomicina (26) – anotação de genes biossintéticos. A. <i>Cluster</i> de genes anotado do genoma de <i>Streptomyces tsusimaensis</i> ATCC 15141 ¹⁶³ codificando para biossíntese da valinomicina. B. BGC anotado do genoma de <i>Streptomyces</i> sp. CBMAI 2042. Os genes 23 e 24 codificam para VLM1 e VLM2, que se subdividem em 4 módulos, sendo os módulos compostos por domínios A, C, PCP, KR, E e TE. Cada módulo é responsável pela incorporação de uma unidade de Hiv (azul), Val (roxo), Lac (laranja) e Val (verde).....	120
Figura 59. Identificação de valinomicina por espectrometria de massas. A. ESI(+)-FTMS obtido por infusão direta do extrato bruto do cultivo de <i>Streptomyces</i> sp. CBMAI 2042 em meio A. B. ESI(+)-MS/MS do íon m/z 1128 $[M+NH_4]^+$ correspondente ao padrão de fragmentação da valinomicina.....	121
Figura 60. Anotação do <i>cluster</i> de genes relacionado a biossíntese de valinomicina em <i>Streptomyces</i> sp. CBMAI 2042. O <i>cluster</i> apresenta 22% de similaridade com o BGC codificando para a biossíntese de valinomicina em <i>Streptomyces tsusimaensis</i> ATCC 15141 (BGC0000453). Anotação obtida pela plataforma antiSMASH (vs 5.0 beta).	122
Figura 61. ESI(+)-FTMS obtidos por infusão direta dos extratos de cultivos em meio A dos exconjugantes contendo o <i>cluster</i> de genes codificando para a biossíntese da valinomicina.....	123
Figura 62. ESI(+)-FTMS obtidos por infusão direta dos extratos de cultivos em meio ISP2 dos exconjugantes contendo o <i>cluster</i> de genes codificando para a biossíntese da valinomicina.....	124
Figura 63. Cromatogramas de bico base obtidos dos extratos de cultivo de <i>S. coelicolor</i> M1146 (abaixo) e M1146/ESAC4H#5 (acima). O pico mais intenso (m/z 1128) corresponde ao aduto amoniado de valinomicina $[M+NH_4]^+$ e os picos correspondentes a m/z 1100, m/z 1114 e m/z 1142 estão relacionados a prováveis análogos estruturais da valinomicina.	126

Figura 64. *Molecular networking* obtido a partir de espectros de fragmentação de extratos de cultivos de *Streptomyces* sp. CBMAI 2042 em meio ISP2. Azul: íons precursores da valinomicina e montanastatina. Verde: íons correspondentes a análogos estruturais. 127

Figura 65. Cromatogramas de íon extraído correspondente a m/z 1128 (valinomicina – **26**). Vermelho: linhagem heteróloga M1146/ESAC4H#5. Azul: linhagem selvagem *Streptomyces* sp. CBMAI 2042. Preto: *S. coelicolor* M1146. **27** – montanastatina (m/z 741 [M+H]⁺), **28** – m/z 781 [M+Na]⁺, **29** – m/z 1100 [M+NH₄]⁺, **30** – m/z 1114 [M+NH₄]⁺, **31** – m/z 1142 [M+NH₄]⁺ e **32** – m/z 1156 [M+NH₄]⁺. 128

Figura 66. Identificação de padrões em espectros MS/MS da valinomicina e análogo **6**. **A.** Espectro MS/MS e atribuições de íons de fragmentação para proposta de estrutura valinomicina. Propostas de fragmentação da valinomicina. **B.** Espectro MS/MS e atribuições de íons de fragmentação para proposta de estrutura do análogo **31** (m/z 1142,6845). **C.** Proposta de biossíntese do análogo **31**. 129

Figura 67. Estrutura da valinomicina (**26**), montanastatina (**27**) e análogos estruturais **28**, **29**, **30**, **31** e **32**. As caixas pontilhadas e resíduos em vermelho destacam as modificações nas estruturas quando comparado a valinomicina. 131

Figura 68. Estruturas da **A.** Taromicina A e B; **B.** Daptomicina. As caixas pontilhadas representam as diferenças entre as estruturas. **C.** *Cluster* de genes anotado de *Saccharomonospora* sp. CNQ-490 codificando para biossíntese de taromicina. ... 134

Figura 69. Via da quinurenina - metabolismo do triptofano. Etapas das reações envolvendo a triptofano dioxigenase (Tar13) e quinurenina formamidase (Tar16). 135

Figura 70. Purificação de Tar13. **A.** Cromatografia de afinidade – Ni: carregando a coluna. **B.** Coluna de níquel carregada. **C.** Diferentes frações obtidas em separação por cromatografia de afinidade. **D.** Coluna de gel-filtração Superdex. **E.** SDS-PAGE (12,5% acrilamida) obtida para proteína após separação por cromatografia de gel-filtração. Peso molecular obtido experimentalmente para Tar13 = ~28 kDa. Linhas I: Tar13 – 0,25 mg.mL⁻¹; II: EZ-Run™ Rec protein Ladder (Fisher Scientific); III: Tar13 – 0,5 mg.mL⁻¹. 138

Figura 71. SDS-PAGE (12,5% acrilamida) para Tar16 purificada por **A.** Cromatografia de afinidade – Ni²⁺. **B.** Cromatografia de troca iônica. **C.** Cromatografia de gel-filtração.

Peso molecular obtido para Tar16 ~33 kDa. Linhas I: TAR16 – 0,25 mg. mL⁻¹; II: EZ-Run™ Rec protein Ladder (Fisher Scientific); III: TAR16 – 0,5 mg.mL⁻¹..... 139

Figura 72. Cromatogramas de pico base obtidos por análise em LC-MS dos ensaios enzimáticos com Tar13 em condições otimizadas. Laranja: padrão comercial de 6-Cl-Trp *m/z* 238,0509 (Sigma-Aldrich). Verde: controle com Tar13 desnaturada. Azul: reação sem Tar13. Cinza: reação sem ácido ascórbico. Amarelo: reação sem hemina. Roxo: reação enzimática otimizada com Tar13, substrato (6-Cl-Trp) e cofatores, hemina e ácido ascórbico. *m/z* 270,0407 produto de reação esperado..... 142

Figura 73. Cromatogramas de pico base obtidos para os ensaios enzimáticos com Tar16 em condições otimizadas. Preto: padrão comercial de 4-Cl-Kyn *m/z* 242,05 (Sigma-Aldrich). Vermelho: padrão comercial de 6-Cl-Trp *m/z* 238,05 (Sigma-Aldrich). Amarelo: sem Tar13 e Tar16. Verde: sem Tar16. Azul: Tar16 desnaturada. Roxo: reação enzimática otimizada com Tar13 e Tar16, substrato (6-Cl-Trp) e cofatores, hemina e ácido ascórbico. *m/z* 242,05 produto de reação esperado. *m/z* 270,04: 4-Cl-formil-quinurenina..... 143

LISTA DE TABELAS

Tabela 1. Principais antibióticos isolados de diferentes organismos durante a era de ouro dos antibióticos (adaptado de Durand et. al ³¹).....	33
Tabela 2. Plasmídeos e linhagens utilizados neste estudo	55
Tabela 3. Reagentes e parâmetros utilizados para PCR em tempo real.....	60
Tabela 4. Protocolos para preparo de reações de PCR para construção do plasmídeo	61
Tabela 5. Condições das reações de PCR para construção do plasmídeo (30 ciclos)	61
Tabela 6. Pares de <i>primers</i> utilizados para amplificar fragmentos de DNA.....	62
Tabela 7. Parâmetros utilizados para reações de PCR para obtenção dos fragmentos	62
Tabela 8. Pares de <i>primers</i> obtidos para amplificação dos fragmentos correspondentes ao <i>cluster</i> t1PKS – <i>trans</i> -AT PKS – NRPS	64
Tabela 9. Condições de preparo das reações de PCR.....	64
Tabela 10. Parâmetros para reação de PCR em termociclador (25 ciclos).....	64
Tabela 11. Parâmetros utilizados para reações de PCR para obtenção de fragmentos para construção de plasmídeo de deleção.....	67
Tabela 12. <i>Clusters</i> biossintéticos anotados do genoma de <i>Streptomyces</i> sp. CBMAI 2042 codificando para o metabolismo secundário.	71
Tabela 13. Anotações das principais orfs dos <i>clusters</i> de genes 03 e 30	72
Tabela 14. Concentração de fragmentos de DNA e do vetor pCAP03_p21 linearizado transformados em esferoplastos de <i>S. cerevisiae</i> VL6-48N	91
Tabela 15. Padrões de restrição obtidos <i>in silico</i> para plasmídeo recombinante pCAP03_p21_BGC03	93
Tabela 16. Dados obtidos por espectrometria de massas de alta resolução para determinação de fórmula molecular	101
Tabela 17. Dados obtidos por RMN para alpiniamida em CD ₃ OD*	103

Tabela 18. Quantificação da produção de valinomicina correspondente ao cultivo de 7 dias em meio A.....	125
--	-----

Tabela 19. Parâmetros testados em ensaios enzimáticos com Tar13. Em amarelo estão as condições ótimas	139
--	-----

LISTA DE GRÁFICOS

Gráfico 1. Curva de calibração obtida com padrão comercial para quantificação de valinomicina. A seta em azul representa o valor de concentração médio correspondente a produção pela linhagem recombinante M1146/ESAC4H#5. 125

Gráfico 2. Representação gráfica dos resultados obtidos a partir de diferentes parâmetros testados para otimização dos ensaios enzimáticos de Tar13. 141

LISTA DE ABREVIações

A	Domínio de adenilação
ACP	Proteína carreadora de grupo acila
AT	Aciltransferase
BAC	Cromossoma artificial bacteriano
BGC	<i>Cluster</i> de genes biossintéticos
BPC	Cromatograma de Pico Base
C	Domínio de condensação
CBMAI	Coleção Brasileira de Microrganismos de Ambiente e Indústria
cMT	Metiltransferase
CoA	Coenzima A
COSY	<i>Correlation Spectroscopy</i>
DH	Desidratase
DMSO	Dimetil sulfóxido
DNA	Ácido desoxirribonucleico
ER	Enoilredutase
ESAC	Cromossomo artificial de <i>E. coli-Streptomyces</i>
ESI	Ionização por electrospray
HMBC	<i>Heteronuclear Multiple Bond Coherence Spectroscopy</i>
HPLC	Cromatografia líquida de alta eficiência
HRMS	Espectrometria de massas de alta resolução
HSQC	<i>Heteronuclear Single Quantum Coherence Spectroscopy</i>
KR	Cetorredutase
KS	Cetossintase
LC	Cromatografia líquida
m/z	Razão massa/carga
MRSA	<i>Staphylococcus aureus</i> resistente à meticilina
MS	Espectrometria de massas
MS/MS	Espectrometria de massas sequencial
NCBI	<i>National Center for Biotechnology Information</i>
NRPS	Sintetase de peptídeo não ribossomal
PAC	Cromossomo artificial P1-derivado

PCP	Proteína carreadora de grupo peptídico
PCR	Reação em cadeia da polimerase
pHMM	<i>Poisson Hidden Markov Models</i>
PKS	Sintase de policetídeo
Q-Tof	Analisador de <i>m/z</i> híbrido do tipo quadrupolo – tempo-de-voo
RMN	Ressonância magnética nuclear
RNA	Ácido ribonucleico
TAR	Transformação associada a recombinação
TE	Tioesterase
VRE	<i>Enterococcus</i> resistente à vancomicina
WT	Linhagem selvagem
YAC	Cromossomo artificial de leveduras

SUMÁRIO

1. INTRODUÇÃO.....	29
1.1. Produtos Naturais.....	29
1.2. <i>Streptomyces</i>: fonte de prospecção de compostos bioativos	32
1.2.1. <i>Resistência microbiana.....</i>	34
1.3. Era pós-genômica.....	35
1.3.1. <i>Ferramentas de bioinformática</i>	38
1.4. Principais enzimas envolvidas na biossíntese de produtos naturais	38
1.4.1. <i>Sintase de Policetídeo (PKS)</i>	39
1.4.2. <i>Sintetase de Peptídeo Não Ribossomal (NRPS)</i>	44
1.5. Clonagem de clusters de genes biossintéticos.....	46
1.5.1. <i>Bibliotecas Genômicas</i>	47
1.5.2. <i>Recombinação de fragmentos de DNA in vivo e in vitro</i>	48
1.5.2.1. <i>Sistemas λ Red e RecET</i>	49
1.5.2.2. <i>Gibson Assembly®</i>	50
1.5.2.3. <i>TAR-cloning</i>	50
1.5.3. <i>Expressão em organismos heterólogos</i>	51
2. OBJETIVOS.....	53
2.1. Objetivos específicos	53
3. MATERIAIS E MÉTODOS.....	54
3.1. Materiais	54
3.1.1. <i>Meios de cultura</i>	54
3.1.2. <i>Microrganismos e plasmídeos.....</i>	55
3.1.3. <i>Enzimas e Primers.....</i>	56
3.2. Métodos de microbiologia e biologia molecular	56
3.2.1. <i>Biblioteca genômica.....</i>	56
3.2.1.1. <i>Extração de DNA de alto peso molecular e Cinética de Digestão.....</i>	56
3.2.1.2. <i>Ligação e Transformação</i>	58
3.2.1.3. <i>Minipreparação de BAC.....</i>	59
3.2.1.4. <i>Pool de placas e Amplificação com Phi29</i>	59
3.2.1.5. <i>PCR em tempo real (qPCR)</i>	60
3.2.1.6. <i>Sequenciamento das pontas de BACs</i>	60
3.2.2. <i>Recombinação de fragmentos de DNA</i>	61
3.2.2.1. <i>Construção do plasmídeo</i>	61

3.2.2.2.	Obtenção dos fragmentos de DNA do <i>cluster</i> biossintético	62
3.2.2.3.	Recombinação via λ -RED.....	63
3.2.2.4.	Recombinação via <i>Gibson Assembly</i> ®.....	63
3.2.3.	<i>TAR-cloning</i>	63
3.2.3.1.	Amplificação dos fragmentos por PCR.....	63
3.2.3.2.	Obtenção de esferoplastos de levedura	64
3.2.3.3.	Transformação de esferoplastos	65
3.2.4.	<i>Construção de plasmídeo de deleção</i>	66
3.2.5.	<i>Conjugação triparental</i>	67
3.2.6.	<i>Conjugação E. coli – Streptomyces</i>	67
3.2.7.	<i>Experimentos com precursores marcados</i>	68
3.3.	Métodos químicos e analíticos.....	69
3.3.1.	<i>Extração dos metabólitos</i>	69
3.3.2.	<i>Análise por LC-MS</i>	69
3.3.3.	<i>Ressonância Magnética Nuclear (RMN)</i>	70
4.	RESULTADOS E DISCUSSÃO.....	71
4.1.	<i>Streptomyces</i> sp. CBMAI 2042: análise <i>in silico</i> das anotações do genoma	71
4.2.	Clonagem de vias de biossíntese	73
4.2.1.	<i>Biblioteca genômica em vetor do tipo BAC</i>	73
4.2.2.	<i>Biblioteca genômica em vetor do tipo ESAC</i>	79
4.2.3.	<i>Recombinação de fragmentos de DNA via λ Red e Gibson Assembly</i> ®	83
4.2.4.	<i>TAR-cloning</i>	87
4.3.	Deleção “<i>in-frame</i>” de genes específicos	93
4.4.	Alpiniamida	97
4.4.1.	<i>Caracterização e determinação estrutural</i>	100
4.4.2.	<i>Elucidação da Biossíntese da Alpiniamida</i>	109
4.5.	Valinomicina.....	119
4.5.1.	<i>Quantificação de valinomicina</i>	124
4.5.2.	<i>Análogos estruturais da valinomicina</i>	126
5.	CONCLUSÕES.....	132
6.	ESTUDO DA BIOSÍNTESE DE 4-CLORO-QUINURENINA MEDIADA POR TAR13 E TAR16	134
6.1.	Objetivos	136
6.2.	Materiais e Métodos	136
6.2.1.	<i>Expressão das proteínas em E. coli BL21 (DE3)</i>	136

6.2.2. Purificação de Tar13 e Tar16	136
6.2.3. Ensaios enzimáticos <i>in vitro</i>	137
6.3. Resultados e Discussão	138
6.4. Conclusões	144
7. REFERÊNCIAS	145
8. APÊNDICE.....	163
8.1. Expansão 1 do espectro de RMN de ^1H da alpiniamida (MetOH- <i>d</i> 4, 500 MHz).	164
8.2. Expansão 2 do espectro de RMN de ^1H da alpiniamida (MetOH- <i>d</i> 4, 500 MHz).	165
8.3. Expansão 3 do espectro de RMN de ^1H da alpiniamida (MetOH- <i>d</i> 4, 500 MHz).	166
8.4. Expansão 4 do espectro de RMN de ^1H da alpiniamida (MetOH- <i>d</i> 4, 500 MHz).	167
8.5. Expansão 5 do espectro de RMN de ^1H da alpiniamida (MetOH- <i>d</i> 4, 500 MHz).	168
8.6. A. MS e MS/MS em baixa resolução da valinomicina $[\text{M}+\text{NH}_4]^+$. B. MS/MS em alta resolução.	169
8.7. A. MS e MS/MS em baixa resolução do análogo de m/z 1100,6418 $[\text{M}+\text{NH}_4]^+$. B. MS/MS alta resolução.	170
8.8. A. MS e MS/MS em baixa resolução do análogo de m/z 1114,6514 $[\text{M}+\text{NH}_4]^+$. B. MS/MS alta resolução.	171
8.9. A. MS e MS/MS em baixa resolução do análogo de m/z 1142,6845 $[\text{M}+\text{NH}_4]^+$. B. MS/MS alta resolução.	172
8.10. A. MS e MS/MS em baixa resolução do análogo de m/z 1142,6845 $[\text{M}+\text{NH}_4]^+$. B. MS/MS alta resolução.	173
9. ANEXOS	174
9.1. Oligonucleotídeos sintéticos (<i>primers</i>)	174
9.2. Certificado da Comissão Interna de Biossegurança – CIBio	176
9.3. Cadastro no Sistema Nacional de Gestão do Patrimônio Genético	177

1. INTRODUÇÃO

1.1. Produtos Naturais

A importância e o uso pela humanidade dos produtos naturais são registrados há milênios por muitas civilizações, sobretudo por meio da aplicação das plantas e seus extratos no tratamento de diversas doenças.¹ Os produtos naturais, também denominados metabólitos secundários, são pequenas moléculas com grande variedade estrutural produzidas por organismos vivos. São codificados por extensas regiões de DNA e não estão envolvidos diretamente no crescimento, desenvolvimento e reprodução do organismo, contudo desempenham um importante papel na defesa, competição, como agentes de transportes de íons e metais, auxiliando na adaptação e sobrevivência em relações competitivas ou em condições hostis.²⁻⁴

Em 1806, Friedrich Sertürner isolou de uma planta denominada *Papaver somniferum* o primeiro produto natural em sua forma pura e com atividade farmacológica. O composto tratava-se da morfina (**1**), um opióide e potente analgésico até hoje muito conhecido e utilizado, que foi largamente empregado durante a Guerra Civil dos Estados Unidos levando a dependência química de milhares de soldados.^{5,6}

Outro produto natural oriundo de planta e de grande impacto para a humanidade foi a quinina (**2**), isolada da casca de cinchona (*C. officinalis*) pelos químicos franceses Pierre-Joseph Pelletier e Joseph Caventou em 1820.⁷ A casca da planta já era utilizada por povos indígenas na América Latina no preparo de infusões para tratamento de febres e foi introduzida na Europa no século XVII para tratamento da malária. Posteriormente, o princípio ativo virou base para a síntese de outros medicamentos antimaláricos, como a cloroquina (**3**) e a mefloquina (**4**), de grande impacto para a saúde pública mundial.¹

O conhecimento oriundo das práticas da Medicina Tradicional Chinesa e de povos indígenas estimulou, até o início do século XX, a investigação e isolamento de produtos naturais a partir das plantas. Entretanto, a partir de 1929, quando Alexander Fleming acidentalmente descobriu o antibiótico penicilina (**5**) produzido pelo fungo *Penicillium notatum*, despertou-se a atenção e a percepção de que microrganismos também seriam uma importante fonte de compostos bioativos.^{8,9}

Posteriormente, no ano de 1939, René J. Dubos isolou pela primeira vez um composto com atividade bactericida de amostras de solo e produzido por *Bacillus*

brevis, a gramicidina (6).^{10,11} Este foi o primeiro antibiótico a ser testado clinicamente e os resultados positivos estimularam os cientistas Ernst Chain, Howard Florey, Norman Heatley e Edward Abraham, da Universidade de Oxford, a retomar o trabalho com a penicilina, na tentativa de isolar o composto e aumentar a produção para testes clínicos e comercialização, feito que rendeu a Fleming, Florey e Chain o prêmio Nobel de Medicina em 1945.¹²⁻¹⁴ A gramicidina e, principalmente, a penicilina foram de grande importância durante a Segunda Guerra Mundial, garantindo a sobrevivência de muitos soldados.^{12,15}

Selman Abraham Waksman, microbiologista de solo, e seu aluno foram impulsionados pelas descobertas de Fleming e Dubos e em 1940 isolou a actinomicina produzida por *Actinomyces antibioticus*, porém altamente tóxica em testes com animais. Já em 1943 isolou a estreptomicina (7), produzida por *Streptomyces griseus*, efetiva no tratamento de infecções por *Mycobacterium tuberculosis*, viabilizando o tratamento de tuberculose pela primeira vez na história. A estreptomicina concedeu ao cientista o prêmio Nobel de Medicina em 1952.^{9,16}

Essas descobertas revolucionaram a história da medicina no tratamento de uma das principais causas de morte na época, as infecções bacterianas, tanto pela alta efetividade no tratamento, quanto pela baixa toxicidade,^{13,17} se comparados aos antibióticos sintéticos disponíveis na época, como o salvarsan (Arsfenamina ou composto 606 - 8), a base de arsênio, e as sulfonamidas, a primeira a ser conhecida como prontasil (9).¹⁵

Os químicos iniciaram, então, uma era marcada pela investigação de microrganismos, especialmente do solo, através de técnicas e estratégias de cultivo designadas “plataforma de Waksman”¹⁸ na busca por novas moléculas. Posteriormente, juntaram-se a eles bioquímicos, microbiologistas, biólogos e indústrias farmacêuticas (como Merck, Sandoz e Takeda), e todos esses esforços somados nos levaram a chamada “Era de Ouro dos Antibióticos” entre 1940 e 1970.^{9,13,19-21} Este período foi caracterizado pela exponencial descoberta de compostos bioativos isolados de microrganismos, sobretudo antibióticos.

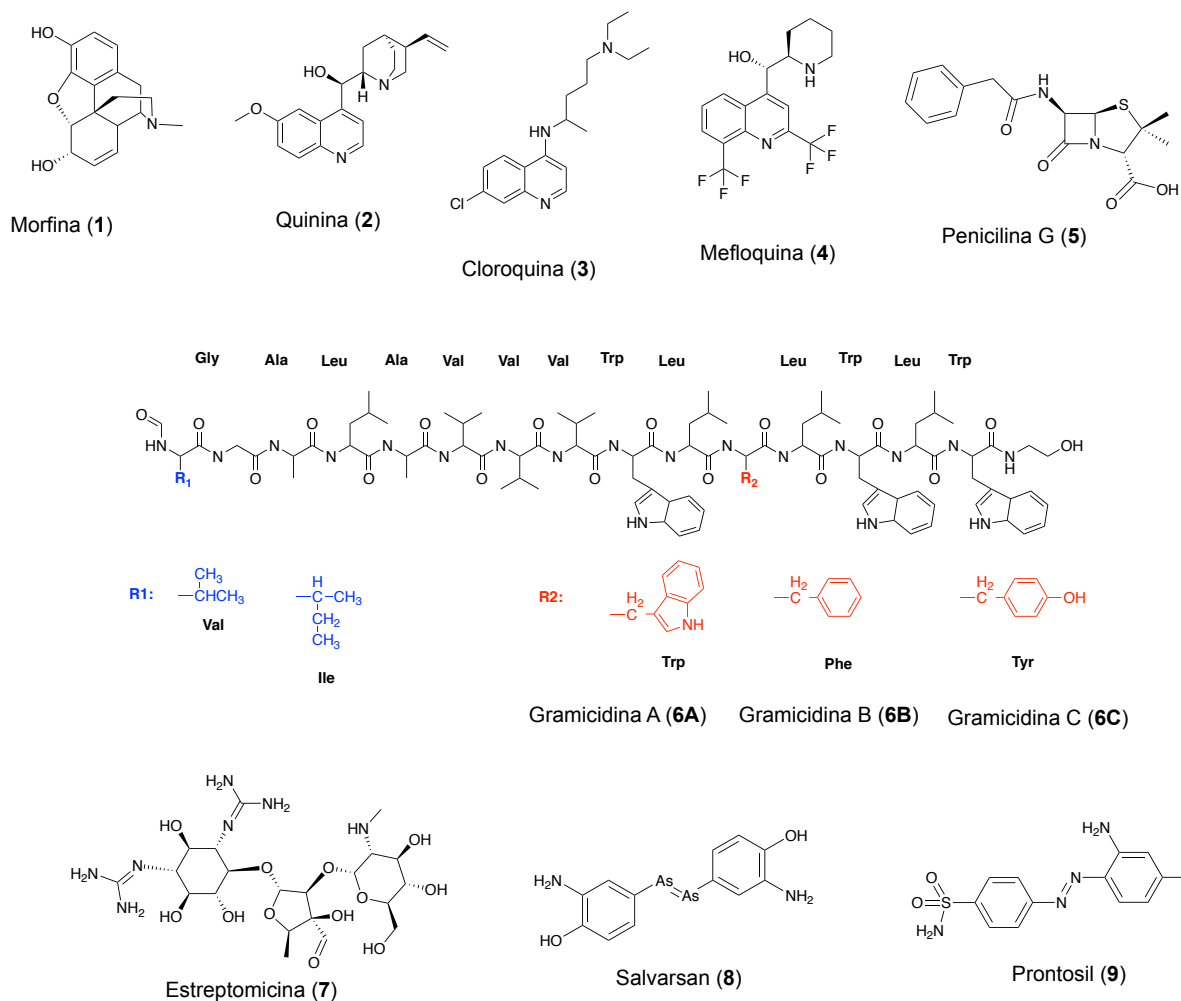


Figura 1. Produtos naturais de grande contribuição para a humanidade isolados na era pré-antibiótica.

Segundo Demain²², já são conhecidos cerca de 1 milhão de produtos naturais (PN), ~25% com alguma atividade biológica, e dentre as classes terapêuticas de maior relevância, os produtos naturais e derivados compreendem 78% dos antibióticos e 74% dos antitumorais. Até 2013, a *US Food and Drug Administration* (FDA) aprovou 1453 novas entidades químicas, sendo 6% PN, 27% derivados de PN e 5% compostos sintéticos contendo grupo farmacóforo de PN, totalizando 38% das drogas aprovadas para uso clínico. Desse total, considerando a fonte natural 43% são oriundos de plantas, 30% de bactérias, 23% de fungos e 4% de outros organismos. Quando considerados a origem de antibióticos aprovados pela FDA, 3% são de plantas, 46% de fungos e 51% de bactérias, predominantemente actinobactérias do gênero *Streptomyces*.^{21–23}

1.2. *Streptomyces*: fonte de prospecção de compostos bioativos

Streptomyces pertencem ao filo Actinobacteria, formado por bactérias filamentosas Gram-positivas com alto conteúdo de citosina e guanina em seu DNA (C+G: 70 – 75%), e desde a descoberta da estreptomicina têm se destacado como maquinários enzimáticos produtores de diversas classes de antibióticos. As actinobactérias em geral são responsáveis pela produção de 2/3 dos antibióticos produzidos por microrganismos atualmente, sendo que ~ 80% foram isolados de *Streptomyces* e, se considerados todos os metabólitos secundários com alguma atividade terapêutica, além da antibiótica, *Streptomyces* contribuem também com 80% dentro do filo.²⁴

As primeiras evidências de *Streptomyces* no planeta datam 450 milhões de anos atrás como resultado da seleção natural estimulada pelo surgimento de plantas superiores, tornando-as bem adaptadas para viver, principalmente, em relações endofíticas de ambiente terrestre. O desenvolvimento de hifas, filamentos lineares ou ramificados que formam os micélios característicos de fungos e actinobactérias, favoreceu a adaptação e sobrevivência no nicho ecológico, viabilizando a captura de nutrientes e decomposição de componentes mais complexos através da produção de enzimas extracelulares, como as xilanases e as celulases.^{24,25}

A formação de esporos foi também uma característica evolutiva importante para a sobrevivência desses microrganismos em ambientes hostis, em situações de escassez de água e nutrientes, e até mesmo no interior de outros animais, como vermes e artrópodes. Ao mesmo tempo, foram importantes para a dispersão das linhagens para diferentes localidades.²⁵ Actinobactérias, incluindo *Streptomyces*, foram isoladas mais recentemente, por exemplo, de solos da Antártida²⁶ ou desérticos²⁷ e até mesmo de sedimentos nas profundezas dos oceanos.²⁸

As relações endofíticas e coevolução de *Streptomyces* e plantas levaram a seleção natural de organismos benéficos ao hospedeiro e o perfil de metabólitos produzidos por *Streptomyces* emergiu dessas interações com as plantas, outros endofíticos, fitopatógenos e invasores, variando desde antimicrobianos até compostos de ativação de vias biossintéticas, produção de precursores, moléculas envolvidas no *quorum sensing*, promoção de crescimento do hospedeiro, moduladores epigenéticos do organismo e até mesmo interações físicas diretas com outros microrganismos.^{29,30}

A complexidade e versatilidade metabólica caracterizam *Streptomyces* por desempenhar um importante papel na degradação e ciclagem de nutrientes no solo,²⁸ além de produzir importantes entidades químicas introduzidas para tratamentos da saúde humana. Na Tabela 1 e Figura 2 são apresentados alguns dos principais antibióticos isolados durante a “Era de Ouro dos Antibióticos”, ficando evidente a importância desses organismos como fonte de prospecção de novas moléculas bioativas.

Tabela 1. Principais antibióticos isolados de diferentes organismos durante a era de ouro dos antibióticos (adaptado de Durand et. al³¹)

Ano*	Antibiótico	Origem	Classe
1946	Cloranfenicol	<i>Streptomyces venezuelae</i>	Cloranfenicol
1947	Polimixina	<i>Paenibacillus polymyxa</i> (NRPS)	Polipeptídeos
1948	Clorotetraciclina	<i>Streptomyces aureofaciens</i> <i>Streptomyces rimosus</i> (PKS)	Tetraciclinas
1948	Eritromicina	<i>Streptomyces erythraeus</i> (PKS)	Macrolídeos
1948	Cefalosporinas	<i>Cephalosporium acremonium</i> (NRPS)	β -lactâmicos
1953	Estreptogramina B	<i>Streptomyces graminofaciens</i> (NRPS)	Estreptograminas
1953	Vancomicina	<i>Amycolatopsis orientalis</i> (NRPS)	Glicopeptídeos
1957	Rifamicina	<i>Streptomyces mediterranei</i> (NRPS/ PKS)	Rifamicinas
1960	Metronidazol	<i>Streptomyces</i> sp. (semi-sintético)	Nitromidazóis
1962	Ácido fusídico	<i>Fusidium coccineum</i> (Terpeno)	Esteróides
1963	Lincomicina	<i>Streptomyces lincolnensis</i> (NRPS)	Lincosamidas
1966	Bleomicina	<i>Streptomyces verticillus</i> (PKS/NRPS)	Glicopeptídeos
1969	Fosfomicina	<i>Streptomyces fradiae</i>	Fosfomicinas
1971	Mupirocina	<i>Pseudomonas fluorescens</i> (PKS)	Ácidos carboxílicos
1986	Daptomicina	<i>Streptomyces roseosporus</i> (NRPS)	Lipopeptídeos

* Ano da descoberta.

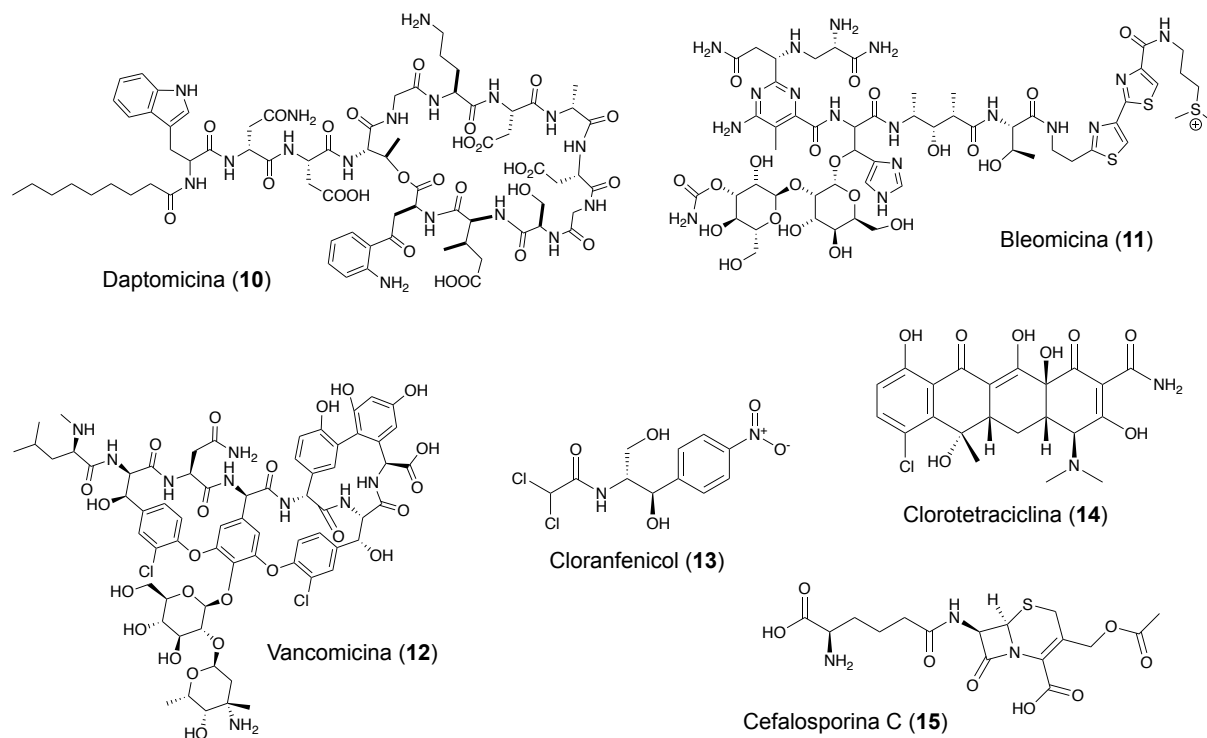


Figura 2. Exemplos das principais classes de antibióticos descobertos durante a “Era de Ouro”.

1.2.1. Resistência microbiana

Organismos produtores de antimicrobianos também apresentam genes de resistência como mecanismo de defesa, tanto aos metabólitos por eles produzidos, como para sobreviver às relações complexas no ambiente. Com o aumento do uso de antibióticos em humanos, os microrganismos, principalmente patógenos, desenvolveram meios de adquirir esses genes de resistência, seja por transferência horizontal ou mutação, e poucos anos depois da descoberta dos primeiros antibióticos já começaram a surgir os primeiros organismos resistentes.^{9,31,32}

Além disso, o uso indiscriminado desses medicamentos acelerou a seleção desses microrganismos resistentes e, nos últimos anos, iniciou-se uma corrida na descoberta de novas moléculas bioativas contra o aumento exponencial de organismos resistentes.^{32,33} Em 2012, o número de organismos patogênicos resistentes aos antibióticos em uso clínico era de 70% e a mortalidade por essas infecções multirresistentes variava entre 50 a 80%. Nos Estados Unidos, 2 milhões de pessoas foram infectadas por bactérias multirresistentes e a mortalidade chegou a 100.000 pessoas por ano.³⁴

Em 2014, a Organização Mundial da Saúde (OMS) reportou a resistência bacteriana como um problema de saúde pública mundial,³⁵ sendo já observada resistência de bactérias para a maioria das classes de antibióticos disponíveis para tratamento clínico. Foi anunciada a urgência de planos de ação para descoberta de novas classes de moléculas para transpor a barreira atual que limita o tratamento de infecções causadas por bactérias super-resistentes. E neste cenário, a importância da investigação e isolamento dos produtos naturais para a descoberta de novos fármacos foi revista.^{36–38}

1.3. Era pós-genômica

As primeiras estratégias de isolamento e identificação de produtos naturais eram focadas principalmente em Actinobacterias, na procura por agentes com nenhuma toxicidade em humanos e amplo espectro de ação, ou seja, ativos no combate de infecções causadas por múltiplas espécies de microrganismos, o que tornou a busca muito seletiva, levando ao reisolamento de vários compostos conhecidos e descartou inúmeros outros de baixo espectro.¹⁵ Somado a esses fatores, a dificuldade de isolamento, purificação, determinação estrutural e alto custo associado as triagens resultaram, ao final dos anos 70, em um cenário frustrante, desacelerando a descoberta de novas moléculas bioativas. As indústrias farmacêuticas reduziram drasticamente ou anularam seus investimentos e esforços na bioprospecção de produtos naturais, sendo impulsionadas, entre os anos 80 e 90, pela relativa rapidez na obtenção de bibliotecas de compostos sintéticos a partir da química combinatória e química computacional.^{15,39}

Contudo, o avanço nas tecnologias de sequenciamento de DNA, a partir do século XXI, viabilizou o rápido acesso ao genoma completo de diversos organismos. Quando o sequenciamento do genoma de *Streptomyces coelicolor* A3(2) foi anunciado, o primeiro do gênero, apenas três antibióticos e um pigmento eram conhecidos, enquanto a informação do seu genoma revelou a presença de pelo menos 20 *clusters* de genes codificando para o metabolismo secundário, incluindo sideróforos, pigmentos, policetídeos e moléculas de sinalização.⁴⁰ Posteriormente, a publicação do genoma completo de *S. avermitilis*, conhecida por produzir o agente antiparasitário avermectina, indicou a presença de 30 *clusters* biossintéticos.⁴¹ Diferentes projetos de sequenciamento de *Streptomyces* revelaram um potencial

biossintético muito mais expressivo do que aquele acessado em cultivos de laboratório esperando para ser descoberto.⁴²

A partir de então, como alternativa aos métodos de triagem tradicionais, surgiu a estratégia de *genome mining*, guiando cientistas a partir da informação do sequenciamento do genoma na identificação de potenciais linhagens e metabólitos secundários não acessados em condições normais de cultivo em laboratório, estimulando o interesse na triagem por novos produtos naturais.^{43,44} Além disso, mesmo a investigação de organismos não sequenciados foi reestabelecida, estimulando cientistas a explorar estratégias de cultivos e co-cultivo para induzir a expressão de genes silenciados ou produzidos em baixa concentração.

O acesso às bases de dados contendo inúmeras sequências depositadas permitiram conectar os genes (ou a informação genética) às moléculas obtidas e, apesar da grande diversidade química obtida, observou-se um repertório limitado de enzimas codificadas por *cluster* (ou conjuntos) de genes responsáveis pela biossíntese de metabólitos secundários com atividade biológica relevante e dentre elas destacam-se as sintases de policetídeos (PKSs), sintetases de peptídeos não ribossomais (NRPSs) e sintases de terpenos.⁴⁵ Todas essas enzimas possuem sequências de DNA conservadas que foram utilizadas como guias nas triagens, inicialmente, experimentais por *Southern blot* ou PCR (reação em cadeia da polimerase), e permitiram que muitos genes ditos “crípticos”, “silenciados” ou “órfãos”, codificando enzimas biossintéticas putativas envolvidas na produção de metabólitos secundários, fossem identificados.^{46–49}

Adicionalmente, o *genome mining* viabilizou o refinamento de metodologias direcionadas a triagem de metabólitos em função, por exemplo, das propriedades físico-químicas ou das especificidades frente a substratos, além de outras como deleção de genes/avaliação do perfil metabólico, expressão heteróloga, reconstituição *in vitro* e aproximações genomisotópicas (Figura 3), que levaram a descoberta de novos policetídeos, terpenóides e peptídeos não-ribossomais.^{50,51,60–62,52–59}

A estrutura de coeliquelina (peptídeo não ribossomal) foi inicialmente sugerida a partir do *cluster* de genes anotado de *S. coelicolor* M145 e um método desenvolvido para prever a especificidade de incorporação do domínio de adenilação (A) de cada módulo.⁵⁰ Em seguida, um experimento de deleção de genes, combinado a análise de propriedades físico-químicas do composto, viabilizaram o isolamento e a determinação estrutural correta.⁵¹ Já a descoberta da orfamida A,⁶² um

lipopeptídeo cíclico produzido por *Pseudomonas fluorescens* Pf-5, foi direcionada pela previsão *in silico* das subunidades incorporadas ao longo da biossíntese e a molécula foi identificada a partir da incorporação de precursores isótopo-marcados estáveis adicionados ao meio de cultivo.⁶³

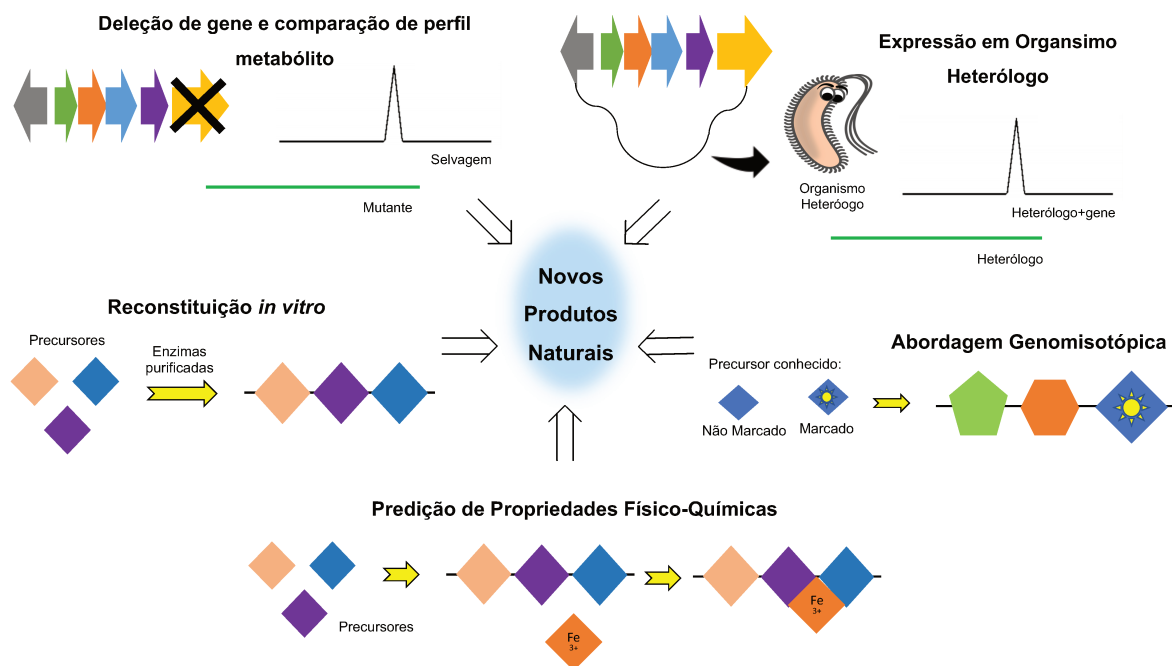


Figura 3. Estratégias via *genome mining* para obtenção de novos produtos naturais bioativos. (Adaptado de Corre & Challis⁶⁴)

Outros exemplos de compostos bioativos derivados das vias NRPS/PKS descobertos por estratégias de *genome mining* incluem a salinilactama A⁶⁵ de *Salinispora tropica* CNB-440, guiada pela combinação de previsões do genoma e propriedades físico-químicas (presença de um grupo cromóforo característico na estrutura); chaxamicinas A–D⁶⁶ via triagem do genoma por PCR; emericelamida A⁶⁷ via deleção de genes; acetilaszonalenina⁶⁸ obtida por reconstituição *in vitro* e, mais recentemente, a taromicina A⁶⁹ identificada após expressão em organismo heterólogo do *cluster* de genes silenciado.

Atualmente, existem 171.328 projetos de sequenciamento de genoma completo de diferentes organismos registrados em base de dados (JGI GOLD – *Genomes Online Database*; disponível em: <https://gold.jgi.doe.gov/index> - data de acesso: 03 de Julho de 2019), sendo 16.778 projetos completos e publicados, 135.440 *drafts* permanentes e 18.349 projetos incompletos. O filo Actinobacteria corresponde

a 16.399 projetos de sequenciamento, sendo 123 referentes a projetos de sequenciamento de genomas de *Streptomyces* completos e publicados.

1.3.1. Ferramentas de bioinformática

A facilidade de acesso e o baixo custo associado ao sequenciamento de genomas levaram ao aumento do número de sequências depositadas e ao acúmulo de informações, dificultando as triagens experimentais e “manuais”. Sendo assim, com o intuito de viabilizar a análise *in silico* de todas essas informações de modo rápido e preciso, ferramentas como BLAST⁷⁰ e DIAMOND⁷¹ foram desenvolvidas para alinhamento de sequências de DNA com base em algoritmos.⁴⁴

Posteriormente, foram desenvolvidas ferramentas de bioinformática que automaticamente identificam e anotam os *clusters* de genes codificando enzimas biossintéticas do metabolismo secundário. O mais conhecido e utilizado atualmente é o antiSMASH (*antibiotics and Secondary Metabolite Analysis SHell*), que foi lançado em 2011⁷² e já está na sua versão 5.0,⁷³ desenvolvido para análise de genomas de bactérias, fungos e plantas. Essa plataforma online e gratuita concentra diferentes ferramentas para alinhamento de sequências (NCBI BLAST+, HMMer⁷⁴, Muscle⁷⁵) e comparação das informações para identificação de enzimas putativas.

O MIBiG (*Minimum Information about a Biosynthetic Gene cluster*)^{76,77} é outra plataforma de grande importância, conectada recentemente ao antiSMASH, que reúne informações depositadas por cientistas sobre anotações de *clusters* biossintéticos e suas entidades químicas associadas. Essa ferramenta auxilia na triagem por *clusters* de genes inéditos, otimiza o processo de *mining* dos genomas, viabiliza avaliação da variabilidade de *clusters* de genes biossintéticos (BGC) de acordo com o ambiente, entre outros.

Outras ferramentas de bioinformática podem ser encontradas em uma comunidade online denominada “*The Secondary Metabolite Bioinformatics Portal*” (disponível em: <http://www.secondarymetabolites.org>).

1.4. Principais enzimas envolvidas na biossíntese de produtos naturais

As informações obtidas através dos sequenciamentos de genomas de diversos organismos, associadas às ferramentas de bioinformática viabilizaram

estudos e caracterizações de enzimas responsáveis pela biossíntese de produtos naturais, os quais compreendem uma grande diversidade de classes de moléculas representadas por terpenos, alcalóides, fenazinas, peptídeos ribossomais modificados pós-traducionalmente (RiPPs), policetídeos (PK) e peptídeos não ribossomais (NRP).

Dentre esses compostos, os policetídeos (PK) e os peptídeos não ribossomais (NRP), biossintetizados por sintases de policetídeos (PKSs) e sintetases de peptídeos não ribossomais (NRPSs), respectivamente, se destacam por apresentarem um grande espectro de atividades farmacológicas (antimicrobianas, imunossupressores, antitumorais). São produzidos por inúmeros organismos, sendo as actinobactérias os mais prolíficos e a enzimologia do arranjo colinear dessas multienzimas é razoavelmente bem conhecida, possibilitando, na maioria dos casos, estabelecer uma correspondência direta entre os genes e o produto final.

1.4.1. Sintase de Policetídeo (PKS)

Sintases de policetídeos são enzimas responsáveis pela biossíntese de compostos com uma grande diversidade estrutural, tais como macrolídeos, policetídeos aromáticos e polienos. Essas enzimas podem ser classificadas de acordo com sua organização estrutural e mecanismo catalítico em três tipos, PKS do tipo I, II e III e muito do que se sabe sobre como elas atuam nas diversas etapas biossintéticas foi impulsionado pelos estudos da sintase do policetídeo eritromicina.

A eritromicina, isolada de *Saccharopolyspora erythraea*⁷⁸, é biossintetizada por ação de uma PKS do tipo I que compreende três genes principais (*eryAI*, *eryAII*, *eryAIII*) codificando para as enzimas DEBS 1, 2 e 3, responsáveis pela produção do 6-deoxieritronolídeo B (6DEB), o qual sofre modificações pós-traducionais para formação da eritromicina A. Estudos com essas enzimas foram fundamentais e possibilitaram a compreensão da organização modular de PKSs do tipo I, sendo cada módulo subdividido em vários domínios enzimáticos e responsável por um ciclo de reação para incorporação de uma unidade na cadeia policetídica. O resultado é a correlação direta entre o número de módulos presentes na eritromicina-PKS e o comprimento da cadeia do policetídeo (Figura 4).⁷⁹

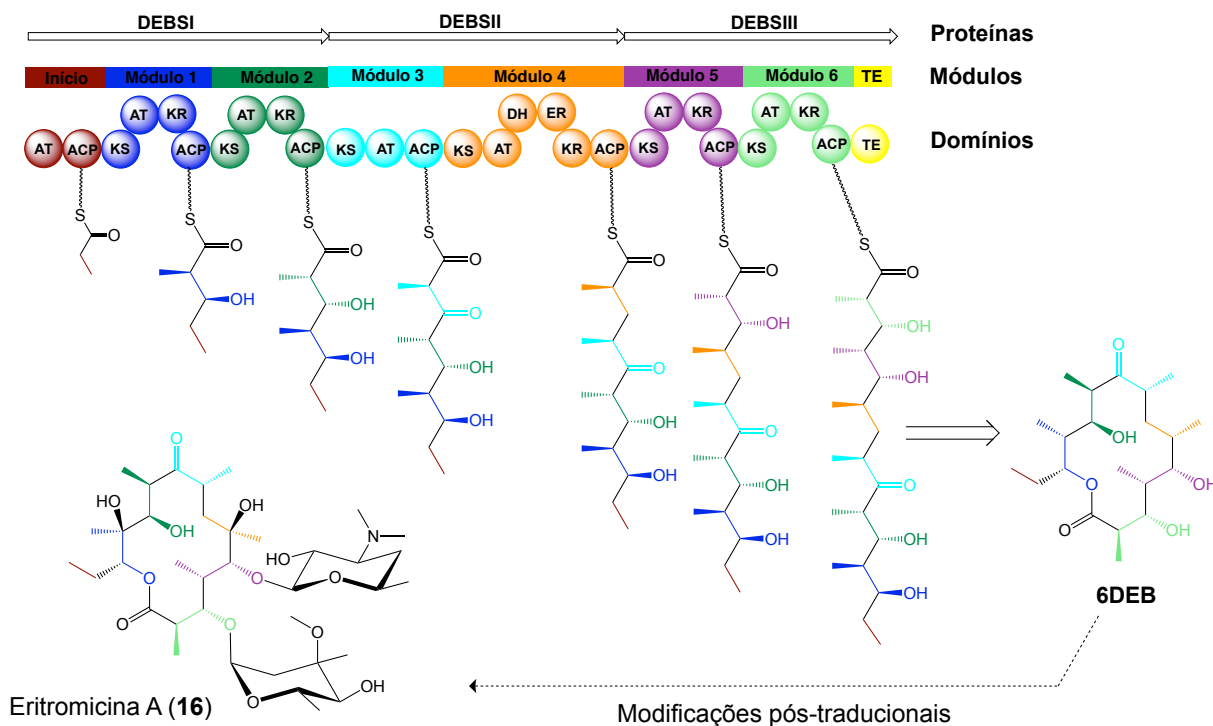


Figura 4. Biossíntese da Eritromicina A. O *cluster* de três genes codifica para as proteínas DEBS1-3 e cada proteína possui dois módulos extensores contendo três domínios obrigatórios (KS, AT e ACP) e domínios redutores opcionais (KR, DH e ER). DEBS1 contém o módulo iniciador de cadeia e DEBS3 contém uma tioesterase (TE), responsável pela ciclização e liberação da cadeia do complexo enzimático (adaptado de Menzella et al⁸⁰).

As PKSs do tipo I são classificadas, portanto, como modulares ou interativas, sendo o primeiro grupo exclusivo de bactérias. As PKSs modulares, como já mencionado, possuem diferentes módulos subdivididos em no mínimo três domínios enzimáticos obrigatórios, representados na via de biossíntese da eritromicina A (**16**) pela cetossintase (KS), aciltransferase (AT) e proteína carreadora de acila (ACP). O domínio AT é responsável pela seleção da unidade extensora da cadeia (malonil- ou metilmalonil-CoA) e transferência para o domínio ACP. O domínio ACP é ativado pelo cofator de fosfopanteteína, onde se liga uma unidade extensora de malonila ou metilmalonila que é incorporada a cadeia policetídis em formação via reação de condensação de Claisen, catalisada pelo domínio KS (Figura 5A). Adicionalmente, podem ser citados os domínios redutores opcionais, responsáveis pela diversidade de estruturas obtidas, cetorreductase (KR), desidratase (DH) e enoilreductase (ER) (Figura 5B), além dos domínios de transferência de metila (-CH₃) a um átomo de carbono ou oxigênio, conhecidos como metiltransferases (MT). Em

fungos, as PKSs do tipo I são monomodulares e interativas, sendo que um único módulo contém minimamente os domínios enzimáticos obrigatórios que catalisam mais de um ciclo de condensação na formação do policetídeo.^{81–83}

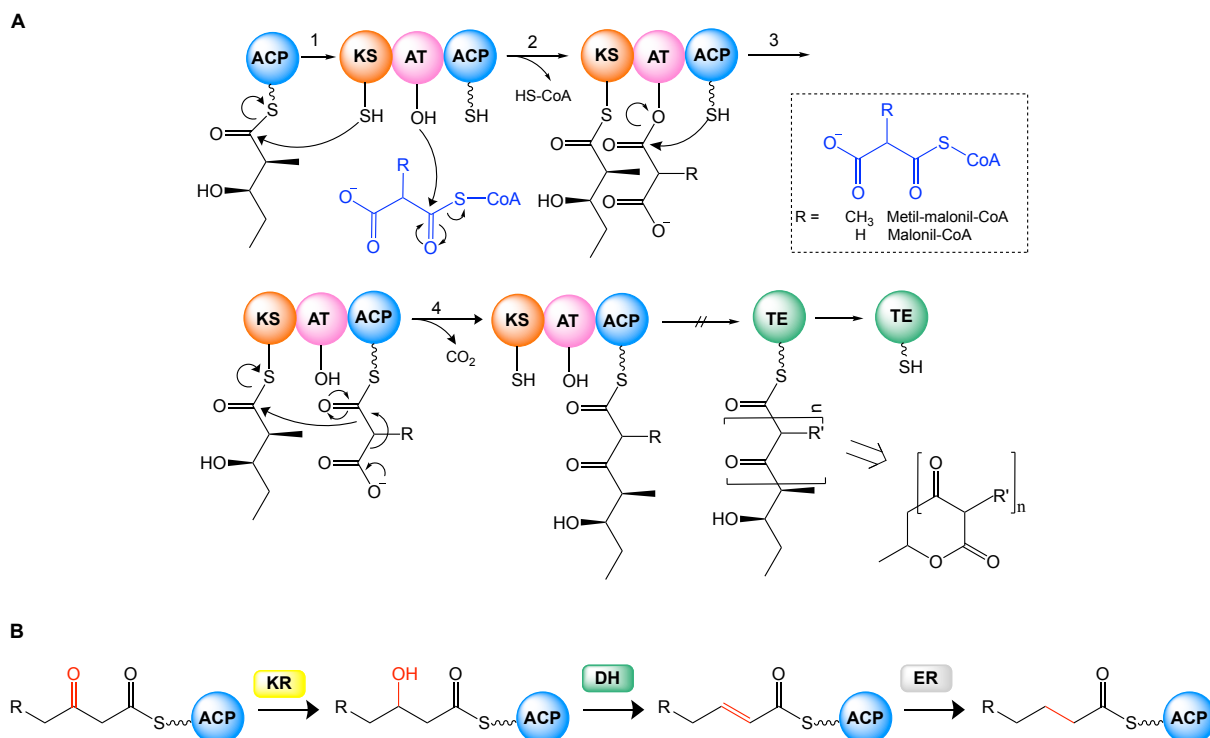


Figura 5. Sintase de policetídeo (PKS) do tipo I. **A.** Mecanismo de reação de incorporação de “n” unidades extensoras na cadeia policetí dica via reação de condensação de Claisen catalisada por um domínio obrigatório de KS em uma PKS do tipo I. Ao final, o policetídeo é ciclizado no domínio TE (Adaptado de Kapur et al⁸²). **B.** Domínios redutivos opcionais. KR: cetorredutase; DH: desidratase; ER: enoilredutase.

A salinomicina (**17**), um antibiótico com atividade terapêutica no tratamento de câncer de mama,⁸⁴ é outro importante exemplo de policetídeo produzido pela ação enzimática de PKS do tipo I modular em *Streptomyces albus*. Já a lovastatina (**18**), pertencente a classe das estatinas e utilizada no tratamento clínico para redução de níveis de colesterol, foi isolado do fungo *Aspergillus terreus* e é um exemplo de policetídeo biossintetizado por uma PKS do tipo I interativa (Figura 6).⁸⁵

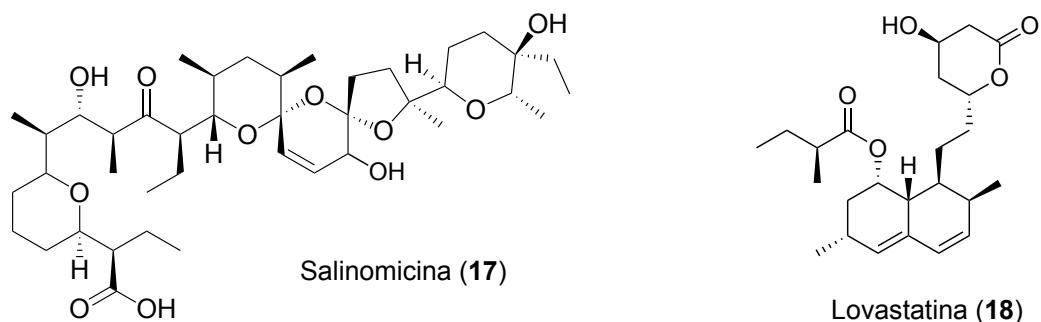


Figura 6. Exemplos de policetídeos biossintetizados por PKS do tipo I modular (salinomicina) e interativa (lovastatina).

PKSs do tipo II são complexos multienzimáticos exclusivos de bactérias, onde cada domínio catalítico compõe proteínas distintas que atuam de forma interativa, formando heterodímeros. Possuem os domínios mínimos de KS e ACP, com a diferença de que existem dois domínios KS (KS_{α} e KS_{β}), sendo que o segundo não possui atividade catalítica de condensação e atua regulando o tamanho da cadeia em formação (Figura 7A). Essas enzimas são responsáveis pela biossíntese de policetídeos aromáticos ou polifenóis e os domínios redutivos não participam da rota biossintética desses policetídeos.^{81,83} A doxorrubicina (19), pertencente à classe das antraciclínas, é um importante exemplo de policetídeo biossintetizado por ação de uma PKS do tipo II em linhagem modificada de *Streptomyces peucetius* subsp. *caesius*. Foi listada pela OMS como um medicamento essencial para a saúde básica da população no tratamento de câncer (Figura 7B).⁸⁶

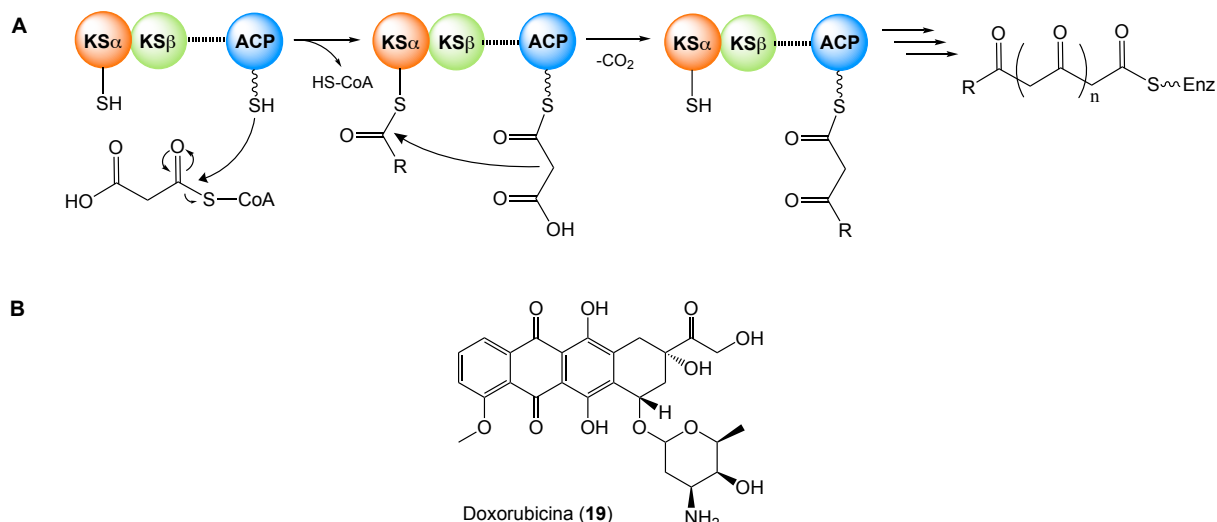


Figura 7. Sintase de policetídeo (PKS) do tipo II. **A.** Domínios catalíticos envolvidos na biossíntese de policetídeos aromáticos e polienos (adaptado de Hertweck et al⁸⁷). **B.** Estrutura do policetídeo doxorubicina biossintetizado via uma PKS do tipo II.

Diferentemente dos outros grupos, as PKSs do tipo III são mais simples, formadas por duas subunidades idênticas de KS e independentes de ACP. O substrato é reconhecido pelo domínio de KS em sua forma livre, ligado a coenzima A (Co-A), o que garante uma menor especificidade do domínio catalítico gerando maior variedade de estruturas, tais como chalconas, pironas, acridonas, floroglucinóis, estilbenos e alquilresorcinóis. Estudos dessas enzimas são bem estabelecidos em plantas e bactérias, e mais recentemente foram também isolados policetídeos do tipo III de fungos (Figura 8).^{88–90}

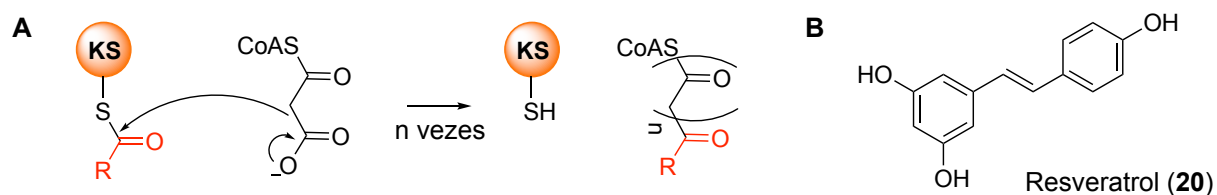


Figura 8. Sintase de policetídeo (PKS) do tipo III. **A.** Domínio catalítico essencial composto apenas por KS **B.** Resveratrol (20),⁹¹ exemplo de policetídeo biossintetizado por PKS do tipo III.

Adicionalmente, os sistemas modulares de PKS podem ser divididos em *cis*-AT e *trans*-AT (ou AT-less). O primeiro sistema compõe os arranjos lineares de PKS, como apresentado para a PKS do tipo I na biossíntese da eritromicina A, o segundo foi descoberto mais recentemente a partir de estudos mais detalhadas dos

complexos enzimáticos. No sistema modular contendo uma *trans*-AT, o domínio AT não se encontra posicionado na sequência linear dos módulos de biossíntese, em geral, é codificado por um gene adjacente aos genes biossintéticos e pode atuar de forma interativa na incorporação das unidades extensoras da cadeia policetídica, sendo específico para substratos do tipo malonil-CoA. São sistemas mais comuns de organismos não-cultiváveis, em relações de mutualismo e patogênicos, indicando um grande potencial para descoberta de novas moléculas bioativas.^{81,92,93}

O primeiro gene codificando para uma *trans*-AT PKS foi reportado em 1993 no genoma de *Bacillus subtilis*.⁹⁴ Contudo, o metabólito bacilaeno (**21**)⁹⁵ associado só foi identificado e caracterizado em 2007, pelo mesmo grupo de pesquisas, coordenado por Jon Clardy.⁹⁶ O mecanismo de ação interativa dessa classe de enzimas foi avaliado e caracterizado em 2003, em estudos *in vitro* para elucidação da biossíntese de leinamicina (**22**), isolada de *Streptomyces atroolivaceus* S-140.⁹⁷ A partir de então, diferentes policetídeos foram caracterizados de vias de *trans*-AT PKS.^{92,98,99}

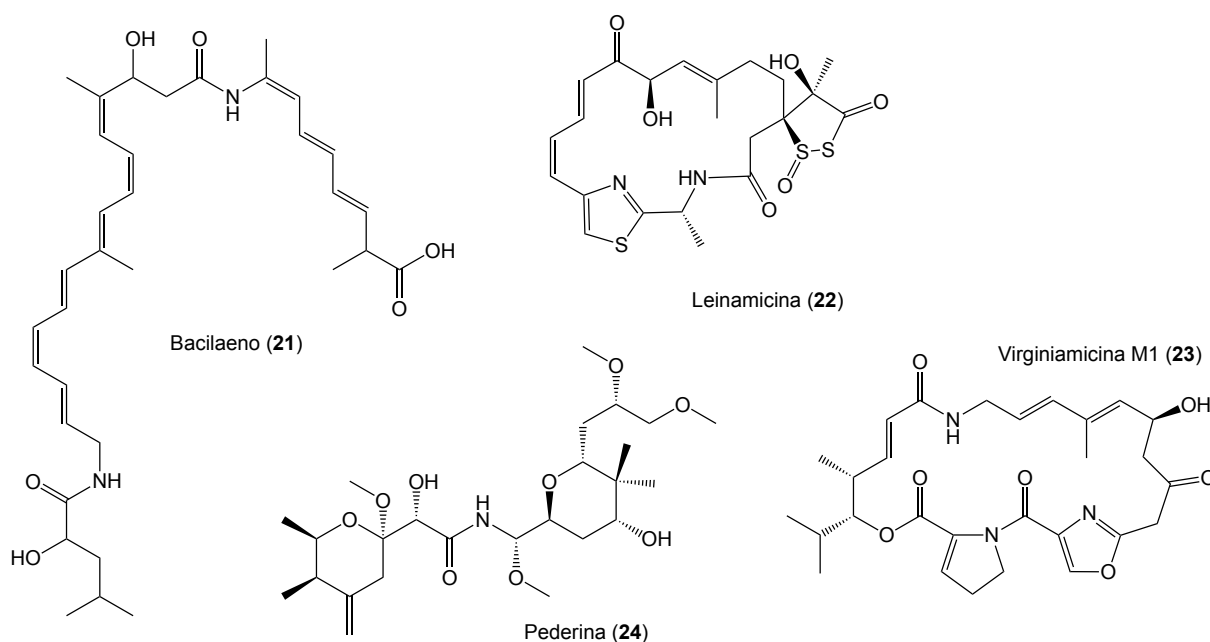


Figura 9. Estruturas de policetídeos derivados de vias biossintéticas de *trans*-AT PKS.

1.4.2. Sintetase de Peptídeo Não Ribossomal (NRPS)

Peptídeos não ribossomais (NRP), como o próprio nome diz, são peptídeos biossintetizados fora do ribossomo por enzimas denominadas sintetases de peptídeos

não ribossomais (NRPS), as quais apresentam uma organização modular semelhante ao apresentado para PKS do tipo I e são constituídas por três domínios catalíticos essenciais, adenilação (A), condensação (C) e tiolação ou proteína carreadora de peptídica (PCP). O domínio de adenilação é responsável pela seleção, ativação do aminoácido extensor da cadeia peptídica a partir de uma molécula de adenosina trifosfato (ATP) e transferência para o domínio PCP posterior adjacente. O domínio de condensação catalisa a formação da ligação peptídica entre o aminoácido a ser incorporado e a cadeia de peptídeos em formação, ligada ao PCP anterior adjacente (Figura 10). O domínio PCP também sofre ativação do resíduo de serina (Ser) no sítio ativo da enzima pela ligação de fosfopanteteína.¹⁰⁰

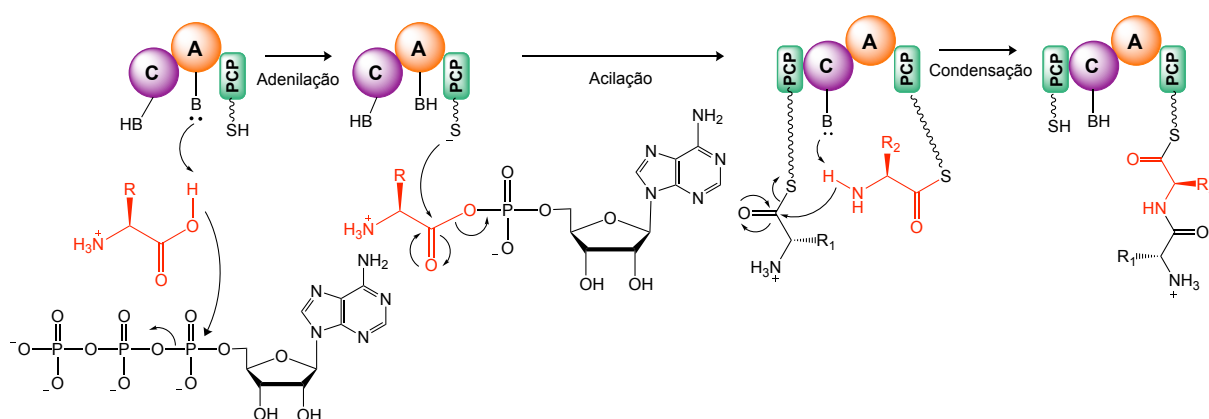


Figura 10. Etapas de biossíntese de peptídeo não ribossomal catalisadas pelos domínios obrigatórios (C, A e PCP) de NRPS. (adaptado de Fischbach & Walsh¹⁰⁰)

Adicionalmente, NRPSs também possuem domínios opcionais que geram maior diversidade estrutural, tais como o domínio de epimerização (E), que catalisa a formação de D-aminoácidos, epímeros dos aminoácidos naturais (L-aminoácidos) (Figura 11A). Os domínios Cy catalisam reações de ciclização intramolecular entre resíduos de Cys, Ser ou Thr e a carbonila do resíduo anterior, formando anéis de tiazolina ou oxazolina (Figura 11D). Esses anéis de cinco membros formados podem sofrer oxidação catalisada de domínio Ox gerando anéis de oxazol e tiazol (Figura 11B). Anéis de tiazolina e oxazolina podem sofrer reações catalisadas por domínios de redução (Re), favorecendo a transferência de grupos metilas ao nitrogênio, por ação de uma metiltransferase (MT) (Figura 11C).¹⁰⁰

A daptomicina (**10**), isolada de *Streptomyces roseosporus*, é um antibiótico produzido por uma NRPS e se destaca por possuir diferentes mecanismos de ação,

sendo de importante aplicação para o tratamento de infecções causadas por bactérias multirresistentes, tais como *Staphylococcus aureus* metilina-resistente (MRSA) e enterococci vancomicina-resistente (VRE).¹⁰¹

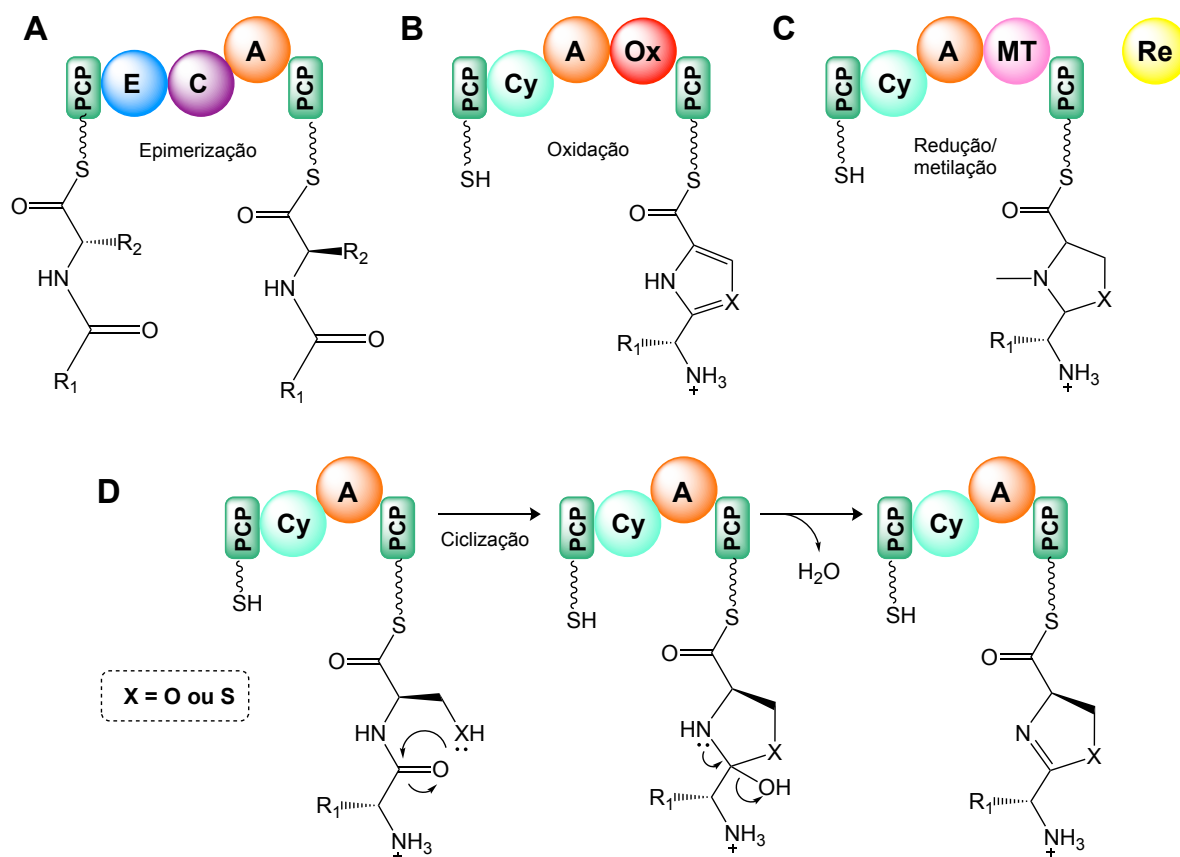


Figura 11. Reações catalisadas por domínios opcionais em NRPS. **A.** Domínio de epimerização. **B.** Domínio de oxidação de anéis tiazolina e oxazolina. **C.** Domínios redutivos e de transferência de metila. **D.** Domínio de ciclização intramolecular e formação de ligação amida.

1.5. Clonagem de *clusters* de genes biossintéticos

A manipulação genética de microrganismos, associada com os avanços de metodologias de clonagem e edição de genes para expressão heteróloga, desenvolveram-se como ferramentas altamente vantajosas, uma vez que a síntese química (ou etapas dela) de fármacos e outros insumos químicos por diversas rotas podem ser substituídas pela fermentação de microrganismos, e juntamente com as estratégias de *mining*, constituem claramente alternativas para a obtenção de novas moléculas, considerando o grande desafio em se encontrar moléculas inéditas candidatas para introdução no tratamento humano e o combate a resistência de

microrganismos patogênicos aos tratamentos com antibióticos existentes. A seguir serão detalhadas algumas estratégias de clonagem de *clusters* biossintéticos, com foco em genes codificando para NRPSs e PKSs, que viabilizaram a obtenção de importantes compostos ao longo das últimas décadas.

1.5.1. Bibliotecas Genômicas

A construção de bibliotecas genômicas foi uma das primeiras estratégias para obtenção de grandes fragmentos de DNA e de grande importância na era do sequenciamento. Essas bibliotecas tem por objetivo representar de maneira significativa o genoma completo de um determinado organismo em diferentes clones, permitindo o mapeamento e estudo de genomas complexos (eucariotos e procariotos), estudo e caracterização de vias de biossíntese de metabólitos secundários, clonagem de *clusters* de genes crípticos ou silenciados e transferência para ativação da expressão em organismo heterólogo ou engenharia genética, entre outras aplicações.

As primeiras construções foram obtidas em cosmídeos, híbridos de bacteriófagos (infectam células bacterianas com seu DNA linear) e plasmídeos (DNA circular extra cromossomal que se replica de forma autônoma), os quais comportam fragmentos de DNA de 35 a 45 kb. Apesar das limitações do tamanho dos fragmentos clonados, os cosmídeos foram largamente utilizados para caracterização de genes biossintéticos, tais como da vancomicina¹⁰² e dos macrolídeos metimicina, neometimicina, narbomicina e picromicina.¹⁰³ As clonagens de fragmentos maiores (100 a 2000 kb) foram obtidas em cromossomos artificiais de levedura (YAC) e ganharam destaque em estudos envolvendo mapeamento de genomas mais complexos, como o humano.¹⁰⁴

Posteriormente, as bibliotecas genômicas em cromossomos artificiais bacterianos (BAC) e cromossomos artificiais P1-derivados (PAC) se destacaram por superar alguns problemas de bibliotecas de YACs, tais como a instabilidade dos clones obtidos e a dificuldade de extrair/isolar o DNA de leveduras. Os vetores do tipo PAC/BAC comportam fragmentos menores que YACs (até 300 kb),¹⁰⁵ uma importante vantagem com relação aos cosmídeos para estudos envolvendo as vias biossintéticas de PKSs e NRPSs, em geral, codificadas em regiões que compreendem até 120 kb.¹⁰⁶ A aplicação de BACs viabilizou a obtenção de um clone contendo o *cluster* completo

relacionado a biossíntese de geldanamicina¹⁰⁷ e a construção de uma biblioteca genômica da esponja *Halichondria* sp.¹⁰⁸

Especialmente após a obtenção de vetores ESAC (cromossomo artificial *Escherichia coli-Streptomyces*), que viabilizaram a transferência direta para *Streptomyces*, a construção de bibliotecas genômicas teve um maior destaque na obtenção de *cluster* biossintéticos e expressão heteróloga para obtenção de novos compostos. O pESAC13,¹⁰⁹ derivado do vetor pPAC-S1,¹¹⁰ é um exemplo de um vetor construído visando reduzir as limitações encontradas para clonagem e manipulação genética em estudos de biossíntese de produtos naturais, bem como para viabilizar a transferência dos clones para linhagens de *Streptomyces*. Possui os elementos necessários para replicação (P1-derivado) e transferência (*oriT*) em *E. coli* e para integração específica no sítio *attB* do cromossomo de *Streptomyces* (ϕ C31 *attP-int*).

A expressão em organismo heterólogo (*S. coelicolor* M512) do clone obtido em pESAC contendo os genes relacionados a biossíntese do importante imunossupressor FK-506 (nome comercial Prograf ou tacrolimus), possibilitou o aumento da produção do composto em 5 vezes comparado a concentração obtida em cultivo da linhagem selvagem.¹⁰⁹ Outro exemplo importante envolve a clonagem do *cluster* de genes *act* de *Streptomyces* sp. T676 e elucidação da biossíntese da antracimicina, um potente antibiótico reportado em 2013 contra MRSA e VRE, produzido por ação de uma *trans*-AT PKS.¹¹¹

Contudo, apesar de bibliotecas genômicas se mostrarem adequadas e eficientes na obtenção de *clusters* biossintéticos e transferência para organismos heterólogos, algumas limitações persistiram e inviabilizaram o fácil acesso devido as dificuldades na construção, relacionadas a obtenção de grandes fragmentos de DNA, clonagem em vetores apropriados, além do custo e tempo associados. Em paralelo, os rápidos avanços nas metodologias de recombinação e clonagem direta de fragmentos do DNA genômico transpuseram, recentemente, as vantagens das bibliotecas genômicas.

1.5.2. Recombinação de fragmentos de DNA *in vivo* e *in vitro*

Os eventos de recombinação homóloga de fragmentos de DNA ocorrem em diferentes organismos (bactérias, fungos, plantas e animais) e são importantes para garantir a evolução dos mesmos, funcionando como um mecanismo de reparo

de DNA ou de troca de informação gênica entre cromossomos.¹¹² O conhecimento sobre esses eventos naturais de recombinação permitiu o desenvolvimento de ferramentas para aplicação em engenharia genética, viabilizando a recombinação e clonagem de DNA em sistemas *in vitro* e *in vivo*.

1.5.2.1. Sistemas λ Red e RecET

Experimentos de recombinação homóloga *in vivo* são utilizados em engenharia genética, com a vantagem de permitir a inserção ou deleção de fragmentos sem a necessidade do uso de endonucleases de restrição, ou seja, localização de sítios específicos de restrição para clonagem.¹¹³ A recombinação pode ocorrer entre fragmentos lineares ou circulares contendo regiões de homologia, sendo promovida por dois sistemas principais o λ Red ou Red $\alpha\beta$ (Gam, Exo, Beta), proveniente do fago (ou bacteriófago) λ , e o RecET (RecE e RecT), proveniente do profago Rac. Ambos são formados por genes que expressam duas enzimas cooperativas, uma exonuclease que degrada dsDNA no sentido 5'→3' (Exo ou RecE) e uma proteína que promove a ligação entre fitas simples de DNA complementares (Beta ou RecT). Ocorre também, a expressão de uma terceira enzima auxiliar (Gam), a qual inibe ação de nucleases RecBCD presentes nas células para digestão de fragmentos lineares exógenos, permitindo a ação das enzimas do sistema de recombinação.^{113–115}

Para viabilizar o uso desses sistemas em *E. coli*, diferentes plasmídeos foram desenvolvidos, os quais podem ser transferidos para outro organismo para expressão dos genes codificando as proteínas necessárias para recombinação de fragmentos de DNA. O pIJ790 (λ Red),¹¹⁶ por exemplo, contém origem de expressão das enzimas do sistema de recombinação controlada por temperatura e, assim como pBAD-ET γ ,¹¹⁷ indução por L-arabinose.

A literatura descreve diferentes aplicações utilizando recombinações homólogas *in vivo*, tais como estudos utilizando fragmentos lineares obtidos por PCR, contendo regiões de diferentes tamanhos de homologia (20 e 1000 pb) flanqueando uma região^{115,118} para modificação/deleção em plasmídeos,¹¹⁸ cosmídeos^{116,119} e inativação de gene em PAC.¹²⁰ Outros estudos apresentam metodologias de clonagem de *cluster* de genes completo, partindo do DNA genômico digerido parcialmente com enzimas de restrição e do plasmídeo contendo regiões de homologia flanqueando o *cluster*.^{121,122} Ficou estabelecido, após diferentes estudos,

que sistemas λ Red são eficientes para recombinação homóloga linear-circular (LCHR) e sistemas RecET para recombinação homóloga linear-linear (LLHR).¹²³

1.5.2.2. Gibson Assembly®

Com base nos sistemas de recombinação descritos anteriormente, Gibson *et al.*, conseguiram reconstituir o genoma da bactéria artificial *Mycoplasma genitalium*, utilizando uma combinação de recombinação *in vitro* e *in vivo*.¹²⁴ Observaram que era possível reconstituir o genoma completo da mesma bactéria a partir de 25 fragmentos contendo sobreposição entre si e transformação em *Saccharomyces cerevisiae*.¹²⁵ A partir dessas observações, os autores realizaram experimentos de recombinação de fragmentos *in vitro* utilizando exonuclease, DNA polimerase e DNA ligase em uma reação isotérmica de etapa única e observaram que o tamanho das regiões de sobreposição necessário para a recombinação pode variar entre 15 e 500 pb, sendo também a quantidade de exonuclease um fator limitante, visto que quanto maior a região de sobreposição, maior a concentração de enzima necessária. Também avaliaram o tamanho dos fragmentos a serem recombinados e obtiveram resultados positivos para recombinação de fragmentos até 150 kb.¹²⁶

Com esses estudos desenvolveram a metodologia do *Gibson Assembly®* utilizando uma mistura de reação com condições ótimas de concentração enzimática para recombinação *in vitro* de 2 a 6 fragmentos em uma reação isotérmica de etapa única, que pode ser obtida comercialmente (*Gibson Assembly® Master Mix*) ou preparada em laboratório,¹²⁶ e confirmaram a viabilidade da metodologia ao reconstituir o genoma mitocondrial de rato (total de 16,5 kb).¹²⁷

1.5.2.3. TAR-cloning

A recombinação associada a transformação (*Transformation-Associated Recombination - TAR*) baseia-se no sistema natural de recombinação *in vivo* de leveduras e os primeiros princípios da técnica foram desenvolvidos no final da década de 90, possibilitando a clonagem direta de grandes fragmentos do DNA genômico de mamíferos e recombinação em levedura para obtenção de cromossomos artificiais YACs.^{128–132} Ao longo dos últimos 10 anos, a técnica foi aperfeiçoada para clonagem de genes específicos com alta reprodutibilidade e eficiência.^{132,133}

Uma das limitações da técnica para clonagem de genes de mamíferos estava relacionada a presença de uma região de sequência de replicação autônoma

(ARS) no fragmento clonado, essencial para a replicação em levedura. Essas regiões são frequentemente encontradas em DNA de mamíferos, mas raramente estão presentes em DNA com alto conteúdo GC. Sendo assim, foi desenvolvido um vetor para clonagem via TAR contendo o elemento ARS.¹³⁴ Posteriormente, Yamanaka e colaboradores¹³⁵ desenvolveram o vetor pCAP01 para clonagem de vias biossintéticas completas diretamente do DNA genômico e identificaram um composto inédito, a taromicina A, análogo da daptomicina.^{135–137} O vetor contém elementos que permitem replicação em levedura (ARSH4/CEN6-TRP1, proveniente de pRS314), replicação e manutenção em *E. coli* (pUC *ori* do SuperCos1) e a região ϕ C31,¹³⁸ derivado de pSET152, que permite a integração sítio-específica do plasmídeo no cromossomo de *Streptomyces*.

Um derivado de pCAP01, o pCAP03-*acc(3)/V*,¹³⁹ foi obtido com o intuito de aumentar a eficiência da triagem de clones recombinantes através da inserção de marcador de seleção (URA3) que codifica para uma descarboxilase que converte ácido 5-fluoro orótico (5-FOA) em 5-fluorouracila, composto tóxico para as células (inibe a síntese de RNA).¹⁴⁰ Adicionalmente, o sítio de restrição de clonagem foi inserido entre o promotor *pADH1* (*Schizosaccharomyces pombe*), o qual não aceita insertos > 130 pb entre a região TATA e o início da transcrição, e URA3. Sendo assim, se o fragmento clonado for maior que o descrito, URA3 não é expresso e as células sobrevivem e na ausência de fragmento, ou recircularização do plasmídeo, o gene é ativado e o crescimento celular inibido na presença de 5-FOA. Com esse novo vetor, promoveram a clonagem direta de *cluster* de genes órfão e identificaram uma série de antibióticos ácidos tiotetrônico, viabilizando a caracterização a nível genético e de biossíntese desses compostos.¹³⁹

1.5.3. Expressão em organismos heterólogos

A expressão heteróloga teve participação fundamental na era pós-genômica, com o sequenciamento de diversos genomas permitindo o acesso ao grande potencial biossintético de diferentes organismos, e tem sido uma das principais estratégias utilizadas para o estudo e obtenção de produtos naturais até anos mais recentes. Em geral é utilizada para obtenção de metabólitos codificados por genes crípticos ou silenciados na linhagem selvagem, para aumentar o nível de expressão de metabólitos em plasmídeos modificados, caracterizar rotas biossintéticas, obter

compostos não-naturais, entre outras aplicações.^{106,141} Diferentes fatores devem ser levados em consideração em experimentos de expressão heteróloga devido a complexidade das enzimas que compõem *clusters* biossintéticos, tais como a escolha de um vetor estável para “carregar” o fragmento de DNA ou o *cluster* de genes e a escolha do organismo heterólogo adequado com o metaboloma bem conhecido e o mais semelhante possível do organismo selvagem.¹⁰⁶

Para a obtenção de produtos naturais, em especial policetídeos e peptídeos não-ribossomais, diferentes organismos vem sendo estudados e aprimorados ao longo das últimas décadas para aplicação em experimentos de expressão heteróloga, de modo a obter, por exemplo, rendimentos em larga escala e com aplicação industrial.¹⁴² *S. coelicolor* é uma das espécies de *Streptomyces* mais estudada e geneticamente bem caracterizada, possuindo também um metabolismo pouco complexo¹⁰⁹ e, em paralelo com a linhagem *S. avermitilis*, foram modificadas geneticamente para obtenção de novas linhagens, a partir da deleção de genes que codificam para o metabolismo secundário, eliminando assim a competição por fontes de carbono e nitrogênio entre as vias biossintéticas inerentes do organismo e dos genes clonados e facilitando também a identificação do produto heterólogo.^{141,143,144}

Por fim, todas essas metodologias combinadas permitem tanto identificar novos compostos, como melhorar os processos de fermentação ou gerar novos produtos “não-naturais”. Metabólitos secundários como ácido artemisínico,¹⁴⁵ precursores de taxol,¹⁴⁶ ivermectina¹⁴⁷ e lovastatina⁸⁵ são exemplos de compostos produzidos atualmente com sucesso aplicando-se algumas dessas estratégias de engenharia metabólica.

Nesse contexto, o presente trabalho visa explorar o genoma sequenciado do organismo endofítico *Streptomyces* sp. CBMAI 2042, isolado de *Citrus sinensis*, com foco em PKSs e NRPSs, na busca por um alvo molecular desconhecido. Com esse objetivo, diferentes metodologias de clonagem disponíveis, associadas às ferramentas de bioinformática, foram testadas e avaliadas, de modo a estabelecer metodologias viáveis no grupo de pesquisas para bioprospecção de produtos naturais e viabilizar a anotação de *cluster* de genes desconhecidos em bancos de dados.

2. OBJETIVOS

A partir da aquisição de dados do sequenciamento do genoma da actinobactéria endofítica *Streptomyces* sp. CBMAI 2042, o estudo aqui proposto teve como principal objetivo a aplicação de *genome mining* e metodologias de clonagem para expressar em organismos heterólogos os genes responsáveis pela biossíntese de valinomicina e um *cluster* de genes órfão constituído por uma PKS do tipo I, *trans*-AT PKS e NRPS.

2.1. Objetivos específicos

- Clonar os *clusters* biossintéticos codificando para t1PKS – *trans*-AT PKS – NRPS e uma NRPS (valinomicina) a partir de biblioteca do genoma de *Streptomyces* sp. CBMAI 2042;
- Clonar o *cluster* t1PKS – *trans*-AT PKS – NRPS a partir de metodologias de recombinação de fragmentos de DNA *in vivo* (λ -RED e TAR) e *in vitro* (Gibson Assembly®);
- Deletar domínio de adenilação anotado no *cluster* t1PKS – *trans*-AT PKS – NRPS e avaliar alterações no perfil metabólico da linhagem *Streptomyces* sp. CBMAI 2042 para identificação do metabólito associado;
- Transferir os *clusters* de genes clonados para linhagens heterólogas (*S. coelicolor* M1146 e M1152) para identificar os metabólitos associados;
- Isolar e caracterizar alpiniamida biossintetizada por ação do *cluster* de genes órfão de t1PKS – *trans*-AT PKS – NRPS;
- Elucidar a biossíntese da alpiniamida.

3. MATERIAIS E MÉTODOS

3.1. Materiais

3.1.1. Meios de cultura

Meio LB (Luria-Bertani)		Meio 2TY	
Extrato de levedura	5 g	Triptona	16 g
NaCl	10 g	Extrato de Levedura	10 g
Triptona	10 g	NaCl	5 g
Ágar	20 g	Água Milli-Q	q.s.p. 1 L
Água Milli-Q	q.s.p. 1 L		
Meio A		Meio MP	
Peptona	4 g	Glicose	10 g
Extrato de levedura	2 g	Amido solúvel	10 g
Extrato de carne	4 g	Peptona	2 g
Extrato de soja	2 g	Extrato de levedura	3 g
Maltose	20 g	Soytone	3 g
Dextrina	10 g	Extrato de malte	2 g
Água Milli-Q	q.s.p. 1 L	TES buffer pH 7,2	25 mM
		Água Milli-Q	q.s.p. 1 L
Meio TSBY		Meio ISP2	
TSB	30 g	Extrato de malte	10 g
Sacarose	103 g	Extrato de levedura	4 g
Extrato de levedura	5 g	Dextrose	4 g
Água Milli-Q	q.s.p. 1 L	Água Milli-Q	q.s.p. 1 L
Meio SFM		GPMY	
Extrato de soja	20 g	Amido de batata	20 g
Manitol	20 g	Extrato de levedura	5 g
Agar	20 g	Extrato de malte	5 g
MgCl ₂ 1 M	Conc. Final 10 mM	Glicerol	20 mL
Água Milli-Q	q.s.p. 1 L	Água Milli-Q	q.s.p. 1 L

3.1.2. Microrganismos e plasmídeos

Tabela 2. Plasmídeos e linhagens utilizados neste estudo

Plasmídeo/ linhagem	Características	Referência
Plasmídeos		
pESAC13A	<i>aac(3)IV, tsr, oriT, attP-int^{φC31}</i>	148
pIndigoBAC-5HindIII	<i>pIndigoBAC-5 digerido no sítio de restrição de HindIII Cloning Read</i>	Epicentre®
pYH7	<i>tsr, bla, aac(3)IV, cos, oriT, T3, T7, rep(pIJ101), ori(ColEI)</i>	149,150
pSET152cos	<i>attP(φC31), int(φC31), lacZ, oriT(RP4), ori(pUC18), aac(3)IV, cos</i>	Prof. Peter Leadlay
pRSalpΔA _{domain}	Plasmídeo de deleção do domínio de adenilação da NRPS	Este estudo
pIJ790	Sistema λ-RED <i>gam, bet, exo, cat, araC, rep101ts</i>	116
pCAP03_p21	Derivado de pCAP03- <i>acc(3)IV</i> (Addgene plasmid #69862) ¹³⁹	Prof. Bradley S. Moore
pUZ8002	Derivado de RK2	151
pUB307	Derivado OriT+ RP1	152
pRSalp_6B e 2C	pCAP03_p21 contendo os genes <i>alp</i>	Este estudo
pRSalp_6B_ermE	Derivado do plasmídeo pRSalp_6B. <i>ermEp upstream</i> ao gene 119	Este estudo
Linhagens		
<i>Streptomyces</i> sp. CBMAI 2042	Linhagem selvagem endofítica isolada de <i>Citrus sinensis</i> . Doadora dos <i>clusters</i> de genes biossintéticos.	153
<i>S. cerevisiae</i> VL6-48N	Derivada de VL6-48. <i>MATalpha, his3-Δ200, trp1-Δ1, ura3-Δ1, lys2, ade2-101, met14</i>	154
<i>E. coli</i> ET12567	Linhagem <i>E. coli</i> metilase deficiente <i>dam, dcm, hsdM, hsdS, hsdR, cat, tet</i>	155
TOPO10/ pR9406	Conjugação triparental	109
<i>E. coli</i> BW25113	<i>lacI^q, rrnB_{T14}, ΔlacZ_{WJ16}, hsdR514, ΔaraBAD_{AH33}, ΔrhaBAD_{LD78}</i>	156
M1146_RSalp	Linhagem heteróloga produtora de alpiniamida	Este estudo
M1146/ESAC4H#5 DSM 32578	Linhagem heteróloga produtora de valinomicina e análogos	Este estudo
<i>S. coelicolor</i> M1152	<i>act Δred Δcpk Δcda rpoB[C1298T]</i>	143
<i>S. coelicolor</i> M1146	<i>Δact Δred Δcpk Δcda</i>	143

3.1.3. Enzimas e Primers

As enzimas de restrição utilizadas neste estudo foram obtidas de Thermo Scientific ou New England Biolabs. *Gibson Assembly*® *Master Mix* foi obtido de New England Biolabs. Os fragmentos para recombinação de DNA foram amplificados com enzima *Phusion Flash High-Fidelity PCR Master Mix* (Thermo Scientific). Os fragmentos para *TAR-cloning* foram amplificados com enzima PrimeSTAR HS DNA Polymerase (Takara). Todos os protocolos foram seguidos de acordo com orientações dos fornecedores, qualquer alteração está descrita nas sessões experimentais.

Os *primers* utilizados neste trabalho foram desenhados com auxílio do programa *Geneious Pro* (versão 5.6.7 e 10.2.3) e produzidos pela Exxtend Biotecnologia Ltda ou IDT® *Integrated DNA Technologies*. Os *primers* foram estocados com concentração de 100 µM em água Milli-Q, sendo a concentração de uso 10 µM. Todas as sequências dos oligonucleotídeos encontram-se no Anexo 9.1.

3.2. Métodos de microbiologia e biologia molecular

3.2.1. Biblioteca genômica

O protocolo para construção da biblioteca genômica de *Streptomyces* sp. CBMAI 2042 em vetor do tipo BAC foi adaptado de Peterson et al.¹⁵⁷

3.2.1.1. Extração de DNA de alto peso molecular e Cinética de Digestão

Para extração de DNA de alto peso molecular uma colônia isolada de *Streptomyces* sp. CBMAI 2042 foi inoculada em 50 mL de meio TSBY, a 30 °C, por 48 h – 72 h, 250 rpm. O *pellet* foi ressuspendido gentilmente em 50 mL de tampão TE 10/1 (10 mM Tris-HCl, pH 8; 1 mM EDTA, pH 8) e lavado repetidamente até obtenção de uma suspensão límpida. Foi adicionado ao *pellet* um volume igual de 1,5% InCert® Agarose LMP (Lonza), seguido de agitação moderada para boa homogeneização e obtenção de *plugs*. A suspensão foi adicionada rapidamente aos moldes (Bio-Rad) estéreis e mantidos em gelo por 20 minutos. Os *plugs* foram desenhados em tubo contendo 50 mL de tampão TE e, posteriormente, incubados em *Solução I* (10 mM Tris-HCl; 50 mM EDTA; 1 mg.mL⁻¹ Lisozima) a 37 °C por 2 h. Após o tempo de incubação, foram transferidos para a *Solução II* (10 mM Tris-HCl; 50 mM EDTA; 300 µg.mL⁻¹ Proteinase K) e incubados a 55 °C, *overnight*, para lise celular.¹⁰⁷

Os *plugs* foram lavados com 50 mL de tampão TE por período de 1 h, sob leve agitação e em gelo, trocando a solução a cada lavagem. Em seguida, foram incubados 3 *plugs* testes em um tubo de 15 mL contendo tampão H3M (10% tampão enzima de restrição; 10% espermidina 40 mM; 1% BSA 100X; 0,1% DTT 1 mM), por 1 hora, sob leve agitação e em gelo. Os *plugs* foram divididos de modo a obter 12 x ¼, os quais foram cortados em 8 pedaços e transferidos para 11 tubos de 1,5 mL nomeados de “a até k”, contendo 250 µL de tampão H3M, sendo que para o tubo “a” foi transferido ½ plug. Foram adicionadas enzima HindIII (New England BioLabs), como segue: a (0 unidades.mL⁻¹); b (0,5 unidades.mL⁻¹); c (1 unidades.mL⁻¹); d (1,5 unidades.mL⁻¹); e (2 unidades.mL⁻¹); f (3 unidades.mL⁻¹); g (4 unidades.mL⁻¹); h (6 unidades.mL⁻¹); i (10 unidades.mL⁻¹); j (20 unidades.mL⁻¹); k (200 unidades.mL⁻¹). Os tubos de reação foram mantidos em gelo, sob leve agitação, por 1 h e, então, colocados em banho-maria a 37 °C por 10 minutos. Após este período, foram mantidos em gelo e foram adicionados 30 µL de EDTA 0,5 M pH 8.

Para visualizar o padrão de digestão enzimática obtido, foi preparado um gel de agarose 0,8% em TAE 1X (0,04 M tris-base, 1 M ácido acético glacial, 50 mM EDTA pH 8). O gel foi submetido a uma corrida em eletroforese de campo pulsado (Bio-Rad CHEF-DR II) utilizando os parâmetros: 6 V.cm⁻¹; P_{Inicial} 1 s; P_{final} 40 s; tempo de corrida 18 h e temperatura 12 °C. O gel foi corado com brometo de etídeo e fotodocumentado sob luz UV em câmara Molecular Imager® Gel Doc™ XR (Bio-Rad).

Foi determinada, de forma aproximada, a concentração ótima de enzima para obtenção de fragmentos entre 100 e kb e 200 kb. Os *plugs* foram incubados em 50 mL de tampão H3M, por 1 h em gelo, sob leve agitação. Após este período, os *plugs* foram divididos em pequenos pedaços (~ 32/40 pedaços por *plug*) e transferidos para tubos de 1,5 mL contendo 1 mL de tampão H3M. A concentração de enzima determinada foi adicionada e os tubos foram incubados em gelo por 1 hora, seguido de adição de 150 µL de EDTA 0,5 M pH 8 em cada tubo.

Para seleção de tamanho dos fragmentos foram preparados dois géis de agarose. Na Seleção I, foi preparado um gel de agarose Seaken Gold (Lonza) 1% TAE 0.75X. A separação do DNA ocorreu em eletroforese de campo pulsado com os seguintes parâmetros: Corrida 1 - 6 V.cm⁻¹; P_{Inicial} 1 s; P_{final} 40 s; tempo de corrida 13 h; Corrida 2 - 5 V.cm⁻¹; P_{Inicial} 4 s; P_{final} 5 s; tempo de corrida 8 h; temperatura 10 °C. A porção do gel correspondente aos fragmentos entre 100 e 350 kb foi excisada e dividida em 3 partes, seguindo para a Seleção II.

Foi preparado um gel de agarose Seaken Gold (Lonza) 1% TAE 1X. Após solidificar, uma porção mediana do gel foi cortada e agarose Sealplaque LMP (Lonza) 1% TAE 1X foi adicionada. Três poços maiores foram preparados de modo a acomodar as 3 partes obtidas da Seleção I. Os parâmetros utilizados para corrida foram: voltagem 5 V.cm^{-1} ; P_{Inicial} 3s; P_{final} 4,5 s; tempo de corrida 13 h e temperatura 12°C . Foram excisados os fragmentos A, B e C ($>100 \text{ kb}$) e armazenados em tampão TAE 1X gelado.

A recuperação do DNA dos blocos de agarose foi feita em eletroeluidor da marca BIORAD e tampão TAE 1X, com os seguintes parâmetros: tempo de corrida 2 h e amperagem 10 mA por cubeta (30 mA), sendo cada cubeta contendo um bloco de agarose (A, B ou C). A voltagem foi controlada entre 80 – 100 V. Ao final, os eletrodos foram invertidos e a fonte permaneceu ligada por 90 segundos. No coletor, foram recuperados os seguintes volumes: A = $60 \mu\text{L}$, B = $70 \mu\text{L}$ e C = $70 \mu\text{L}$. O DNA foi quantificado por eletroforese em gel de agarose 0,8% TAE 1X, utilizando como padrão diferentes concentrações de solução Lambda DNA (*New England's BioLabs*): 5, 10, 15, 20, 30, 40 e $50 \text{ ng.}\mu\text{L}^{-1}$. Os parâmetros de corrida utilizados foram: voltagem 110 V, por 45 – 60 min. O gel foi corado com brometo de etídeo e fotodocumentado sob luz UV.

3.2.1.2. Ligação e Transformação

A biblioteca foi construída utilizando o plasmídeo *pIndigoBAC-5* - HindIII *Cloning Read* (Epicentre®). Para a reação de ligação foram misturados volumes de vetor e inserto na razão de 1:3 (em ng). As reações de ligação foram preparadas em gelo e homogeneizadas por inversão. Para cada fração, foram adicionados os seguintes volumes: Fração A: $30 \mu\text{L}$ inserto e $4 \mu\text{L}$ vetor; Fração B e C: $35 \mu\text{L}$ inserto e $4,7 \mu\text{L}$ vetor. Enzima e Buffer: $7 \mu\text{L}$ Buffer T4 ligase 10X, $2 \mu\text{L}$ T4 DNA ligase (Thermo Scientific) e água até $70 \mu\text{L}$. Reações incubadas em termociclador a 16°C por 10 h, seguido de inativação da enzima a 65°C por 20 min, e mantidas a 4°C . As ligações foram transferidas para tubos de dessanilização (10% glicose 1 M, 1% agarose), os quais foram mantidos em repouso por 1 h e 30 min, em gelo.

Para transformação foi utilizado eletroporador MicroPulser (Bio-Rad) 3 kV, 1 Pulso, cubeta de 0,2 cm. Foram adicionados $13 \mu\text{L}$ de ligação em $100 \mu\text{L}$ de *E. coli* DH10B eletrocompetente, a mistura foi mantida em gelo por 10 min e $22,5 \mu\text{L}$ de

células foram adicionados em 5 cubetas, reunidas posteriormente em 2 mL de meio SOC (10 g.L⁻¹ triptona; 5 g.L⁻¹ extrato de levedura; 10 g.L⁻¹ cloreto de sódio; 20 mL glicose 1 M e 10 mL de solução Mg²⁺ (204 g.L⁻¹ MgCl₂.6H₂O; 246 g.L⁻¹ MgSO₄.7H₂O)) e incubadas a 37 °C, por 1 hora, 250 rpm. Foram plaqueados 100 µL de cada amostra em meio LB contendo 20 mg.mL⁻¹ XGAL, 2% IPTG e cloranfenicol (Chl – 12,5 µg.mL⁻¹). As placas foram incubadas a 37 °C por 18 – 20 h.

Transformantes obtidos foram plaqueados em Q-Tray contendo meio LB + Chl + IPTG + XGAL. As colônias recombinantes foram selecionadas e repicadas em duplicata para placa de 384 poços contendo 80 µL de meio líquido LB glicerol 6% + Chl. A repicagem foi feita com auxílio de robô modelo Qpix2 (Genetix). As placas foram incubadas a 37 °C, sob agitação de 40 rpm, por 18 h, e então, conservadas a -80 °C.

3.2.1.3. Minipreparação de BAC

Para verificar o tamanho de fragmentos inseridos, foi feita minipreparação com kit para miniprep de BACs da Macherey-Nagel 740618.24, seguida de digestão enzimática com enzima HindIII – 20 unidades/µL, em placa de 96 poços. Foram adicionados 15 µL de minipreparação e 5 µL do mix de digestão (1 reação = 2 µL Tampão NE 2.1; 0,2 µL BSA 100X; 0,06 µL HindIII; 2,75 µL H₂O Milli-Q). Condições de reação: 37 °C por 3 horas.

O DNA foi aplicado em gel de agarose 0,8% TBE 0,5X e eluído por eletroforese de campo pulsado (6 V.cm⁻¹, tempo de corrida 16 horas, P_{inicial} 5 s, P_{final} 15 s). O gel foi corado com brometo de etídeo e fotodocumentado sob luz UV.

3.2.1.4. Pool de placas e Amplificação com Phi29

Pools de placas foram preparados carimbando cada placa de 384 poços, previamente descongelada, e repicando em placa de *pool* contendo 20 mL de LB glicerol 6% e antibiótico. As placas foram seladas e incubadas a 37 °C, sob agitação de 40 rpm, por 20 – 22 h. Posteriormente, 2 mL de cada pool foram transferidos para tubos criogênicos, identificados e preservados a -80 °C.

Para reação de amplificação com Phi 29 foi utilizado kit *illustra GenomePhi V2 DNA Amplification* (GE Healthcare). Em placa de 96 poços foram adicionados 9 µL de *Sample Buffer* por poço, incluindo controle negativo e positivo. Foi adicionado 1 µL de cultura previamente descongelada e no poço de controle positivo, 1 µL do controle

fornecido pelo kit. A placa foi selada e as amostras desnaturadas a 95 °C por 5 min em termociclador C1000 Touch™ (Bio-Rad) e imediatamente transferidas para gelo, centrifugadas por 10 s a 1500 rpm e colocadas em gelo novamente. Foram adicionados em cada poço 10 µL do mix de reação (9 µL de *Reaction Buffer* + 1 µL de enzima Phi29). A placa foi selada, centrifugada por 1 min a 1500 rpm e incubada a 30 °C por 2 h. A enzima foi inativada a 65 °C por 10 min. Para reações de PCR, foram feitas diluições de 1/400 (em água Milli-Q) para cada amostra. Os produtos de amplificação foram verificados por eletroforese em gel de agarose 1% TAE 1X, corado com brometo de etídeo e fotodocumentado sob luz UV.

3.2.1.5. PCR em tempo real (qPCR)

Foram preparadas reações de PCR em tempo real (qPCR) utilizando os produtos obtidos pela amplificação com Phi29, em equipamento CFX384 Touch™ *Real-Time PCR Detection System* (Bio-Rad), facilitando o processo de triagem. Todas as reações foram preparadas utilizando *Maxima SYBR Green/ROX qPCR Master Mix* (Thermo Scientific) e foram utilizados os pares de *primers*: Primer_Inicio_F_1 e Primer_Inicio_R_1; Contig51_Final_F3 e Contig51_Final_R3. Os produtos de reação foram analisados por eletroforese em gel de agarose 2% TAE 1X, 45 min a 120 V, corado com brometo de etídeo e fotodocumentado sob luz UV.

Tabela 3. Reagentes e parâmetros utilizados para PCR em tempo real

Reagentes		Condições de reação		
		Etapa	Temperatura	Tempo
<i>SYBR Green</i>	5,0 µL			
Primer_Inicio_F_1	0,3 µL	Desnaturação inicial	95 °C	10 min
Primer_Inicio_R_1	0,3 µL	Desnaturação*	95 °C	10 s
DNA (BAC)	4,0 µL	Anel./Extensão*	60 °C	30 s
Água	0,4 µL	Curva dissociação	65 – 95 °C (rampa: 0,5 °C)	-
Volume final	10 µL	* 55 ciclos		

3.2.1.6. Sequenciamento das pontas de BACs

Reações de sequenciamento foram preparadas a partir do DNA obtido do experimento de minipreparação, utilizando *Life BigDye terminator v.3.1* e purificadas com EDTA/etanol. Para sequenciamento foi utilizado equipamento *Life 3500xL Genetic Analyzer*.

3.2.2. Recombinação de fragmentos de DNA

3.2.2.1. Construção do plasmídeo

O plasmídeo pSET152cos foi extraído com kit Purificação de DNA Plasmidial AP-MN-P-250 (AXYGEN) e digerido com BamHI (Thermo Scientific), seguindo protocolo do fabricante. Posteriormente, foi adicionado 1 µL de CIP (New England BioLabs) e a reação incubada a 37 °C, 1 h. O plasmídeo foi concentrado com *Pellet Paint® Co-Precipitant* (Novagen®), segundo protocolo recomendado pelo fabricante. Para incorporação das regiões homólogas foram preparadas 3 reações de PCR: 1) A e B – Obtenção de fragmentos de homologia a partir do DNA genômico de CBMAI 2042; 2) Inserção do fragmento final e 3) Inserção do fragmento inicial (Tabela 4 e Tabela 5).

Tabela 4. Protocolos para preparo de reações de PCR para construção do plasmídeo

Reação 1A (50 µL)		Reação 1B (50 µL)	
<i>Phusion HF Master Mix</i>	25 µL	<i>Phusion HF Master Mix</i>	25 µL
DNA Genômico	2,5 µL	DNA Genômico	2,5 µL
DMSO	3,0 µL	DMSO	3,0 µL
H ₂ O	12 µL	H ₂ O	12 µL
pSET_contig51_Inicio_F	2,5 µL	Contig51_Final_F	2,5 µL
Contig51_Inicio_R	2,5 µL	Contig51_pSET_Final_R	2,5 µL
Reação 2 (50 µL)		Reação 3 (50 µL)	
<i>Phusion HF Flash Master Mix</i>	25 µL	<i>Phusion HF Flash Master Mix</i>	25 µL
DMSO	1,5 µL	DMSO	1,5 µL
H ₂ O	16 µL	H ₂ O	15 µL
Contig51_Final_F	2,5 µL	Contig51_Final_F	2,5 µL
pSET_contig51_Inicio_R	2,5 µL	Contig51_Inicio_R	2,5 µL
DNA 1 (Reação 1B)	1,2 µL	DNA 1 (Reação 1A)	1,2 µL
DNA 2 (pSET152cos/ BamHI)	1,2 µL	DNA 2 (Reação 2)	1,2 µL

Tabela 5. Condições das reações de PCR para construção do plasmídeo (30 ciclos)

Etapa	Reação 1 – A e B		Reação 2 e 3	
Desnaturação inicial	98 °C	30 s	98 °C	10 s
Desnaturação	98 °C	10 s	98 °C	5 s
Anelamento/ Extensão	72 °C	180 s	72 °C	180 s
Extensão final	72 °C	300 s	72 °C	60 s

Foram preparadas 4 reações de 50 µL para cada etapa e as mesmas foram concentradas com *Pellet Paint* e purificadas em gel de agarose 0,8% TAE1X. Para extração do DNA da agarose, foi utilizado kit *Invisorb® Fragment CleanUp* (Invitex), segundo protocolo recomendado pelo fabricante. As construções foram confirmadas por padrão de digestão com enzima de restrição HindIII.

3.2.2.2. Obtenção dos fragmentos de DNA do *cluster* biossintético

O *cluster* foi dividido em 9 fragmentos (6,5 kb a 9 kb) contendo sobreposição entre si, obtidos por reações de PCR de 20 µL: Fragmentos 3, 4, 5, 6, 8, 9 e 10 – 10 µL *Phusion Flash High Fidelity*, 2 µL DNA, 4 µL de betaína, 2 µL de cada *primer*; Fragmentos 3B e 7 – 10 µL *Phusion Flash High Fidelity*, 2 µL DNA, 2 µL DMSO, 2 µL de cada *primer* e 2 µL de H₂O (Tabela 6 e

Tabela 7). Após confirmação dos produtos de amplificação, foram preparadas 5 reações x 50 µL para cada fragmento, concentradas com *Pellet Paint*, purificados em gel de agarose 0.8% TAE1X e extração com kit *Invisorb® Fragment CleanUp* (Invitex).

Tabela 6. Pares de *primers* utilizados para amplificar fragmentos de DNA

Fragmento	Primer Forward	Primer Reverse
3	Contig51_3_F	Contig51_3_R
3B	Contig51_pSET_Inicial_F	Contig51_3B_R
4	Contig51_4_F	Contig51_4_R
5	Contig51_5_F	Contig51_5_R
6	Contig51_6_F	Contig51_6_R
7	Contig51_7_F	Contig51_7_R
8	Contig51_8_F	Contig51_8_R
9	Contig51_9_F	Contig51_9_R
10B	Contig51_10_F	Contig51_pSET_Final_R

Tabela 7. Parâmetros utilizados para reações de PCR para obtenção dos fragmentos

Fragmento	3B		3 a 9		10B	
Desnaturação inicial	98 °C	10 s	98 °C	10 s	99 °C	10 s
Desnaturação**	98 °C	5 s	98 °C	5 s	99 °C	5 s
Anelamento**	72 °C	150 s	72 °C	180 s	69,5 °C	15 s
Extensão**	-	-	-	-	72 °C	210 s

Extensão Final	72 °C		60 s
** N° Ciclos	25	30	25

3.2.2.3. Recombinação via λ -RED

A recombinação homóloga *in vivo* foi feita utilizando uma linhagem de *E. coli*/pIJ790, induzida por adição de L-arabinose. Foram preparadas células eletrocompetentes *E. coli* pIJ790 de acordo com protocolo estabelecido.¹⁵⁸ Em 50 μ L de células eletrocompetentes, foram adicionados 10 μ L de mistura composta por 1 μ L de cada fragmento. As células foram transferidas para cubeta de eletroporação e submetidas a um pulso de 1800V. Foi adicionado 1 mL de meio LB e incubado a 37 °C, por 1 h. 100 μ L foram plaqueados em meio LB sólido contendo apramicina (50 μ g.mL⁻¹) e as placas incubadas a 37 °C, 18 h. A triagem por recombinantes foi feita por PCR de colônia.

3.2.2.4. Recombinação via *Gibson Assembly*®

Os experimentos de recombinação *in vitro* foram realizados utilizando *Gibson Assembly*® *Master Mix* (New England BioLabs), conforme protocolo estabelecido pelo fabricante. Foram utilizados fragmentos obtidos para o *cluster* de interesse em proporção equimolar. Foram adicionados 1 μ L de cada fragmento (~15 ng) e 10 μ L de *Master Mix* (fornecido pelo kit) em microtubo de 0,6 mL. A reação foi incubada a 50 °C por 1 hora. 2 μ L da reação foram transformados em 50 μ L de células de DH10B eletrocompetentes. 100 μ L da transformação incubada por 1 h a 37 °C em 1 mL de LB líquido foram plaqueados em meio LB sólido contendo apramicina (50 μ g.mL⁻¹) e as placas incubadas a 37 °C, 18 h. A triagem por recombinantes foi feita por PCR de colônia.

3.2.3. *TAR-cloning*

Os procedimentos para *TAR-cloning* foram feitos seguindo protocolo estabelecido por Zhang et al.,¹⁵⁹ com algumas adaptações e modificações.

3.2.3.1. Amplificação dos fragmentos por PCR

DNA genômico de *Streptomyces* sp. CBMAI 2042 foi extraído de 2 mL de pré-inóculo em TSBY e usado nas reações de PCR para obtenção dos 7 fragmentos correspondentes ao *cluster* t1PKS-*trans*AT PKS-NRPS. Os primers utilizados estão

listados na Tabela 8. A condição de reação A foi utilizada para obter os fragmentos 2, 3, 4, 5 e 6 com enzima *PrimeSTAR* e a condição de reação B foi utilizada para obter os fragmentos 1 e 7 com enzima *Phusion* (Tabela 9 e Tabela 10).

Tabela 8. Pares de *primers* obtidos para amplificação dos fragmentos correspondentes ao *cluster* t1PKS – *trans*-AT PKS – NRPS

Fragmento	Par de <i>Primers</i>	Fragmento	Par de <i>Primers</i>
1	TAR_BGC3_01_F TAR_BGC3_01_R	5	TAR_BGC3_05_F TAR_BGC3_05_R
2	TAR_BGC3_02_F TAR_BGC3_02_R	6	TAR_BGC3_06_F TAR_BGC3_06_R
3	TAR_BGC3_03_F TAR_BGC3_03_R	7	TAR_BGC3_07_F TAR_BGC3_07_R
4	TAR_BGC3_04_F TAR_BGC3_04_R		

Tabela 9. Condições de preparo das reações de PCR

Reagente	A	B
Enzima	0,2 µL	0,5 µL
Buffer	25,0 µL	10,0 µL
DNA	2,5 µL	2,5 µL
Primer (F/R)	2,5 µL	2,5 µL
dNTP	4,5 µL	4,0 µL
Água Milli-Q	15,3 µL	30,5 µL
Volume total	50 µL	50 µL

Tabela 10. Parâmetros para reação de PCR em termociclador (25 ciclos)

Etapas	Fragmentos 2 – 6		Fragmentos 1 e 7	
Desnaturação inicial	98 °C	2 min	98 °C	30 s
Desnaturação	98 °C	10 s	98 °C	10 s
Anelamento	67,5 °C	5 s	72 °C (- 0.3 °C/ ciclo)	5 s
Extensão	72 °C	4,5 min	72 °C	2,5 min
Extensão Final	72 °C	5 min	72 °C	5 min

3.2.3.2. Obtenção de esferoplastos de levedura

Uma colônia isolada de *S. cerevisiae* VL6-48N foi inoculada em 10 mL de meio YPD líquido, contendo 1X adenina (solução estoque 100X – 0,01 g.mL⁻¹), *overnight*, 30 °C e 220 rpm. 1 mL do pré-inóculo foi transferido para 50 mL de meio YPD contendo 1X adenina em frasco de 250 mL e incubado por 5 horas a 30 °C, 220

rpm, até A_{600} 0,7 – 1,0. O cultivo foi centrifugado por 3 min a 1800 x g e 4 °C, sobrenadante descartado e células lavadas com 50 mL de água estéril (gelada), seguido novamente de centrifugação e ressuspensão em 50 mL de sorbitol 1 M (gelado). A suspensão foi mantida em geladeira, *overnight*, centrifugada a 1800 x g e 4 °C, sobrenadante descartado (invertendo o tubo em um papel toalha) e as células ressuspendidas em 20 mL de solução SPE (HEPES 10 mM pH 7,5; EDTA 5 mM pH 8,0; sorbitol 1 M). Posteriormente, foram adicionados 40 µL de 2-mercaptoetanol (2-ME) e 80 µL de solução zimoliase-20T (zimoliase-20T 10 mg.mL⁻¹; glicerol 25% (w/v); Tris-HCl 50 mM pH 7,5), seguido de incubação a 30 °C até A_{600} atingir a diferença de 5 – 20 vezes (~ 40 min). O nível de esferoplastos foi medido checando a densidade óptica da suspensão celular (600 nm). Foram preparadas duas soluções: 1) 200 µL de células tratadas com zimoliase foram diluídas em 800 µL de sorbitol 1 M e 2) 200 µL de células tratadas com zimoliase foram diluídas em 800 µL de SDS 2% (em água). A_{600} de ambas as suspensões foram medidas e comparadas.

3.2.3.3. Transformação de esferoplastos

Após obtido o nível de esferoplastos adequado, foi adicionado volume suficiente de sorbitol 1 M para atingir 50 mL e a suspensão homogeneizada gentilmente por inversão, seguida de centrifugação por 10 min a 600xg, 4 °C. A lavagem com sorbitol 1 M foi repetida mais uma vez. Os esferoplastos obtidos após lavagem foram ressuspendidos em 2-3 mL de solução STC (Tris buffer 10 mM pH 7,5; CaCl₂ 10 mM; Sorbitol 1 M) e incubados por 10 min a temperatura ambiente (~22 °C). 200 µl foram adicionados lentamente em um tubo de 1,5 mL contendo ~20 µL (≤ 40 µL) de solução de DNA a ser transformada (100 ng vetor linearizado e 100 ng cada fragmento) e incubados por 10 min a temperatura ambiente. 800 µL de PEG 20% foram adicionados e a suspensão homogeneizada por inversão do tubo (10 vezes). O tubo foi incubado por 20 min em temperatura ambiente seguido de centrifugação por 10 min a 700xg, 4 °C. O sobrenadante foi cuidadosamente removido com pipeta de 1 mL e as células foram ressuspendidas em 800 µL de solução SOS (CaCl₂ 6,5 mM; extrato de levedura 0,25%; sorbitol 1 M, peptona 0,5%) e incubadas por 30-40 min a 30 °C.

A suspensão final obtida foi adicionada em 8 mL de *Top selective agar* (Sorbitol 1 M; dextrose 2,2%; ágar 3%), preparado em tubo de 15 mL e mantido a 60 °C. imediatamente a mistura foi invertida em placa preparada previamente, contendo

Bottom selective agar (Sorbitol 1 M; dextrose 2,2%; agar 2%) contendo 1X 5-FOA [ácido trifluoroacético – 100 mg.mL⁻¹ em DMSO (100X solução estoque)]. As placas foram incubadas por 4 a 7 dias a 30 °C. Colônias isoladas foram transferidas para uma nova placa de meio ágar seletivo (*Bottom selective agar*) contendo 5-FOA e incubadas por 2 dias a 30 °C. As colônias que cresceram após segunda seleção, foram incubadas em meio YPD líquido por 24 h a 30 °C, 200 rpm para extração do DNA com kit comercial Zymoprep™ Yeast Plasmid Miniprep II (ZYMO RESEARCH). Amostras de DNA obtidas foram triadas por PCR com *primers* BGC03_117_176bp_R e BGC03_117_176bp_F. As amostras de DNA que apresentaram um resultado positivo na triagem por PCR foram transformadas por eletroporação em TransforMax™ EPI300™ *E. coli* (Epicentre) e, após nova extração, os plasmídeos obtidos foram transformados em *E. coli* DH10B. Para extração dos plasmídeos de *E. coli* foi utilizado o kit comercial Qiagen Plasmid Extraction.

3.2.4. Construção de plasmídeo de deleção

Para a construção do plasmídeo de deleção, foram desenhados *primers* para amplificar fragmentos de ~2000 pb flanqueando a região de deleção. Para obtenção dos fragmentos foram preparadas reações de PCR de 10 µL: 5 µL *Phusion Flash High Fidelity*, 1 µL DNA, 1 µL de cada *primer*, 2 µL de betaína ou 1 µL H₂O e 1 µL DMSO. Pares de *primers* – fragmento I: B1_cluster14_FI e B1_cluster14_RI (amplicon = 1884 pb); fragmento II: B1_cluster14_FII e B1_cluster14_RII (amplicon = 2039 pb). Os fragmentos foram concentrados com *Pellet Paint* e purificados por eletroforese em gel de agarose 0,5% TAE1X. Para extração do DNA da agarose, foi utilizado kit *Invisorb® Fragment CleanUp* (Invitex), segundo protocolo recomendado pelo fabricante. Para ligação dos fragmentos com o plasmídeo pYH7 foi utilizado *Gibson Assembly® Master Mix*, seguindo protocolo estabelecido pelo fabricante para cálculo de concentração dos fragmentos e a reação foi incubada a 50 °C, 1 h. 2 µL de reação foram transformados em DH10B eletrocompetente.

Tabela 11. Parâmetros utilizados para reações de PCR para obtenção de fragmentos para construção de plasmídeo de deleção

Fragmento	I e II	
** N° Ciclos	25	
Desnaturação inicial	98 °C	10 s
Desnaturação**	98 °C	5 s
Anelamento**	70 °C	5 s
Extensão**	72 °C	40 s
Extensão Final	72 °C	60 s

3.2.5. Conjugação triparental

Os clones de ESACs obtidos em *E. coli* DH10B foram transferidos para *E. coli* ET12567 por experimento de conjugação triparental. Colônias isoladas de *E. coli* ET12567 (*chl^R*), TOPO10/pR9406 (*amp^R*) e DH10B/ESAC (*apr^R*) foram inoculados em 5 mL de meio LB suplementado com os respectivos antibióticos a 37 °C, 200 rpm *overnight*. 500 µL foram transferidos em 10 mL de meio LB contendo metade da concentração inicial de antibiótico e então incubados a 37 °C, 200 rpm até A_{600} 0,4 – 0,6. As células foram separadas por centrifugação a 2200 x g por 5 minutos e lavadas duas vezes com 20 mL de meio LB. O sobrenadante foi descartado e o *pellet* ressuspenso em 500 µL de meio LB. 20 µL de cada linhagem foram misturados e 60 µL foram inoculados na mesma região de uma LB-ágar na ausência de antibióticos e incubada a 37 °C, *overnight*. Para seleção de *E. coli* ET12567 contendo o ESACs e o plasmídeo pR9406, colônias obtidas foram estriadas uma segunda placa LB-ágar contendo cloranfenicol (25 µg.mL⁻¹), ampicilina (100 µg.mL⁻¹) e apramicina (50 µg.mL⁻¹) e placa incubada a 37 °C, *overnight*. Colônias isoladas obtidas foram triadas por PCR de colônia.

3.2.6. Conjugação *E. coli* – *Streptomyces*

Os diferentes plasmídeos em ET12567 obtidos para deleção, por TAR-cloning ou biblioteca genômica, foram transferidos para *Streptomyces* por evento de conjugação. Uma colônia isolada de *Streptomyces* foi inoculada em 25 mL de meio TSBY, em erlenmeyer de 250 mL com mola para aeração, e incubado a 30 °C, 200 rpm, por 48 h. O cultivo foi centrifugado a 4000 rpm por 10 min, o sobrenadante descartado e o *pellet* lavado 2 vezes com 20 mL de meio 2TY ou LB. Posteriormente,

200 μ L do *pellet* foram aliquotados (A), para preparo de sucessivas diluições. Foram adicionados 500 μ L de 2TY a alíquota A, obtendo-se a diluição B, da qual 200 μ L foram diluídos em 500 μ L de 2TY, obtendo-se a diluição C.

Colônia isolada de ET12567 contendo os respectivos plasmídeos de interesse, foi pré-inoculada, por 20 h, a 37 °C, 250 rpm, em falcon de 15 mL contendo 5 mL de meio 2TY e os respectivos antibióticos de seleção. 500 μ L do pré-inóculo foram inoculados em 10 mL de meio 2TY em falcon de 50 mL, contendo metade da concentração dos antibióticos, e incubados a 37 °C, 250 rpm, até atingir A_{600} 0,4 – 0,6 (aproximadamente 3,5 h). O cultivo foi centrifugado a 4000 rpm por 10 min, o sobrenadante descartado e o *pellet* foi lavado 2 vezes em 20 mL de meio 2TY. Posteriormente, 200 μ L de ET12567 contendo os plasmídeos foram misturados com 200 μ L da diluição C de *Streptomyces*. Então, 100 – 200 μ L foram plaqueados em meio SFM contendo 10 mM $MgCl_2$. Após incubação a 30 °C por 16 horas, 1 mL de solução contendo 100 μ g.mL⁻¹ de antibiótico de seleção dos respectivos plasmídeos e 500 μ g.mL⁻¹ de ácido nalidíxico foram adicionadas por toda superfície das placas. Após incubação a 30 °C, por 3 a 5 dias, colônias isoladas foram transferidas para meio SFM contendo antibióticos de seleção de cada plasmídeo e ácido nalidíxico, para confirmar a integração dos plasmídeos ao cromossomo de *Streptomyces*. Em experimentos de deleção, colônias obtidas foram novamente estriadas em placa SFM sem antibiótico para promover o 2° cruzamento e as colônias obtidas foram transferidas para placa de meio SFM com e sem apr/kan. Colônias que cresceram apenas na presença de antibióticos foram triadas por PCR para confirmar a integração da informação genética em *Streptomyces* sp. CBMAI 2042.

3.2.7. Experimentos com precursores marcados

Colônias isoladas de M1146_RS_{alp} foram pré-cultivadas a 30 °C, 48h e 250 rpm, em meio TSBY com metade da concentração de canamicina e ácido nalidíxico (25 μ g.mL⁻¹). 1% do pré-cultivo foi inoculado em 25-30 mL de meio MP sem antibiótico, em triplicata para cada precursor marcado, e após 10-12h o cultivo foi suplementado com solução estoque de precursores a 5 mg.mL⁻¹ em água Milli-Q, o procedimento se repetiu por 4 vezes, a cada 12 h, até uma concentração final de 1 mM (metionina) ou 12,9 mM (propanoato de sódio). Depois de 7 dias os cultivos foram extraídos com resina XAD7HP, como descrito no item 3.3.1, e analisados por LC-MS.

3.3. Métodos químicos e analíticos

3.3.1. Extração dos metabólitos

Os cultivos dos organismos heterólogos obtidos por ESAC foram centrifugados (4000 rpm, 10-15 min) e sobrenadantes foram extraídos com mesmo volume de acetato de etila. O solvente foi removido e o extrato ressuspendido em 1 mL de metanol para análise por LC-MS (Q-TOF). Cultivos de organismos obtidos via *TAR-cloning* foram extraídos com resina não-iônica amberlite XAD7HP (Sigma), adicionando $\sim 20 \text{ g.L}^{-1}$ de resina ao cultivo e agitação (30 – 60 min), seguido de centrifugação e extração da resina com metanol. Os extratos foram concentrados para análise por LC-MS.

3.3.2. Análise por LC-MS

- Análises de experimentos com clones obtidos por biblioteca genômica ESAC: HPLC (Hewlett Packard, Agilent Technologies 1290 series) acoplado ao espectrômetro de massas 6550 iFunnel Q-TOF LC/MS (Agilent Technologies); fonte de ionização por electrospray (ESI) no modo positivo; coluna de fase reversa: Zorbax Extend C-18 $1.8 \mu\text{m}$ (2,1 x 50 mm, Agilent Thechnologies). Gradiente de eluição: 0-1 min 90% A/10% B; 1-10 min, 90% A/10% B a 30% A/70% B; 10-16 min 30% A/70% B a 0.1% A/99,9% B, 16-17 min 0.1% A/99,9% B a 90% A/10% B. Fluxo: $0,6 \text{ mL.min}^{-1}$. Os dados foram analisados em programa Agilent MassHunter B.05.0.
- Análises de experimentos de *TAR-cloning* e deleção de domínio de adenilação:

Baixa resolução: LC Agilent 1260 Infinity acoplado a espectrômetro de massas Bruker Amazon Ion Trap. Alta resolução: HPLC Agilent Technologies 1200 Series acoplado a espectrômetro de massas 6530 Accurate-Mass Q-TOF (Agilent Technologies); fonte de ionização por electrospray (ESI) no modo positivo; coluna de fase reversa: Luna® $5 \mu\text{m}$ C18(2) 100\AA (150 x 4,60mm, Phenomenex); gradiente de eluição: 0-2 min 95% A/5% B; 2-27 min, 95% A/5% B a 0% A/100%

B; 27-32 min 0% A/100% B, 32-32,5 min 0% A/100% B a 95% A/5% B, 32,5-35 min 95% A/5% B. Fluxo: 0,7 mL.min⁻¹. MS^E *continuum* durante 35 min, intervalo de m/z 50–1000; scan time: 0,5 s. Os dados foram analisados em programa Agilent MassHunter B.05.0.

- Análises de resultados dos experimentos com metionina marcada: HPLC (Hewlett Packard, Agilent Technologies 1290 series) acoplado ao espectrômetro de massas 6550 iFunnel Q-TOF LC/MS (Agilent Technologies); fonte de ionização por electrospray (ESI) no modo positivo; coluna de fase reversa: InfinityLab Poroshell 120 C-18 2,7µm; 4.6x100mm (Agilent). Gradiente de eluição: 0-2 min 95% A/5% B; 2-27 min, 95% A/5% B a 0% A/100% B; 27-32 min 0% A/100% B, 32-32,5 min 0% A/100% B a 95% A/5% B, 32,5-35 min 95% A/5% B. Fluxo: 0,5 mL.min⁻¹. Os dados foram analisados em programa Agilent MassHunter B.05.0.
- Análises de resultados dos experimentos com propanoato marcado: Nano LC Acquity (Waters) acoplado a espectrômetro de massas q-ToF maxis 3G (Bruker Daltonics); fonte de ionização por electrospray (ESI) no modo positivo; coluna de fase reversa: Acquity UPLC BEH C18 (1,7 µm, 100 x 2,1 mm); gradiente de eluição: 0-6 min 95% A/5% B a 50% A/50% B; 6-9 min, 50% A/50% B a 0% A/100% B. Fluxo: 0,5 mL.min⁻¹. MS^E *continuum* durante 10 min, intervalo de m/z 50–800; scan time: 0,2 s. Os dados foram analisados em programa MassLynx 4.1.

3.3.3. Ressonância Magnética Nuclear (RMN)

Para análise estrutural dos produtos obtidos, utilizou-se a técnica de Ressonância Magnética Nuclear de ¹H, COSY, HSQC e HMBC à temperatura ambiente e em equipamento JEOL ECA 500 MHz. Os deslocamentos químicos (δ) são referenciados em relação ao resíduo do solvente não-deuterado. As constantes de acoplamento (*J*) são reportadas em Hz. As multiplicidades são apresentadas como: s, singleto, d, duplete, t, tripleto, q, quarteto, m, multiplete; dd, duplo duplete; dt, duplo tripleto. Os espectros de RMN foram tratados utilizando o Software MestreNova 6.0.

4. RESULTADOS E DISCUSSÃO

4.1. *Streptomyces* sp. CBMAI 2042: análise *in silico* das anotações do genoma

A linhagem selvagem objeto de estudo desta Tese de Doutorado foi depositada na Coleção Brasileira de Microrganismos de Ambiente e Indústria (CBMAI) sob o registro de *Streptomyces* sp. CBMAI 2042 com acesso aberto. A linhagem foi isolada de *Citrus sinensis* e doada pelo Professor Dr. Wellington Araújo. O genoma foi sequenciado pela tecnologia V2 Illumina (colaboração com Markyian Samborsky, Universidade de Cambridge, UK), sendo representado em 8.211.326 pares de base (pb), dividido em 3 *contigs* e com conteúdo G+C de 68,4% (dados do sequenciamento disponíveis em: GenBank RCOL000000000, BioProject PRJNA487706, BioSample SAMN09908177 e SRA PRJNA487706).¹⁶⁰ A análise *in silico* do sequenciamento via antiSMASH 4.2.0¹⁶¹ revelou a presença de 35 *clusters* biossintéticos codificando para a biossíntese de metabólitos secundários, conforme apresentado na Tabela 12.

Tabela 12. *Clusters* biossintéticos anotados do genoma de *Streptomyces* sp. CBMAI 2042 codificando para o metabolismo secundário.

<i>Cluster</i>	Tipo	<i>Cluster</i>	Tipo
Cluster 01	Butirolactona	Cluster 19	Lantipeptídeo
Cluster 02	Terpeno	Cluster 20	Terpeno
Cluster 03	t1PKS- <i>trans</i> ATPKS-NRPS	Cluster 21	Sideróforo
Cluster 04	NRPS-Melanina	Cluster 22	Bacteriocina
Cluster 05	t1PKS-NRPS	Cluster 23	t3PKS-NRPS
Cluster 06	t3-PKS-NRPS	Cluster 24	Butirolactona
Cluster 07	Outros	Cluster 25	Lantipeptídeo
Cluster 08	Terpeno	Cluster 26	Terpeno
Cluster 09	Ectoína	Cluster 27	t1PKS-NRPS
Cluster 10	Lantipeptídeo	Cluster 28	Bacteriocina
Cluster 11	Sideróforo	Cluster 29	NRPS
Cluster 12	Outros	Cluster 30	NRPS
Cluster 13	Lantipeptídeo	Cluster 31	Melanina
Cluster 14	Melanina	Cluster 32	t3PKS
Cluster 15	t1PKS	Cluster 33	Terpeno
Cluster 16	Lantipeptídeo	Cluster 34	Tiopeptídeo-t1PKS-NRPS
Cluster 17	Lasso-peptídeo	Cluster 35	t2PKS-Butirolactona
Cluster 18	Bacteriocina		

A análise dos *clusters* de genes biossintéticos (BGC) disponíveis despertou o interesse em estudar o *cluster* 03 (~64 kb), identificado como uma t1PKS – *trans*-AT PKS – NRPS, por três características principais: a) o BGC não estava associado a nenhum metabólito já conhecido/descrito em banco de dados, sendo então um *cluster* órfão; b) os genes codificam para um complexo multienzimático híbrido composto por uma sintase de policetídeo (PKS) do tipo I, uma sintetase de peptídeo não ribossomal (NRPS) e uma *trans*-AT PKS, o que pode aumentar a diversidade química; c) a presença de *trans*-AT PKS em actinobactérias⁹³ foram descritas apenas recentemente, portanto, estudos envolvendo essa classe de enzimas contribuem para compreensão de vias de biossíntese de outros metabólitos. O *cluster* 30, codificando para a biossíntese da valinomicina,^{162,163} um metabólito já descrito em diferentes linhagens de *Streptomyces*,¹⁶⁴ também foi selecionado como prova de conceito de clonagem via biblioteca genômica e para estudos envolvendo a possível obtenção de análogos estruturais. Os principais genes relacionados à biossíntese foram analisados com auxílio da ferramenta BLAST®, disponível no NCBI (*National Center for Biotechnology Information*) e as proteínas homólogas foram identificadas em bancos de dados, confirmando a correlação do *cluster* 30 com a biossíntese de valinomicina e evidenciando o potencial biossintético do *cluster* 03, codificando para um metabólito desconhecido (Tabela 13).

Tabela 13. Anotações das principais orfs dos *clusters* de genes 03 e 30

ORF	Microrganismo	Proteína homóloga	Nº acesso	e-value	I (%)
Cluster 03 – t1PKS–<i>trans</i>-AT PKS–NRPS					
115	Múltiplas espécies <i>Streptomyces</i>	[ACP] S-maloniltransferase	WP_026242027.1	0.0	99
116	<i>Streptomyces</i> sp. ScaeMP-e10	3-oxoacil-ACP sintase	WP_018515077.1	0.0	99
117	<i>Streptomyces</i> sp. ScaeMP-e10	PKS tipo I	WP_018515076.1	0.0	99
118	<i>Streptomyces</i> sp. ScaeMP-e10	Metiltransferase/ PKS	WP_018515075.1	0.0	99
119	<i>Streptomyces</i> sp. ScaeMP-e10	NRPS	WP_018515074.1	0.0	99
120	<i>Streptomyces</i> sp. CS090A	acil-CoA desidrogenase	WP_109981366.1	0.0	99
121	<i>Streptomyces</i> sp. ScaeMP-e10	Metiltransferase	WP_018515072.1	0.0	99
Cluster 30 – valinomicina					
24	<i>Streptomyces</i> sp.	vIm2 (NRPS)	AWD39038.1	0.0	100
23	<i>Streptomyces</i> sp.	vIm1 (NRPS)	AWD39037.1	0.0	100

Portanto, com objetivo de caracterizar os BGCs codificando para PKSs e NRPSs e de aprofundar estudos das vias de biossíntese, foram testadas diferentes metodologias de clonagem para obtenção dos *clusters* completos e posterior transferência para organismos heterólogos, para expressão e identificação dos metabólitos. As metodologias utilizadas, bem como os resultados obtidos e dificuldades limitantes de cada técnica, serão detalhadas nas próximas sessões.

4.2. Clonagem de vias de biossíntese

4.2.1. Biblioteca genômica em vetor do tipo BAC

O primeiro passo para obtenção do *cluster* de genes codificando para o complexo híbrido de enzimas t1PKS – *trans*-AT PKS – NRPS foi a tentativa de construção de uma biblioteca do genoma de *Streptomyces* sp. CBMAI 2042 em vetor do tipo BAC (pIndigoBAC-5 HindIII), de forma a obter em um único clone o *cluster* completo de genes compreendido em uma região de ~ 64 kb. Os experimentos foram realizados em colaboração com a Profa. Dra. Anete Pereira de Souza e Dr. Danilo Augusto Sforça (CBMEG – UNICAMP). Considerando o tamanho de insertos de DNA necessários para construção de uma biblioteca genômica representativa, a extração de DNA de alto peso molecular é uma das etapas mais críticas e importantes para o resultado final. Por esta razão, para extração do DNA de *Streptomyces* sp. CBMAI 2042, foram preparados *plugs* de agarose, os quais tem por objetivo preservar o DNA no processo de lise celular, e todo o procedimento foi realizado utilizando ponteiras com pontas cortadas, para minimizar quebra do DNA por ação mecânica e obter um DNA de melhor qualidade.

O experimento de cinética de digestão determinou a concentração ideal de enzima para digestão parcial do DNA genômico e obtenção de fragmentos entre 100 e 350 kb. O gel de agarose obtido por eletroforese de campo pulsado está representado na Figura 12, sendo a coluna “a” representada pelo DNA não digerido, o qual apresentou uma pequena degradação (rastros de DNA no gel). Adicionalmente, observam-se rastros de DNA oriundos da alta concentração de DNA digerido, especialmente nas colunas “g” a “j”. A presença de DNA na parte superior do gel demonstra que as concentrações de enzima HindIII não foram suficientes para

digestão completa do DNA contido nos *plugs*. A concentração em “k” excedeu o ideal, sendo observado DNA de baixo peso moléculas (abaixo de 50 kb). A concentração estimada ideal para digestão dos *plugs* foi de 100 unidades.mL⁻¹, compreendendo a região entre o mínimo e máximo de digestão. Foram também selecionadas concentrações margeando o ideal (50 unidades.mL⁻¹ e 150 unidades.mL⁻¹).

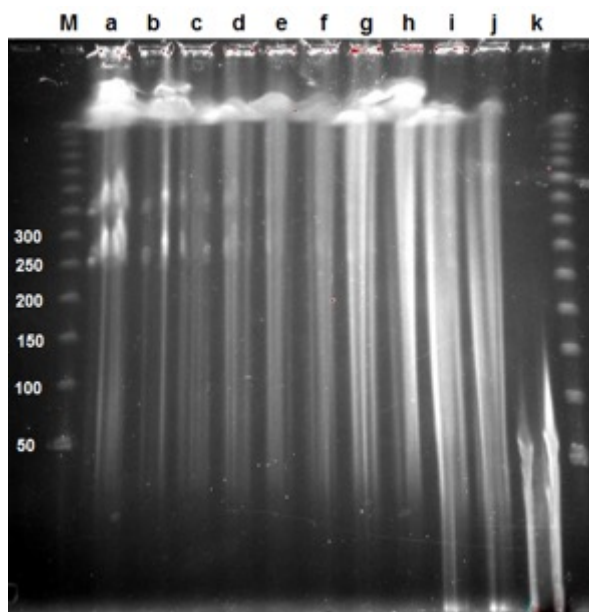


Figura 12. Determinação da concentração de HindIII para digestão parcial de DNA genômico. Eletroforese de campo pulsado (6 V.cm⁻¹; P_{Inicial} 1 s; P_{final} 40 s; 18 h; 12 °C) em gel de agarose 0,8% TAE1X. M: PFGE Lambda *Ladder* (New England BioLabs). Colunas “a” ao “k”: *plugs* digeridos. a: 0 unidades.mL⁻¹; b: 0,5 unidades.mL⁻¹; c: 1 unidades.mL⁻¹; d: 1,5 unidades.mL⁻¹; e: 2 unidades.mL⁻¹; f: 3 unidades.mL⁻¹; g: 4 unidades.mL⁻¹; h: 6 unidades.mL⁻¹; i: 10 unidades.mL⁻¹; j: 20 unidades.mL⁻¹; k: 200 unidades.mL⁻¹.

Determinado o padrão de digestão, foram selecionados 8 *plugs*, divididos ao meio, totalizando 16 *plugs*, para digestão parcial do DNA e seleção de tamanho dos fragmentos. Foram feitos dois ciclos de seleção, a primeira seleção para remover a maior parte de fragmentos <100 kb e a segunda seleção foi feita para garantir que qualquer pequeno fragmento preso ao DNA de alto peso molecular fosse removido.¹⁵⁷ O resultado da primeira seleção de tamanho, apresentado na Figura 13, foi satisfatório e a porção retirada do gel compreende a região contendo DNA entre 100 e 350 kb. O gel removido foi dividido em três partes: A (~100-150 kb), B (~150-250 kb) e C (~250-350 kb). Os três blocos de DNA obtidos na primeira seleção foram colocados em outro gel para segunda seleção de tamanho (Figura 14).

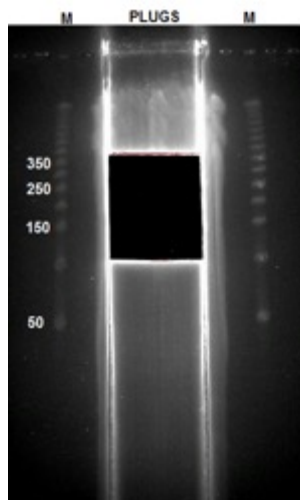


Figura 13. Seleção I de fragmentos de DNA. *Plugs* de DNA parcialmente digeridos com enzima HindIII. Eletroforese de campo pulsado em agarose Seaken Gold (Lonza) 1% TAE 0.75X. Parâmetros: 1) 6 V.cm⁻¹; P_{Inicial} 1 s; P_{final} 40 s; 13 h; 2) 5 V.cm⁻¹; P_{Inicial} 4 s; P_{final} 5 s; 8 h; 10 °C para 1 e 2. M: PFGE Lambda *Ladder*. “PLUGS”: contem todos os *plugs* digeridos. A porção removida do gel corresponde aos fragmentos entre 100 e 350 kb. Antes do corte da porção contendo os fragmentos de interesse, apenas as partes do gel contendo os marcadores foram coradas em brometo de etídeo.

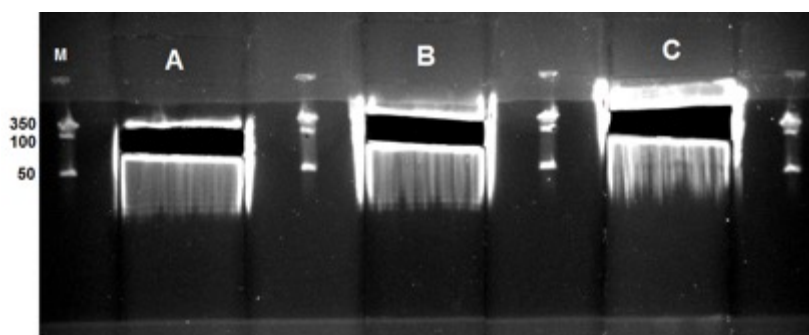


Figura 14. Seleção II de fragmentos de DNA. Eletroforese de campo pulsado em agarose Sealplaque LMP (Lonza) 1% TAE 1X. Parâmetros: 5 V.cm⁻¹; P_{Inicial} 3 s; P_{final} 4,5 s; 13 h; 12 °C. A, B e C: blocos obtidos da seleção I. M: PFGE Lambda *Ladder*. Antes dos cortes no gel contendo os fragmentos de interesse, apenas as partes dos marcadores foram coradas em brometo de etídeo.

Após segunda seleção de tamanho, o DNA foi extraído por eletroeluição, por esta razão a agarose utilizada no gel de seleção II é menos densa, de modo que o DNA possa migrar mais facilmente, sem degradar. A eletroeluição também contribui para obter um DNA mais preservado, quando comparado com kits de extração, e com maior rendimento. As concentrações obtidas para cada bloco de fragmento foram: A = 60 ng.μL⁻¹, B = 70 ng.μL⁻¹ e C = 70 ng.μL⁻¹.

Os experimentos de ligação dos fragmentos obtidos foram realizados conforme protocolos padrões com vetor pIndigoBAC-5 HindIII, um vetor BAC que comporta insertos até 300 kb e obtido comercialmente pronto para uso, ou seja, linearizado em sítio único de restrição de HindIII e purificado. Na primeira tentativa, a reação de ligação da fração C (fragmentos maiores) apresentou alta viscosidade, portanto foi descartada. Apesar da eficiência de transformação, os recombinantes obtidos com as frações A e B, após minipreparação e digestão, apresentaram média de tamanho dos insertos muito baixa, principalmente para fração A (< 45 kb). Sendo assim, foi repetida a reação de ligação e transformação em DH10B para fração C e de transformação para fração B. O resultado da transformação foi plaqueado em LB sólido contendo cloranfenicol e X-gal.

O vetor utilizado na clonagem possui o gene *lacZ* codificando para a enzima β -galactosidase localizado no sítio de ligação e quando fragmentos são inseridos a enzima fica inativa, já quando ocorre recircularização do plasmídeo, o gene é expresso. Na presença de X-gal (5-bromo-4-cloro-3-indoxil- β -D-galactopiranosídeo) no meio de cultura a enzima expressa hidrolisa o composto, formando galactose e 5-bromo-4-cloro-3-indoxil, o segundo é oxidado na presença de oxigênio e dimeriza formando o cromóforo de cor azul, 5,5'-dibromo-4,4'-dicloro-índigo.¹⁶⁵ O sistema “branco-azul” facilita na triagem por recombinantes, que devem apresentar coloração branca. Sendo assim, foram selecionadas aleatoriamente 20 colônias brancas de cada placa (B e C) para minipreparação e digestão com enzima HindIII.

Os insertos da fração B apresentaram-se, em sua maioria, abaixo da linha de 48 kb, então, apenas a fração C foi utilizada para os experimentos posteriores. A média de fragmentos obtida foi de ~69 kb (Figura 15). Toda reação de ligação do fragmento C foi transformada em células eletrocompetentes e plaqueadas em Q-Trays. As colônias foram repicadas e selecionadas pelo sistema “branco-azul” com auxílio de robô. Assim, a biblioteca do genoma de *Streptomyces* sp. CBMAI 2042 foi obtida em vetor pIndigoBAC-5 HindIII, sendo composta por 23 placas de 384 poços totalizando ~8500 clones. Foram feitas cópias de cada placa as quais foram armazenadas a -80 °C.

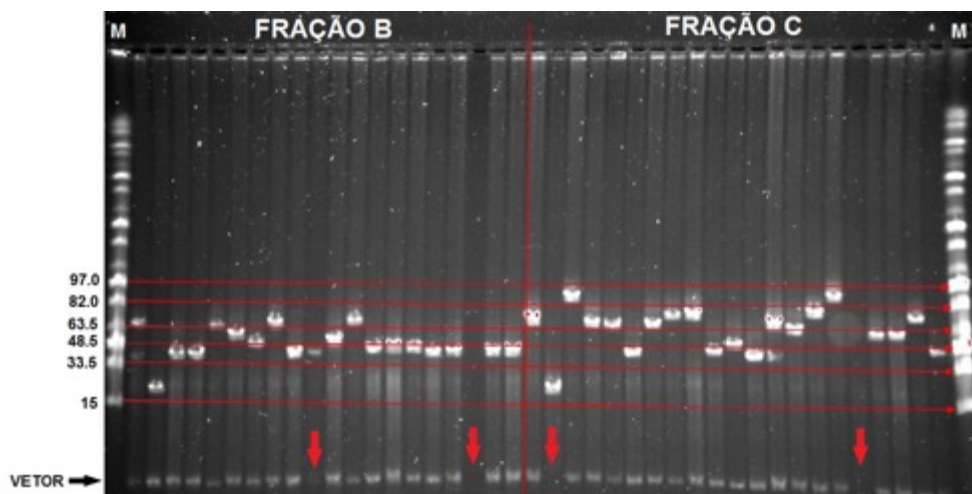


Figura 15. Média de insertos da biblioteca genômica em vetor BAC. Eletroforese de campo pulsado em gel de agarose 0.8% TBE1X. Parâmetros: 6 V.cm⁻¹; 16 h; Pinicial 5 s e Pfinal 15 s). M: Marcador MidiRange PFG Marker (New England's BioLabs). Demais colunas: BACs após digestão com enzima HindIII. A banda ~7,5 kb representa o vetor pIndigoBAC-5. A seta vermelha indica colônias “vazias”.

Devido ao grande número de placas obtidas, algumas delas foram triadas aleatoriamente por PCR em tempo real, utilizando os pares de *primers*: Primer_Início_F_1 e Primer_Início_R_1; Primer_Final_F_4 e Primer_Final_R_4. Porém, devido incoerências de resultados, foram preparados pools das placas, onde todos os clones de uma placa de 384 poços foram reunidos em uma única placa e, então, após incubação a 37 °C, 18 h, o cultivo foi utilizado para reação de amplificação com enzima Phi29 (Figura 16), com o objetivo de aumentar a concentração de DNA dos BACs para reações de qPCR e triagem dos clones. As amostras foram diluídas em água Milli-Q (1/400) para utilizar no preparo das reações de PCR com primers amplificando o início (Primer_Início_F_1 e Primer_Início_R_1) e final do gene (Contig51_Final_F3 e Contig51_Final_R3). Apenas a placa 21 apresentou amplificação para ambas as regiões de fronteira do *cluster* (Figura 17).

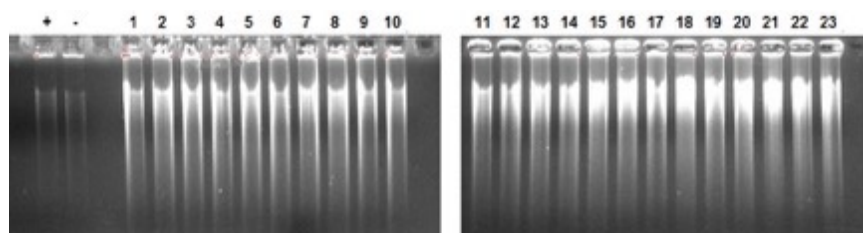


Figura 16. Produto de amplificação com enzima Phi29. Eletroforese em gel de agarose 1% TAE 1X. controle positivo (+) e negativo (-). Colunas 1 a 23: cada uma representa o *pool* de cada placa.

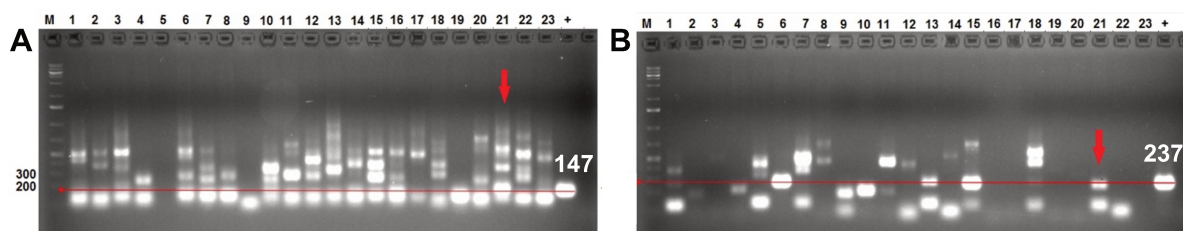


Figura 17. Triagem de clones em biblioteca BAC por qPCR. Eletroforese em gel de agarose 2% TAE 1X. **A.** Reação amplificando início do *cluster* (Primer_Início_F_1 e Primer_Início_R_1). **B.** Reação amplificando final do *cluster* (Contig51_Final_F3 e Contig51_Final_R3). A seta vermelha representa a placa (21) com produto de amplificação para ambas regiões. M: Marcador GeneRuler™ 1 kb Plus DNA Ladder (Thermo Scientific). (+): Controle positivo (gDNA).

A placa 21 (384 clones) foi então selecionada para reação de qPCR, 13 clones (P1, K2, O6, I7, L7, C11, F15, M15, G24, K17, H19, N20) apresentaram produto de amplificação em reação com par de *primers* do final do *cluster* (Contig51_Final_F3 e Contig51_Final_R3) (Figura 18), onde letra representa linha e número representa coluna na placa de 384). Contudo, a triagem para o início do *cluster* não apresentou produtos de amplificação na região desejada e nas condições testadas. As reações foram repetidas, mas não apresentaram reprodutibilidade até mesmo para os clones que apresentaram anteriormente produto de amplificação na região final do *cluster*.

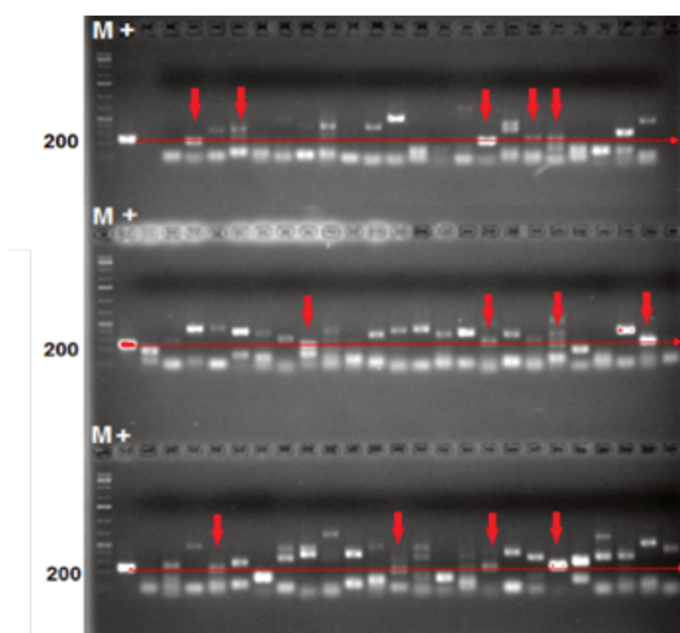


Figura 18. Produtos de amplificação da região final do *cluster* obtidos por qPCR. Eletroforese em gel de agarose 2% TAE1X. Colunas marcadas representam clones positivos. M: Marcador GeneRuler™ 1 kb Plus DNA Ladder (Thermo Scientific). Segunda coluna (+): controle positivo.

Optou-se, então pelo sequenciamento das pontas de BACs de 96 clones da placa 21, incluindo aqueles que apresentaram resultado positivo para a região final. Os resultados de sequenciamento e triagens por PCR não confirmaram a presença do *cluster* de genes em nenhum clone. Portanto, apesar de obtida com sucesso ao final de diversas tentativas, a biblioteca genômica não foi representativa o suficiente para identificação de clones contendo o *cluster* de genes em estudo. Em função desses resultados e entendendo a dificuldade advinda dos experimentos, optou-se por encomendar a construção da biblioteca genômica com a empresa Bio S&T Inc. (Montreal, Canadá).

4.2.2. Biblioteca genômica em vetor do tipo ESAC

A biblioteca genômica foi construída pela empresa Bio S&T Inc., especializada na construção de bibliotecas genômicas de *Streptomyces*, obtida em *E. coli* DH10B e os fragmentos clonados em vetor pESAC13A (Figura 19). O vetor pESAC13A¹⁴⁸ é uma versão modificada do pESAC13¹⁰⁹ (derivado do vetor pPAC-S¹¹⁰), contendo o gene de resistência a apramicina (*aac(3)/IV*), ao invés de kanamicina (*aph(3)/II*). Ambos foram construídos visando reduzir algumas limitações encontradas para clonagem e manipulação genética, viabilizando a transferência direta para linhagens de *Streptomyces*. Possuem origem de replicação (P1-derivado) e transferência (*oriT*) em *E. coli* e o sistema ϕ C31 *attP-int*, o qual permite a integração específica do ESAC no sítio *attB* do cromossomo de *Streptomyces*. Portanto, viabiliza a transferência direta do vetor obtido em *E. coli* e contendo o *cluster* de genes, para expressão heteróloga em *Streptomyces*.

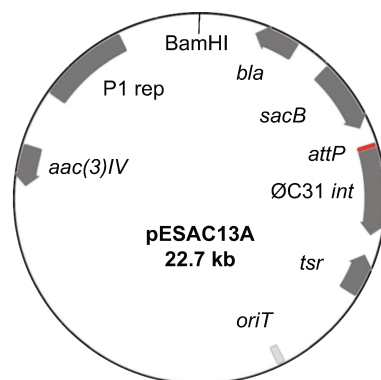


Figura 19. Mapa do vetor pESAC13A. P1 rep: origem de replicação; *oriT*: origem de transferência; ϕ C31 *attP-int*: sistema para integração sítio-específica em *Streptomyces*; *aac(3)IV*: resistência apramicina; BamHI: sítio de restrição e clonagem.

A biblioteca de ESAC do genoma de *Streptomyces* sp. CBMAI 2042 foi obtida com uma média de fragmentos de 92 kb (Figura 20A), sendo composta por 14 placas de 384 poços, totalizando 5376 clones, os quais foram triados com pares de *primers* amplificando regiões central, direita e esquerda delimitando os *clusters* de interesse. Foram triados clones contendo *cluster* de genes codificando para a biossíntese de valinomicina (*primers* início: Valinomycin_Left_F/R; *primers* centro: Valinomycin_Center_F/R; *primers* final: Valinomycin_Right_F/R) e para o *cluster* t1PKS – *trans*-AT PKS – NRPS (*primers* início: Ctg51_PAC_Left_R/F; *primers* centro: Ctg51_PAC_Center_R/F; *primers* final: Ctg51_PAC_Right_R/F), sendo obtidos dois clones para cada (Figura 20B).

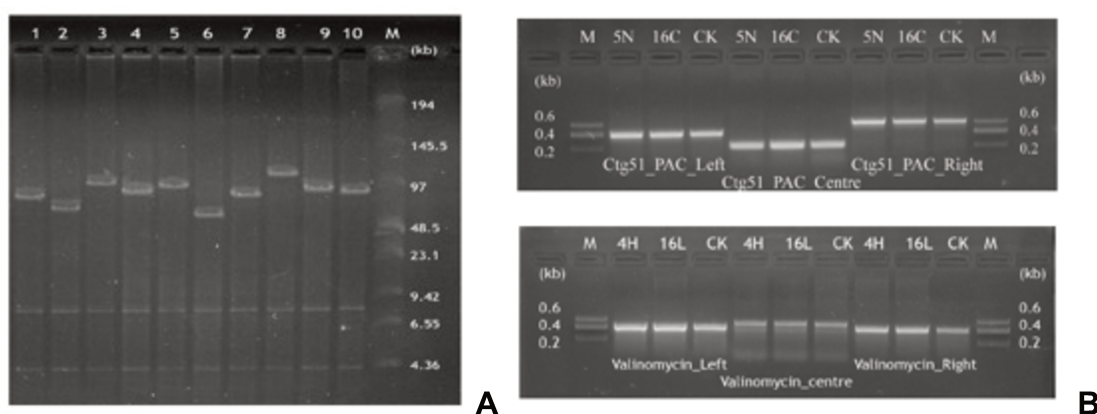


Figura 20. Biblioteca genômica em vetor ESAC de *Streptomyces* sp. CBMAI 2042. **A.** Média de insertos clonados. **B.** Triagem de clones. 5N: clone 5076 – placa9-5N; 16C: clone 5076 – placa11-16C; 4H: Clone 5076 – placa11-4H; 16L: clone 5076 – placa13-16L; CK: controle (gDNA). M: Marcador. Acima: clones t1PKS – *trans*-AT PKS – NRPS e abaixo clones da valinomicina. Os experimentos de triagem foram realizados por Bio S&T Inc.

Os ESACs dos clones positivos foram extraídos e digeridos com *DraI* para confirmar o tamanho de inserto. Os clones referentes ao *cluster* da valinomicina apresentaram insertos de 97 kb (clone 5076 – placa11-4H) e 105 kb (clone 5076 – placa13-16L). Os clones referentes ao *cluster* t1PKS-*trans*ATPKS-NRPS apresentaram insertos de 130 kb (clone 5076 – placa9-5N) e 150 kb (clone 5076 – placa11-16C).

Os clones enviados para o Brasil pela empresa foram cultivados em meio LB líquido contendo apramicina ($50 \mu\text{g.mL}^{-1}$), a 37°C , 250 rpm, *overnight*, para preservação em glicerol 10% e, para obtenção de colônias isoladas em meio LB sólido contendo apramicina para os experimentos de conjugação triparental e transferência para organismo heterólogo.

O vetor pESAC13A, como citado anteriormente viabiliza a transferência direta dos ESACs obtidos em *E. coli* para *Streptomyces* por conjugação. Contudo, como a construção é feita em *E. coli* DH10B os vetores devem ser primeiramente transferidos para uma linhagem metilase deficiente, *E. coli* ET12567,¹⁶⁶ pois o sistema de defesa de *Streptomyces* reconhece DNA exógeno metilado proveniente de *E. coli* e degradaria os vetores se transferidos direto de DH10B.¹⁶⁷

Os métodos tradicionais de transferência de DNA, como a eletroporação, apresentam baixa eficiência quando utilizando grandes fragmentos de DNA,¹⁶⁸ como as construções em ESAC. Para esses casos, utiliza-se o sistema bacteriano de conjugação e troca de informação genética, denominado conjugação triparental. Para este procedimento, três linhagens de *E. coli* foram inoculadas em um mesmo halo em meio LB sólido sem antibiótico para viabilizar a conjugação. A linhagem auxiliar (*E. coli* TOPO10), contendo o plasmídeo pR9406 (Amp^{R}) promove a transferência do ESAC da linhagem doadora (DH10B/ ESAC – Apr^{R}) para a linhagem aceptora, *E. coli* ET12567 (Chl^{R}). Posteriormente, os clones contidos no mesmo halo em meio sólido, são transferidos para meio LB contendo o antibiótico de resistência dos 2 plasmídeos e da linhagem aceptora (Apr^{R} , Chl^{R} e Amp^{R}). As colônias isoladas obtidas dos experimentos de conjugação triparental foram triadas com os *primers* da triagem da biblioteca genômica (Figura 21).

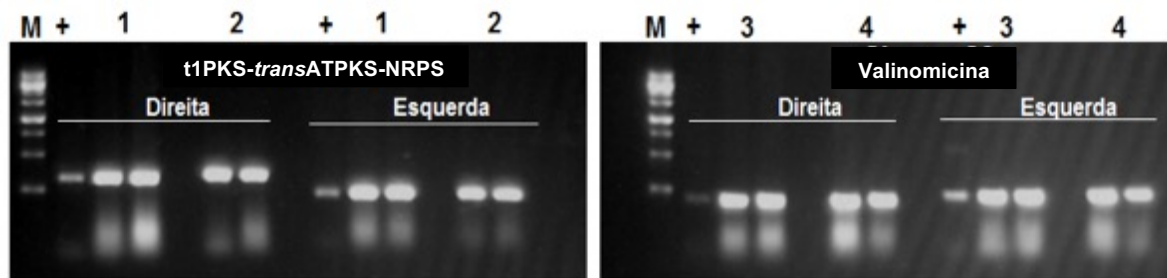


Figura 21. Triagem de clones obtidos por conjugação triparental. Produtos de amplificação por PCR de colônia. Eletroforese em gel de agarose 1% TAE 1X. Linhas 1, 2, 3 e 4 correspondem aos clones 5076 – placa9-5N, 5076 – placa11-16C, 5076 – placa11-4H e 5076 – placa13-16L, respectivamente.

Confirmada a presença dos ESACs em linhagem ET12567, os mesmos foram transferidos por conjugação para os organismos heterólogos *S. coelicolor* M1146 e M1152. Após *overlay* com antibióticos de seleção (apramicina e ácido nalidíxico) e incubação por 16 h, as placas correspondentes ao ESAC16C não apresentaram exconjugantes e foram desconsideradas. Para facilitar a nomenclatura os clones serão referidos como, por exemplo, M1146/ESAC4H, correspondente ao ESAC placa9-4H em *S. coelicolor* M1146.

Os exconjugantes obtidos foram selecionados aleatoriamente e transferidos em quadrantes para placas de SFM sólido contendo tioestreptona e ácido nalidíxico e incubados a 30 °C por 3 a 5 dias. Colônias isoladas de cada quadrante foram cultivadas em TSBY e o DNA genômico extraído foi analisado por reação de PCR, com os mesmos *primers* utilizados anteriormente, para confirmar a integração dos *clusters* biossintéticos no genoma do organismo heterólogo (Figura 22). Após a confirmação da transferência dos *clusters* biossintéticos para as linhagens heterólogas, os mesmos foram preservados em glicerol 20% para posterior cultivo e identificação/ isolamento de metabólitos.

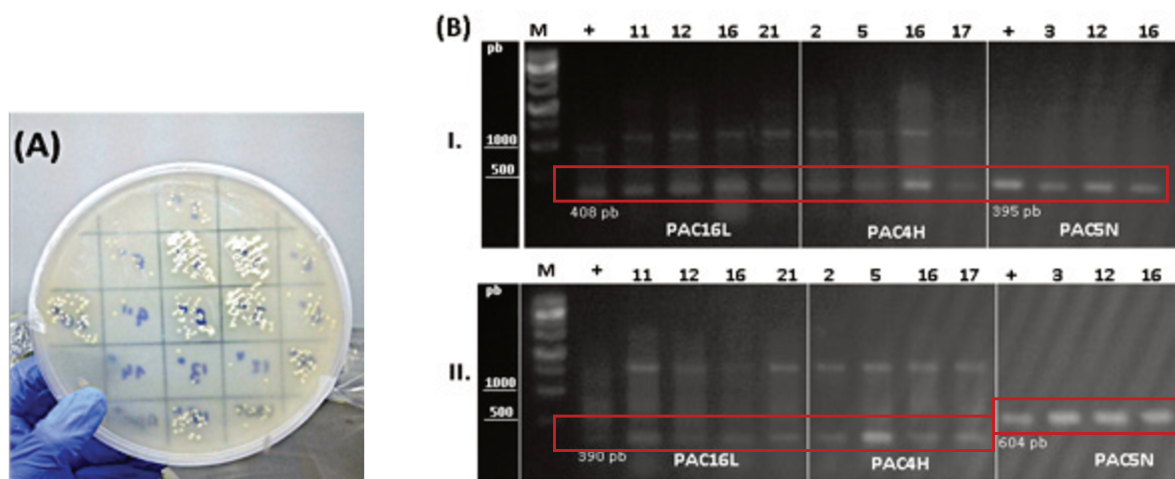


Figura 22. Experimentos de conjugação com *Streptomyces*. **A.** Placa de SFM-Nal/Thio com exconjugantes selecionados separados em quadrantes, para confirmação da integração do ESAC no cromossomo de *Streptomyces*. **B.** Confirmação do evento de conjugação por PCR com os *primers* correspondentes às extremidades do *cluster* biossintético. I = início do *cluster* e II = final do *cluster*; exconjugante #11 e #21 - M1146/ESAC16L; exconjugante #12 e #16 - M1152/ESAC16L; exconjugante #5 e #17 - M1146/ESAC4H; exconjugante #2 e #16 - M1152/ESAC4H; exconjugante #12 e #16 - M1146/ESAC5N; exconjugante #3 e #18 - M1152/ESAC5N. (+): controle positivo (gDNA); M: Marcador de peso molecular 1 kb (Sinapse).

4.2.3. Recombinação de fragmentos de DNA via λ Red e Gibson Assembly®

Em paralelo aos experimentos com biblioteca genômica e como alternativa de clonagem do BGC t1PKS – *trans*-AT PKS – NRPS para expressão em organismo heterólogo, foram testadas metodologias de recombinação de DNA. Para isso, o BGC foi dividido em 9 fragmentos com sobreposição de ~200 pb entre si (Figura 23), os quais foram amplificados por reação de PCR e purificados em gel de agarose 0,5% (Figura 24). Diversos plasmídeos foram desenvolvidos para facilitar a utilização de sistemas de recombinação de DNA *in vivo*, os quais podem ser transferidos para o organismo hospedeiro para expressar as proteínas necessárias para o evento de recombinação e para o estudo foi selecionado o plasmídeo pIJ790.

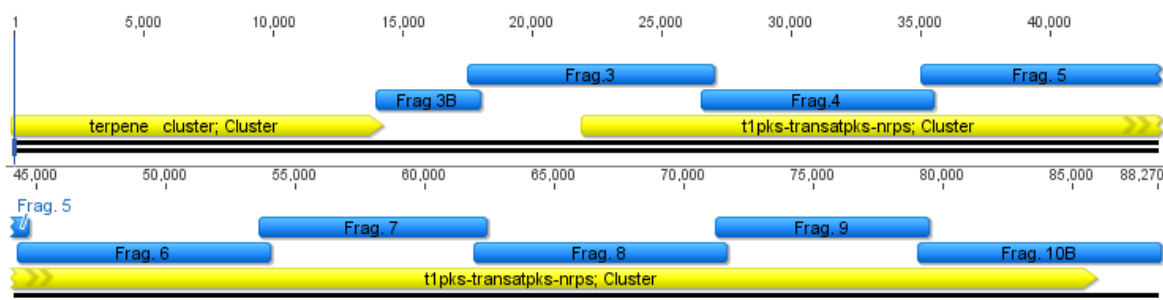


Figura 23. Representação *in silico* dos 9 fragmentos do *cluster* de genes t1PKS–trans-AT PKS–NRPS contendo sobreposição entre si. Imagem obtida pelo Programa Geneious Pro 5.6.7

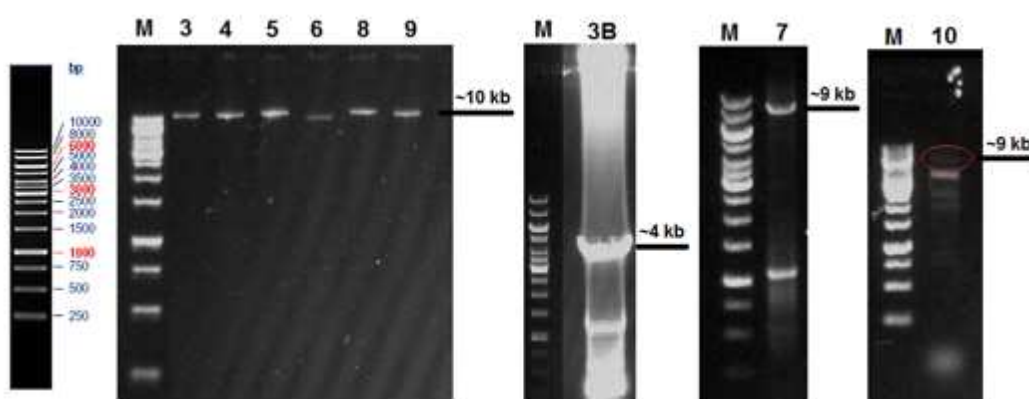


Figura 24. Produtos de amplificação dos fragmentos correspondentes ao *cluster* de genes t1PKS – trans-AT PKS – NRPS. Eletroforese em gel de agarose 0,5%. M: GeneRuler™ 1 kb DNA Ladder.

O plasmídeo pset152cos foi selecionado para comportar o *cluster* completo, ao qual foram incorporadas regiões de homologia flanqueando o BGC. Foram desenhados *primers* contendo 20 nt correspondentes ao pSET152cos e 25 nt correspondentes ao *cluster* para amplificar regiões de ~1000 pb de cada extremidade do BGC. Primeiro foram amplificados dois fragmentos: inicial (1A – *primers*: Contig51_Início_R e pSET_contig51_Inicio_F) e final (1B – *primers*: Contig51_Final_F e Contig51_pSET_Final_R), a partir do DNA genômico de CBMAI 2042 e os fragmentos foram purificados em gel de agarose. O fragmento 1B e o plasmídeo linear foram utilizados como moldes em uma PCR do tipo “*overlap*” para obtenção de R2 (~7 kb) (*primers*: Contig51_Final_F e pSET_contig51_Inicio_R). Posteriormente, R2 e 1A foram utilizados como moldes para um novo “*overlap*” e obtenção de R3 (~8 kb) (*primers*: Contig51_Final_F e Contig51_Inicio_R) (Figura 25).

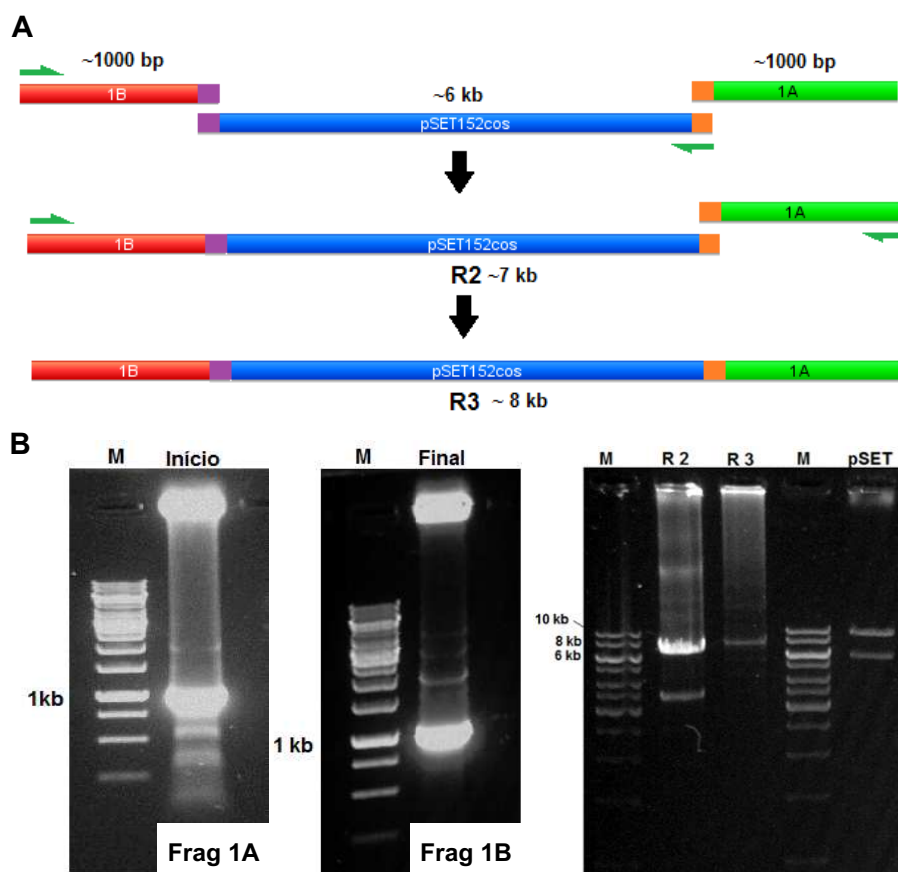


Figura 25. Construção do plasmídeo de recombinação. **A.** Esquema para reações de PCR e construção do plasmídeo com regiões de homologias. Regiões de homologia do BGC: roxo e laranja; fragmento 1B: vermelho; fragmento 1A: verde; fragmento azul: pSET152cos linear. **B.** Produtos de amplificação 1A, 1B, R2 e R3 (plasmídeo para recombinação – pRS_recBGC03). Eletroforese em gel de agarose 0,7% TBE1X. pSET: pSET152cos digerido com BamHI; M: GeneRuler™ 1 kb DNA Ladder (Thermo Scientific).

Obtido o plasmídeo pRS_recBGC03 ($517 \text{ ng} \cdot \mu\text{L}^{-1}$), uma mistura foi preparada contendo $1 \mu\text{L}$ de cada fragmento do *cluster* (3B a 10 - concentração $\sim 15 \text{ ng} \cdot \mu\text{L}^{-1}$) e do plasmídeo, para transformação em $50 \mu\text{L}$ de DH10B/ pIJ790 eletrocompetente. Após transformação e incubação a 37°C em meio LB (apr) foram selecionadas colônias para triagem por PCR utilizando os pares de *primers* Contig51_5R e 6F (Figura 26A) e, posteriormente, com os *primers* pSET_Contig51_inicio_F e Contig51_Inicio_R; Contig51_Final_F e Contig51_pSET_Final_R (Figura 26B), os quais não apresentaram produto de amplificação. Os experimentos foram repetidos diversas vezes, mas não foram confirmados clones contendo o *cluster* completo de genes.

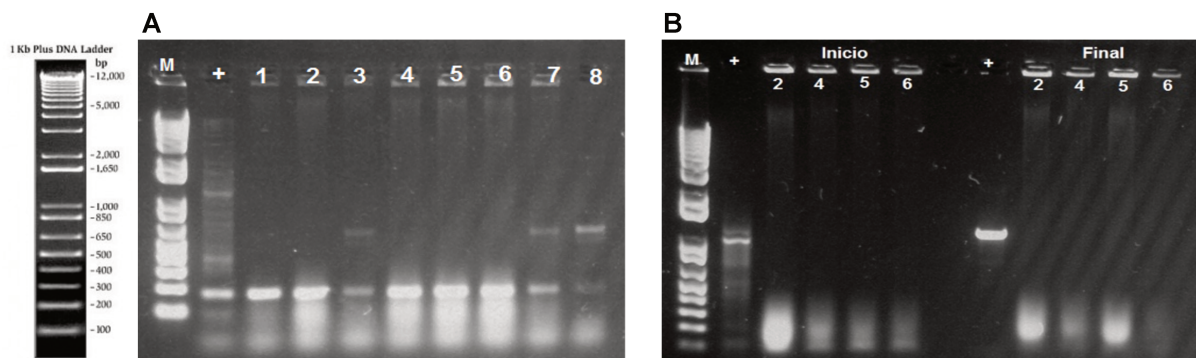


Figura 26. Triagem de clones obtidos após recombinação. Eletroforese em gel de agarose 1% TAE1X dos produtos de amplificação por PCR de colônia. **A.** PCR de colônia com *primers* Contig51_5R e 6F. **B.** PCR de colônia com *primers* pSET_Contig51_início_F e Contig51_início_R; Contig51_Final_F e Contig51_pSET_Final_R. M: 1 kb Plus DNA Ladder (Invitrogen™). Controle (+): gDNA.

Um dos possíveis problemas para as tentativas de recombinação *in vivo* pode estar relacionado com o sistema λ Red escolhido e os fragmentos utilizados. Segundo a literatura,¹¹³ sistema de recombinação λ Red é bastante eficiente para promover modificações de BACs ou outros plasmídeos, ou seja, um fragmento linear contendo a modificação desejada, o qual é transformado em *E. coli* contendo, por exemplo, o BAC (circular) e a recombinação ocorre. No caso de experimentos de recombinação entre múltiplos fragmentos lineares (plasmídeo linear e fragmentos lineares) o sistema RecET¹¹⁴ apresenta maior eficiência e reprodutibilidade nos resultados, principalmente considerando a quantidade e tamanho dos fragmentos utilizados (quanto maiores os fragmentos, menor a eficiência de recombinação), porém o mesmo não estava disponível para realização de testes.

O sucesso da aplicação da metodologia de *Gibson Assembly*® para obtenção de genes biossintéticos¹²⁷ e reorganização de fragmentos de até ~300 kb,¹²⁶ descritos na literatura, despertou nosso interesse por testar a metodologia, partindo dos mesmos fragmentos obtidos anteriormente para a recombinação *in vivo*. Sendo assim, foi preparada uma mistura contendo 1 μ L de cada fragmento (15 ng. μ L⁻¹) e do plasmídeo pRS_recBGC03 (517 ng. μ L⁻¹), adicionados a 10 μ L de *Gibson Assembly*® Master Mix. A reação foi incubada a 50 °C por 1 h e, em seguida, 2 μ L foram transformados em DH10B eletrocompetente. As colônias obtidas foram triadas por PCR de colônia e nenhuma apresentou resultado positivo de recombinação.

Foram feitas diferentes tentativas, com o objetivo de otimizar a reação, e utilizando novas células competentes, porém sem sucesso.

Uma possível explicação deve-se a quantidade de fragmentos utilizados, visto que o protocolo foi otimizado para até 6 fragmentos. Sendo assim, foram preparadas reações de recombinação em 4 etapas, sendo que em cada uma foram adicionados apenas 3 fragmentos (etapa 1: frag 3B, 3 e 4; etapa 2: 5, 6 e 7; etapa 3: 8, 9 e 10). Os produtos obtidos foram purificados e concentrados com QIAquick PCR Purification Kit e utilizados em uma quarta reação com o plasmídeo linearizado, sem a região de homologia, considerando que foram obtidos novos fragmentos correspondentes a extremidade inicial (3B) e final (10) com homologia de ~20 nt correspondentes ao plasmídeo. Porém, as poucas colônias obtidas não continham o *cluster* completo. As limitações da metodologia podem também estar relacionadas com a problemática de manipulação de grandes fragmentos de DNA, que podem facilmente degradar com a manipulação.

4.2.4. TAR-cloning

A identificação do metabólito associado ao *cluster* 03 não foi trivial pelos cultivos do clone obtido da biblioteca genômica. Sendo assim, foi elaborada uma proposta de colaboração com o grupo de pesquisas do professor Bradley S. Moore no Centro de Biotecnologia e Biomedicina Marinha (CMBB), do Instituto de Oceanografia Scripps (SIO), Universidade da Califórnia San Diego (UCSD), aprovada pela FAPESP (BEPE Proc Nº 2016/25735-1) durante o período de Abril/ 2017 a Março/ 2018. A proposta tinha como objetivo aplicar a metodologia de captura direta de BGC via TAR-cloning a partir do DNA genômico digerido com enzima de restrição.

O sistema de clonagem direta de BGC via TAR (*direct cloning by TAR* ou *TAR-cloning*) foi desenvolvido pelo grupo do professor Moore com base na construção do vetor pCAP01¹³⁵ para clonagem de vias biossintéticas completas, diretamente do DNA genômico e transferência para organismos heterólogos de *Streptomyces*.^{135–137} A metodologia foi otimizada com a construção de um segundo vetor de captura, o pCAP03-*acc(3)/IV*¹³⁹ (Figura 27A), derivado do pCAP01 e como já descrito, foi obtido pela inserção do marcador de seleção URA3, otimizando o sistema de triagem por recombinantes, diminuindo o resultado de falsos positivos em placas de meio seletivo. Outra modificação foi feita pela Dra. Hanna Luhavaya (não publicado) no vetor

pCAP03 para inserção do promotor p21¹⁶⁹ *upstream* ao sítio de clonagem (Figura 27B).

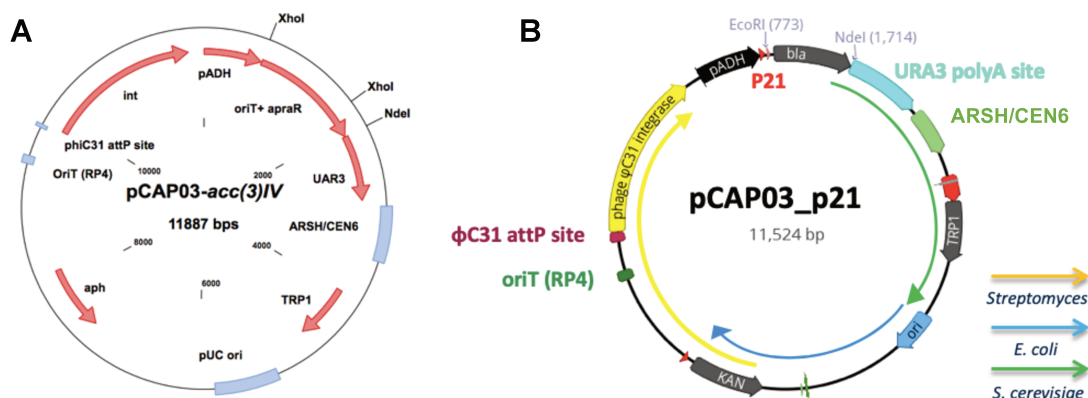


Figura 27. Vetores obtidos para experimentos de TAR-cloning. **A.** pCAP03-acc(3)/IV, derivado de pCAP01, contendo gene de resistência a apramicina entre os sítios de restrição XhoI e NdeI, sítio de clonagem.¹³⁹ **B.** pCAP03_p21, derivado de pCAP03 com promotor p21 *upstream* ao sítio de clonagem EcoRI e resistência a ampicilina entre o sítios EcoRI e NdeI, obtido por Dra. Hanna Luhavaya.

Originalmente, o TAR-cloning foi desenvolvido para captura BGC diretamente do DNA genômico, deste modo a enzima de restrição selecionada deve possuir sítios de restrição nas regiões de fronteira do gene ou *cluster* de genes e não fragmentar a região de interesse. Neste estudo, conduzido sob supervisão da pós-doutoranda Dra. Hanna Luhavaya, optou-se por, além de capturar o *cluster*, utilizar o promotor p21 para regulação de expressão gênica *upstream* ao gene 122, conforme anotação do antiSMASH referente ao *cluster* de genes t1PKS-*trans*AT-PKS-NRPS (~ 64 kb), que em uma análise mais detalhada possibilitou a minimização dos possíveis genes essenciais à biossíntese (~ 30,5 kb) – (Figura 28A). Sendo assim, não foi possível identificar uma enzima de restrição específica a região próxima ao gene e optou-se por obter o BGC em sete fragmentos contendo região de homologia de ~150 nt entre si e ~45 nt de homologia nas extremidades com o vetor pCAP03_p21 (Figura 28B).

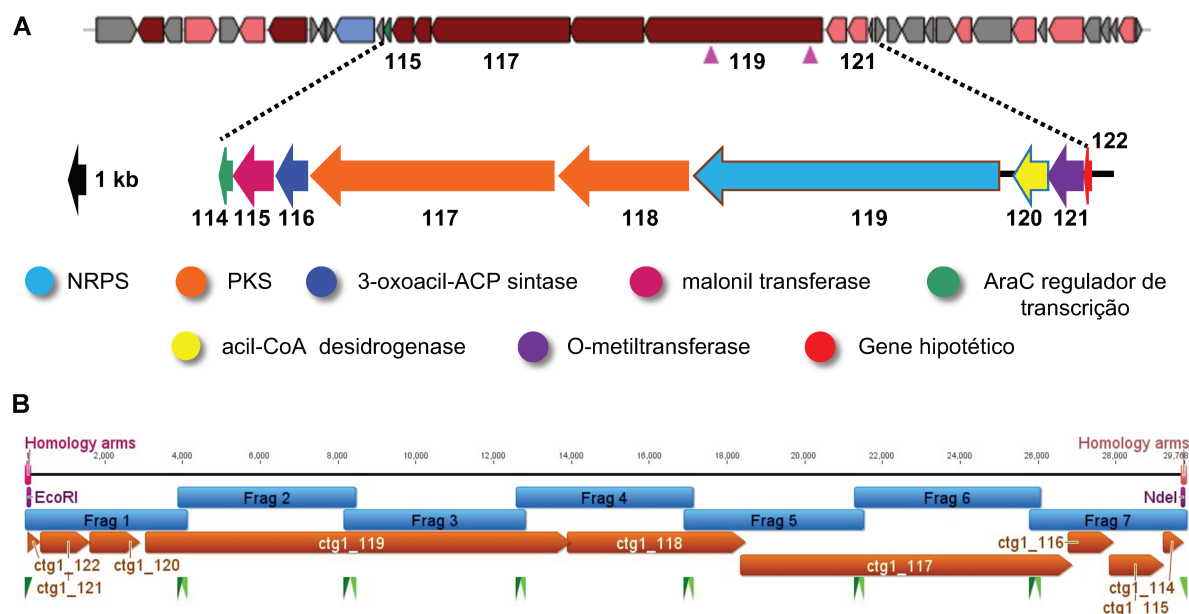


Figura 28. Representação dos genes clonados por TAR-cloning. **A.** Cluster de genes biossintéticos de *Streptomyces* sp. CBMAI 2042 codificando para o complexo enzimático t1PKS-trans-AT PKS-NRPS (acima), representado pelos genes biossintéticos essenciais (abaixo). **B.** Esquema representando a divisão do cluster biossintético em 7 fragmentos contendo ~150 pb de homologia entre si e ~45 nt de homologia com o vetor pCAP03_p21.

Os fragmentos foram amplificados em reação de PCR (~ 5 reações de 50 μ L) de modo a obter uma concentração mínima de 30 ng. μ L⁻¹ para cada fragmento, após purificação em gel de agarose e extração com Kit Zymoclean Gel DNA Recovery. Para confirmar a correspondência dos fragmentos com as regiões de DNA esperadas, foram feitas análises com enzimas de restrição. Os fragmentos 2 e 7 foram confirmados com enzimas *Stu*I (1512 pb; 2975 pb) e *Pst*I (1622 pb; 2168 pb; 150 pb), respectivamente. Os fragmentos 1, 4 e 5 foram concentrados e confirmados com enzima *Nco*I (2409 pb; 1498 pb; 154 pb) e *Sac*I (frag 4: 3710 pb; 750 pb e frag 5: 2017 pb; 1765 pb; 751 pb). Para os fragmentos 3 e 6 foram preparadas novas reações de PCR, visto que os primeiros produtos de PCR apresentaram padrão inesperado, e a confirmação foi feita com enzimas *Stu*I (2687 pb; 1865 pb) e *Pst*I (3087 pb; 1586 pb), respectivamente (Figura 29B).

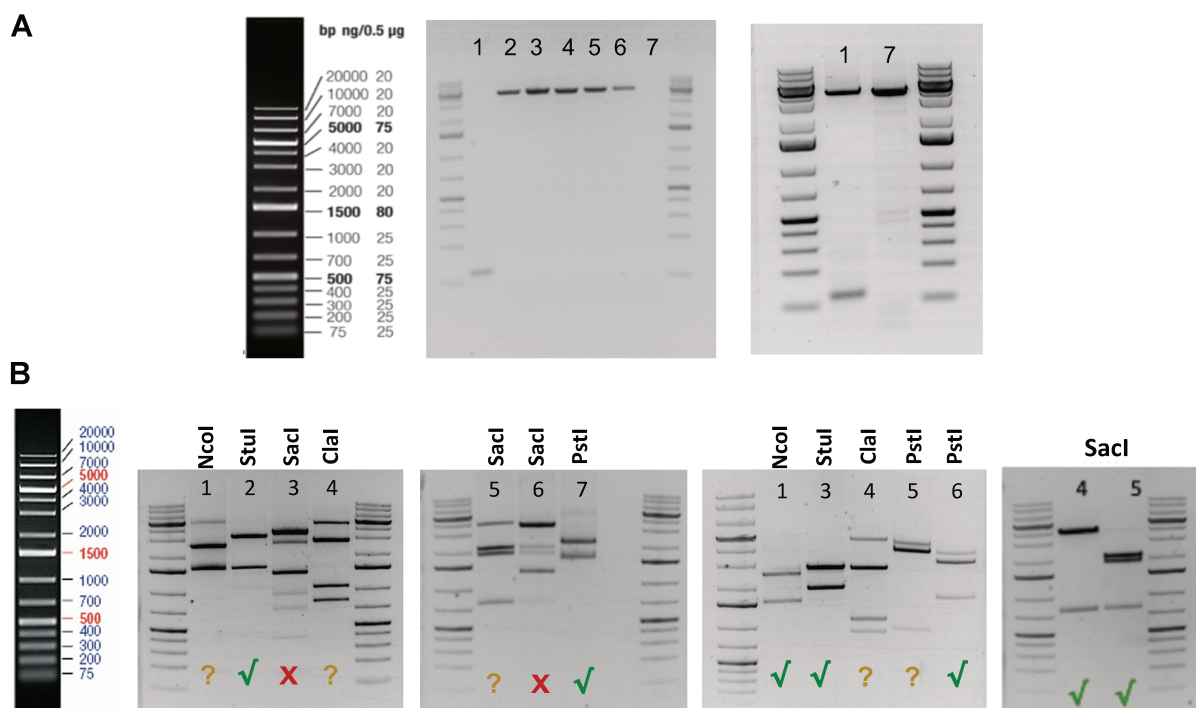


Figura 29. Fragmentos obtidos por PCR para clonagem do *cluster* t1PKS – *trans*-AT PKS – NRPS. **A.** Eletroforese em gel de agarose 0,5% TAE 1X dos produtos de amplificação por PCR dos fragmentos 1 a 7 (3,5 – 5 kb). **B.** Confirmação dos fragmentos obtidos por análise de restrição.

A linhagem *Saccharomyces cerevisiae* VL6-48N (doada por Dr. Vladimir Larionov – *National Cancer Institute, National Institutes of Health, Bethesda, EUA*), derivada da linhagem VL6-48,¹³⁵ foi utilizada para recombinação dos fragmentos em vetor pCAP03_p21. Foram preparados esferoplastos de levedura, ou seja, células com formato esférico geradas após ação de lisozima ou outro agente químico/físico, para remoção parcial da parede celular, tornando a célula permeável para transformação de DNA. Uma solução contendo o vetor pCAP03_p21 linearizado nos sítios de restrição de NdeI e EcoRI e os 7 fragmentos de DNA foi preparada, de modo a obter uma concentração de ~100 ng para cada fragmento em um volume total <40 µL (Tabela 14).

Tabela 14. Concentração de fragmentos de DNA e do vetor pCAP03_p21 linearizado transformados em esferoplastos de *S. cerevisiae* VL6-48N

Fragmento	Conc. Inicial (ng.µL ⁻¹)	Vol (µL)	Conc. Final (ng.20 µL ⁻¹)
1	33,0	3,3	108,0
2	46,9	2,2	103,2
3	74,0	1,4	103,6
4	36,0	2,8	100,8
5	35,0	2,9	101,5
6	30,0	3,4	102,0
7	27,8	3,6	100,0
pCAP03_p21	60,5	2,0	121,0

A mistura de DNA foi transformada em suspensão contendo os esferoplastos da levedura, porém a primeira tentativa de transformação não rendeu resultados satisfatórios, portanto o procedimento para obtenção de novos esferoplastos foi repetido e após segunda transformação e incubação em meio ágar seletivo, contendo 5-FOA, a 30 °C, por 5 dias, obteve-se colônias, conforme apresentado na Figura 30A. Colônias isoladas foram transferidas para nova placa de ágar contendo 5-FOA e incubadas por 48 h. As colônias que cresceram foram selecionadas para extração do DNA e análise por PCR. Foram selecionados 4 de 6 clones (lanes 1, 3, 5 e 6, Figura 30B) os quais apresentaram resultado positivo para amplificação de um fragmento de 176 pb da região central do gene 117 do BGC.

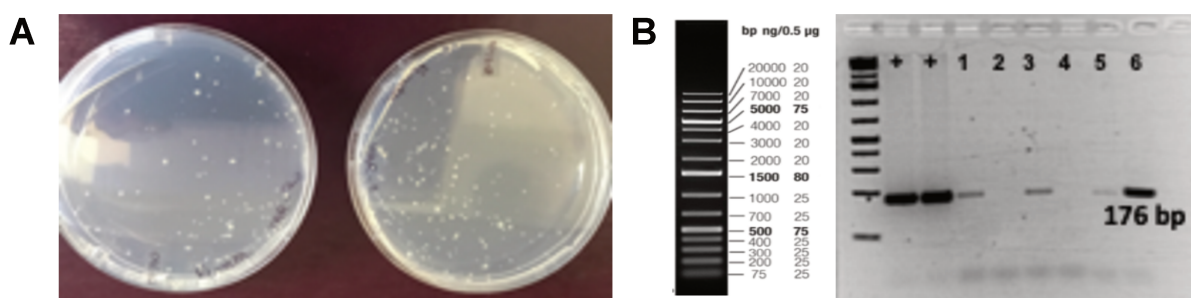


Figura 30. Triagem de colônias de levedura após recombinação. **A.** Colônias de *Saccharomyces cerevisiae* VL6-48N obtidas em meio ágar seletivo contendo 5-FOA após recombinação dos fragmentos, incubadas por 5 dias. **B.** Triagem de colônias recombinantes. Eletroforese em gel de agarose 1,5% de produtos de amplificação por PCR com o par de *primer* BGC03_117_176bp_R/F. (+): controle positivo (gDNA). Lane 1 a 6: colônias selecionadas.

As amostras obtidas após extração de DNA das colônias de levedura e que apresentaram resultados positivos em triagem por PCR foram utilizadas para transformação em *E. coli* TransforMax™ EPI300™ (Epicentre) para replicação e

manutenção dos plasmídeos obtidos após recombinação dos fragmentos de DNA, sendo importante ressaltar que foram feitas várias tentativas de transformação em célula competente preparada no laboratório, mas a transformação só foi bem sucedida após o uso de célula comercial. A dificuldade associada a transformação de DNA de levedura se dá pela complexidade do material genético obtido. Diferentes colônias isoladas obtidas em meio LB ágar contendo canamicina foram selecionadas para nova extração dos plasmídeos e análise com enzimas de restrição. A construção do plasmídeo após recombinação de todos os fragmentos corresponde a ~ 40,3 kb (Figura 31A). Os plasmídeos 2C e 6B foram confirmados com 4 combinações de enzimas de restrição (Figura 31B), a partir da comparação com o padrão *in silico* obtido (Tabela 15). Já o plasmídeo 1A apresentou dois fragmentos sem correspondência com o padrão esperado para enzima BamHI e foi desconsiderado. A estabilidade dos plasmídeos 2C e 6B foi confirmada após transformação em DH10B, sendo então nomeados como pRSalp_2C e pRSalp_6B (Kan^R).

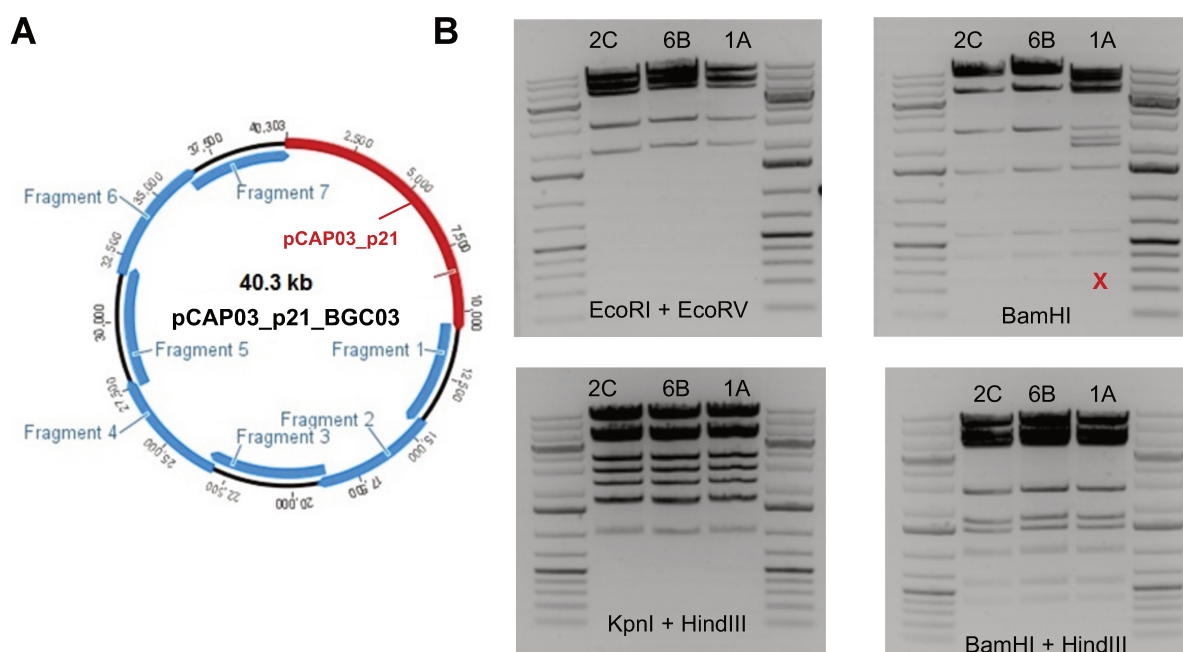


Figura 31. Confirmação da construção dos plasmídeos obtidos por TAR-cloning. **A.** Construção do plasmídeo pCAP03_p21_BGC03 (40,3 kb) esperada após recombinação dos fragmentos. **B.** Análise de restrição dos plasmídeos 2C, 6B e 1A. Análise de restrição com enzimas EcoRI + EcoRV; KpnI + HindIII; HindIII + BamHI e BamHI.

Tabela 15. Padrões de restrição obtidos *in silico* para plasmídeo recombinante pCAP03_p21_BGC03

Enzima	Sítios de restrição	Fragmentos (pb)
BamHI	7	16778; 11862; 6505; 2685;1464; 598; 411
BamHI+HindIII	10	11862; 8156; 6505; 5957; 2685; 1696; 1464; 969; 598; 411
KpnI+HindII	9	14313; 6845; 5957; 3733; 2871; 2257; 1730; 1628; 969
EcoRI+EcoRV	6	12822; 8374; 7670; 6312; 3105; 2020

O plasmídeo pRSalp_6B foi transferido para *E. coli* ET12567 (Cam^R) para experimento de conjugação entre gêneros (*E. coli* – *Streptomyces*). A conjugação foi feita entre as linhagens: ET12567/ pRSalp_6B (doadora), ET12567/pUB307 (auxiliar) e *S. coelicolor* M1146 (aceptora). Os exconjugantes obtidos foram transferidos para meio ágar SFM contendo ácido nalidíxico e canamicina, antibiótico de seleção para o plasmídeo pRSalp_6B. O DNA genômico foi extraído para confirmar a integração do plasmídeo no cromossomo do organismo heterólogo *S. coelicolor* M1146 utilizando 3 pares de *primers* flanqueando as regiões centrais dos genes biossintéticos principais (ctg1_117, 118 e 119 - Figura 32).

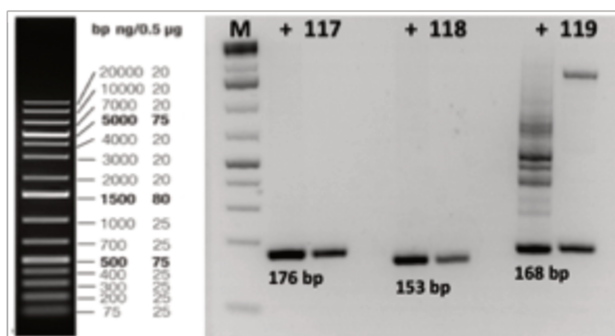


Figura 32. Confirmação da integração do plasmídeo pRSalp_6B ao cromossomo de *S. coelicolor* M1146. Eletroforese em gel de agarose 1%. Pares de *primers*: BGC03_117_176bp_F/R, BGC03_118_153bp_F/R; BGC03_119_168bp_F/R.

4.3. Deleção “*in-frame*” de genes específicos

Experimentos de deleção de genes específicos foram conduzidos com objetivo de alterar o perfil metabólico da linhagem selvagem, ou seja, verificar se ocorreria a supressão da produção de algum produto devido a deleção de determinado gene, podendo assim associá-lo ao *cluster* biossintético em análise. O

domínio de adenilação (A), presente na porção correspondente a NRPS do complexo enzimático, foi selecionado para o experimento de deleção na linhagem *Streptomyces* sp. CBMAI 2042 e para promover a deleção foi construído um plasmídeo, a partir de pYH7, contendo fragmentos que flanqueiam a região de deleção. O plasmídeo pYH7 não possui uma origem de replicação estável em *Streptomyces* e degrada facilmente na ausência do antibiótico apramicina, sendo por isso umas das principais estratégias para experimentos de deleção por cruzamento.^{149,150} Além disso, possui sítio de integração em *Streptomyces*, permitindo a transferência da informação para a linhagem selvagem.

Os *primers* para amplificação dos fragmentos I e II (~ 2000 pb) flanqueando a região de deleção do gene foram desenhados de modo a obter dois fragmentos contendo 25 nt de sobreposição entre si (laranja) e 24 nt de homologia ao plasmídeo pYH7 (azul), sendo que a região de deleção (2991 pb) selecionada manteve a regra de 3n nucleotídeos, para conservar o *frame* de leitura (Figura 33).

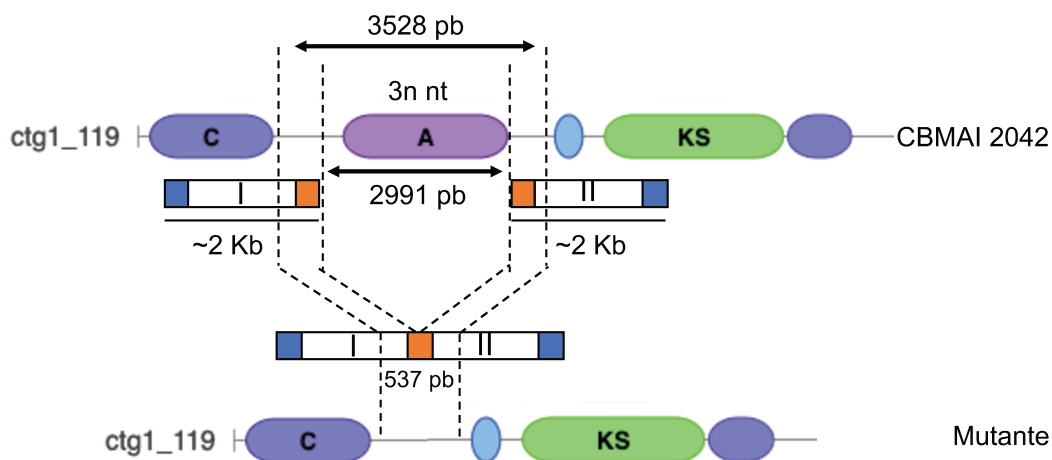


Figura 33. Construção do plasmídeo de deleção. Representação dos fragmentos I e II flanqueando a deleção *in-frame* do domínio de adenilação (A) no gene *ctg1_119*. As regiões em laranja representam a região de homologia entre os fragmentos I e II; em azul são as regiões de homologia ao plasmídeo pYH7.

Os fragmentos I (1884 pb) e II (2039 pb) foram amplificados por reação de PCR com os pares de *primers* B1_cluster14_FI/RI e B1_cluster14_FII/RII, purificados em gel de agarose 0,5% e extraídos com kit *Invisorb® Fragment CleanUp* (Invitex). O plasmídeo de deleção pRSalpΔA_{domain} foi obtido por reação recombinação com *Gibson Assembly®* a partir do plasmídeo pYH7, linearizado no sítio de *NdeI*, e dos fragmentos

obtidos, seguindo protocolo do fabricante. O produto resultante da reação de recombinação isotérmica (2 μ L) foi transformado em DH10B eletrocompetente e recombinantes foram selecionados em meio LB contendo apramicina. Colônias isoladas foram triadas por PCR de colônia com os pares de *primer* utilizados para amplificar os fragmentos I e II e os plasmídeos dos clones que apresentaram produto de amplificação positivo foram extraídos para confirmação da construção por análise de restrição com enzima NdeI, sendo esperados fragmentos de ~10,5 kb (referente ao plasmídeo pYH7) e de ~4 kb (referente a região do DNA amplificada). Em primeira análise, foram confirmados 3 de 6 plasmídeos triados, os quais foram posteriormente digeridos com uma segunda enzima (SmaI), sendo dois plasmídeos confirmados por apresentarem a construção correta, representada principalmente pela presença dos fragmentos que compreendem as regiões de ligação entre fragmento I e vetor (3360 pb), ligação entre fragmentos I e II (911 pb) e fragmento II com vetor (2916 pb) (Figura 34).

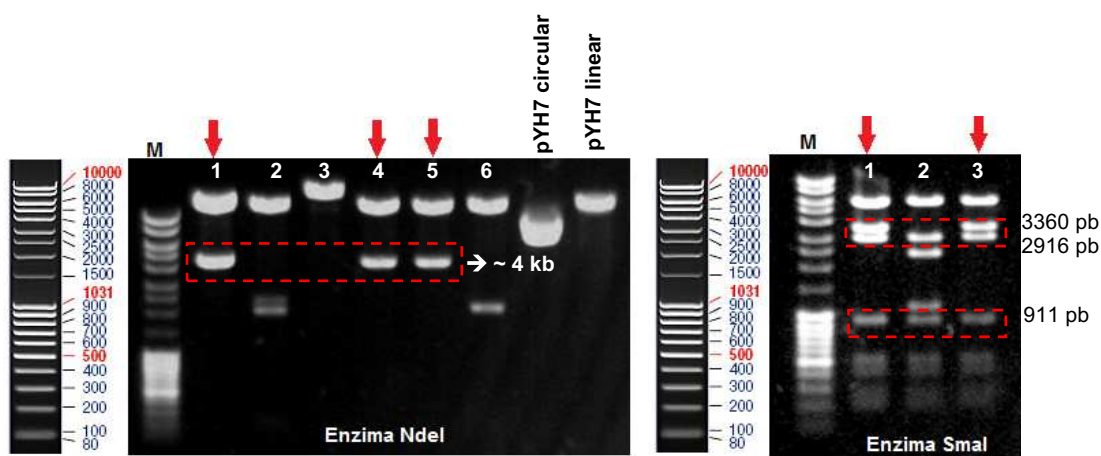


Figura 34. Padrão de digestão obtidos por restrição com NdeI e SmaI dos plasmídeos de deleção. Eletroforese em gel de agarose 0,5%. Foram observados três possíveis clones positivos quando digeridos com NdeI, dos quais dois foram confirmados com a enzima SmaI.

Um dos plasmídeos de deleção obtido, denominado pRSalp Δ A_{domain}, foi transformado em *E. coli* ET12567/pUZ8002 quimicamente competente e depois transferido por conjugação para *Streptomyces* sp. CBMAI 2042. Os exconjugantes obtidos foram transferidos para meio SFM contendo apramicina (Apr) e ácido nalidíxico (Nal), para promover o primeiro cruzamento (etapa 1), correspondente a incorporação do plasmídeo de deleção ao genoma da CBMAI 2042. Colônias obtidas

foram transferidas para placas de SFM (sem apramicina), para promover o segundo cruzamento (etapa 2) e incorporação da deleção. Colônias isoladas obtidas no segundo cruzamento, foram transferidas para placas de SFM com e sem apramicina, e apenas colônias obtidas em SFM sem apramicina foram triadas por PCR de colônia, aquelas que cresceram em SFM com apramicina foram submetidas novamente as etapas 1 e 2, até obter pelo menos um mutante.

Após seis meses de experimento, dois mutantes foram obtidos e confirmados em análise por PCR de colônia (Figura 35A) e com DNA genômico (Figura 35B). A triagem foi feita utilizando par de *primers* desenhados de modo a obter um fragmento de 537 pb para o mutante e um fragmento de 3528 pb para a linhagem selvagem (Figura 33). Os dois mutantes obtidos foram preservados em glicerol 10% e cultivados para avaliação dos perfis metabólicos em comparação com a linhagem selvagem. Nesse momento, após vários experimentos de cultivo e avaliação do perfil metabólico nenhuma alteração evidente foi observada, porém após clonagem dos *cluster* de genes via *TAR-cloning* e expressão em heterólogo, foi possível confirmar a correlação da via de biossíntese com o novo metabólito identificado, a partir do cultivo paralelo da linhagem mutante (ver item 4.4).

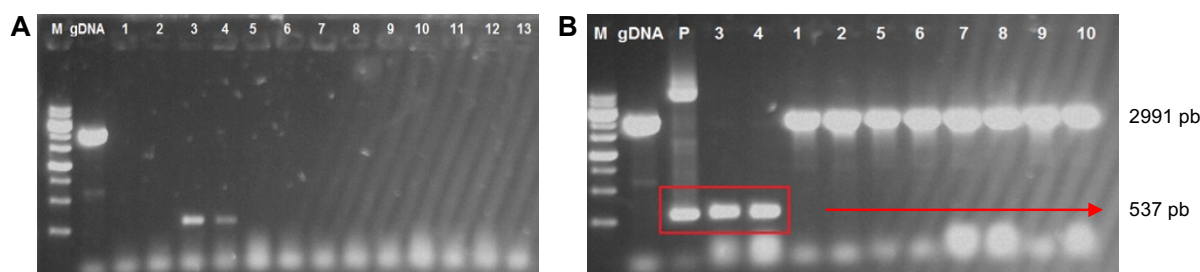


Figura 35. Triagem de mutantes. Eletroforese em gel de agarose 0,7% TAE1X. **A.** PCR de colônia. Linhas 1 a 13: colônias isoladas; gDNA: controle negativo. **B.** PCR com DNA genômico. Linhas 1 a 10: gDNA de cada colônia isolada; gDNA: controle negativo; P (plasmídeo de deleção): controle positivo. M: marcador 1 kb Sinapse.

4.4. Alpiniamida

Estudos relacionados a alpiniamida foram reportados pela primeira vez em 2013¹⁷⁰ e 2018¹⁷¹, sendo o primeiro relacionado ao isolamento e caracterização por RMN do composto e o segundo a identificação, caracterização e anotação do *cluster* de genes responsável pela biossíntese. Para identificação e anotação da alpiniamida como metabólito correspondente produzido pelo *cluster* de genes selecionado para este estudo, o organismo heterólogo M1146_RSalp obtido por clonagem via TAR, conforme detalhado na sessão 4.2.4, foi inicialmente cultivado em paralelo com a linhagem selvagem *Streptomyces* sp. CBMAI 2042 e linhagem heteróloga *S. coelicolor* M1146 (sem o BGC) em meios TSBY e MP para comparação dos perfis metabólicos.

Os cultivos obtidos em triplicata para cada linhagem foram extraídos por dois métodos, extração com resina não-iônica amberlite XAD7HP (Sigma-Aldrich) e extração com acetato de etila. Com base nos resultados preliminares obtidos, optou-se por adotar a extração com resina, pois este apresentou maior reprodutibilidade dos resultados dos cultivos em meio MP. Deste modo, para confirmar a produção dos metabólitos (íons) identificados, os experimentos de cultivos e extratos das três linhagens foram repetidos mais duas vezes.

A análise comparativa dos cromatogramas obtidos por cromatografia líquida (LC) evidenciou a produção de alpiniamida, identificada no tempo de retenção entre 13 e 14 min, em amostras correspondentes a linhagem heteróloga M1146_RSalp (vermelho) e a linhagem selvagem *Streptomyces* sp. CBMAI 2042 (WT – azul) (Figura 36).

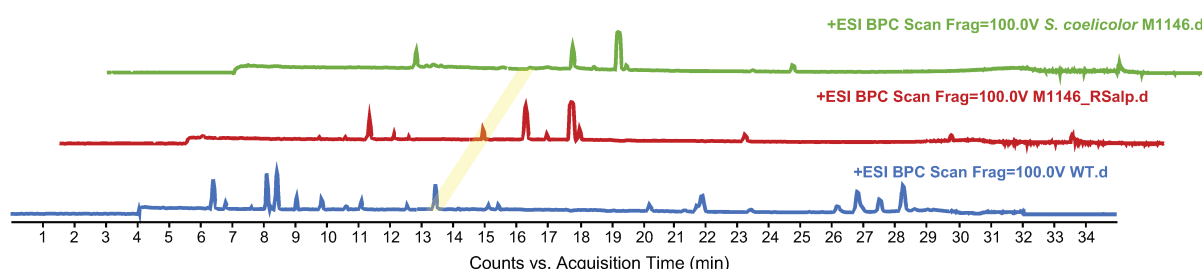


Figura 36. Cromatogramas de pico base do extrato metanólico dos cultivos em meio MP após análise por LC-ESI(+)-MS. *S. coelicolor* M1146 (verde) – 1; M1146_RSalp (vermelho) – 2; *Streptomyces* sp. CBMAI 2042 (azul) – 3. Em amarelo destaca-se o pico correspondente a alpiniamida e observado apenas nas amostras de 2 e 3.

O espectro de massas obtido referente ao tempo de retenção 13,442 min apresentou um conjunto de íons principais correspondentes a m/z 326, m/z 344, m/z 366 e m/z 709 (Figura 37).

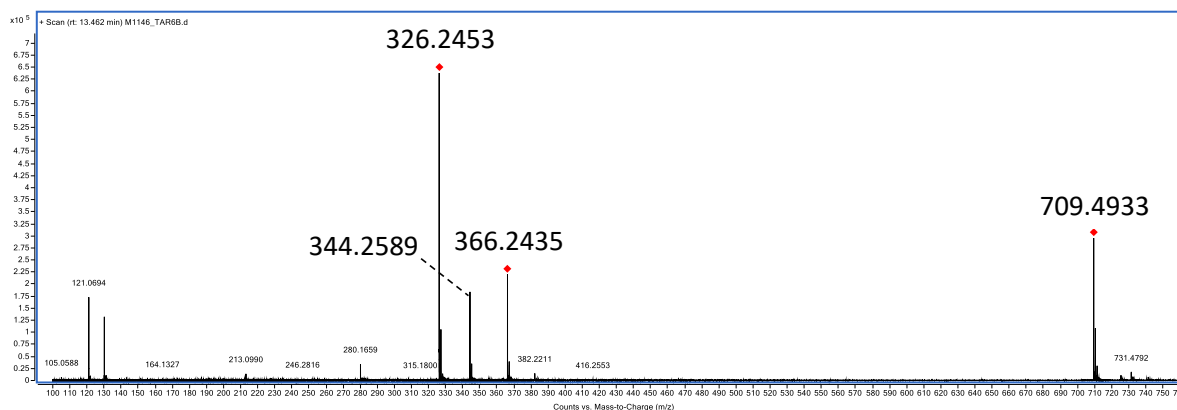


Figura 37. Espectro de massas correspondente ao pico com tempo de retenção em 13,442 min.

De modo a confirmar se a origem biossintética da alpiniamida era correspondente ao *cluster* de genes clonado, a linhagem mutante RSalp Δ A_{domain} obtida a partir da deleção do domínio de adenilação (ver item 4.3) foi cultivada nas mesmas condições estabelecidas anteriormente. A deleção de um domínio essencial para a biossíntese pode levar a supressão da produção de um determinado metabólito pelo organismo, facilitando a identificação dos genes correlacionados com a produção do mesmo. Adicionalmente, a linhagem M1146/ESAC5N obtida pela clonagem dos genes a partir da biblioteca genômica (ver item 4.2.2) também foi cultivada nas mesmas condições descritas, para verificar se a mesma produzia o composto agora detectado. Após análises por LC-MS, os cromatogramas obtidos (Figura 38) foram comparados e foi observada a supressão do sinal em 13,4 min para o mutante RSalp Δ A_{domain}, ou seja, não houve produção da alpiniamida, não deixando dúvidas sobre a origem biossintética da molécula.

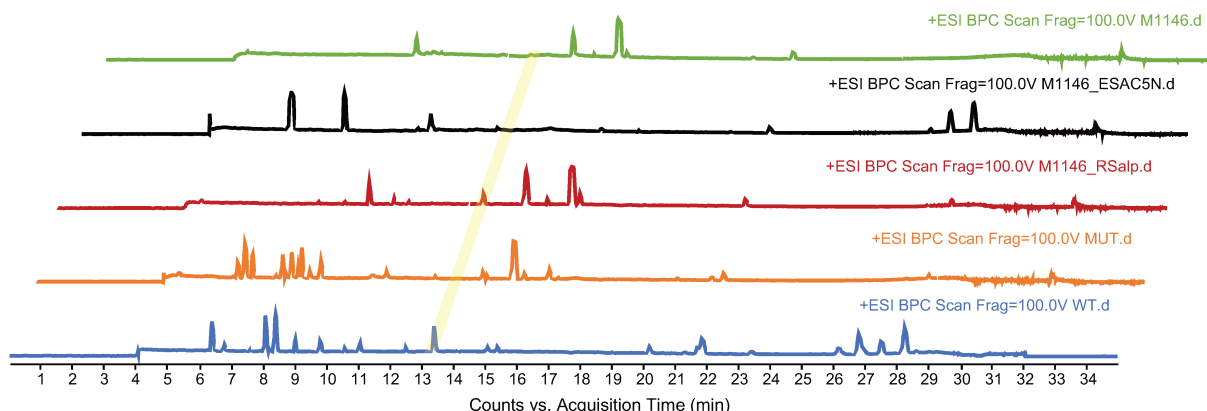


Figura 38. Cromatogramas de pico base obtidos para todas as linhagens no estudo de identificação da alpiniamida após análise por LC-MS. As amostras são resultantes de extração por resina dos cultivos em meio MP. *S. coelicolor* M1146 (verde), M1146/ESAC5N (preto), M1146_RSalp (vermelho), RSalp Δ A_{domain} (laranja) e *Streptomyces* sp. CBMAI 2042 (azul).

Os cromatogramas de íon extraído correspondentes ao íon majoritário de m/z 326,19 permitiram identificar a produção de alpiniamida na linhagem M1146/ESAC5N, não identificada pelos métodos utilizados anteriormente, embora aparentemente em pequena quantidade (Figura 39).

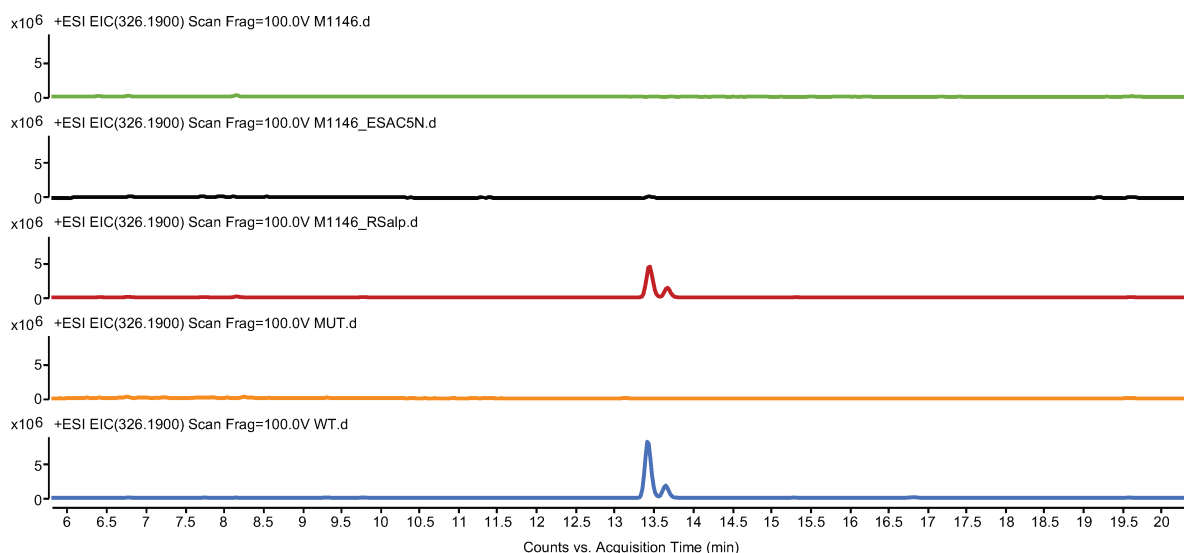


Figura 39. Cromatogramas de íon extraído para m/z 326,19 obtidos para todas as linhagens no estudo de identificação da alpiniamida após análise por LC-MS. As amostras são resultantes da extração dos cultivos em meio MP com resina. *S. coelicolor* M1146 (verde), M1146/ESAC5N (preto), M1146_RSalp (vermelho), RSalp Δ A_{domain} (laranja) e *Streptomyces* sp. CBMAI 2042 (azul).

A partir dos resultados de análises por espectrometria de massas de alta resolução e identificado o composto alvo, a linhagem M1146_RSalp foi cultivada em maior escala (500 mL de meio MP) para purificação e caracterização da alpinimida. O extrato bruto obtido por extração com resina/metanol foi previamente purificado em cartuchos C18 (HyperSep™ C18 Cartridges – Thermo Scientific) por eluição com água:metanol (95:5; 90:10, 50:50 e 20:80) e coletada em fração correspondente a 90:10. A fração foi concentrada em rotaevaporador e o produto obtido foi purificado em HPLC preparativo e as frações correspondentes a alpiniamida foram coletadas, concentradas e água totalmente removida sob vácuo, resultando em ~3 mg de produto. A amostra foi submetida a análise por RMN e espectrometria de massas de alta resolução (HRESIMS) para a determinação estrutural.

4.4.1. Caracterização e determinação estrutural

A determinação da estrutura da alpiniamida isolada e purificada do cultivo em meio MP da linhagem M1146_RSalp, resultante da clonagem do *cluster* biossintético t1PKS – *trans*-AT PKS – NRPS anotado no genoma de *Streptomyces* sp. CBMAI 2042, foi baseada nos resultados obtidos por RMN de ^1H , HSQC, HMBC e COSY, associados aos dados obtidos por espectrometria de massas de alta resolução (HR-ESI-TOFMS), para determinação da fórmula molecular.

A partir da análise por espectrometria de massas de alta resolução (+)-HRESIMS (Figura 40) foi possível atribuir a fórmula molecular $\text{C}_{17}\text{H}_{29}\text{NO}_6$ (grau de instauração 4) ao íon de m/z 366,1886 $[\text{M}+\text{Na}]^+$ (calcd 366,1887) e a fórmula molecular $\text{C}_{17}\text{H}_{27}\text{NO}_5$ (grau de instauração 5) ao íon de m/z 326,1960 $[\text{M}'+\text{H}]$ ou $[\text{M}-\text{H}_2\text{O}+\text{H}]^+$ (calcd 326,1962) – (Tabela 16).

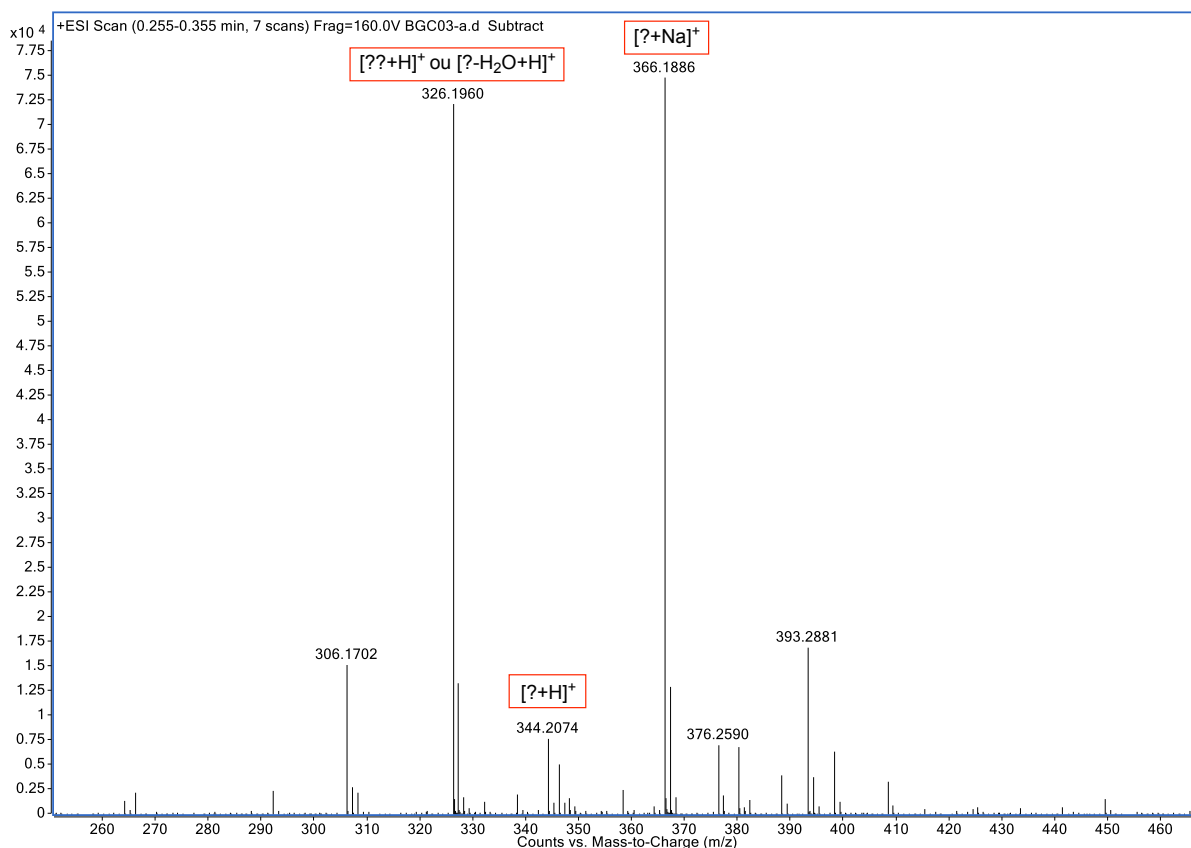


Figura 40. (+)-HRESIMS da alpiniamida isolada por HPLC-preparativo. m/z 366,1886 $[M+Na]^+$ (calcd 366,1887); m/z 326,1960 $[M'+H]$ ou $[M-H_2O+H]^+$ (calcd 326,1962); m/z 344,2074 $[M+H]^+$ (calcd 344,1994).

Tabela 16. Dados obtidos por espectrometria de massas de alta resolução para determinação de fórmula molecular

<i>Fórmula (M)</i>	<i>Massa</i>	<i>Massa Calc.</i>	<i>Calc m/z</i>	<i>Δ (ppm)</i>	<i>m/z</i>
$C_{16}H_{21}N_8$	325,1887	325,1889	326,1962	-0,6	326,1960
$C_{17}H_{27}NO_5$	325,1887	325,1889	326,1962	-0,61	326,1960
$C_{15}H_{25}N_4O_4$	325,1887	325,1876	326,1949	3,52	326,1960
$C_{18}H_{23}N_5O$	325,1887	325,1903	326,1975	-4,73	326,1960
$C_{16}H_{23}N_8O$	343,1994	343,1995	366,1887	-0,3	366,1886
$C_{17}H_{29}NO_6$	343,1994	343,1995	366,1887	-0,32	366,1886
$C_{15}H_{27}N_4O_5$	343,1994	343,1981	366,1874	3,6	366,1886
$C_{18}H_{25}N_5O_2$	343,1994	343,2008	366,1900	-4,21	366,1886

O espectro de RMN de ^1H (500 MHz) em metanol- d_4 (Figura 41) apresentou sinais adicionais aos sinais principais referentes a presença de compostos que estão possivelmente associadas a isômeros estruturais da alpiniamida e, por isso, para alguns multipletos não foi possível determinar a constante de acoplamento (J). Pelo espectro de RMN de ^1H foram observados grupos de sinais relacionados a hidrogênios de cinco grupos metílicos, sendo: um grupo metila em δH 1,60 (H-9, s) ligado ao carbono olefínico C-4; três dubletos dos grupos metila vicinais a C2', C4' e C2: δH 1,15 referente a H-7' (d, $J_{\text{H}7'-\text{H}2'}$ 7,1 Hz), δH 1,12 referente a H-8' (d, $J_{\text{H}8'-\text{H}4'}$ 7,2 Hz) e δH 0,93 referente a H-8 (d, $J_{\text{H}8-\text{H}2}$ 7,0 Hz); e um tripleto do grupo metila H-7 (vicinal a C6) em δH 0,97 (t, $J_{\text{H}7-\text{H}6}$ 7,6 Hz). Uma dupla ligação foi indicada na estrutura pela presença de um tripleto em δH 5,42 (H-5, t, $J_{\text{H}5-\text{H}6}$ 7,4 Hz) o qual corresponde a um hidrogênio vinílico. Sinais correspondentes aos hidrogênios da posição H6' do aminoácido glicina foram observados em δH 4,12 (H-6a', d, $J_{\text{H}6a'-\text{H}6b'}$ = 18,7 Hz) e 4,26 (H-6b', d, $J_{\text{H}6b'-\text{H}6a'}$ = 18,7 Hz). A expansão dos sinais atribuídos por RMN de ^1H estão apresentados no Apêndice 8 (itens 8.1 a 8.5). Adicionalmente, foram também obtidos mapas de correlação por HSQC (Figura 43) e HMBC (Figura 44), corroborando para elucidar a estrutura da alpiniamida (Tabela 17). Os deslocamentos químicos de ^{13}C foram atribuídos por correlação a partir do espectro de RMN do tipo HMBC, apresentando deslocamentos característicos de carbono carbonílico de ácidos/amidas (C-1, 177,2 ppm; C-1', 177,1 ppm) e carbono carbonílico de cetonas (C-5', 208,8 ppm), além de carbonos saturados sp^3 – CH (C-3, 80,0 ppm; C-3', 73,2 ppm) ligados aos grupos hidroxila.

Tabela 17. Dados obtidos por RMN para alpiniamida em CD₃OD*

Posição	δ_c^1 ; tipo	$\delta_c^{a,b}$	δ_H (J em Hz)	δ_H^a (J em Hz)	HMBC
1'	177,1; C	179,03			
2'	43,6; CH	42,71	2,53; m	2,5; quint (7,0)	3', 7'
3'	73,2; CH	73,54	4,04; m	4,01; ddd (7,0; 5,3; 1,5)	
4'	46,6; CH	47,23	2,94; m	2,92; qd (7,0; 5,3)	5'
5'	208,8; C	209,27			
6'	47,8; CH ₂	48,81	4,12; d (18,68) 4,26; d (18,71)	4,13; d (18,5) 4,25; d (18,5)	5', 1
7'	13,5; CH ₃	14,49	1,15; d (7,10)	1,16; d (7,0)	1', 2', 3'
8'	9,4; CH ₃	9,79	1,12; d (7,20)	1,13; d (7,0)	3', 5', 6'
1	177,2; C	176,75			
2	43,9; CH	44,28	2,53; m	2,54; dq (9,5; 7,0)	1, 3
3	80,0; CH	80,40	3,99; d (9,63)	3,98; d (9,5)	5, 9
4	134,0; C	133,13			
5	130,8; CH	131,94	5,42; t (7,37)	5,41; t (7,2)	9, 6, 3
6	20,4; CH ₂	20,85	2,06; m	2,06; quint (7,5)	7, 5, 4
7	13,0; CH ₃	13,88	0,97; t (7,59)	0,97; t (7,5)	6, 5
8	13,9; CH ₃	14,38	0,93; d (6,98)	0,93; d (7,0)	1, 2, 3
9	9,1; CH ₃	10,41	1,60; s	1,60; s	3, 4, 5

* os espectros de RMN de ¹H foram adquiridos a 500 MHz em metanol-*d*₄.

¹ deslocamentos de ¹³C foram obtidos de mapas de correlação HSQC/HMBC.

^a dados de RMN adquiridos para alpiniamida A isolada por Paulus et. al (2018)¹⁷¹.

^b dados de RMN adquiridos em CDCl₃.

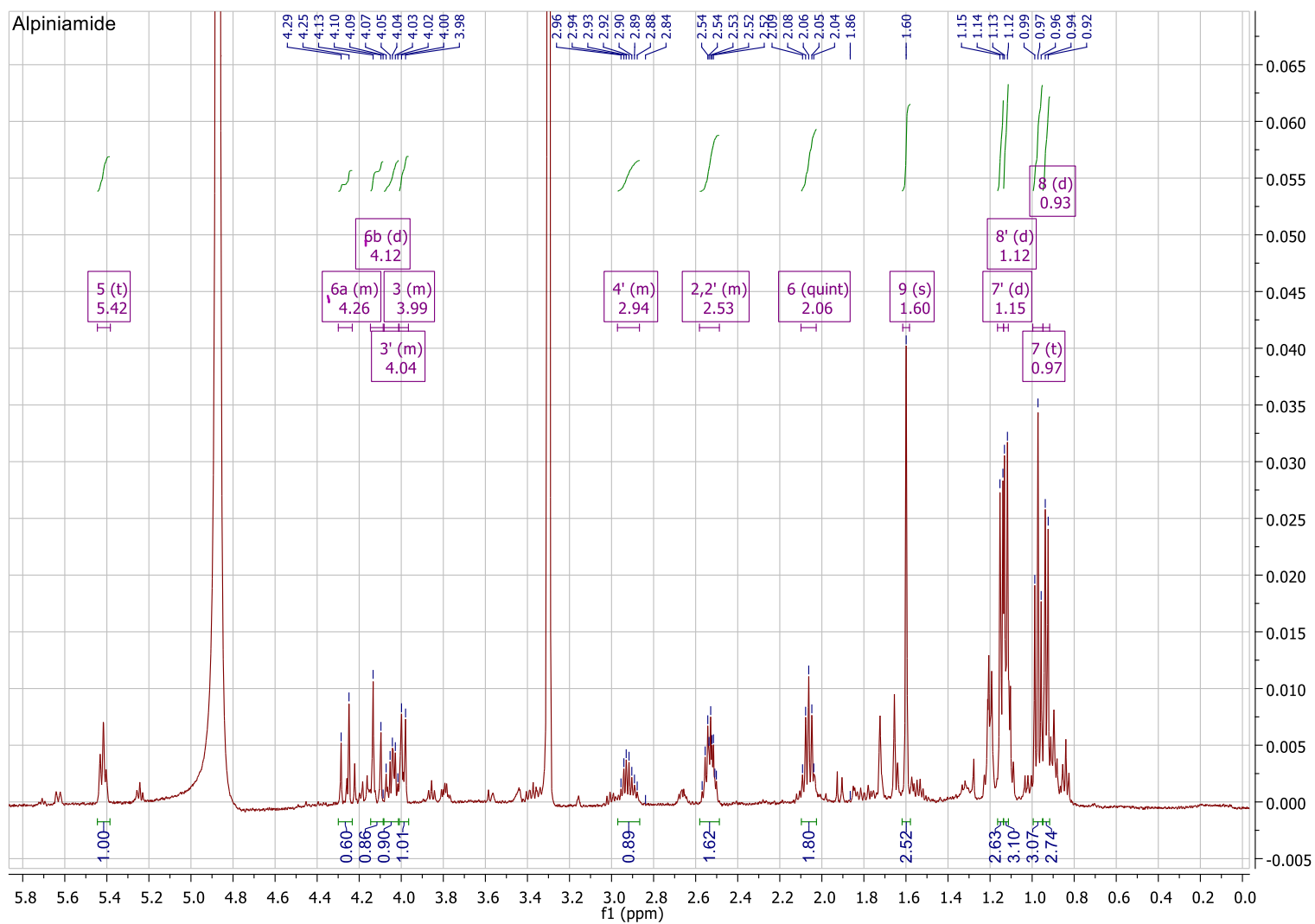


Figura 41. Espectro de RMN de ^1H adquirido para alpiniamida (CD_3OD , 500 MHz).



Figura 42. Espectro de RMN do tipo COSY adquirido para alpiniamida (CD₃OD, 500 MHz).

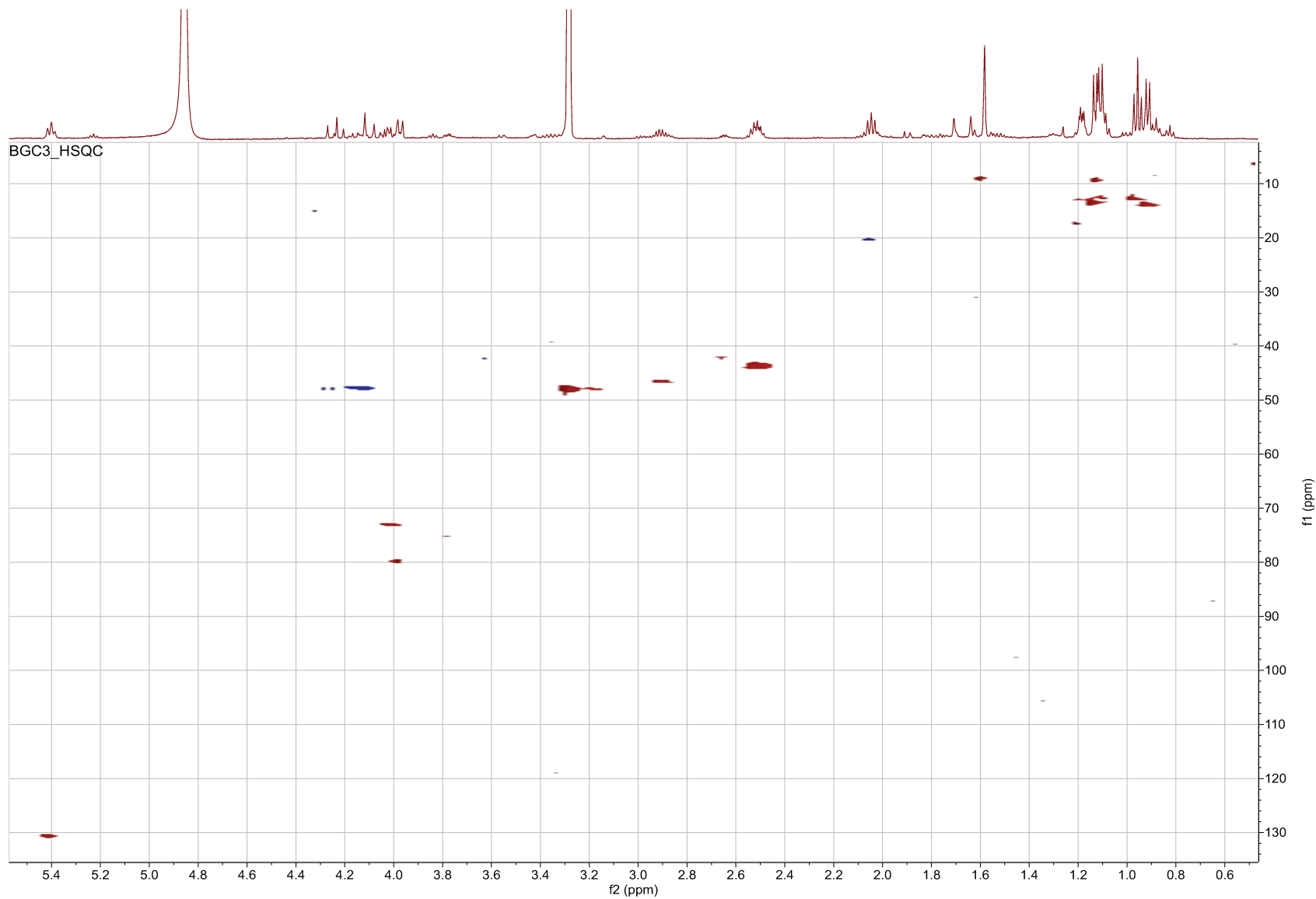


Figura 43. Espectro de RMN do tipo HSQC adquirido para alpiniamida (CD_3OD , 500 MHz).

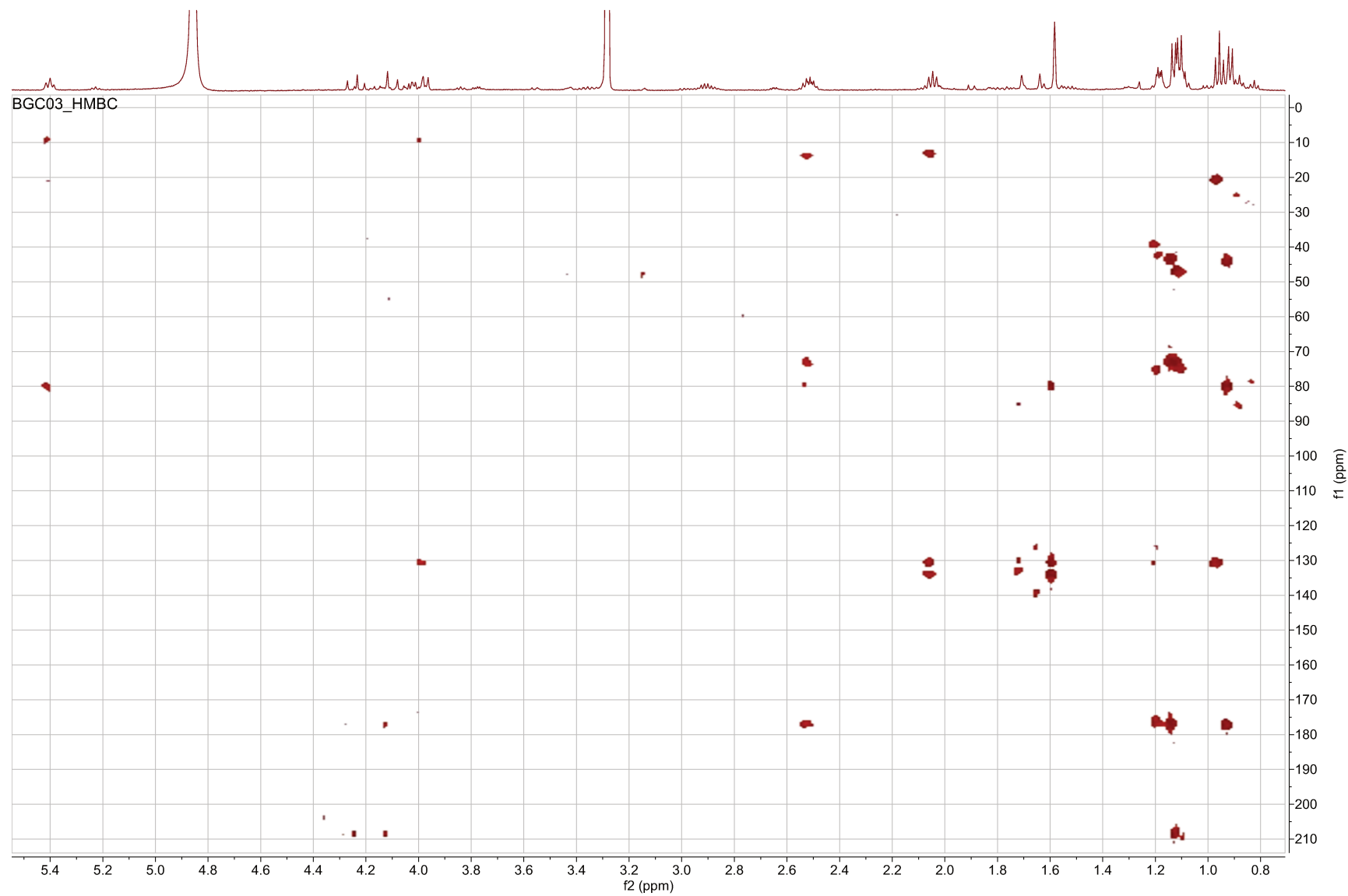


Figura 44. Espectro de RMN do tipo HMBC (500 MHz) adquirido para alpiniamida (CD_3OD , 500 MHz).

A partir da fórmula molecular obtida e dos dados de RMN, pode-se propor uma estrutura para a alpiniamida codificada pelo BGC t1PKS – *trans*-AT PKS – NRPS contendo 17 carbonos, 27 hidrogênios, 1 nitrogênio oriundo do aminoácido glicina incorporado pela NRPS e 5 oxigênios, conforme representado na Figura 45.

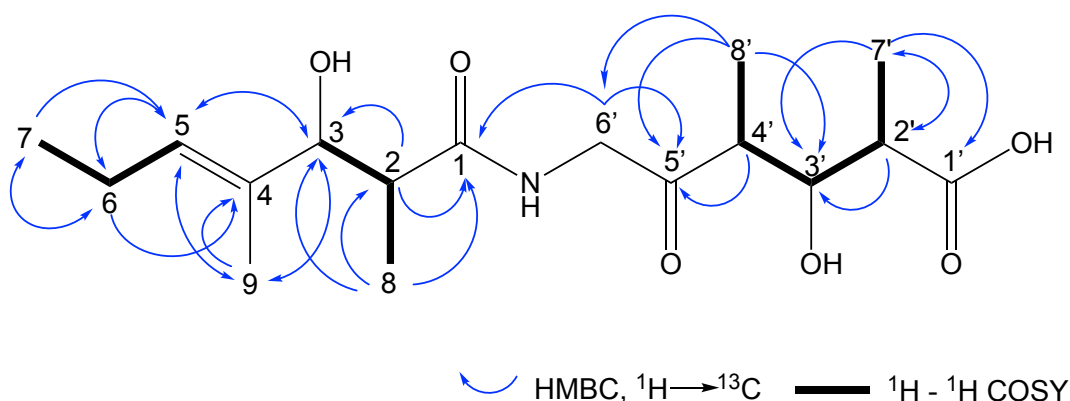


Figura 45. Atribuição de dados obtidos por espectroscopia de RMN para alpiniamida (**25**). Representação da correlações obtidas por COSY (negrito) e HMBC (setas em azul).

A partir da estrutura proposta viabilizada pelos métodos espectrométricos e espectroscópicos de análise partiu-se para a busca em base de dados como o SciFinder® e o StreptomeDB¹⁷² e foi possível, então, identificar a molécula isolada como sendo a alpiniamida (**25**). Como mencionado anteriormente, a alpiniamida foi isolada pela primeira vez em 2013 do organismo endofítico *Streptomyces* YIM66017,¹⁷⁰ isolado de *Alpinia oxyphylla*, dando origem ao nome da molécula. Contudo, embora já reportada sua estrutura, a biossíntese e os *clusters* de genes não haviam sido descritos ou anotados na literatura, por isso os mesmos não estavam anotados em bancos de dados, inviabilizando a correlação com os genes deste estudo.

Posteriormente, em 2018, quando estavam sendo finalizados os estudos deste projeto de doutorado, Paulus et al.¹⁷¹ relataram também o isolamento da alpiniamida A e quatro análogos estruturais (B1, B2, C e D) biossintetizados por *Streptomyces* sp. IB2014/011-12. No estudo, os autores descrevem a origem biossintética dos compostos e, portanto, os clusters de genes relacionados a produção da alpiniamida na linhagem IB2014/011-12, bem como foi feita uma proposta de biossíntese e determinada a estereoquímica relativa da alpiniamida A (Figura 46).

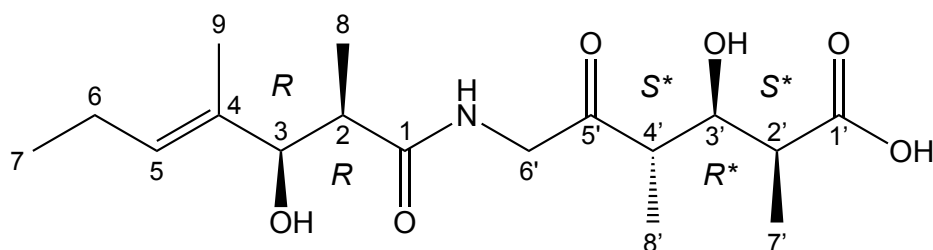


Figura 46. Estereoquímica sugerida para alpiniamida (**25**) corresponde aquela descrita por Paulus et al.¹⁷¹

4.4.2. Elucidação da Biossíntese da Alpiniamida

A alpiniamida isolada de *Streptomyces* sp. CBMAI 2042 (este estudo) não estava diretamente relacionada com anotação dos genes e números de módulos conforme a análise *in silico*. O BGC clonado e anotado em *Streptomyces* sp. CBMAI 2042 (Figura 47) possui 3 genes biossintéticos principais (ctg1_117, 118 e 119), divididos em módulos e domínios obrigatórios para PKS (AT, ACP, KS) e NRPS (A, PCP, C), apresenta domínios “acessórios” de KR, DH e cMT, além de um gene (ctg1_115) codificando para uma *trans*-AT PKS, classe de PKSs identificada e caracterizada mais recentemente.⁹³

A análise *in silico* do *cluster* de genes, via antiSMASH, permitiu inferir algumas informações importantes sobre como a biossíntese da alpiniamida poderia ocorrer. A primeira, se refere a especificidade dos domínios das aciltransferases, ou seja, as unidades extensoras incorporadas na cadeia do policetídeo. O antiSMASH gera a informação a partir de uma sequência (assinatura) de 24 aminoácidos do sítio ativo das enzimas,¹⁷³ além de utilizar o método de Minowa et al¹⁷⁴ baseado em pHMMs (modelo estatístico composto por parâmetros desconhecidos, utilizado para identificar padrões ocultos a partir de outros observáveis, ou seja, para reconhecimento de padrões). O *cluster* em estudo apresenta dois domínios AT essenciais para a biossíntese, um localizado no gene 117, aqui denominado de AT₁, cujo substrato específico corresponde ao metilmalonato (MMal), e outro presente no gene ctg1_115, aqui denominado AT_{trans}, específico para o substrato malonato (Mal). Uma segunda predição importante, através do NRPSPredictor2^{175,176} e do método de Minowa et al,¹⁷⁴ trata da incorporação específica do aminoácido glicina (Gly) pelo domínio de adenilação no gene ctg1_119, o que pôde ser confirmado pela estrutura

obtida. E a terceira, pela organização dos genes e o composto obtido, pode-se ressaltar que a biossíntese não apresenta a linearidade comum de PKSs do tipo I, mas o resultado de uma ação interativa característica de *trans*-AT PKS (Figura 47).

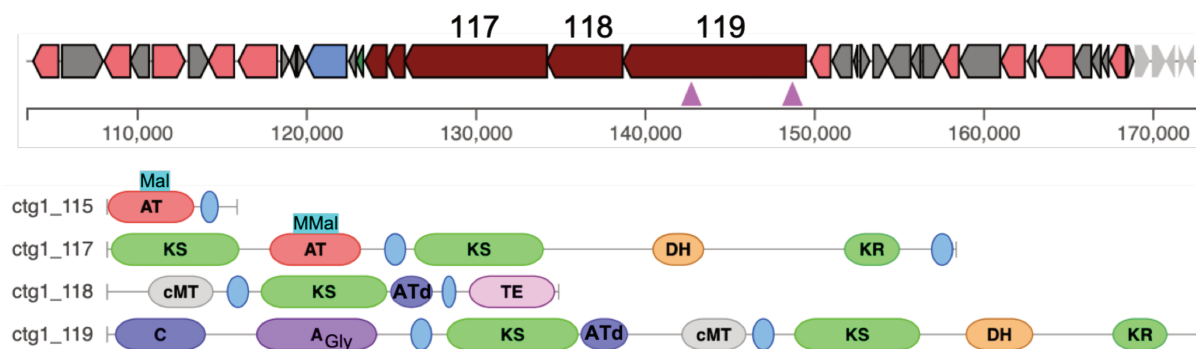


Figura 47. Anotação do *cluster t1PKS – trans-AT PKS – NRPS* no genoma da linhagem *Streptomyces sp. CBMAI 2042*.

Quando Paulus et al.¹⁷¹ relataram também o isolamento da alpiniamida A e análogos estruturais de *Streptomyces sp. IB2014/011-12*, foi feita uma comparação com o *cluster* de genes reportado pelo autores e anotado no NCBI com o registro contig00008 (sequência QEIK01000008.1 obtida via BioProject: PRJNA453997 – NCBI). A sequência apresentou 90,31% de similaridade e 99% de cobertura aos genes descritos para a linhagem *Streptomyces sp. CBMAI 2042* e com base na anotação dos genes apresentados por Paulus et al., combinados aos experimentos de expressão heteróloga desse projeto, foi possível estabelecer e delimitar os genes essenciais para produção de apiniamida àqueles contidos entre ctg1_114 e ctg1_119.

Os autores apresentaram uma proposta de biossíntese, representada na Figura 48, na qual descreveram o gene *alp1* (correspondente ao gene ctg1_117 em CBMAI 2042), específico para o substrato metilmalonato, como sendo o iniciador da cadeia. As demais unidades extensoras da cadeia policetílica, segundo os autores, são derivadas do malonato e incorporadas por ação do domínio *trans*-AT, sendo as metilações resultantes da ação de dois domínios de metiltransferases (cMT), anotadas nos genes *alpA2* e *alpA3*. Por fim, atribuíram uma interação de parte do módulo 2, contendo uma TE, na etapa final de biossíntese, onde ocorre uma reação de hidrólise e liberação da cadeia do complexo enzimático, dando origem a alpiniamida A. Os análogos são resultantes da ativação/ inativação de domínios acessórios.

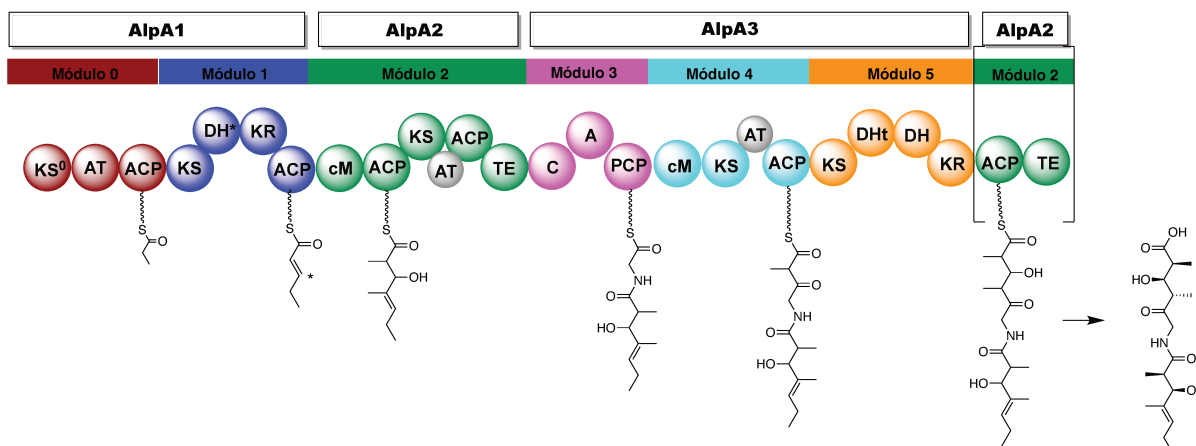


Figura 48. Proposta de biossíntese da alpiniamida A envolvendo AlpA1, AlpA2 e AlpA3, apresentada por Paulus e colaboradores (adaptado de Paulus et al.¹⁷¹).

Contudo, os experimentos de incorporação de metionina marcada (-metil-¹³C) e a proposta de biossíntese apresentados pelos autores não corroboram entre si, visto que foi observada a incorporação de apenas uma metila proveniente de metionina marcada, ao invés de quatro como poderia ser esperado caso a *trans*-AT atuasse na incorporação de todas as unidades extensoras. Posto isso, compreendemos a importância de realizar experimentos adicionais para elucidar a biossíntese da alpiniamida. Foram realizados os experimentos de incorporação de metionina marcada (L-metionina-(metil-¹³C)), bem como precursores adicionais, tais como acetato de sódio-1-¹³C, propanoato de sódio-1-¹³C, D-glicose-1-¹³C e D-glicose-6-¹³C. Dos experimentos realizados, apenas aqueles com metionina e propanoato marcados apresentaram resultados conclusivos de incorporação. Os demais precursores não foram incorporados na cadeia policetídica, talvez pela concentração utilizada e competição pelos mesmos substratos em outras vias metabólicas primárias.

O espectro de massas de alta resolução obtido a partir de extratos do cultivo da linhagem M1146_RSalp em meio MP com adição de metionina marcada (1 mM) em comparação com o cultivo padrão, sem adição de precursor marcado, apresentou um perfil de incorporação surpreendente (Figura 49). Foi obtido alpiniamida, resultante da incorporação do grupo metila derivado da L-metionina-(metil-¹³C) por ação de uma metiltransferase (cMT) na presença do cofator SAM (S-adenosilmetionina), em duas posições. Observou-se um incremento de 96,6% no sinal correspondente ao isótopo +1 e 84,5% no sinal correspondente ao isótopo +2, o que

não deixa dúvidas sobre a evidência de incorporação de apenas duas metilas na cadeia policetídica via ação de cMT.

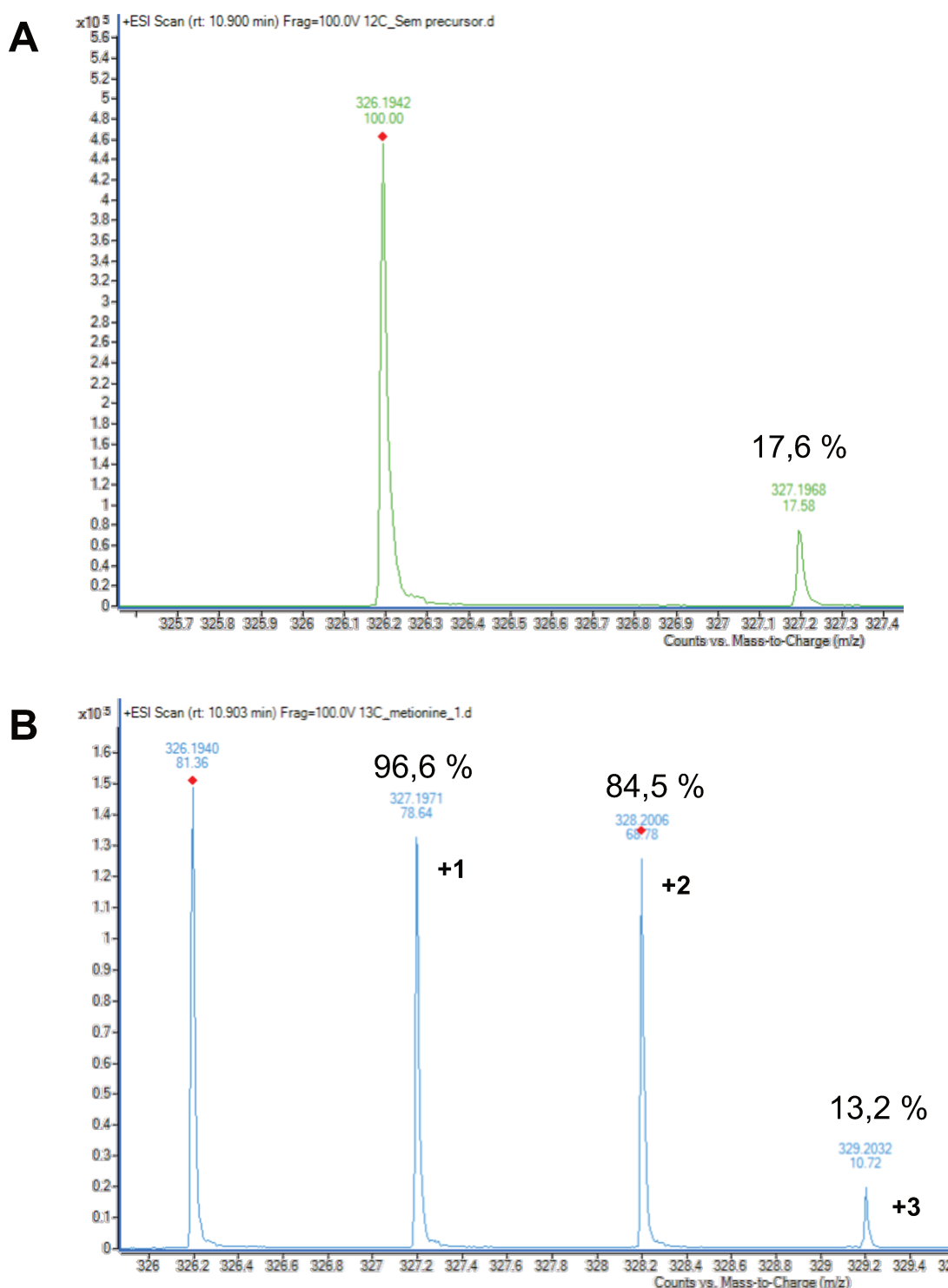


Figura 49. Experimentos com metionina marcada. Espectros de massas de alta resolução obtidos a partir de extratos de cultivo da linhagem M1146_RSalp em meio MP **A.** sem adição de precursor marcado e **B.** com adição de L-metionina-(metil- ^{13}C).

Os experimentos realizados com adição de propanoato-1- ^{13}C de sódio marcado (12,9 mM) nos cultivos de M1146_RSalp em meio MP, também apresentaram resultado de incorporação do substrato marcado durante a biossíntese, sugerindo que três unidades metiladas incorporadas a estrutura da alpiniamida são provenientes do propanoato, ou metilmalonato (Figura 50). Foram observados incrementos de sinais para os isótopos +1 (80,9%), +2 (66,9%) e +3 (47,6%).

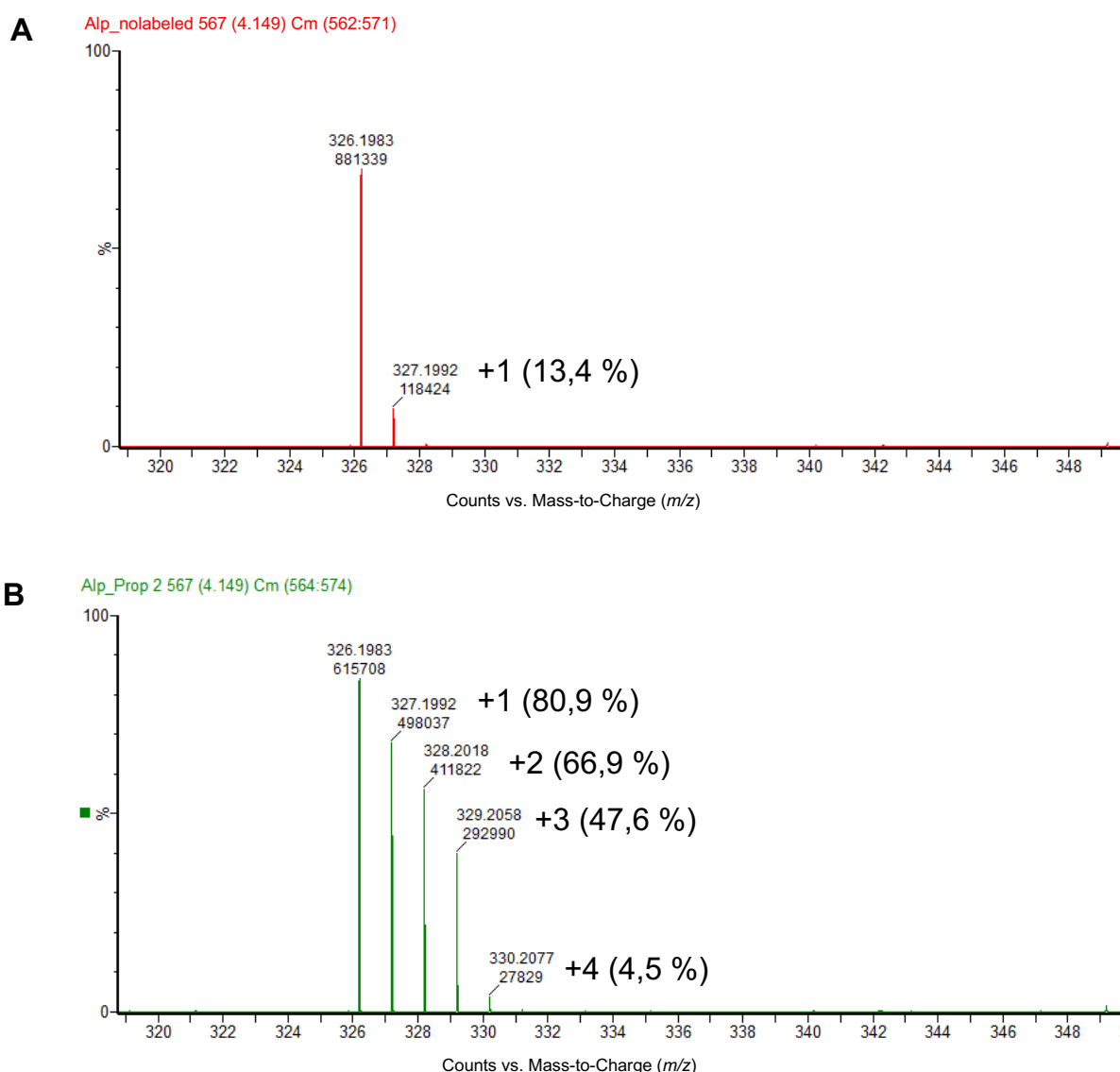


Figura 50. Experimentos com propanoato marcado. Espectros de massas de alta resolução obtidos a partir de extratos de cultivo da linhagem M1146_RSalp em meio MP **A.** sem adição de precursor marcado e **B.** com adição de propanoato-1- ^{13}C de sódio.

De forma complementar, os espectros de fragmentação MS/MS obtidos em análises com metionina marcada foram avaliados e pode-se observar íons de fragmentos específicos que mantiveram o padrão isotópico. Dentre todos os sinais, julgou-se criterioso o conjunto de íons de fragmentos de m/z 137,0962 (calcd 137,0961; erro 0,73 ppm); m/z 137,0594 (calcd 137,0597; erro 2,19 ppm); m/z 154,0820 (calcd 154,0863; erro 27,65 ppm) e m/z 154,1213 (calcd 154,1226; erro 8,43 ppm). Os mesmos correspondem a íons de fragmentos característicos da divisão da molécula em região pré (preto) e pós-NRPS (vermelho). Nos espectros MS/MS ampliados (Figura 51, Figura 52, Figura 53, Figura 54) fica evidente pelo padrão isotópico observado a incorporação de metionina marcada apenas na porção pós-NRPS, representado pelos íons fragmentos de m/z 137,0594 e m/z 154,0820 (vermelho). Adicionalmente, o íon fragmento de m/z 109,1002 (calcd 109,10118; erro 8,98 ppm) intenso e oriundo da porção pré-NRPS, pode ser considerado como uma evidência adicional da região de incorporação de metionina marcada ao não apresentar padrão isotópico (Figura 55).

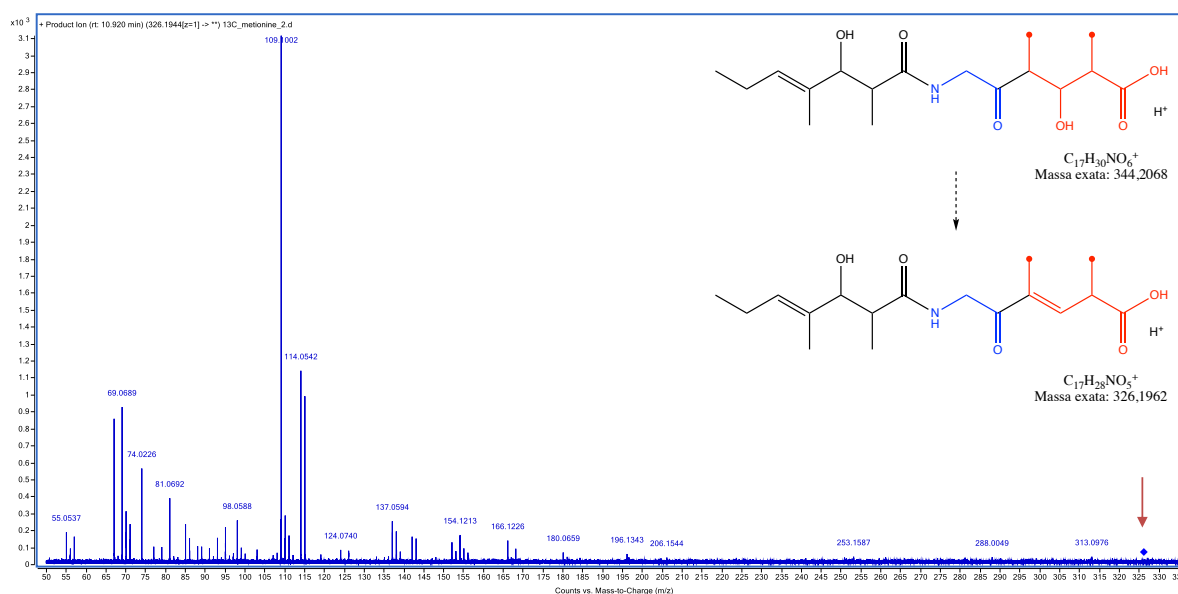


Figura 51. Espectro MS/MS obtido para íon de m/z 326,1944 $[M-H_2O+H]^+$.

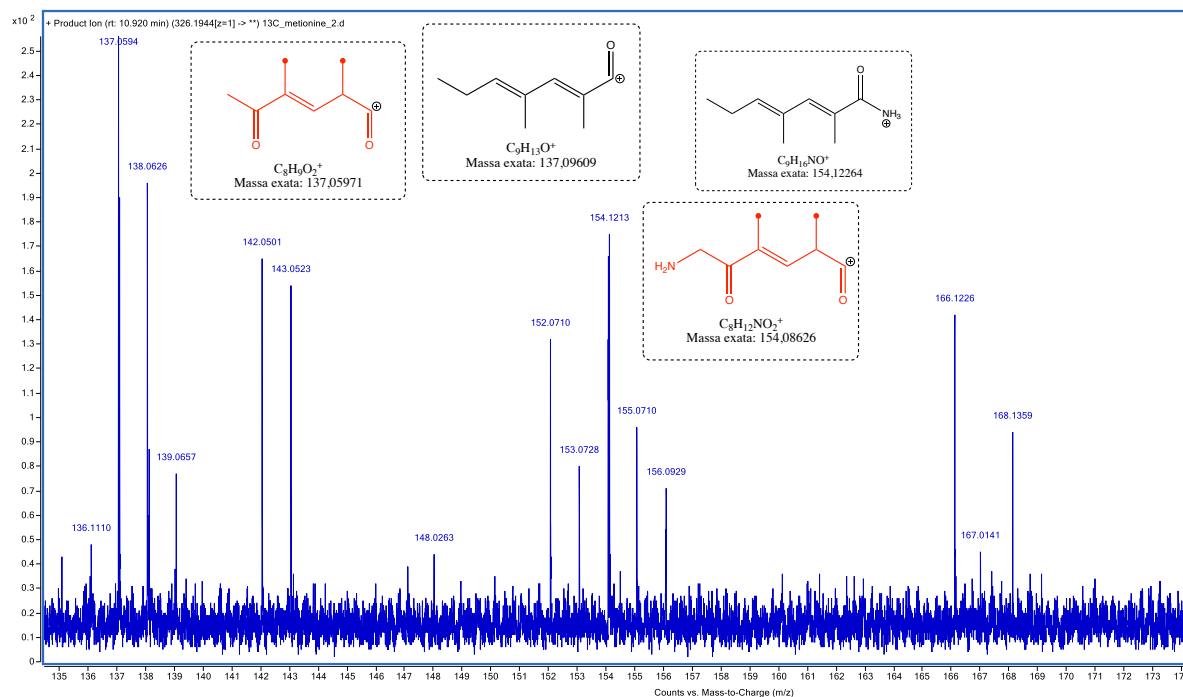


Figura 52. Espectro MS/MS obtido para íon de m/z 326,1944 $[M-H_2O+H]^+$ ampliado para destacar os íons de fragmentos de m/z 137 e m/z 154.

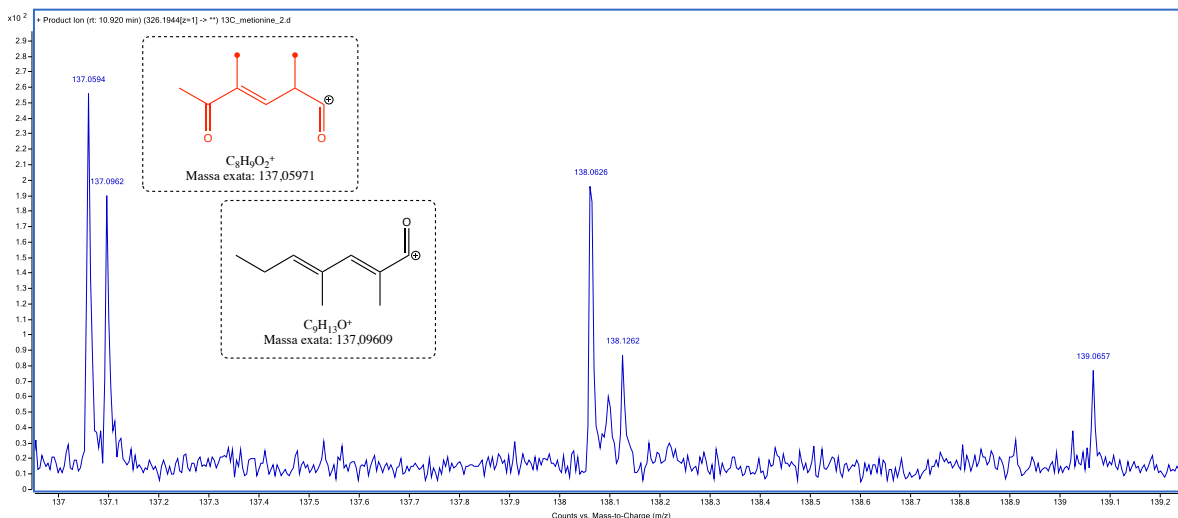


Figura 53. Espectro MS/MS obtido para íon de m/z 326,1944 $[M-H_2O+H]^+$ ampliado para destacar os íons de fragmentos de m/z 137,05971 e m/z 137,09609 e o padrão isotópico.

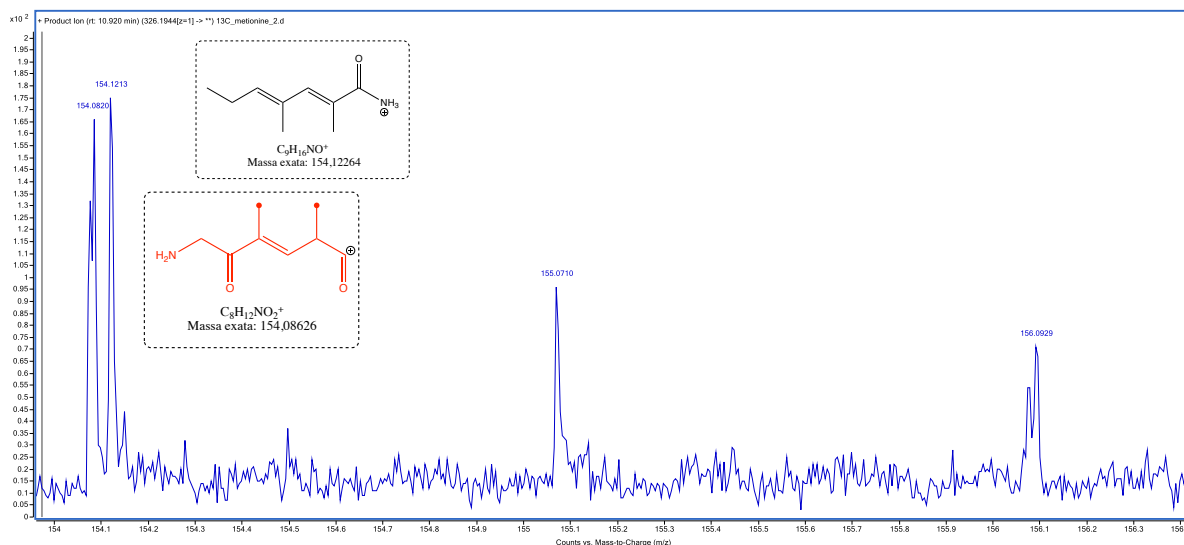


Figura 54. Espectro MS/MS obtido para íon de m/z 326,1944 $[\text{M}-\text{H}_2\text{O}+\text{H}]^+$ ampliado para destacar os íons de fragmentos de m/z 154,08626 e m/z 154,12264 e o padrão isotópico.

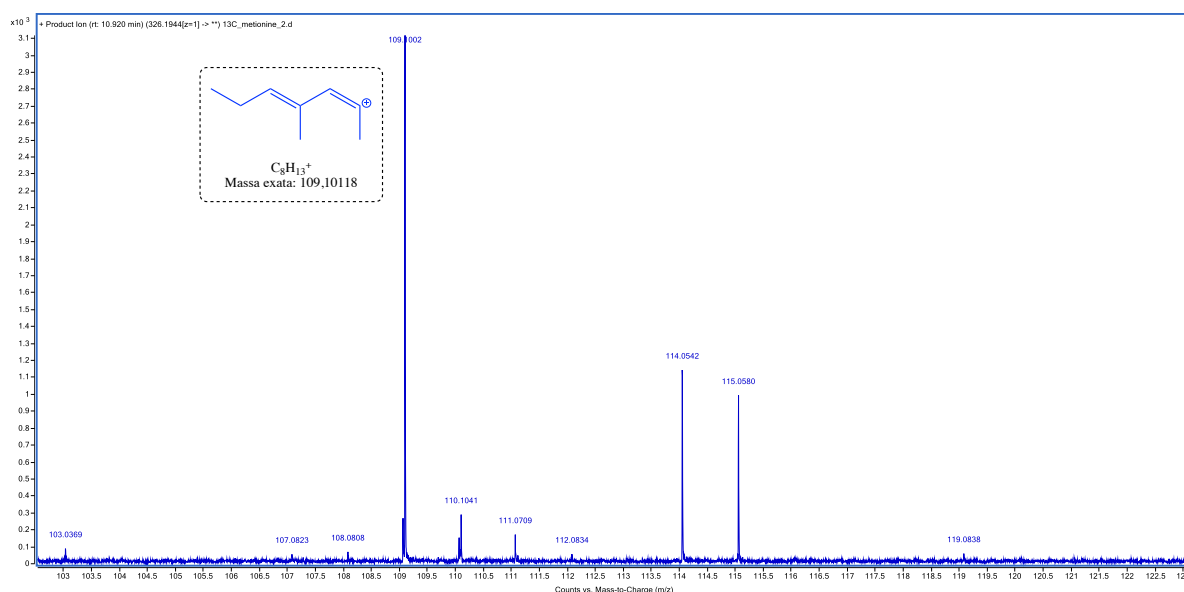


Figura 55. Espectro MS/MS obtido para íon de m/z 326,1944 $[\text{M}-\text{H}_2\text{O}+\text{H}]^+$ ampliado para destacar o íon de fragmentos de m/z 109,10118.

Sendo assim, os resultados de incorporação de precursores marcados apresentaram fortes evidências de interação da *trans*-AT, codificada pelo gene 115, com apenas dois módulos de biossíntese. Pela anotação do genoma são identificados dois domínios de docagem de *trans*-AT, um no gene 119 e outro no 118, sugerindo onde os resíduos de malonato poderiam ser incorporados, seguido de metilação pela ação de metiltransferase, o que pode ser confirmado pelo padrão do espectro MS/MS obtido. Deste modo, o domínio AT do módulo 0 (gene 117) catalisaria a incorporação

da unidade extensora de metilmalonato no módulo 1 A¹⁷⁷, o qual é reutilizado (módulo 1B ou módulo 2 – DH inativo), corroborando com a incorporação de uma terceira unidade de propionato marcado. Nesta etapa o domínio DH estaria inativo. O mesmo exemplo de “reciclagem” de módulos durante a biossíntese de produtos naturais pode ser observado para as lancacidinas.^{92,178}

Por fim, nós sugerimos que o módulo 5 é um módulo não-elongador (*non-elongating*), comum em *trans*-AT PKS¹⁷⁹, onde a KS⁰ não possui a atividade catalítica de condensação e promove apenas a transacilação entre os domínios ACP, devido a uma mutação no resíduo de His do motif conservado “HGTGT”.¹⁸⁰ O alinhamento das sequências de KS anotados do BGC *alp*-CBMAI2042 não evidenciaram a mutação recorrente, entretanto o mesmo padrão é proposto na biossíntese da gladiolina (Figura 56).¹⁸¹

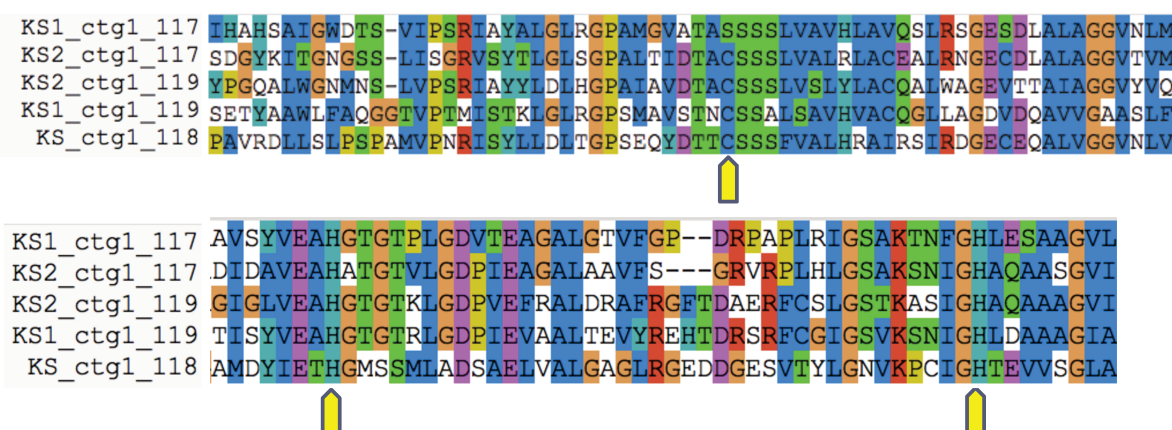


Figura 56. Alinhamento de sequências dos domínios de cetossintase (KS) anotados do genoma de *Streptomyces* sp. CBMAI 2042 (ClustalX). Setas amarelas indicam os resíduos conservados do sítio catalítico.

Com base nos resultados obtidos em experimentos com precursores marcados e nas análises dos genes, ficou evidente a interação da *trans*-AT na etapa de biossíntese pós-NRPS e a atuação de uma PKS do tipo I na formação da cadeia policetílica inicial e, por isso, a rota de biossíntese proposta como resultado deste estudo segue de acordo com aquela apresentada na Figura 57.

Portanto, a principal evidência apresentada para esta proposta de biossíntese que difere daquela relatada por Paulus et al. está relacionada com a interação da *trans*-AT com os diferentes módulos durante a biossíntese, sendo que propomos um híbrido de PKS tipo I e *trans*-AT atuando pré e pós-NRPS,

respectivamente, na formação da alpiniamida, enquanto os demais autores sugerem a incorporação de todos os resíduos derivados de malonil-CoA por ação interativa, o que não poderia ser explicado pela incorporação de unidades decorrentes de propanoato, visto a alta especificidade de domínio *trans*-AT por unidades não metiladas. Os experimentos com precursores marcados, bem como a anotação do BGC, que não apresenta domínio de docagem para reconhecimento e interação da *trans*-AT com os módulos do gene *alp1* (ou *ctg1_117*), corroboraram de forma mais consistente com a nova proposta apresentada. Adicionalmente, a interação do gene *alp2* para incorporação do terceiro resíduo pré-NRPS e, posteriormente, parte do gene atuando ao final da rota biossintética não tem nenhuma evidência experimental, parecendo mais coerente a sugestão do módulo 6 atuando como não-elongador, baseando-se em vias de biossíntese de outras *trans*-AT PKSs descritas recentemente na literatura. Esses resultados apresentados estão sendo elaborados para publicação.

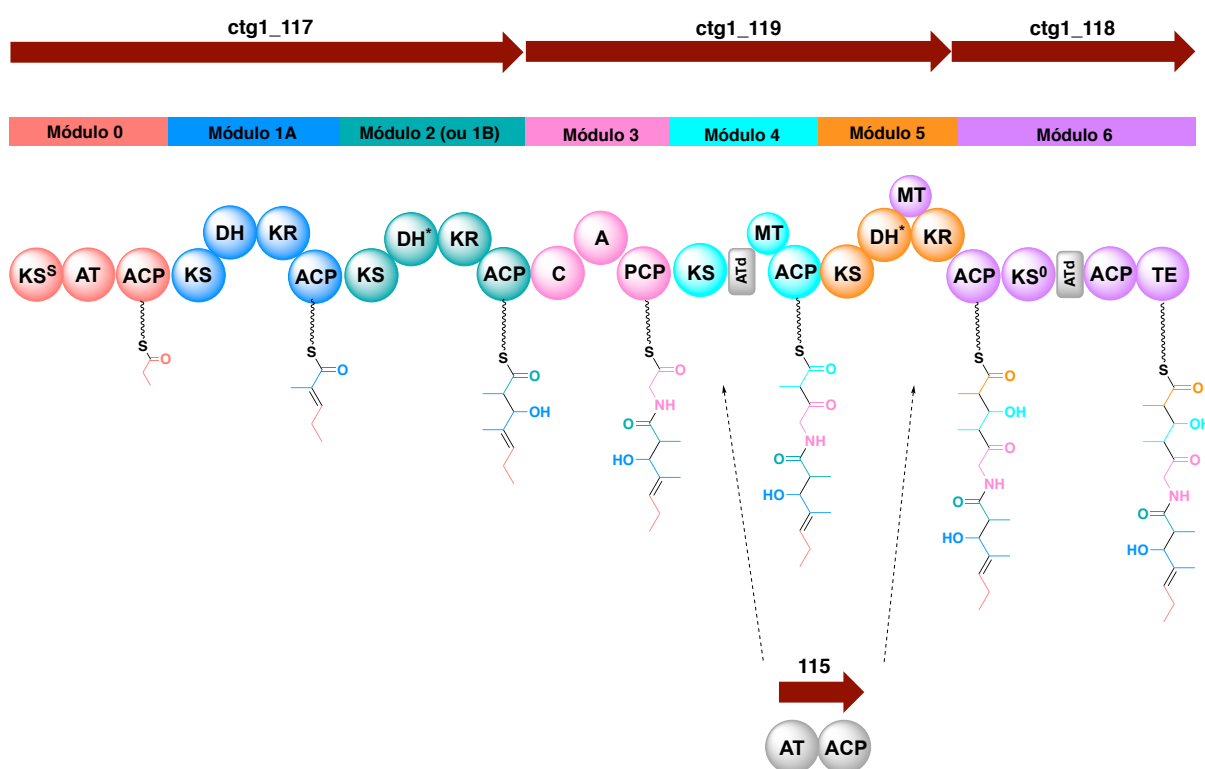


Figura 57. Proposta de biossíntese da alpiniamida baseada nos resultados experimentais obtidos pelo presente estudo.

4.5. Valinomicina

A valinomicina (**26**) (Figura 58) é um dodecadepsipeptídeo não ribossomal produzido por diferentes linhagens de *Streptomyces*¹⁶⁴ e um ionóforo com afinidade ao íon potássio (K⁺). Apresenta potencial atividade biológica contra coronavírus (CoV) associado a síndromes respiratórias agudas graves (SARS), demonstrado em células Vero E6 infectadas (EC₅₀ = 0.85 µM)¹⁸², além de atividade antibiótica¹⁸³, antitumoral¹⁸⁴, antiparasitária¹⁸⁵, antifúngica¹⁸⁶ e bioinseticida¹⁸⁷. Contudo, apesar do amplo espectro de atividades, o uso clínico de valinomicina é inviabilizado devido sua alta toxicidade, por isso a obtenção de análogos estruturais ou diferentes formas de administração tem sido motivo de estudo com o objetivo de minimizar os efeitos adversos.^{163,188,189}

A biossíntese ocorre via NRPS codificada pelos genes *vlm1* e *vlm2* (Figura 58A e B), os quais se subdividem em quatro módulos, sendo cada um responsável pela incorporação de uma unidade de ácido α-D-hidroxi-isovalérico (Hiv), D-valina (Val), L-lactato (Lac) e L-valina (Val). O ciclo de incorporação das unidades se repete três vezes, levando a formação de um dodecadepsipeptídeo cíclico constituído por ligações ésteres e peptídicas alternadas (Figura 58C).¹⁶³ Em 2014, o grupo do pesquisador Neubauer expressou os genes codificando a biossíntese de valinomicina em *E. coli* como organismo heterólogo (0,3 mg.L⁻¹), otimizaram a produção para 10 mg.L⁻¹ e ao coexpressar uma tioesterase (TEII), adjacente ao *cluster* biossintético na linhagem selvagem, obtiveram rendimento de 13 mg.L⁻¹.^{190–192} Posteriormente, otimizaram a produção em biorreator de 10 L com rendimento final de ~ 2 mg.L⁻¹.¹⁹³

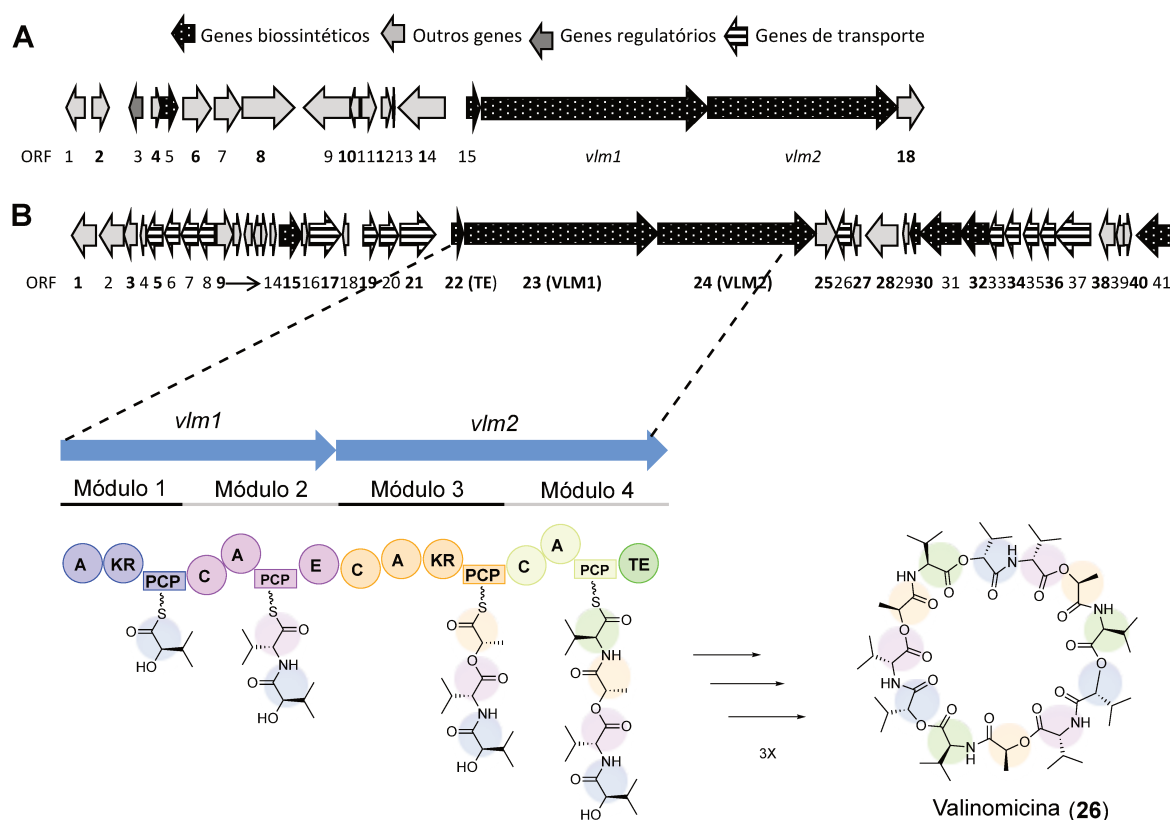


Figura 58. Valinomicina (26) – anotação de genes biossintéticos. **A.** Cluster de genes anotado do genoma de *Streptomyces tsusimaensis* ATCC 15141¹⁶³ codificando para biossíntese da valinomicina. **B.** BGC anotado do genoma de *Streptomyces* sp. CBMAI 2042. Os genes 23 e 24 codificam para VLM1 e VLM2, que se subdividem em 4 módulos, sendo os módulos compostos por domínios A, C, PCP, KR, E e TE. Cada módulo é responsável pela incorporação de uma unidade de Hiv (azul), Val (roxo), Lac (laranja) e Val (verde).

A desreplicação dos extratos de fermentação de *Streptomyces* sp. CBMAI 2042 em meio A e GPMY por *molecular networking*¹⁹⁴ identificou a valinomicina como um dos principais metabólitos produzidos pela linhagem. No espectro de massas de varredura obtido por infusão direta e ESI(+) pode-se observar um conjunto de sinais correspondentes aos adutos de valinomicina na forma potassíada m/z 1149 $[M+K]^+$, sodíada m/z 1133 $[M+Na]^+$ e amoniada m/z 1128 $[M+NH_4]^+$ e a abundância relativa dos adutos confirma a afinidade do ionóforo por cátions, principalmente potássio (K) (Figura 59A).¹⁹⁵ O padrão de fragmentação obtido é caracterizado pelos íons majoritários referentes a clivagem das ligações peptídicas e evidencia perdas de massa de 199 Da, correspondente ao dímero formado por Hiv e Val, e 171 Da correspondente ao dímero formado por Lac e Val, corroborando com a estrutura da valinomicina (Figura 59B).

A

7_151118142049 #253 RT: 0.71 / IL: 1.34E8
T: FTMS + p ESI Full ms [500.00-2000.00]

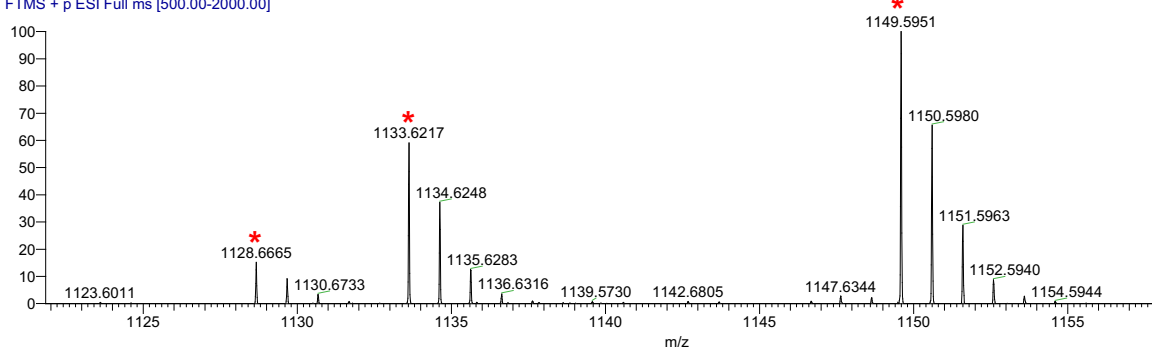
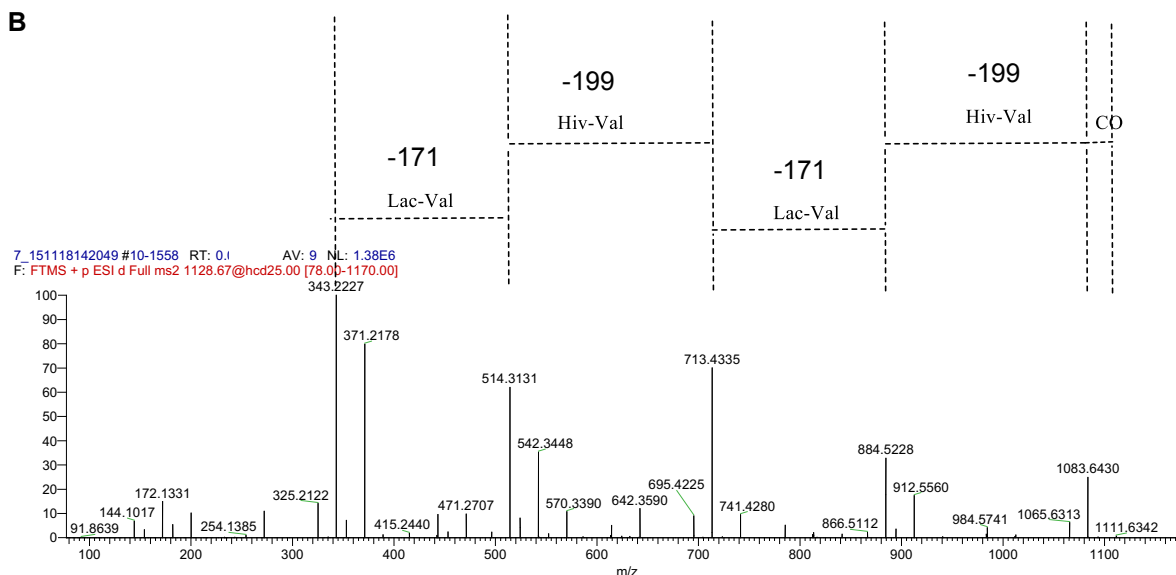
**B**

Figura 59. Identificação de valinomicina por espectrometria de massas. **A.** ESI(+)-FTMS obtido por infusão direta do extrato bruto do cultivo de *Streptomyces* sp. CBMAI 2042 em meio A. **B.** ESI(+)-MS/MS do íon m/z 1128 $[M+NH_4]^+$ correspondente ao padrão de fragmentação da valinomicina.

A análise detalhada das anotações do sequenciamento do genoma da linhagem CBMAI 2042 via antiSMASH, permitiu identificar o *cluster* de genes codificando para a biossíntese de valinomicina (Figura 60) sendo que 22% dos genes totais no *cluster* apresentaram similaridade com o BGC depositado em bancos de dados referente a valinomicina¹⁶³ e os genes essenciais para biossíntese apresentaram 85% de identidade. Optou-se, então, pela triagem da biblioteca genômica ESAC para obter o clone correspondente ao BGC da valinomicina em CBMAI 2042, de modo a implementar metodologias que viabilizassem expressão

heteróloga de genes codificando para metabolismo secundário no grupo de pesquisa e, posteriormente, realizar experimentos na tentativa de obter análogos estruturais.

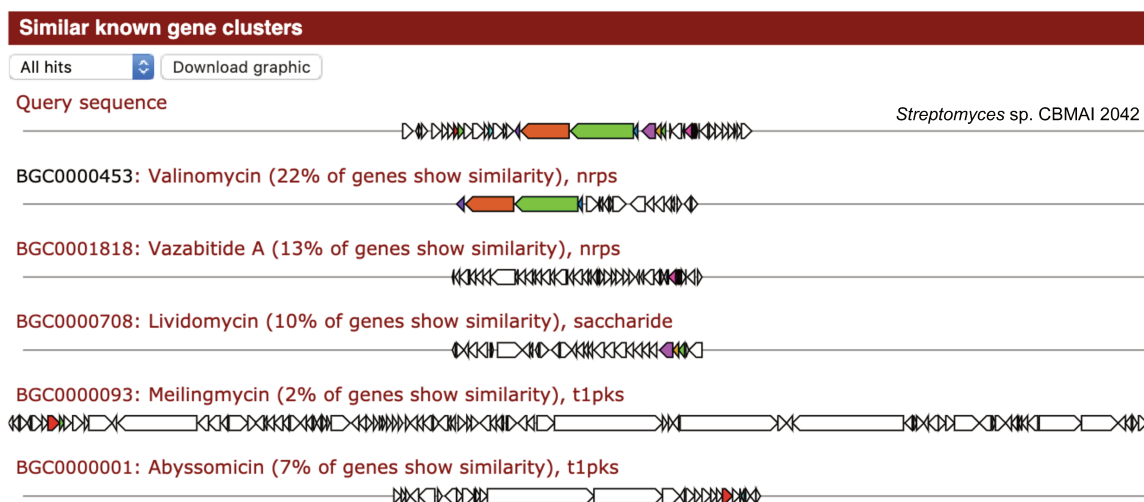
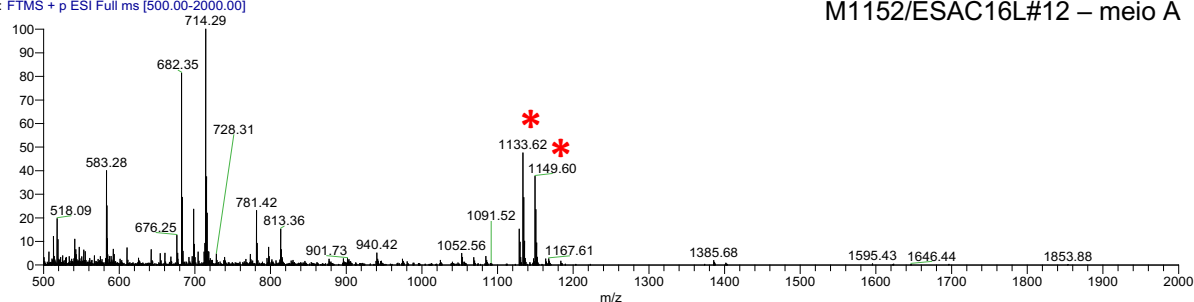


Figura 60. Anotação do *cluster* de genes relacionado a biossíntese de valinomicina em *Streptomyces* sp. CBMAI 2042. O *cluster* apresenta 22% de similaridade com o BGC codificando para a biossíntese de valinomicina em *Streptomyces tsusimaensis* ATCC 15141 (BGC0000453). Anotação obtida pela plataforma antiSMASH (vs 5.0 beta).

Os exconjugantes M1146/ESAC16L, M1152/ ESAC16L, M1146/ESAC4H e M1152/ ESAC4H obtidos e confirmados por PCR (Figura 22B), foram cultivados em meio A e ISP2. As análises dos extratos foram feitas por infusão direta em modo positivo e, a partir dos espectros de massas obtidos, pode-se observar a presença dos íons correspondentes aos adutos sodiado e potassiado da valinomicina (m/z 1133 e m/z 1149) em todas as amostras, exceto para o clone M1152/ESAC4H#16 em meio A (Figura 61 e Figura 62). O organismo heterólogo *S. coelicolor* M1146 demonstrou-se como a melhor plataforma para expressão deste metabólito, notadamente aquele contendo o clone ESAC4H#5 cultivado em meio A, já que o espectro de massas apresentou maior intensidade relativa de íons.

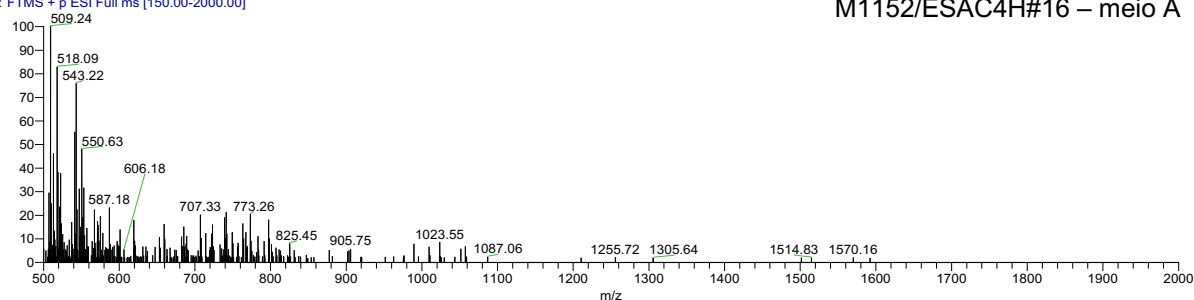
7_151118142601 #1 RT: 0.00 AV 4.01E7
T: FTMS + p ESI Full ms [500.00-2000.00]

M1152/ESAC16L#12 – meio A



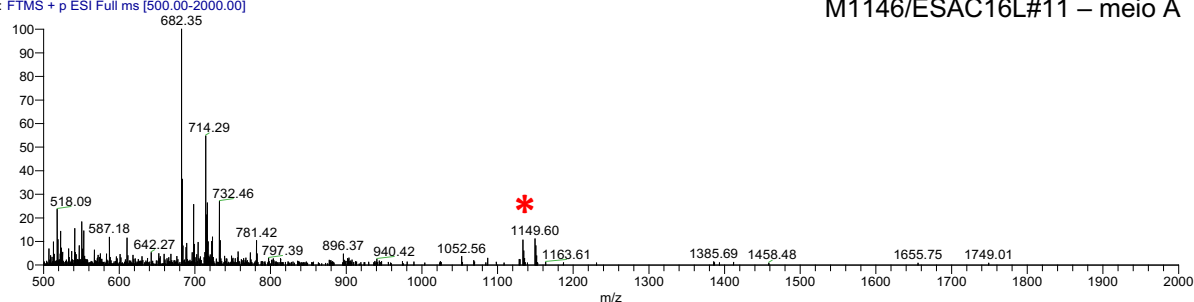
8 #4 RT: 0.05 AV: 1 NL: 6.24E6
T: FTMS + p ESI Full ms [150.00-2000.00]

M1152/ESAC4H#16 – meio A



11_151118150922 #871 RT: 2.44 NL: 8.08E6
T: FTMS + p ESI Full ms [500.00-2000.00]

M1146/ESAC16L#11 – meio A



10_151118145401 #967 RT: 2.71 NL: 9.71E7
T: FTMS + p ESI Full ms [500.00-2000.00]

M1146/ESAC4H#5 – meio A

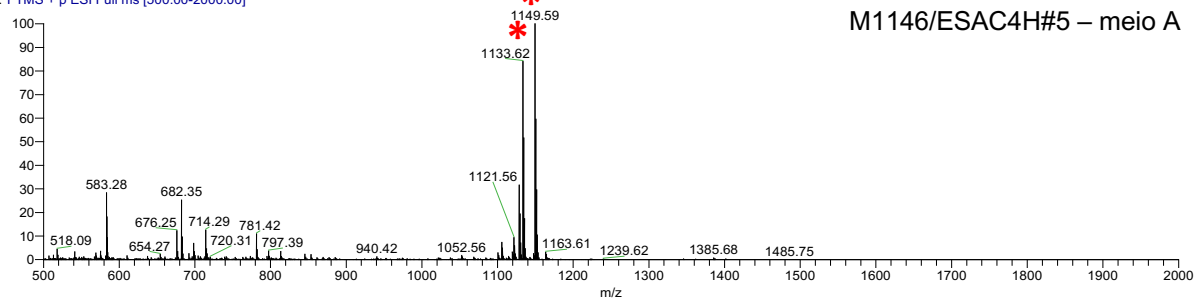
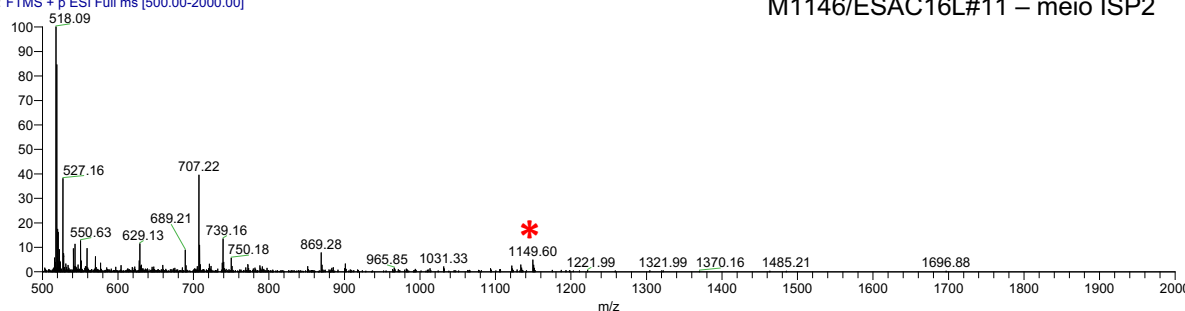


Figura 61. ESI(+)-FTMS obtidos por infusão direta dos extratos de cultivos em meio A dos exconjugantes contendo o *cluster* de genes codificando para a biossíntese da valinomicina.

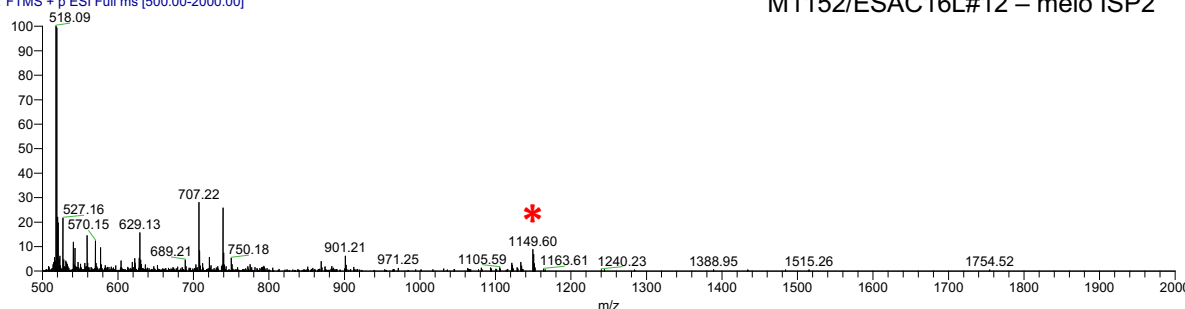
3 #685 RT: 1.92 AV: 1 NL: 9.27E
T: FTMS + p ESI Full ms [500.00-2000.00]

M1146/ESAC16L#11 – meio ISP2



4_151118133950 #817 RT: 2.28 / IL: 1.47E7
T: FTMS + p ESI Full ms [500.00-2000.00]

M1152/ESAC16L#12 – meio ISP2



5_151118135103 #1141 RT: 3.18 NL: 2.21E7
T: FTMS + p ESI Full ms [500.00-2000.00]

M1152/ESAC4H#16 – meio ISP2

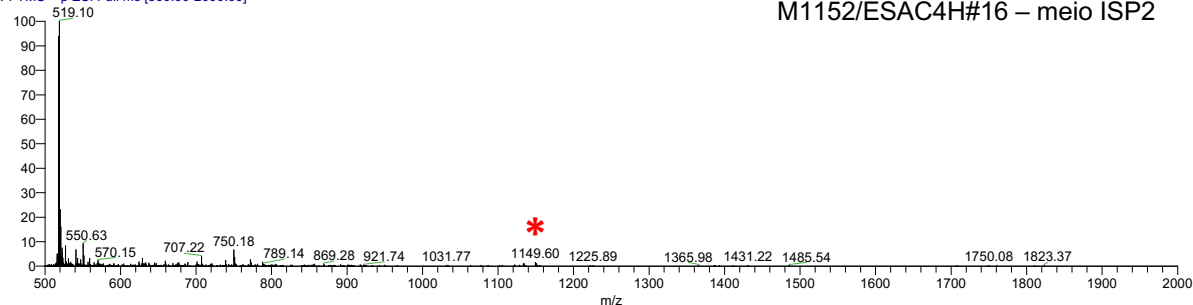


Figura 62. ESI(+)-FTMS obtidos por infusão direta dos extratos de cultivos em meio ISP2 dos exconjugantes contendo o *cluster* de genes codificando para a biossíntese da valinomicina.

4.5.1. Quantificação de valinomicina

A produção de valinomicina pela linhagem heteróloga M1146/ESAC4H#5 e pela linhagem selvagem *Streptomyces* sp. CBMAI 2042 foi quantificada a partir de cultivos de 7 dias em meio de cultura A. A curva linear de calibração foi obtida a partir de soluções padrão de valinomicina comercial (V0627 Sigma-Aldrich®) na faixa de 0,05 a 4,00 $\mu\text{g} \cdot \text{mL}^{-1}$ com coeficiente de regressão linear equivalente a 0,99567489 e equação da reta representada por $y = 99360790,837138x - 1328935,084794$ (Gráfico 1). A quantificação foi feita em quintuplicata por diluição de 10^4 dos extratos obtidos com acetato de etila e por interpolação dos dados na equação de regressão linear resultante.

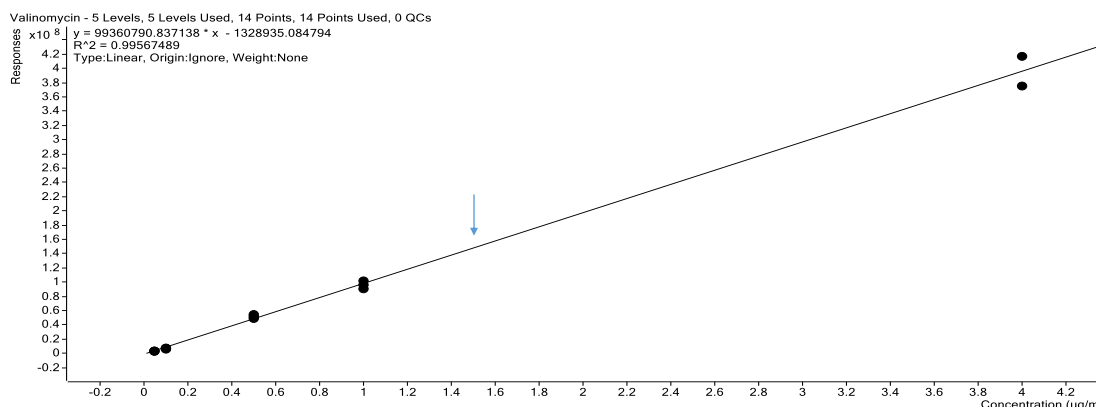


Gráfico 1. Curva de calibração obtida com padrão comercial para quantificação de valinomicina. A seta em azul representa o valor de concentração médio correspondente a produção pela linhagem recombinante M1146/ESAC4H#5.

Os experimentos apontaram produção média de 0,49 mg.L⁻¹ de valinomicina pela linhagem selvagem (*Streptomyces* sp. CBMAI 2042) e 1,47 mg.L⁻¹ pela linhagem recombinante M1146/ESAC4H#5, representando um incremento de três vezes na produção do metabólito em organismo heterólogo (Tabela 18).

Tabela 18. Quantificação da produção de valinomicina correspondente ao cultivo de 7 dias em meio A.

Amostra	Concentração (mg.L ⁻¹)	Média (mg.L ⁻¹)	Desvio padrão
M1146/ESAC4H#5 DSM 32578	1,838	1,47	0,24
	1,425		
	1,155		
<i>Streptomyces</i> sp. CBMAI 2042	0,372	0,49	0,17
	0,746		
	0,357		

A linhagem M1146/ESAC4H#5 foi depositada no banco internacional de microrganismos, Instituto Leibniz DSMZ (Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures GmbH), sob o registro DSM 32578. Adicionalmente, foi enviado um pedido de registro de patente correspondente a produção de valinomicina em organismo heterólogo *S. coelicolor* M1146 (de Oliveira, L. G.; **Sigrist, R.** Método de obtenção de microrganismo recombinante, microrganismo recombinante, método de obtenção da valinomicina, valinomicina assim obtida e uso da valinomicina. **2017**, Brasil. *Pedido de Patente*: Privilégio de Inovação. Número do registro: BR102017080239. Instituição de registro: INPI – Instituto Nacional da Propriedade Industrial. Depósito: 23/08/2017).

4.5.2. Análogos estruturais da valinomicina

O perfil cromatográfico comparativo obtido para as linhagens *S. coelicolor* M1146 sem a informação genética correspondente a biossíntese de valinomicina e M1146/ESAC4H#5 (DSM 32578) evidenciou a produção majoritária de valinomicina (3 – m/z 1128 $[M+NH_4]^+$) no heterólogo e indicou a produção de três outras moléculas, 1 – m/z 1100 $[M+NH_4]^+$, 2 – m/z 1114 $[M+NH_4]^+$ e 4 – m/z 1142 $[M+NH_4]^+$, as quais não foram observadas em *S. coelicolor* M1146, podendo estar associados a mesma via de biossíntese e a prováveis análogos estruturais da valinomicina (Figura 63).

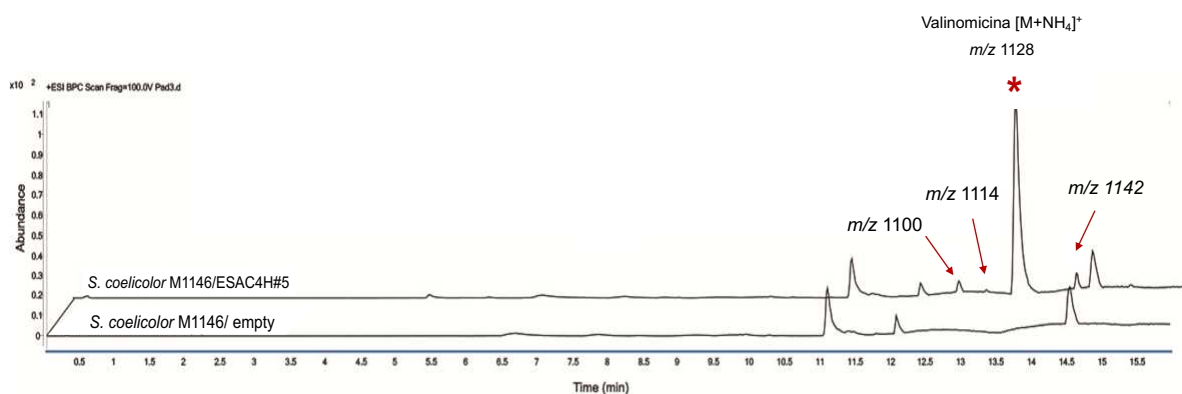


Figura 63. Cromatogramas de bico base obtidos dos extratos de cultivo de *S. coelicolor* M1146 (abaixo) e M1146/ESAC4H#5 (acima). O pico mais intenso (m/z 1128) corresponde ao aduto amoniado de valinomicina $[M+NH_4]^+$ e os picos correspondentes a m/z 1100, m/z 1114 e m/z 1142 estão relacionados a prováveis análogos estruturais da valinomicina.

Estudos de *molecular networking* realizados em projeto de doutorado do aluno MSc. Bruno S. Paulo para desreplacção do metabolismo de *Streptomyces* sp. CBMAI 2042, correlacionaram esses três íons identificados ao precursor da valinomicina, além de três outros íons. A rede molecular, traduzido do inglês *molecular networking*, utiliza-se da informação obtida por fragmentação MS/MS de alta resolução para cruzar com informações de padrões de fragmentação de moléculas já conhecidas contidas em bancos de dados globais, tais como o GNPS (*Global Natural Products Social Molecular Networking*)¹⁹⁶, e cria uma rede de correlação ao agrupar entidades químicas que compartilham de um padrão de fragmentação similar. Deste modo, a identificação de novos compostos ou análogos de compostos já conhecidos produzidos pelo mesmo *cluster* de genes é facilitada.^{194,197}

Sendo assim, com os resultados obtidos por *molecular networking* foram correlacionados ao íon precursor de m/z 1128.6668 $[M+NH_4]^+$ (**26**), os íons correspondentes a montanastatina (**27**) (m/z 741,4286 $[M+H]^+$), já descrito na literatura¹⁸³, e seu análogo linear (**28**) (m/z 759,4386 $[M+H]^+$), além dos íons de m/z 1100,6418 $[M+NH_4]^+$ (**29**), m/z 1114,6514 $[M+NH_4]^+$ (**30**), m/z 1142,6845 $[M+NH_4]^+$ (**32**) e m/z 1156,6985 $[M+NH_4]^+$ (**32**), correspondentes a análogos da valinomicina (Figura 64).

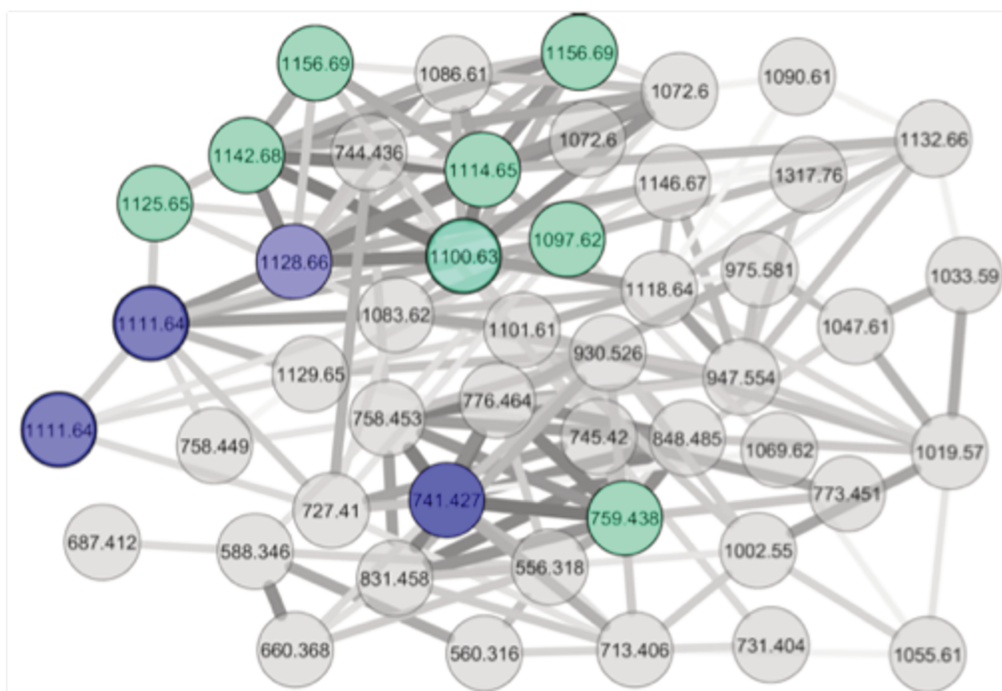


Figura 64. *Molecular networking* obtido a partir de espectros de fragmentação de extratos de cultivos de *Streptomyces* sp. CBMAI 2042 em meio ISP2. Azul: íons precursores da valinomicina e montanastatina. Verde: íons correspondentes a análogos estruturais.

A montanastatina (**27**) é um octadepsipeptídeo isolado de *Streptomyces anulatus*,¹⁸³ análogo estrutural da valinomicina, e, embora já tenha sido relatado na literatura, nunca teve sua via de biossíntese descrita ou associada diretamente ao mesmo *cluster* de genes da via de biossíntese de valinomicina. Portanto, para confirmar a origem biossintética deste e dos demais metabólitos identificados, uma nova análise comparativa de extratos de cultivos da linhagem selvagem CBMAI 2042, M1146/ESAC4H#5 e *S. coelicolor* M1146 foi realizada e observou-se a produção de todos os íons correspondentes na linhagem heteróloga (Figura 65). Sendo assim, os experimentos de *molecular networking*, associados a clonagem e expressão do BGC

em heterólogo foram essenciais para conectar todas essas moléculas a mesma via de biossíntese.

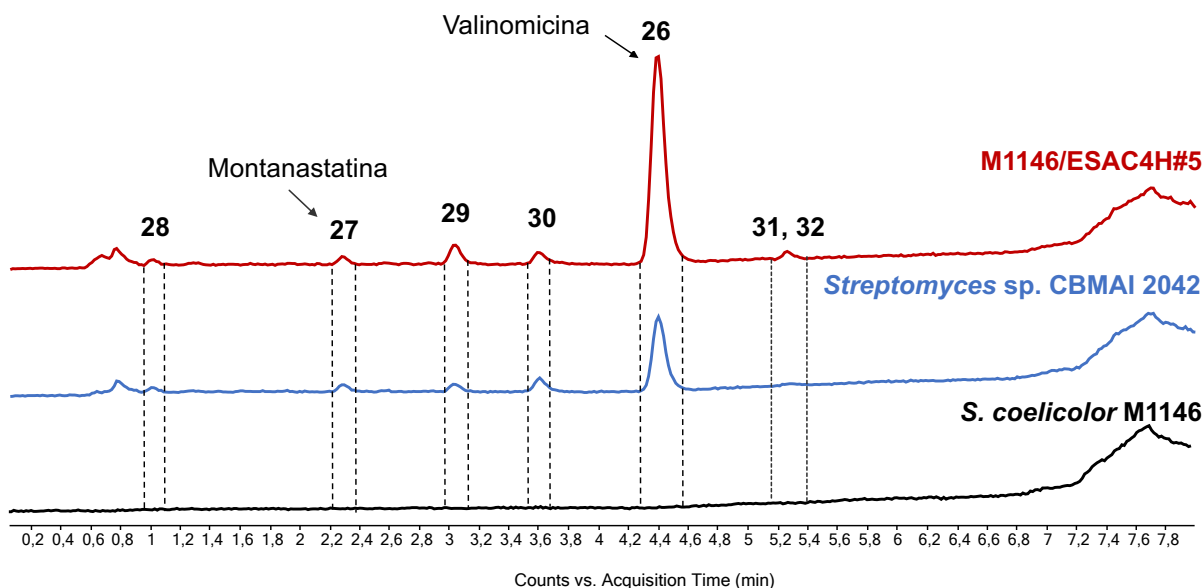


Figura 65. Cromatogramas de íon extraído correspondente a m/z 1128 (valinomicina – 26). Vermelho: linhagem heteróloga M1146/ESAC4H#5. Azul: linhagem selvagem *Streptomyces* sp. CBMAI 2042. Preto: *S. coelicolor* M1146. 27 – montanastatina (m/z 741 $[M+H]^+$), 28 – m/z 781 $[M+Na]^+$, 29 – m/z 1100 $[M+NH_4]^+$, 30 – m/z 1114 $[M+NH_4]^+$, 31 – m/z 1142 $[M+NH_4]^+$ e 32 – m/z 1156 $[M+NH_4]^+$.

Em um trabalho de colaboração, MSc. Bruno S. Paulo propôs as estruturas dos novos compostos (Figura 67) a partir do padrão de fragmentação obtido por espectrometria de massas de alta resolução de cada íon, corroborando com nossas propostas de biossíntese baseadas nas anotações do BGC. Assim como a valinomicina, os análogos também possuem um padrão de fragmentação característico, sendo que as alterações nas estruturas puderam ser identificadas justamente nas diferenças de massas correspondentes a troca de unidades de Val (99 Da) por Leu ou Ile (113 Da) e de Hiv (100 Da) por Lac (72 Da), quando comparadas a valinomicina.

De modo a ilustrar a metodologia utilizada, o íon m/z 1142 (análogo 31) será tomado como exemplo e comparado com a valinomicina. Na Figura 66A é representado o espectro de fragmentação da valinomicina e na Figura 66B é representado o espectro de fragmentação para o dodecadepsipeptídeo 31, onde podem ser observados diferentes íons de fragmentos oriundos de duplas clivagens em diferentes posições na estrutura cíclica dos dodecapeptídeos. As sequências de

íons fragmentos representadas pelas letras na parte superior do espectro representam as clivagens possíveis nas estruturas, sendo que para o análogo **31**, por não ser uma molécula simétrica como a valinomicina, podemos observar mais de uma possibilidade de clivagem e formação de diferentes íons de fragmentos. E, principalmente, observa-se no espectro MS/MS B uma alteração de massa característica de 113 Da (E) entre resíduos de Lac (C) e Hiv (A), evidenciando, por exemplo, a incorporação dos aminoácidos Ile ou Leu entre os resíduos ácidos e não de Val. As demais possibilidades de fragmentação para essa estrutura, bem como espectros MS/MS dos demais análogos encontram-se no Apêndice 8 (8.6 a 8.10).

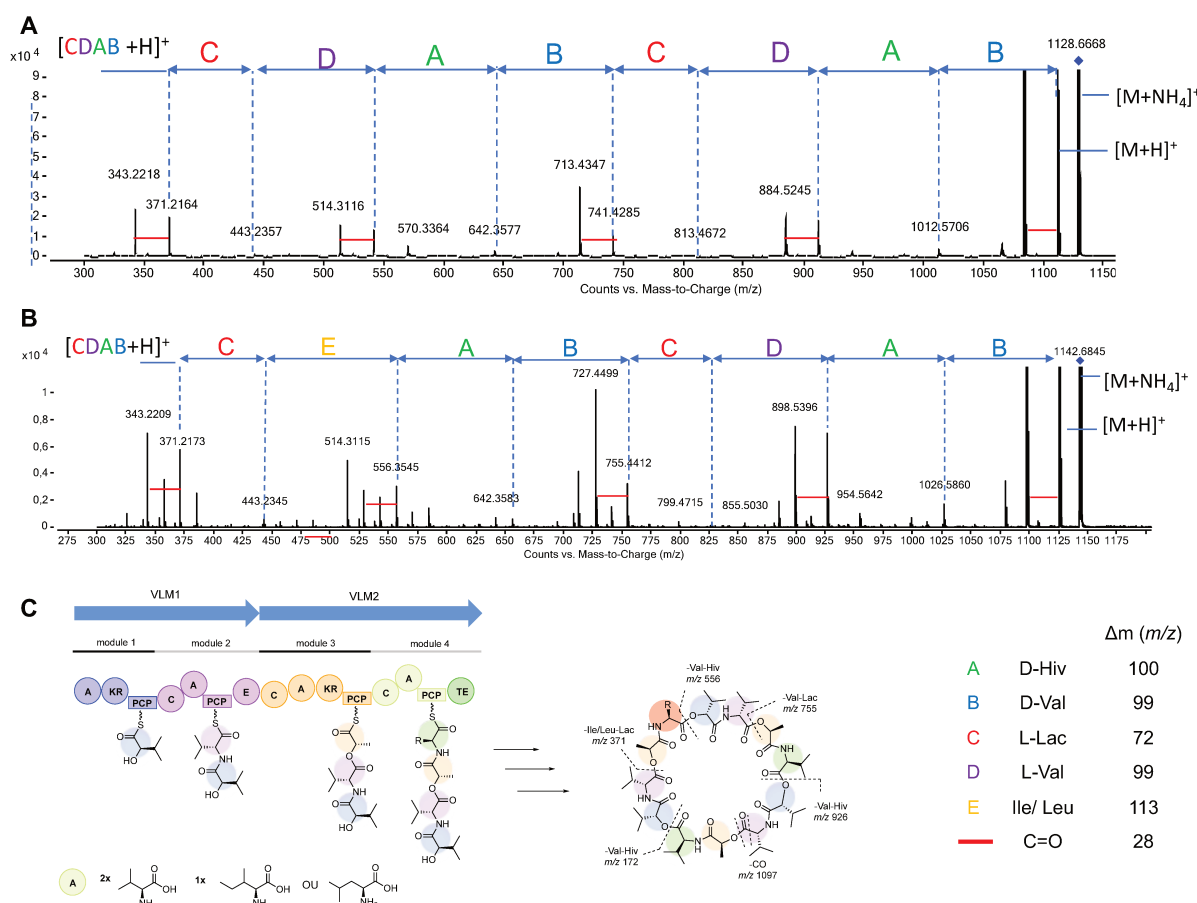


Figura 66. Identificação de padrões em espectros MS/MS da valinomicina e análogo **6**. **A.** Espectro MS/MS e atribuições de íons de fragmentação para proposta de estrutura valinomicina. Propostas de fragmentação da valinomicina. **B.** Espectro MS/MS e atribuições de íons de fragmentação para proposta de estrutura do análogo **31** (m/z 1142,6845). **C.** Proposta de biossíntese do análogo **31**.

A Figura 66C representa o cluster de genes biossintéticos e proposta de biossíntese, sendo anotada a promiscuidade de incorporação de resíduos de Val ou

Ile/Leu pelo domínio de adenilação (A) do módulo 4. A estrutura apresentada possui linhas tracejadas indicando os fragmentos correspondentes aos íons apontados no espectro de massas, sendo que para todas as estruturas optou-se por adotar o início da clivagem sempre na mesma posição para pelo menos uma proposta de fragmentação e facilitar a elucidação estrutural.

A via de biossíntese da valinomicina segue a ordem de módulos (Mod1-Mod2-Mod3-Mod4) e três ciclos de incorporação, já a montanastatina e o seu análogo linear correspondem a mesma ordem de módulos, mas apenas dois ciclos de incorporação. Os análogos **29**, **30**, **31** e **32** apresentam alteração nas unidades incorporadas e na ordem de interação dos módulos. O análogo **29** apresenta três ciclos de incorporação, mas a ordem dos módulos é alterada e proposta como segue: (Mod1-Mod2-Mod3-Mod4) x 2 e (Mod3-Mod4-Mod3-Mod4) x 1. Ao fim, obtém-se uma estrutura análoga a valinomicina com a substituição de um resíduo de Hiv por Lac, bem como a incorporação de um 4º resíduo de L-Val. O análogo **30** apresenta a mesma ordem de incorporação apresentada para o composto **29**, porém em um ciclo é proposta incorporação de Ile ou Leu pelo o domínio de adenilação do módulo 4. Essa promiscuidade na aceitação de outros substratos é predita pela análise *in silico* no antiSMASH. Os compostos **31** e **32** apresentam a mesma ordem na sequência de módulos da valinomicina, com a diferença que apresentam em suas estruturas uma ou duas Ile/Leu, respectivamente, ao invés de Val.

O estudo completo foi aceito para publicação e destacado como artigo de capa na revista *Chemistry Select* (Paulo, B. S.; **Sigrist, R.**; Angolini, C. F. F.; de Oliveira, L. G. *New Cyclodepsipeptide Derivatives Revealed by Genome Mining and Molecular Networking*. *ChemistrySelect*. **2019**. DOI: 10.1002/slct.201902252) e reforça a importância da aplicação de abordagens genômicas associadas ao *molecular networking* para descoberta de análogos estruturais correlacionados a mesma via de biossíntese e produzidos, muitas vezes, em menores quantidades que podem não ser evidentes em cultivos de linhagens selvagens, devido a complexidade de metabólitos produzidos, ou mesmo de organismos heterólogos.

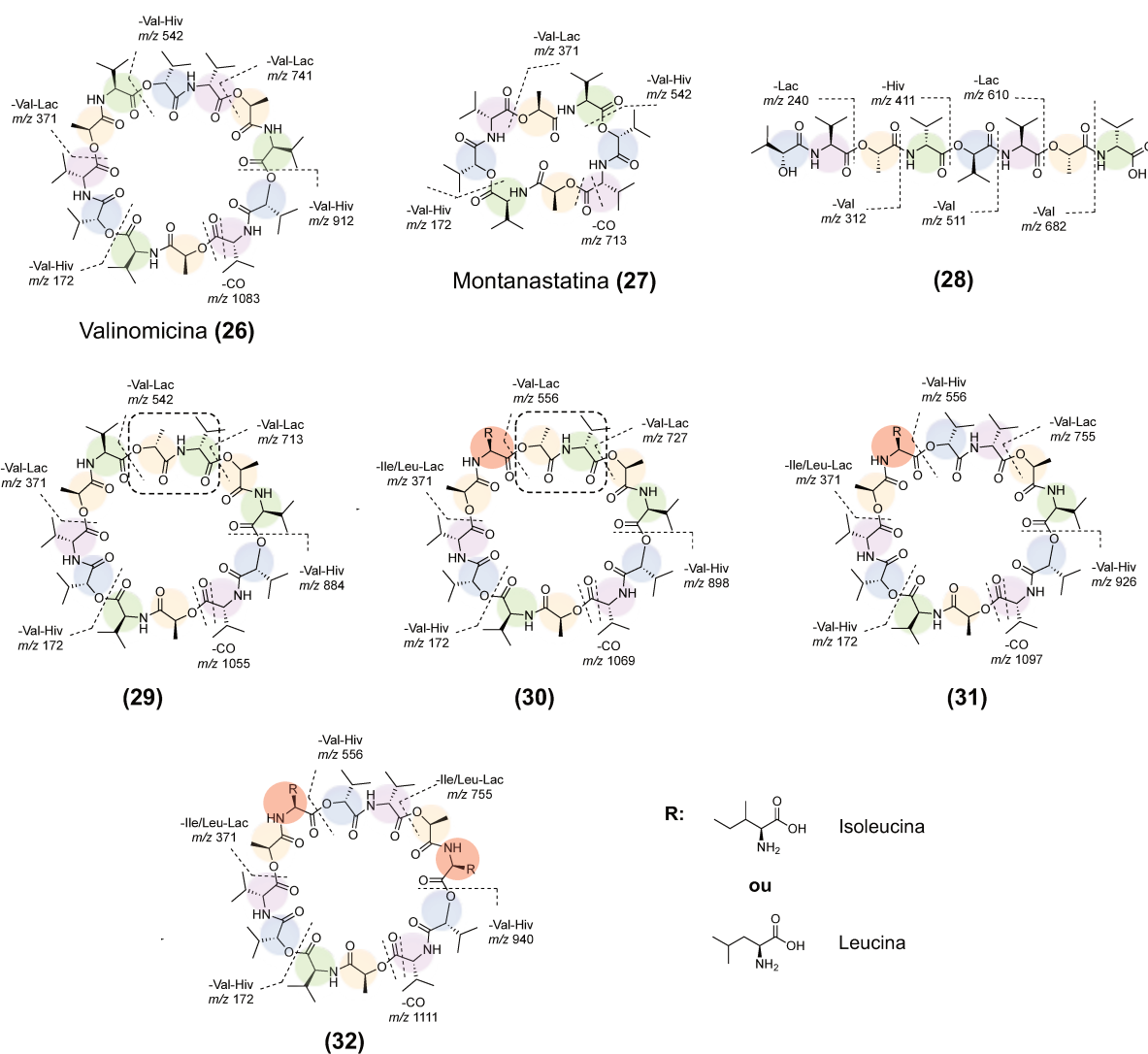


Figura 67. Estrutura da valinomicina (26), montanastatina (27) e análogos estruturais 28, 29, 30, 31 e 32. As caixas pontilhadas e resíduos em vermelho destacam as modificações nas estruturas quando comparado a valinomicina.

5. CONCLUSÕES

Uma biblioteca genômica de *Streptomyces* sp. CBMAI 2042 foi construída em vetor BAC, porém não foi suficientemente representativa e a triagem por clones contendo o *cluster* completo codificando para a t1PKS – *trans*-AT PKS – NRPS não foi bem sucedida. Contudo, os experimentos foram extremamente relevantes para compreender o processo e as limitações nas construções de bibliotecas genômicas e manipulação de grandes fragmentos de DNA. Por outro lado, a biblioteca em vetor pESAC13A obtida pela empresa Bio S&T Inc. foi triada com sucesso e os clones ESACs obtidos identificados como pl9-5H e pl11-16C, correspondentes ao *cluster* órfão t1PKS – *trans*-AT PKS – NRPS, e pl11-4H e pl13-16L, correspondentes ao *cluster* da valinomicina. Os experimentos de conjugação e transferência da informação genética para as linhagens heterólogas de *S. coelicolor* foram conduzidos e a produção de valinomicina foi reconstituída em *S. coelicolor* M1146 e M1152, sendo que a linhagem M1146 se apresentou como uma melhor plataforma para expressão deste BGC. A linhagem heteróloga M1146/ESAC4H#5 obtida foi depositada no Instituto Leibniz DSMZ (*Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures GmbH*), sob o registro DSM 3257, e a produção de valinomicina foi quantificada, correspondendo ao incremento de 3 vezes (1,5 mg.L⁻¹) no rendimento quando comparado com a linhagem selvagem *Streptomyces* sp. CBMAI 2042. Um pedido de patente foi solicitado para esta linhagem produtora de valinomicina. De modo complementar, seis análogos estruturais produzidos pela mesma via de biossíntese da valinomicina foram identificados a partir da associação das metodologias de *molecular networking* e expressão em organismo heterólogo.

As metodologias de recombinação de *Gibson Assembly*® e λ -RED não se mostraram metodologias de clonagem apropriadas para a clonagem do BGC órfão, talvez por fatores relacionados ao tamanho dos fragmentos e quantidade dos fragmentos a serem recombinados, bem como a metodologia utilizada, no caso do λ -RED. A triagem do metabólito desconhecido a partir de cultivos do heterólogo obtido pela conjugação com clones da biblioteca genômica não possibilitou a identificação do composto. No entanto, após clonagem via *TAR-cloning*, expressão em *S. coelicolor* M1146 e extração de cultivos com resina não-iônica o metabólito correspondente ao *cluster* órfão foi identificado. Após caracterização por técnicas de RMN e espectrometria de massas, uma estrutura foi proposta e a busca em bases de dados,

identificou e correlacionou o composto com a alpiniamida, isolada em 2013. Experimentos de deleção do domínio de adenilação foram de extrema importância para confirmar a correlação da molécula com a via biossintética. A alpiniamida, apesar de já isolada, não tinha os genes associados e sua via de biossíntese descritos. Portanto, iniciou-se uma investigação para estabelecer uma proposta de biossíntese visto que os genes anotados não correspondem diretamente com o metabólito obtido. Durante o desenvolvimento desses estudos um outro grupo de pesquisas publicou em agosto de 2018 um trabalho relatando a descoberta da mesma molécula, atribuindo os genes biossintéticos e uma proposta para a biossíntese do composto. Contudo, devido algumas inconsistência experimentais, e baseado nos resultados que havíamos obtido, alcançamos fortes evidências para propor uma nova rota biossintética para obtenção de alpiniamida.

6. ESTUDO DA BIOSÍNTESE DE 4-CLORO-QUINURENINA MEDIADA POR TAR13 E TAR16

Durante o período de Estágio de Pesquisa no Exterior, no laboratório de pesquisas do Prof. Bradley S. Moore (UCSD, San Diego, CA), foi desenvolvido um projeto de pesquisa adicional sob supervisão da Dra. Hanna Luhavaya e do Prof. Moore. Nesta sessão serão apresentados os resultados obtidos durante o estudo envolvendo duas enzimas anotadas no *cluster* de genes codificando para a biossíntese de taromicinas A e B (Figura 68 A), os quais são complementares ao trabalho completo publicado: Luhavaya, H., Sigrist, R., Chekan, J. R., McKinnie, S. M. K., Moore, B. S. *Biosynthesis of 1-4-Chlorokynurenine, an Antidepressant Prodrug and a Non-Proteinogenic Amino Acid Found in Lipopeptide Antibiotics. Angew. Chem., Int. Ed.* **2019**, 58, 8394 – 8399. DOI: 10.1002/anie.201901571.

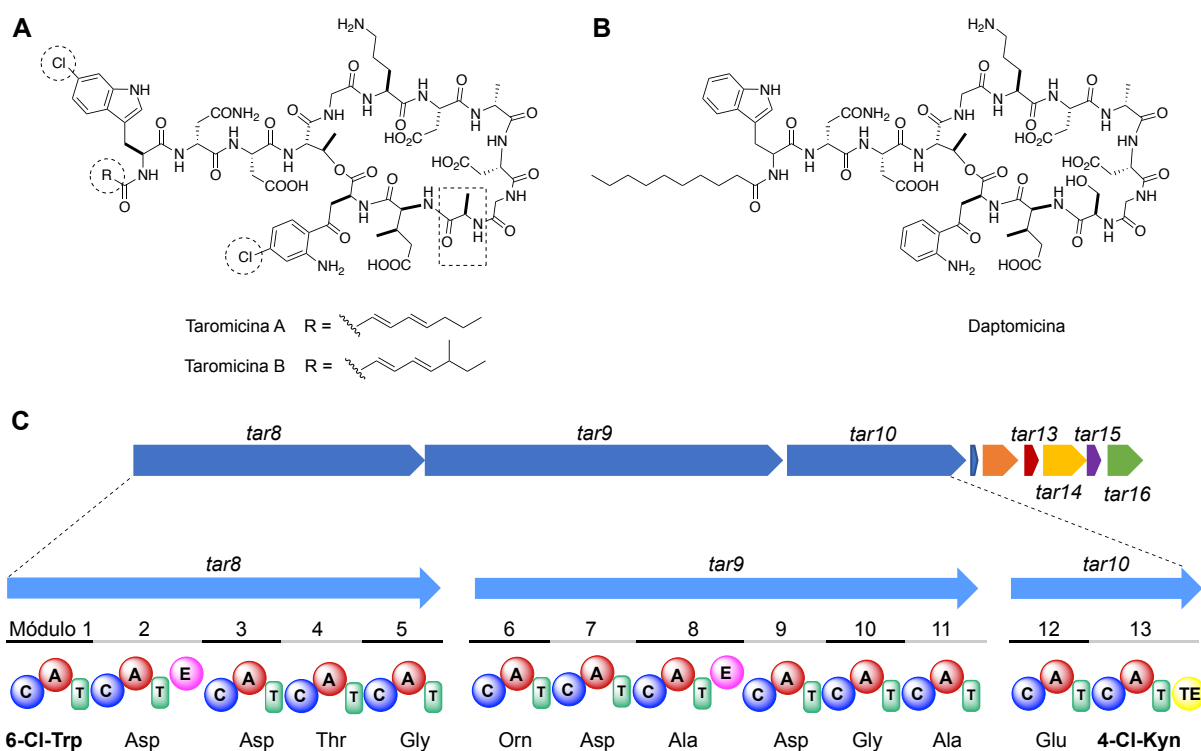


Figura 68. Estruturas da **A.** Taromicina A e B; **B.** Daptomicina. As caixas pontilhadas representam as diferenças entre as estruturas. **C.** *Cluster* de genes anotado de *Saccharomonospora* sp. CNQ-490 codificando para biossíntese de taromicina.

Taromicina A/ B foram isoladas de isolados de *Saccharomonospora* sp. CNQ-490 após clonagem direta do BGC via TAR-cloning, utilizando plasmídeo pCAP01^{135,198}. Os compostos possuem estruturas similares a daptomicina¹⁰¹ (Figura 68B), um potente antibiótico contra bactérias multirresistentes. As principais diferenças entre as moléculas estão relacionadas com a incorporação de um triptofano clorado (6-Cl-Trp) e uma quinurenina clorada (4-Cl-Kyn) pelos módulos 1 e 13, respectivamente, bem como incorporação de alanina (Ala) pelo módulo 11, ao invés de serina (Ser), e formação de uma cadeia alifática insaturada e/ou ramificada (Figura 68C). As enzimas Tar13, Cl-triptofano-2,3-dioxigenase (TDO), e Tar16, *N'*-formil-quinurenina formamidase (FK), foram selecionadas para o estudo por estarem envolvidas na via de modificação do triptofano (via da quinurenina – Figura 69) e formação dos precursores halogenados a serem incorporados durante a rota de biossíntese.

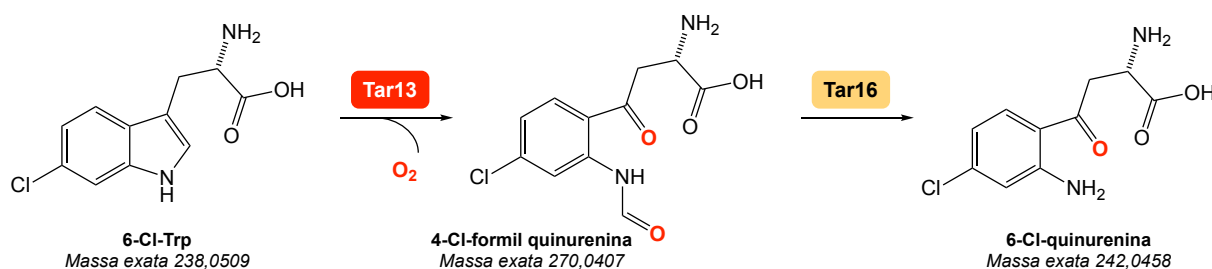


Figura 69. Via da quinurenina - metabolismo do triptofano. Etapas das reações envolvendo a triptofano dioxigenase (Tar13) e quinurenina formamidase (Tar16).

O produto de reação obtido pela ação de Tar13 e Tar16 é a 4-cloro-quinurenina, um pró-fármaco que ao ultrapassar a barreira hematoencefálica é metabolizado ao ácido 7-cloro quinurênico, um antagonista que interage com receptores NMDA (*N*-metil-D-aspartato) inibindo a dor, bem como é utilizado em tratamento clínico de depressão. Alguns estudos demonstram, também, seu potencial efeito na proteção neurodegenerativa.¹⁹⁹ O estudo foi, até a presente data, o primeiro apresentado envolvendo enzimas TDO e KF com ação enzimática em substratos halogenados e descrevendo as três etapas para obtenção, via reação enzimática, do pró-fármaco.

6.1. Objetivos

Expressão em *E. coli* BL21 (DE3) das proteínas Tar13 e Tar16 anotadas no *cluster* biossintético de Taromicina A/B e clonadas de *Saccharomonospora* sp. CNQ-490. Purificação e ensaios enzimáticos para avaliação da atividade catalítica em substrato de triptofano halogenado.

6.2. Materiais e Métodos

6.2.1. Expressão das proteínas em *E. coli* BL21 (DE3)

E. coli BL21(DE3) transformadas com os plasmídeos de expressão de Tar13 e Tar16, construídos pela Dra. Hanna Luhavaya, foram inoculadas em 10 mL de meio LB contendo 50 $\mu\text{g.mL}^{-1}$ de canamicina e incubadas a 37 °C, 220 rpm, 18 h. Os pré-cultivos foram transferidos para 4 ou 5 frascos de 2,8 L contendo 1 L de meio TB (1,2% triptona, 2,4% extrato de levedura, 0,4% (v/v) glicerol, 2,31% KH_2PO_4 , 12,54% K_2HPO_4) e 50 $\mu\text{g.mL}^{-1}$ de canamicina. As culturas foram incubadas a 37 °C, 220 rpm até atingir A_{600} 0,6–0,8, seguido da adição de 0,15 mM de isopropil- β -D-tiogalactopiranosídeo (IPTG). Para o caso da enzima Tar13, antes de induzir com IPTG e ao atingir A_{600} 0,4–0,6, foram adicionados 7 μM de hemina e 1 mM de ácido D-aminolevulínico (precursor de hemina). Após indução, os cultivos foram incubados a 18 °C, 120 rpm, por 18 h. As células foram centrifugadas a 12.000 x g, 5 min, 4 °C e o pellet obtido foi ressuspensionado em tampão gelado (40 mM Tris, 0,1 M NaCl, 20 mM imidazol, pH 7,4), mantido em gelo e lisado por sonicação (Qsonica sonicator, CL-334 com 40% de amplitude, 15 s pulso *On*/ 45 s pulso *Off*, por 5 min). O lisado celular foi centrifugado a 35.000 x g, 45 min, 4 °C e sobrenadante submetido a diferentes métodos para purificação.

6.2.2. Purificação de Tar13 e Tar16

Os procedimentos de purificação das proteínas foram realizados em equipamento ÄKTApurifier (GE Healthcare) com módulos Box-900, UPC-900, R-900 e Frac-900 acoplados, sendo todos os tampões previamente filtrados com membrana

de náilon 0,2 μ M GDWP (Merck). As proteínas foram purificadas a 4 °C por cromatografia de afinidade - Ni^{2+} em coluna de 5 mL HisTrap HP (GE Healthcare), conforme segue: tampão Ni A (50 mM Tris, 500 mM NaCl, 10% glicerol, pH 8,0) e tampão Ni B (50 mM Tris, 500 mM NaCl, 300 mM imidazol, 10% glicerol, pH 8,0); gradiente de eluição: tampão Ni B de 10 a 100% por 40 min, fluxo 2,5 mL.min⁻¹.

Após purificação em coluna de níquel, Tar13 foi mantida em tampão (~ 60% Ni B) e concentrada em filtro Amicon Ultra MWCO 10 kDa (Millipore) a um volume final de 2 mL. A proteína concentrada foi submetida a separação por cromatografia de gel-filtração com coluna Supedex 200 column (16 cm \times 60 cm, GE Healthcare) e tampão 50 mM Tris, 500 mM NaCl, 180 mM imidazol, 10% glicerol, pH 8,0. As frações contendo Tar13 foram combinadas, concentradas e imediatamente congeladas em banho de etanol/ água/gelo seco.

As frações da proteína Tar16, após purificação em coluna de níquel, foram concentradas em filtro Amicon Ultra MWCO 10 kDa (Millipore) a um volume final de 5 mL. O concentrado foi ressuspendido em 50 mL de tampão IEx A (20 mM Tris, 10% glicerol, pH 8,0) e eluído em coluna HiTrap Q HP 5 mL (GE Healthcare) com tampões IEx A e B e gradiente de 5 a 100% de tampão IEx B (20 mM Tris, 1 M KCl, 10% glicerol, pH 8,0) por 40 min, fluxo 2,5 mL.min⁻¹. As frações coletadas contendo proteína foram concentradas a um volume final de 2 mL e purificadas por cromatografia de gel-filtração com coluna Supedex 75 column (16 cm \times 60 cm, GE Healthcare) e tampão 40 mM Tris, 100 mM NaCl, 10% glicerol, pH 8,0. A enzima Tar16 foi concentrada, aliquoteada e imediatamente congelada em banho de etanol/água/gelo. Todos os procedimentos de purificação foram feitos no mesmo dia, devido a instabilidade da proteína.

6.2.3. *Ensaio enzimático in vitro*

As reações foram preparadas em volume final de 50 μ L e em tubos de 1,5 mL. Foi preparado um mix com 25 μ L de tampão de reação (Tris 50 mM, NaCl 50 mM e glicerol 10%), ácido ascórbico, hemina e enzimas, seguido de adição do substrato, previamente adicionado ao volume de tampão restante. As concentrações ótimas estão descritas na sessão de Resultados e Discussão. Após o tempo reacional, foi adicionado metanol grau HPLC em todas as reações, centrifugadas e analisadas em

LC Agilent 1260 Infinity acoplado a espectrômetro de massas Bruker Amazon Ion Trap.

6.3. Resultados e Discussão

As enzimas Tar13 (TDO) e Tar16 (KF) foram clonadas em vetor pET28 pela Dra. Hanna Luhavaya. Os experimentos de expressão foram conduzidos em *E. coli* BL21 (DE3) e induzidos por adição de IPTG. Após indução de expressão das proteínas, os cultivos foram centrifugados e o *pellet* celular foi lisado, seguido de purificação por diferentes métodos cromatográficos.

A enzima Tar13 é uma oxigenase heme-dependente, portanto, para obtenção da enzima na sua forma ativa, hemina foi adicionada ao cultivo antes da indução com IPTG e as frações contendo a proteína em maior concentração apresentaram uma coloração vermelha/marrom, indicando a ligação da proteína ao grupo heme. Todas as tentativas de troca de tampão por cromatografia de troca iônica ou diálise, para reduzir a concentração de sal, levaram a precipitação imediata da proteína. Deste modo, como o método de purificação por cromatografia de afinidade – Ni^{2+} apresentou bom rendimento e eficiência de purificação, os outros métodos foram desconsiderados, sendo adicionalmente submetida apenas ao método de cromatografia por gel-filtração para confirmar o tamanho (~ 28 kDa), sendo ao final obtida uma concentração de 42 mg.L^{-1} de Tar13 (Figura 70).

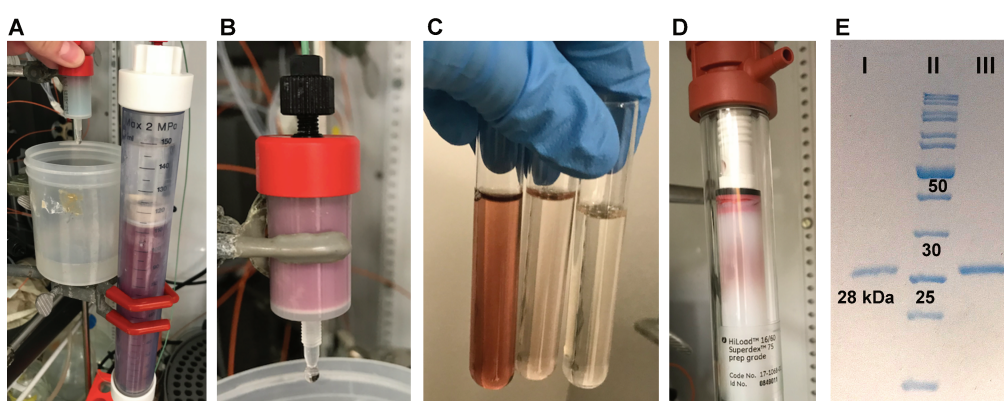


Figura 70. Purificação de Tar13. **A.** Cromatografia de afinidade – Ni: carregando a coluna. **B.** Coluna de níquel carregada. **C.** Diferentes frações obtidas em separação por cromatografia de afinidade. **D.** Coluna de gel-filtração Superdex. **E.** SDS-PAGE (12,5% acrilamida) obtida para proteína após separação por cromatografia de gel-filtração. Peso molecular obtido experimentalmente para Tar13 = ~ 28 kDa. Linhas I: Tar13 – $0,25 \text{ mg.mL}^{-1}$; II: EZ-Run™ Rec protein Ladder (Fisher Scientific); III: Tar13 – $0,5 \text{ mg.mL}^{-1}$.

A proteína Tar16 foi purificada por cromatografia de afinidade – Ni^{2+} ; cromatografia de troca iônica e cromatografia de gel-filtração e todas as etapas de purificação tiveram que ser realizadas no mesmo dia, para minimizar os fatores de precipitação. Contudo, mesmo com todo o cuidado nos procedimentos, devido as várias etapas de purificação, como a enzima é muito instável, o rendimento final obtido foi de $4,7 \text{ mg.L}^{-1}$ para uma proteína de 33 kDa (Figura 71).

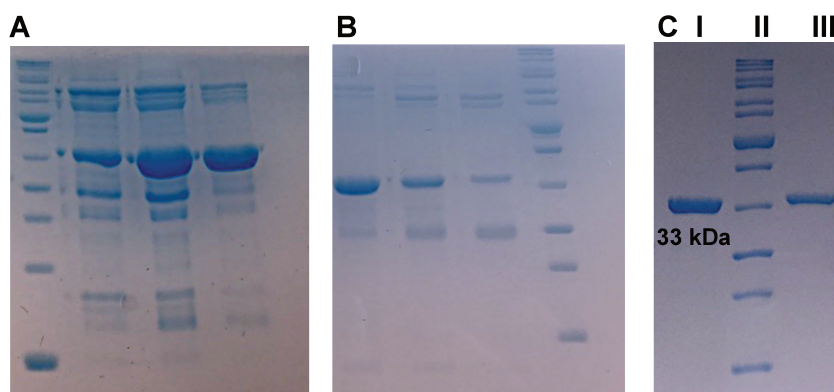
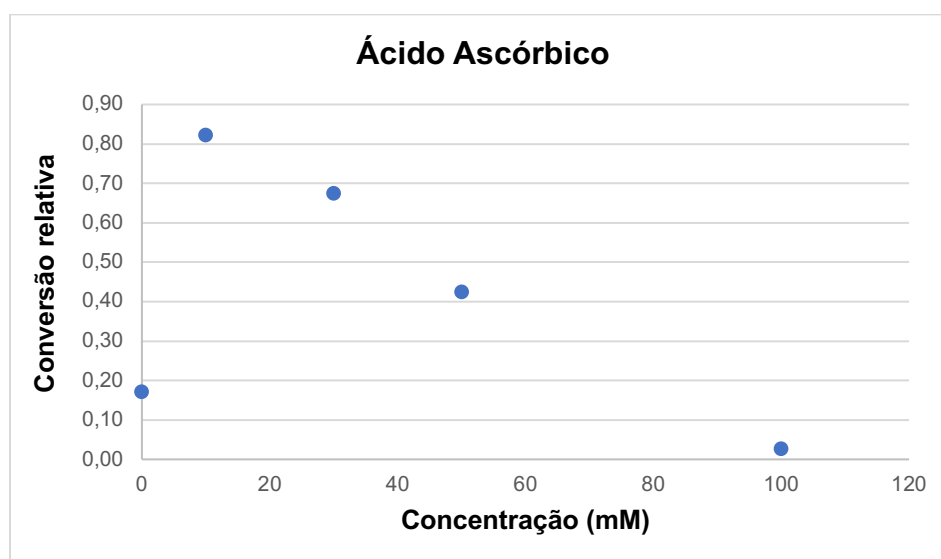
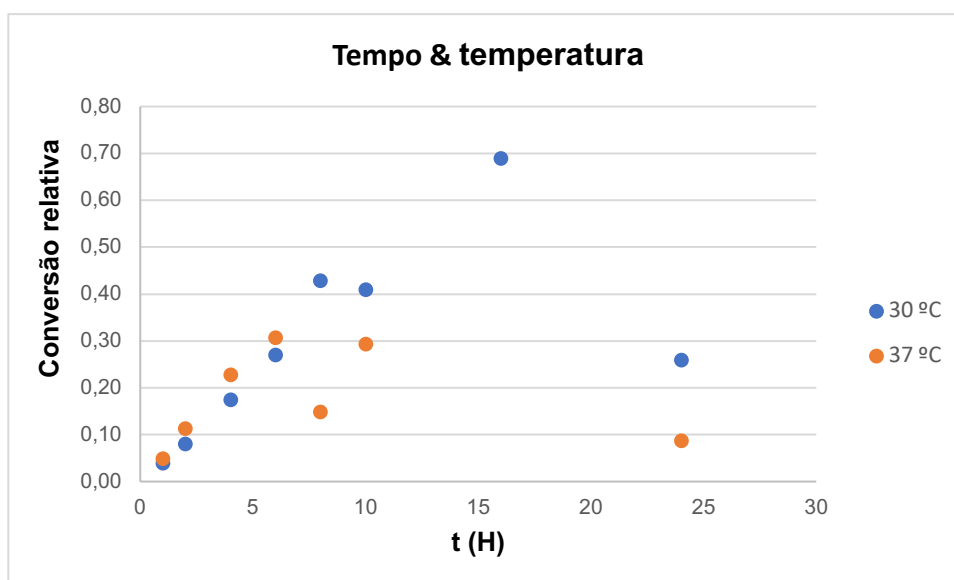
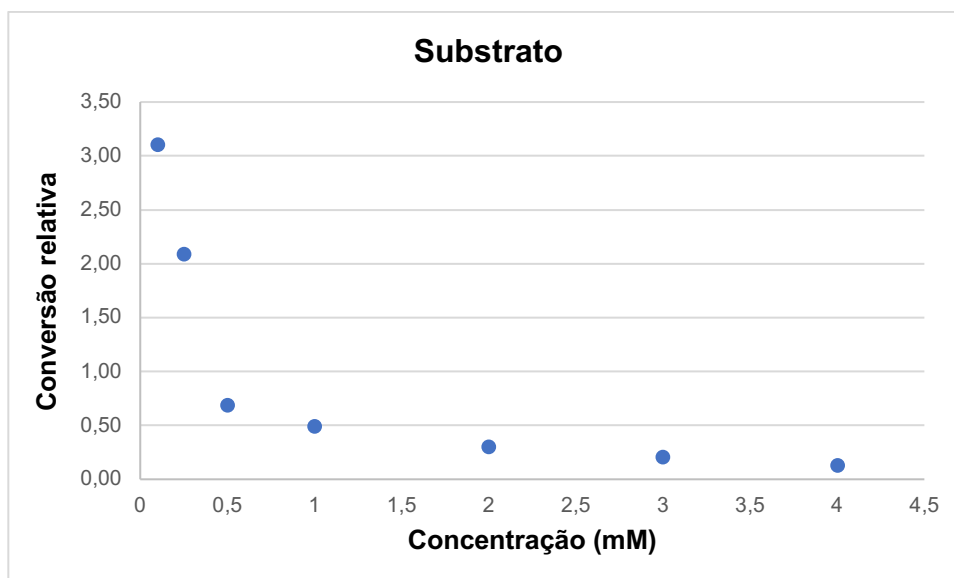


Figura 71. SDS-PAGE (12,5% acrilamida) para Tar16 purificada por **A.** Cromatografia de afinidade – Ni^{2+} . **B.** Cromatografia de troca iônica. **C.** Cromatografia de gel-filtração. Peso molecular obtido para Tar16 ~33 kDa. Linhas I: TAR16 – $0,25 \text{ mg. mL}^{-1}$; II: EZ-Run™ Rec protein Ladder (Fisher Scientific); III: TAR16 – $0,5 \text{ mg.mL}^{-1}$.

Após purificação, a atividade catalítica de cada enzima purificada foi testada em ensaios *in vitro* e as condições de reações foram otimizadas considerando a concentração de cofatores, enzimas e substrato. Além disso, foram testados tempo, temperatura de incubação e pH, em tampão Tris 50 mM, NaCl 50 mM e glicerol 10%,. Todas as condições testadas para a enzima Tar13 estão descritas na Tabela 19 (condições ótimas destacadas em amarelo) e representadas no Gráfico 2.

Tabela 19. Parâmetros testados em ensaios enzimáticos com Tar13. Em amarelo estão as condições ótimas

Parâmetro								
Tempo (h)	1	2	4	6	8	10	16	24
Substrato (mM)	0,1	0,25	0,5	1	2	3	4	
Ác. ascórbico mM	0	10	30	50	100			
Hemina (μM)	0	0,6	2,4	6	12			
Enzima (μM)	10	25	50	100				
pH	6,2	7,0	8,0					
Temperatura ($^{\circ}\text{C}$)	30	37						



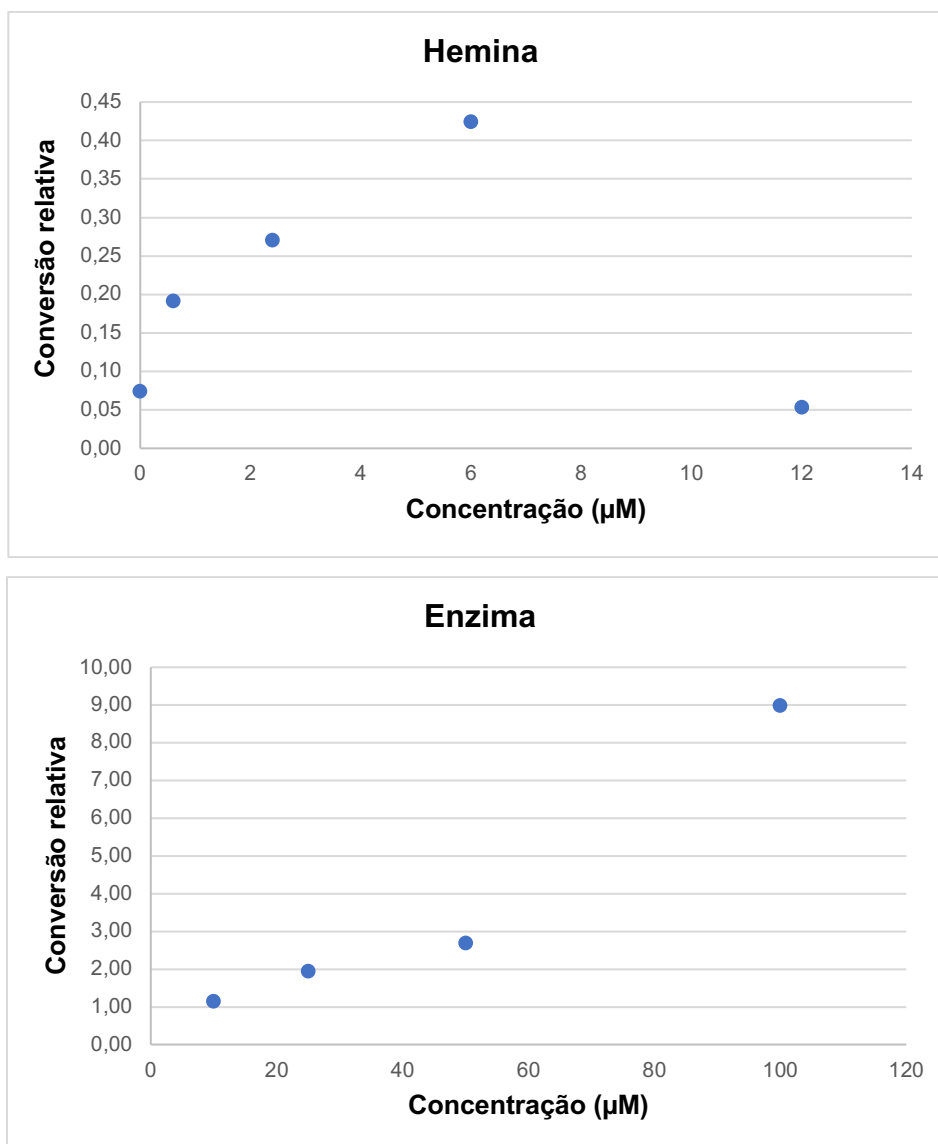


Gráfico 2. Representação gráfica dos resultados obtidos a partir de diferentes parâmetros testados para otimização dos ensaios enzimáticos de Tar13.

Todos os ensaios foram analisados por LC-MS, utilizando controles: sem enzima, enzima desnaturada e sem cofatores (ácido ascórbico e hemina). O padrão comercial de 6-Cl-triptofano (Sigma-Aldrich) foi utilizado como referência e controle positivo para o produto de reação esperado. Os cromatogramas obtidos estão representados na Figura 72 e como esperado, as reações controle (sem enzima ou com enzima desnaturada) não apresentaram conversão do substrato 6-Cl-Trp (m/z 238,05) em 4-Cl-formil-quinurenina (m/z 270,04), já as demais reações, mesmo sem cofatores, confirmaram a atividade catalítica de Tar13 como uma triptofano dioxigenase. Estes experimentos demonstram também a importância da adição de cofatores para otimizar a atividade enzimática.

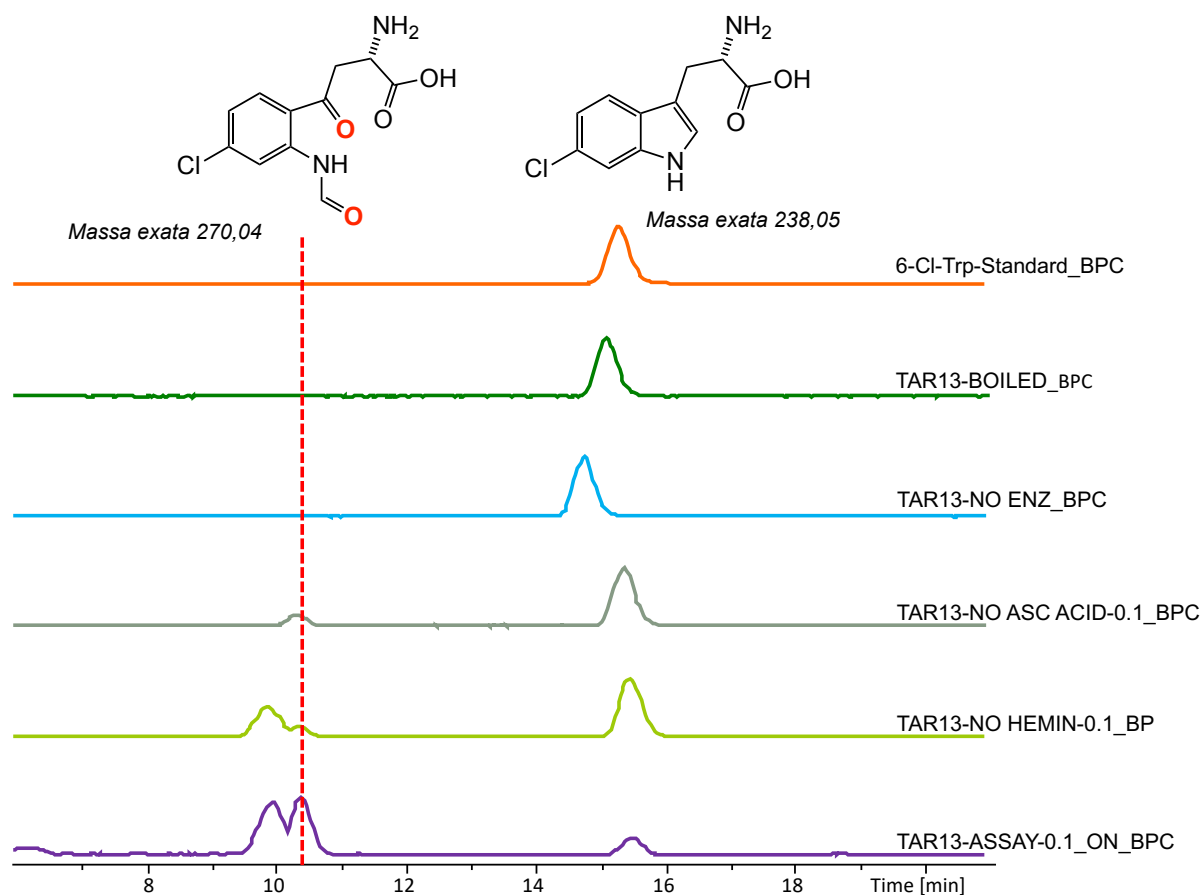


Figura 72. Cromatogramas de pico base obtidos por análise em LC-MS dos ensaios enzimáticos com Tar13 em condições otimizadas. Laranja: padrão comercial de 6-Cl-Trp m/z 238,0509 (Sigma-Aldrich). Verde: controle com Tar13 desnaturada. Azul: reação sem Tar13. Cinza: reação sem ácido ascórbico. Amarelo: reação sem hemina. Roxo: reação enzimática otimizada com Tar13, substrato (6-Cl-Trp) e cofatores, hemina e ácido ascórbico. m/z 270,0407 produto de reação esperado.

Para os ensaios enzimáticos com Tar16 não foram adicionados cofatores, sendo testadas apenas diferentes concentrações da enzima (10 μ M, 20 μ M, 35 μ M e 50 μ M) e diferentes tempos de adição da mesma na reação, conforme segue:

- A – Tar13 e Tar16 adicionadas ao mesmo tempo e reação incubada overnight;
- B – Tar13 adicionada 3 p.m. + Tar16 adicionada 6 p.m. e reação incubada overnight;
- C – Tar13 adicionada e reação incubada overnight + Tar16 adicionada 9 a.m. e reação incubada por 3h.

A adição de Tar16 antes ou depois da reação com Tar13 não apresentou variações significativas na conversão do substrato 4-Cl-formil-quinurenina, portanto optou-se por adicionar Tar13 e Tar16 ao mesmo tempo, em incubação *overnight*. Os ensaios foram analisados por LC-MS, utilizando controles: reação sem enzimas Tar13 e Tar16, sem Tar16 e com Tar16 desnaturada. Foram utilizados padrões comerciais de 6-Cl-Trp (Sigma-Aldrich) e 4-Cl-quinurenina (Sigma-Aldrich) como referência e controle positivo para produto de reação esperado. Os cromatogramas obtidos estão representados na Figura 73.

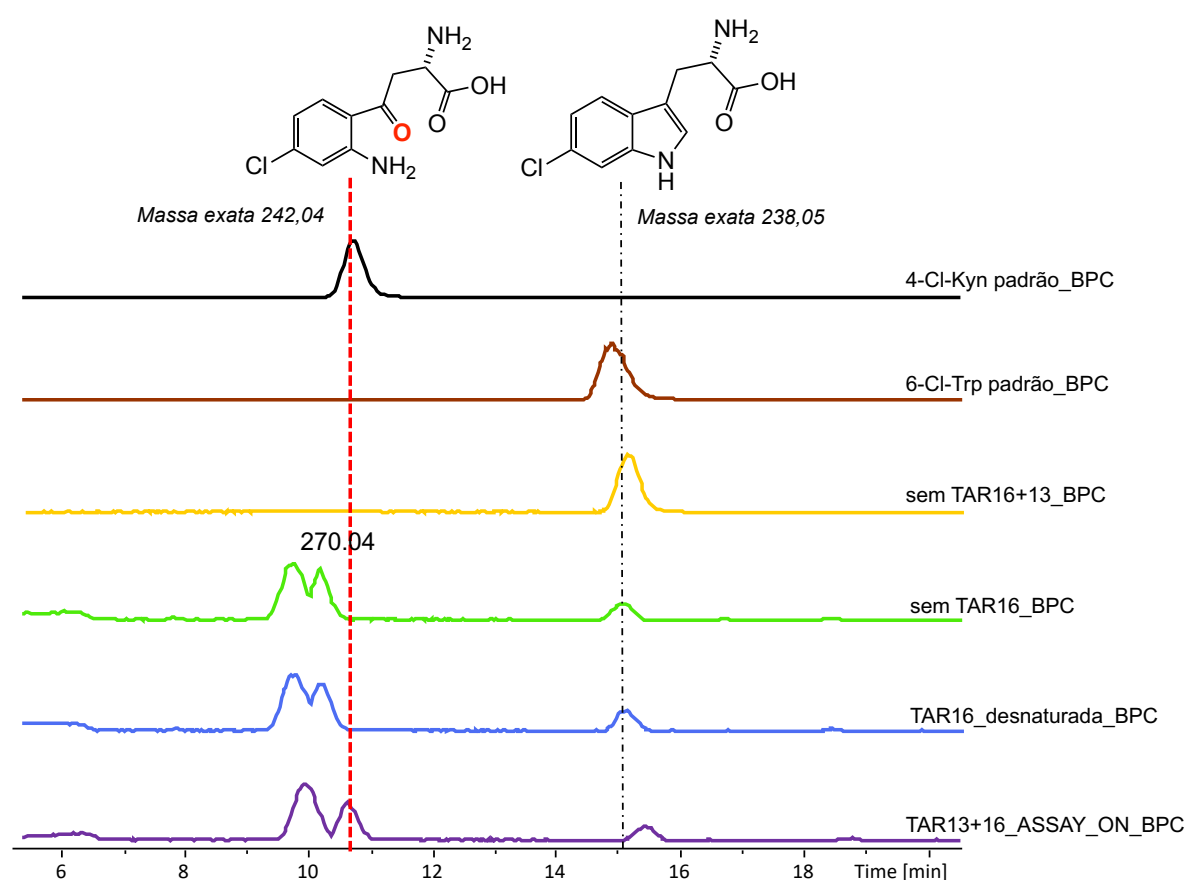


Figura 73. Cromatogramas de pico base obtidos para os ensaios enzimáticos com Tar16 em condições otimizadas. Preto: padrão comercial de 4-Cl-Kyn m/z 242,05 (Sigma-Aldrich). Vermelho: padrão comercial de 6-Cl-Trp m/z 238,05 (Sigma-Aldrich). Amarelo: sem Tar13 e Tar16. Verde: sem Tar16. Azul: Tar16 desnaturada. Roxo: reação enzimática otimizada com Tar13 e Tar16, substrato (6-Cl-Trp) e cofatores, hemina e ácido ascórbico. m/z 242,05 produto de reação esperado. m/z 270,04: 4-Cl-formil-quinurenina.

6.4. Conclusões

Os experimentos para expressão das enzimas foram conduzidos com sucesso, sendo fundamental a adição de hemina no cultivo de expressão de Tar13. Testes realizados previamente no grupo sem adição do cofator Durante a purificação, as enzimas se mostraram bastante instáveis, sendo importante manter a enzima Tar13 em tampão salino e nos ensaios enzimáticos a adição da mesma sem cofatores adicionados previamente resultou em precipitação imediata. Os ensaios enzimáticos *in vitro* com Tar13 e Tar16 demonstraram, pela primeira vez, a ação de uma TDO ativa em substrato triptofano halogenado e foi possível concluir que a enzima Tar13 é o fator limitante da via da quinurenina para completa conversão do substrato 6-Cl-triptofano. Mesmo em baixas concentrações, Tar16 apresentou completa conversão do substrato 4-Cl-formil-quinurenina em 4-Cl-quinurenina.

Os resultados obtidos neste trabalho fizeram parte de um projeto completo desenvolvido pela Dra. Hanna Luhavaya, o qual apresentou, pela primeira vez, de forma detalhada o estudo envolvendo a biossíntese de 4-Cl-quinurenina, um pró-fármaco (VistaGen Therapeutics, Inc, San Francisco, CA) em testes clínicos para tratamento de casos de depressão e doenças neurodegenerativas.

7. REFERÊNCIAS

1. Cragg, G. M. & Newman, D. J. Natural products: A continuing source of novel drug leads. *Biochim. Biophys. Acta - Gen. Subj.* **1830**, 3670–3695 (2013).
2. Williams, D. H., Stone, M. J., Hauck, P. R. & Rahman, S. K. Why are Secondary Metabolites (Natural Products) Biosynthesized? *J. Nat. Prod.* **52**, 1189–1208 (1989).
3. O'Brien, J. & Wright, G. D. An ecological perspective of microbial secondary metabolism. *Curr. Opin. Biotechnol.* **22**, 552–558 (2011).
4. Wink, M. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry* **64**, 3–19 (2003).
5. Li, J. W.-H. & Vederas, J. C. Drug discovery and natural products: End of an era or an endless frontier? *Science* **325**, 161–165 (2009).
6. Hartmann, T. From waste products to ecochemicals: Fifty years research of plant secondary metabolism. *Phytochemistry* **68**, 2831–2846 (2007).
7. Achan, J. *et al.* Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. *Malar. J.* **10**, 144 (2011).
8. Fleming, A. On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to their Use in the Isolation of *B. influenzae*. *Br J Exp Pathol.* **10**, 226–236 (1929).
9. Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336–343 (2016).
10. Hotchkiss, R. D. & Dubos, R. J. Bactericidal fractions sporulating from an aerobic *Bacillus*. *J. Biol. Chem.* **136**, 803–804 (1940).
11. Dubos, R. J. & Cattaneo, C. Studies on a Bactericidal Agent Extracted from a Soil *Bacillus*: III. Preparation and Activity of a Protein-Free Fraction. *J. Exp. Med.* **70**, 249–56 (1939).
12. Van Epps, H. L. René Dubos: unearthing antibiotics. *J. Exp. Med.* **203**, 259 (2006).

13. Drews, J. Drug discovery: a historical perspective. *Science* **287**, 1960–4 (2000).
14. Kelkar, D. A. & Chattopadhyay, A. The gramicidin ion channel: A model membrane protein. *Biochim. Biophys. Acta - Biomembr.* **1768**, 2011–2025 (2007).
15. Wright, G. D. Opportunities for natural products in 21st century antibiotic discovery. *Nat. Prod. Rep.* **34**, 694–701 (2017).
16. Waksman, S. A. Streptomycin: Background, Isolation, Properties, and Utilization. *Science (80-.).* **118**, 259–266 (1953).
17. Butler, M. S. & Buss, A. D. Natural products — The future scaffolds for novel antibiotics? *Biochem. Pharmacol.* **71**, 919–929 (2006).
18. Lewis, K. Recover the lost art of drug discovery. *Nature* **485**, 439–440 (2012).
19. Davies, J. Where have All the Antibiotics Gone? *Can J Infect Dis Med Microbiol Vol* **17**, 287–290 (2006).
20. Aminov, R. I. A Brief History of the Antibiotic Era: Lessons Learned and Challenges for the Future. *Front. Microbiol.* **1**, 1–7 (2010).
21. Katz, L. & Baltz, R. H. Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.* **43**, 155–176 (2016).
22. Demain, A. L. Importance of microbial natural products and the need to revitalize their discovery. *J. Ind. Microbiol. Biotechnol.* **41**, 185–201 (2014).
23. Patridge, E., Gareiss, P., Kinch, M. S. & Hoyer, D. An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov. Today* **21**, 204–207 (2016).
24. Hopwood, D. A., Kieser, T., Buttner, M. J., Bibb, M. J. & Chater, K. F. *Practical Streptomyces Genetics*. (The John Innes Foundation, 2000).
25. Chater, K. F. *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philos. Trans. R. Soc. B Biol. Sci.* **361**, 761–768 (2006).
26. Moncheva, P. *et al.* Characteristics of Soil Actinomycetes from Antarctica. *J. Culture Collections.* **3**, 3-14 (2000-2002).

27. Kurapova, A. I. *et al.* Thermotolerant and thermophilic actinomycetes from soils of Mongolia desert steppe zone. *Microbiology* **81**, 98–108 (2012).
28. Bhatti, A. A., Haq, S. & Bhat, R. A. Actinomycetes benefaction role in soil and plant health. *Microb. Pathog.* **111**, 458–467 (2017).
29. Hardoim, P. R., van Overbeek, L. S. & Elsas, J. D. van. Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol.* **16**, 463–471 (2008).
30. Hardoim, P. R. *et al.* The Hidden World within Plants: Ecological and Evolutionary Considerations for Defining Functioning of Microbial Endophytes. *Microbiol. Mol. Biol. Rev.* **79**, 293–320 (2015).
31. Durand, G. A., Raoult, D. & Dubourg, G. Antibiotic discovery: history, methods and perspectives. *Int. J. Antimicrob. Agents* **53**, 371–382 (2019).
32. Levy, S. B. & Marshall, B. Antibacterial resistance worldwide: causes, challenges and responses. *Nat. Med.* **10**, S122–S129 (2004).
33. Schäberle, T. F. & Hack, I. M. Overcoming the current deadlock in antibiotic research. *Trends Microbiol.* **22**, 165–167 (2014).
34. Bérdy, J. Thoughts and facts about antibiotics: Where we are now and where we are heading. *J. Antibiot. (Tokyo)*. **65**, 385–395 (2012).
35. WHO. *Antimicrobial resistance - Global Report on Surveillance*. World Health Organization (2014). doi:10.1007/s13312-014-0374-3
36. Newman, D. J., Cragg, G. M. & Snader, K. M. Natural products as sources of new drugs over the period 1981-2002. *J. Nat. Prod.* **66**, 1022–1037 (2003).
37. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* **70**, 461–477 (2007).
38. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* **75**, 311–335 (2012).
39. Zerikly, M. & Challis, G. L. Strategies for the Discovery of New Natural Products by Genome Mining. *ChemBioChem* **10**, 625–633 (2009).
40. Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete

- Streptomyces coelicolor* A3(2). *Nature* **417**, 141–147 (2002).
41. Ikeda, H. *et al.* Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat. Biotechnol.* **21**, 526–531 (2003).
 42. Harrison, J. & Studholme, D. J. Recently published *Streptomyces* genome sequences. *Microb. Biotechnol.* **7**, 373–380 (2014).
 43. Bachmann, B. O., Van Lanen, S. G. & Baltz, R. H. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J. Ind. Microbiol. Biotechnol.* **41**, 175–184 (2014).
 44. Ziemert, N., Alanjary, M. & Weber, T. The evolution of genome mining in microbes – a review. *Nat. Prod. Rep.* **33**, 988–1005 (2016).
 45. Corre, C. & Challis, G. L. New natural product biosynthetic chemistry discovered by genome mining. *Nat. Prod. Rep.* **26**, 977 (2009).
 46. Omura, S. *et al.* Genome sequence of an industrial microorganism *Streptomyces avermitilis*: Deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci.* **98**, 12215–12220 (2001).
 47. Paulsen, I. T. *et al.* Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. *Nat. Biotechnol.* **23**, 873–878 (2005).
 48. Bok, J. W. *et al.* Genomic Mining for *Aspergillus* Natural Products. *Chem. Biol.* **13**, 31–37 (2006).
 49. Fazio, G. C., Xu, R. & Matsuda, S. P. T. Genome Mining to Identify New Plant Triterpenoids. *J. Am. Chem. Soc.* **126**, 5678–5679 (2004).
 50. Challis, G. L. & Ravel, J. Coelichelin, a new peptide siderophore encoded by the *Streptomyces coelicolor* genome: structure prediction from the sequence of its non-ribosomal peptide synthetase. *FEMS Microbiol. Lett.* **187**, 111–4 (2000).
 51. Lautru, S., Deeth, R. J., Bailey, L. M. & Challis, G. L. Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nat. Chem. Biol.* **1**, 265–269 (2005).
 52. Tohyama, S. *et al.* Genome-inspired search for new antibiotics. Isolation and structure determination of new 28-membered polyketide macrolactones,

- halstoctacosanolides A and B, from *Streptomyces halstedii* HC34. *Tetrahedron* **60**, 3999–4005 (2004).
53. McAlpine, J. B. *et al.* Microbial Genomics as a Guide to Drug Discovery and Structural Elucidation: ECO-02301, a Novel Antifungal Agent, as an Example. *J. Nat. Prod.* **68**, 493–496 (2005).
 54. Banskota, A. H. *et al.* Isolation and Identification of Three New 5-Alkenyl-3,3(2H)-furanones from Two *Streptomyces* species using a Genomic Screening Approach. *J. Antibiot. (Tokyo)*. **59**, 168–176 (2006).
 55. Banskota, A. H. *et al.* Genomic Analyses Lead to Novel Secondary Metabolites. *J. Antibiot. (Tokyo)*. **59**, 533–542 (2006).
 56. Xiong, Q., Wilson, W. K. & Matsuda, S. P. T. An *Arabidopsis* Oxidosqualene Cyclase Catalyzes Iridal Skeleton Formation by Grob Fragmentation. *Angew. Chemie Int. Ed.* **45**, 1285–1288 (2006).
 57. Lin, X., Hopson, R. & Cane, D. E. Genome Mining in *Streptomyces coelicolor*: Molecular Cloning and Characterization of a New Sesquiterpene Synthase. *J. Am. Chem. Soc.* **128**, 6022–6023 (2006).
 58. Scherlach, K. & Hertweck, C. Discovery of aspoquinolones A–D, prenylated quinoline-2-one alkaloids from *Aspergillus nidulans*, motivated by genome mining. *Org. Biomol. Chem.* **4**, 3517–3520 (2006).
 59. Fischbach, M. a & Walsh, C. T. Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *Chem. Rev.* **106**, 3468–3496 (2006).
 60. Wenzel, S. C., Meiser, P., Binz, T. M., Mahmud, T. & Müller, R. Nonribosomal Peptide Biosynthesis: Point Mutations and Module Skipping Lead to Chemical Diversity. *Angew. Chemie Int. Ed.* **45**, 2296–2301 (2006).
 61. Song, L. *et al.* Type III Polyketide Synthase β -Ketoacyl-ACP Starter Unit and Ethylmalonyl-CoA Extender Unit Selectivity Discovered by *Streptomyces coelicolor* Genome Mining. *J. Am. Chem. Soc.* **128**, 14754–14755 (2006).
 62. Gross, H. *et al.* The Genomisotopic Approach: A Systematic Method to Isolate Products of Orphan Biosynthetic Gene Clusters. *Chem. Biol.* **14**, 53–63 (2007).

63. de Oliveira, L. G., Pupo, M. T. & Vieira, P. C. Explorando produtos naturais microbianos nas fronteiras da química e da biologia. *Quim. Nova* **36**, 1577–1586 (2013).
64. Corre, C. & Challis, G. L. Heavy Tools for Genome Mining. *Chem. Biol.* **14**, 7–9 (2007).
65. Udworthy, D. W. *et al.* Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc. Natl. Acad. Sci.* **104**, 10376–10381 (2007).
66. Rateb, M. E. *et al.* Chaxamycins A–D, Bioactive Ansamycins from a Hyper-arid Desert *Streptomyces* sp. *J. Nat. Prod.* **74**, 1491–1499 (2011).
67. Chiang, Y.-M. *et al.* Molecular Genetic Mining of the *Aspergillus* Secondary Metabolome: Discovery of the Emericellamide Biosynthetic Pathway. *Chem. Biol.* **15**, 527–532 (2008).
68. Yin, W.-B., Grundmann, A., Cheng, J. & Li, S.-M. Acetylaszonalenin Biosynthesis in *Neosartorya fischeri*. *J. Biol. Chem.* **284**, 100–109 (2009).
69. Huang, T. *et al.* Identification and Characterization of the Pyridomycin Biosynthetic Gene Cluster of *Streptomyces pyridomyceticus* NRRL B-2517. *J. Biol. Chem.* **286**, 20648–20657 (2011).
70. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
71. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
72. Medema, M. H. *et al.* antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **39**, W339–W346 (2011).
73. Blin, K. *et al.* antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* (2019). doi:10.1093/nar/gkz310
74. Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29–W37 (2011).
75. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high

- throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
76. Epstein, S. C., Charkoudian, L. K. & Medema, M. H. A standardized workflow for submitting data to the Minimum Information about a Biosynthetic Gene cluster (MIBiG) repository: prospects for research-based educational experiences. *Stand. Genomic Sci.* **13**, 16 (2018).
 77. Medema, M. H. *et al.* Minimum Information about a Biosynthetic Gene cluster. *Nat. Chem. Biol.* **11**, 625–631 (2015).
 78. Weber, J., Leung, J., Swanson, S., Idler, K. & McAlpine, J. An erythromycin derivative produced by targeted gene disruption in *Saccharopolyspora erythraea*. *Science* **252**, 114–117 (1991).
 79. Bevitt, D. J., Cortes, J., Haydock, S. F. & Leadlay, P. F. 6-Deoxyerythronolide-B synthase 2 from *Saccharopolyspora erythraea*. *Eur. J. Biochem.* **204**, 39–49 (1992).
 80. Menzella, H. G. *et al.* Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. *Nat. Biotechnol.* **23**, 1171–1176 (2005).
 81. Jenke-Kodama, H., Sandmann, A., Müller, R. & Dittmann, E. Evolutionary Implications of Bacterial Polyketide Synthases. *Mol. Biol. Evol.* **22**, 2027–2039 (2005).
 82. Kapur, S. *et al.* Reprogramming a module of the 6-deoxyerythronolide B synthase for iterative chain elongation. *Proc. Natl. Acad. Sci.* **109**, 4110–4115 (2012).
 83. Chan, Y. A., Podevels, A. M., Kevany, B. M. & Thomas, M. G. Biosynthesis of polyketide synthase extender units. *Nat. Prod. Rep.* **26**, 90–114 (2009).
 84. Zhou, S. *et al.* Salinomycin: a novel anti-cancer agent with known anti-coccidial activities. *Curr. Med. Chem.* **20**, 4095–4101 (2013).
 85. Askenazi, M. *et al.* Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. *Nat. Biotechnol.* **21**, 150–156 (2003).
 86. WHO. *19th WHO Model List of Essential Medicines*. WHO (2015).

87. Hertweck, C., Luzhetskyy, A., Rebets, Y. & Bechthold, A. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.* **24**, 162–190 (2007).
88. Yu, D., Xu, F., Zeng, J. & Zhan, J. Type III polyketide synthases in natural product biosynthesis. *IUBMB Life* **64**, 285–295 (2012).
89. Kirimura, K., Watanabe, S. & Kobayashi, K. Heterologous gene expression and functional analysis of a type III polyketide synthase from *Aspergillus niger* NRRL 328. *Biochem. Biophys. Res. Commun.* **473**, 1106–1110 (2016).
90. Li, J., Luo, Y., Lee, J.-K. & Zhao, H. Cloning and characterization of a type III polyketide synthase from *Aspergillus niger*. *Bioorg. Med. Chem. Lett.* **21**, 6085–6089 (2011).
91. Baur, J. A. & Sinclair, D. A. Therapeutic potential of resveratrol: the in vivo evidence. *Nat. Rev. Drug Discov.* **5**, 493–506 (2006).
92. Piel, J. Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat. Prod. Rep.* **27**, 996 (2010).
93. Helfrich, E. J. N. & Piel, J. Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat. Prod. Rep.* **33**, 231–316 (2016).
94. Scotti, C. *et al.* A *Bacillus subtilis* large ORF coding for a polypeptide highly similar to polyketide synthases. *Gene* **130**, 65–71 (1993).
95. PATEL, P. S. *et al.* Bacillaene, a Novel Inhibitor of Prokaryotic Protein Synthesis Produced by *Bacillus subtilis*: Production, Taxonomy, Isolation, Physico-chemical Characterization and Biological Activity. *J. Antibiot. (Tokyo)*. **48**, 997–1003 (1995).
96. Butcher, R. A. *et al.* The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. *Proc. Natl. Acad. Sci.* **104**, 1506–1509 (2007).
97. Cheng, Y.-Q., Tang, G.-L. & Shen, B. Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc. Natl. Acad. Sci.* **100**, 3149–3154 (2003).
98. Piel, J., Hui, D., Fusetani, N. & Matsunaga, S. Targeting modular polyketide synthases with iteratively acting acyltransferases from metagenomes of

- uncultured bacterial consortia. *Environ. Microbiol.* **6**, 921–927 (2004).
99. Pulsawat, N., Kitani, S. & Nihira, T. Characterization of biosynthetic gene cluster for the production of virginiamycin M, a streptogramin type A antibiotic, in *Streptomyces virginiae*. *Gene* **393**, 31–42 (2007).
 100. Fischbach, M. a. & Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic machinery, and mechanisms. *Chem. Rev.* **106**, 3468–3496 (2006).
 101. Tally, F. P. & DeBruin, M. F. Development of daptomycin for Gram-positive infections. *J. Antimicrob. Chemother.* **46**, 523–526 (2000).
 102. van Wageningen, A. A. *et al.* Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chem. Biol.* **5**, 155–162 (1998).
 103. Xue, Y., Zhao, L., Liu, H. W. & Sherman, D. H. A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: architecture of metabolic diversity. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 12111–12116 (1998).
 104. Larin, Z. YACs , BACs , PACs and MACs: artificial chromosomes as research tools. *Trends Biotechnol.* **12**, 280–286 (1994).
 105. Shizuya, H. *et al.* Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8794–8797 (1992).
 106. Ongley, S. E., Bian, X., Neilan, B. A. & Müller, R. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Nat. Prod. Rep.* **30**, 1121 (2013).
 107. Dai, S., Ouyang, Y., Wang, G. & Li, X. *Streptomyces autolyticus* JX-47 large-insert bacterial artificial chromosome library construction and identification of clones covering geldanamycin biosynthesis gene cluster. *Curr. Microbiol.* **63**, 68–74 (2011).
 108. Ouyang, Y. *et al.* Isolation of high molecular weight DNA from marine sponge bacteria for BAC library construction. *Mar. Biotechnol. (NY)*. **12**, 318–25 (2010).
 109. Jones, A. C. *et al.* Phage P1-Derived Artificial Chromosomes Facilitate Heterologous Expression of the FK506 Gene Cluster. *PLoS One* **8**, e69319

- (2013).
110. Sosio, M. *et al.* Artificial chromosomes for antibiotic-producing actinomycetes. *Nat. Biotechnol.* **18**, 343–345 (2000).
 111. Alt, S. & Wilkinson, B. Biosynthesis of the Novel Macrolide Antibiotic Anthracimycin. *ACS Chem. Biol.* **10**, 2468–2479 (2015).
 112. Li, X. & Heyer, W.-D. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res.* **18**, 99–113 (2008).
 113. Thomason, L. C., Sawitzke, J. a, Li, X., Costantino, N. & Court, D. L. Recombineering: genetic engineering in bacteria using homologous recombination. in *Current protocols in molecular biology* **106**, 1.16.1-1.16.39 (JohnWiley & Sons, Inc., 2014).
 114. Fu, J. *et al.* Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.* **30**, 440–446 (2012).
 115. Murphy, K. C. Use of bacteriophage λ Recombination recombination functions to promote gene replacement in *Escherichia coli*. *J. Bacteriol.* **180**, 2063–71 (1998).
 116. Gust, B., Challis, G. L., Fowler, K., Kieser, T. & Chater, K. F. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci.* **100**, 1541–1546 (2003).
 117. Zhang, Y., Buchholz, F., Muylers, J. P. & Stewart, a F. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.* **20**, 123–8 (1998).
 118. Yu, D. *et al.* An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5978–5983 (2000).
 119. Gust, B., Rourke, S. O., Bird, N., Kieser, T. & Chater, K. Recombineering in *Streptomyces coelicolor*. 1–22 (2004).
 120. Luhavaya, H., Williams, S. R., Hong, H., Gonzaga de Oliveira, L. & Leadlay, P. F. Site-Specific Modification of the Anticancer and Antituberculosis Polyether

- Salinomycin by Biosynthetic Engineering. *ChemBioChem* **15**, 2081–2085 (2014).
121. Zhang, Y., Muirers, J. P. P., Testa, G. & Stewart, A. F. DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.* **18**, 1314–1317 (2000).
 122. Rodríguez, C., de Gonzalo, G., Fraaije, M. W. & Gotor, V. Enzymatic kinetic resolution of racemic ketones catalyzed by Baeyer-Villiger monooxygenases. *Tetrahedron Asymmetry* **18**, 1338–1344 (2007).
 123. Li, L., Jiang, W. & Lu, Y. New strategies and approaches for engineering biosynthetic gene clusters of microbial natural products. *Biotechnol. Adv.* **35**, 936–949 (2017).
 124. Gibson, D. G. *et al.* Complete Chemical Synthesis, Assembly, and Cloning of a *Mycoplasma genitalium* Genome. *Science* (80-.). **319**, 1215–1220 (2008).
 125. Gibson, D. G. *et al.* One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20404–20409 (2008).
 126. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
 127. Gibson, D. G., Smith, H. O., Hutchison, C. a, Venter, J. C. & Merryman, C. Chemical synthesis of the mouse mitochondrial genome. *Nat. Methods* **7**, 901–903 (2010).
 128. Kouprina, N. *et al.* Functional copies of a human gene can be directly isolated by transformation-associated recombination cloning with a small 3' end target sequence. *Proc. Natl. Acad. Sci.* **95**, 4469–4474 (1998).
 129. Larionov, V., Kouprina, N., Solomon, G., Barrett, J. C. & Resnick, M. A. Direct isolation of human BRCA2 gene by transformation-associated recombination in yeast. *Proc. Natl. Acad. Sci.* **94**, 7384–7387 (1997).
 130. Larionov, V. *et al.* Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc. Natl. Acad. Sci.* **93**, 491–496 (1996).

131. Kouprina, N. & Larionov, V. Exploiting the yeast *Saccharomyces cerevisiae* for the study of the organization and evolution of complex genomes. *FEMS Microbiol. Rev.* **27**, 629–649 (2003).
132. Kouprina, N. & Larionov, V. Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. *Chromosoma* **125**, 621–632 (2016).
133. Kouprina, N. & Larionov, V. TAR cloning: insights into gene function, long-range haplotypes and genome structure and evolution. *Nat. Rev. Genet.* **7**, 805–812 (2006).
134. Noskov, V. N. *et al.* A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic region in yeast. *BMC Genomics* **4**, 16 (2003).
135. Yamanaka, K. *et al.* Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc. Natl. Acad. Sci.* **111**, 1957–1962 (2014).
136. Bonet, B., Teufel, R., Crüsemann, M., Ziemert, N. & Moore, B. S. Direct Capture and Heterologous Expression of *Salinispora* Natural Product Genes for the Biosynthesis of Enterocin. *J. Nat. Prod.* **78**, 539–542 (2015).
137. Ross, A. C., Gulland, L. E. S., Dorrestein, P. C. & Moore, B. S. Targeted Capture and Heterologous Expression of the *Pseudoalteromonas alterochromide* Gene Cluster in *Escherichia coli* Represents a Promising Natural Product Exploratory Platform. *ACS Synth. Biol.* **4**, 414–420 (2015).
138. Hillman, R. T. & Calos, M. P. Site-Specific Integration with Bacteriophage ϕ C31 Integrase. *Cold Spring Harbor Protocols* doi:10.1101/pdb.prot069211 (2012). doi:10.1101/pdb.prot069211
139. Tang, X. *et al.* Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-directed Genome Mining. *ACS Chem. Biol.* **10**, 2841–2849 (2015).
140. Boeke, J. D., LaCroute, F. & Fink, G. R. A positive selection for mutants lacking 5' phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol. Gen. Genet.* **197**, 345–346 (1984).
141. Gomez-Escribano, J. P. & Bibb, M. J. Heterologous expression of natural

- product biosynthetic gene clusters in *Streptomyces coelicolor*: from genome mining to manipulation of biosynthetic pathways. *J. Ind. Microbiol. Biotechnol.* **41**, 425–431 (2014).
142. Baltz, R. H. *Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters. *J. Ind. Microbiol. Biotechnol.* **37**, 759–772 (2010).
 143. Gomez-Escribano, J. P. & Bibb, M. J. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb. Biotechnol.* **4**, 207–215 (2011).
 144. Komatsu, M., Uchiyama, T., Omura, S., Cane, D. E. & Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl. Acad. Sci.* **107**, 2646–2651 (2010).
 145. Ro, D.-K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
 146. DeJong, J. M. *et al.* Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **93**, 212–224 (2006).
 147. Zhang, X. *et al.* Construction of ivermectin producer by domain swaps of avermectin polyketide synthase in *Streptomyces avermitilis*. *Appl. Microbiol. Biotechnol.* **72**, 986–994 (2006).
 148. Tocchetti, A., Donadio, S. & Sosio, M. Large inserts for big data: artificial chromosomes in the genomic era. *FEMS Microbiol. Lett.* **365**, 64 (2018).
 149. Sun, Y. *et al.* ‘*Streptomyces nanchangensis*’, a producer of the insecticidal polyether antibiotic nanchangmycin and the antiparasitic macrolide meilingmycin, contains multiple polyketide gene clusters. *Microbiology* **148**, 361–371 (2002).
 150. Sun, Y. *et al.* Organization of the biosynthetic gene cluster in *Streptomyces* sp. DSM 4137 for the novel neuroprotectant polyketide meridamycin. *Microbiology* **152**, 3507–3515 (2006).
 151. Paranthaman, S. & Dharmalingam, K. Intergeneric Conjugation in *Streptomyces peucetius* and *Streptomyces* sp. Strain C5: Chromosomal Integration and

- Expression of Recombinant Plasmids Carrying the *chiC* Gene. *Appl. Environ. Microbiol.* **69**, 84–91 (2003).
152. Bennett, P. M., Grinsted, J. & Richmond, M. H. Transposition of TnA does not generate deletions. *MGG Mol. Gen. Genet.* **154**, 205–211 (1977).
 153. Araujo, W. L. *et al.* Diversity of Endophytic Bacterial Populations and their Interaction with *Xylella fastidiosa* in *Citrus* Plants. *Appl. Environ. Microbiol.* **68**, 4906–4914 (2002).
 154. Noskov, V. *et al.* A genetic system for direct selection of gene-positive clones during recombinational cloning in yeast. *Nucleic Acids Res.* **30**, E8 (2002).
 155. MacNeil, D. J. *et al.* Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**, 61–68 (1992).
 156. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97**, 6640–6645 (2000).
 157. Peterson, D. G., Tomkins, J. P., Frisch, D. A., Wing, R. A. & Paterson, A. H. Construction of plant bacterial artificial chromosome (BAC) libraries: An illustrated guide. *J. Agric. Genomics* **5**, 1–100 (2000).
 158. Luhavaya, H., Williams, S. R., Hong, H., de Oliveira, L. G. & Leadlay, P. F. Site-specific modification of the anti-cancer and anti-tuberculosis polyether salinomycin by biosynthetic engineering. *ChemBioChem* **15**, 2081–2085 (2014).
 159. Zhang, J. J., Yamanaka, K., Tang, X. & Moore, B. S. Direct cloning and heterologous expression of natural product biosynthetic gene clusters by transformation-associated recombination. in *Methods in Enzymology* 87–110 (Elsevier Inc., 2019). doi:10.1016/bs.mie.2019.02.026
 160. Gonzaga de Oliveira, L., Sigrist, R., Sachetto Paulo, B. & Samborsky, M. Whole-Genome Sequence of the Endophytic *Streptomyces* sp. Strain CBMAI 2042, Isolated from *Citrus sinensis*. *Microbiol. Resour. Announc.* **8**, 1–2 (2019).
 161. Blin, K. *et al.* antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **45**, W36–W41 (2017).

162. Ristow, H., Salnikow, J. & Kleinkauf, H. Biosynthesis of valinomycin. *FEBS Lett.* **42**, 127–130 (1974).
163. Cheng, Y.-Q. Deciphering the Biosynthetic Codes for the Potent Anti-SARS-CoV Cyclodepsipeptide Valinomycin in *Streptomyces tsusimaensis* ATCC 15141. *ChemBioChem* **7**, 471–477 (2006).
164. Matter, A. M., Hoot, S. B., Anderson, P. D., Neves, S. S. & Cheng, Y.-Q. Valinomycin Biosynthetic Gene Cluster in *Streptomyces*: Conservation, Ecology and Evolution. *PLoS One* **4**, e7194 (2009).
165. Julin, D. A. Blue/ White Selection. in *Molecular Life Sciences* (eds. Wells, R. D., Bond, J. S., Klinman, J. & Masters, B. S. S.) 72–73 (Springer New York, 2018). doi:10.1007/978-1-4614-1531-2
166. MacNeil, D. J. *et al.* Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**, 61–68 (1992).
167. MacNeil, D. J. Characterization of a unique methyl-specific restriction system in *Streptomyces avermitilis*. *J. Bacteriol.* **170**, 5607–5612 (1988).
168. Sheng, Y., Mancino, V. & Birren, B. Transformation of *Escherichia coli* with large DNA molecules by electroporation. *Nucleic Acids Res.* **23**, 1990–6 (1995).
169. Myronovskiy, M. & Luzhetskyy, A. Native and engineered promoters in natural product discovery. *Nat. Prod. Rep.* **33**, 1006–1019 (2016).
170. Zhou, H. *et al.* Alkaloids from an endophytic *Streptomyces* sp. YIM66017. *Nat. Prod. Commun.* **8**, 1393–6 (2013).
171. Paulus, C. *et al.* New Alpiniamides from *Streptomyces* sp. IB2014/011-12 Assembled by an Unusual Hybrid non-ribosomal Peptide Synthetase *trans*-AT Polyketide Synthase Enzyme. *Front. Microbiol.* **9**, 1–15 (2018).
172. Klementz, D. *et al.* StreptomeDB 2.0—an extended resource of natural products produced by streptomycetes. *Nucleic Acids Res.* **44**, D509–D514 (2016).
173. Yadav, G., Gokhale, R. S. & Mohanty, D. Computational Approach for Prediction of Domain Organization and Substrate Specificity of Modular Polyketide Synthases. *J. Mol. Biol.* **328**, 335–363 (2003).

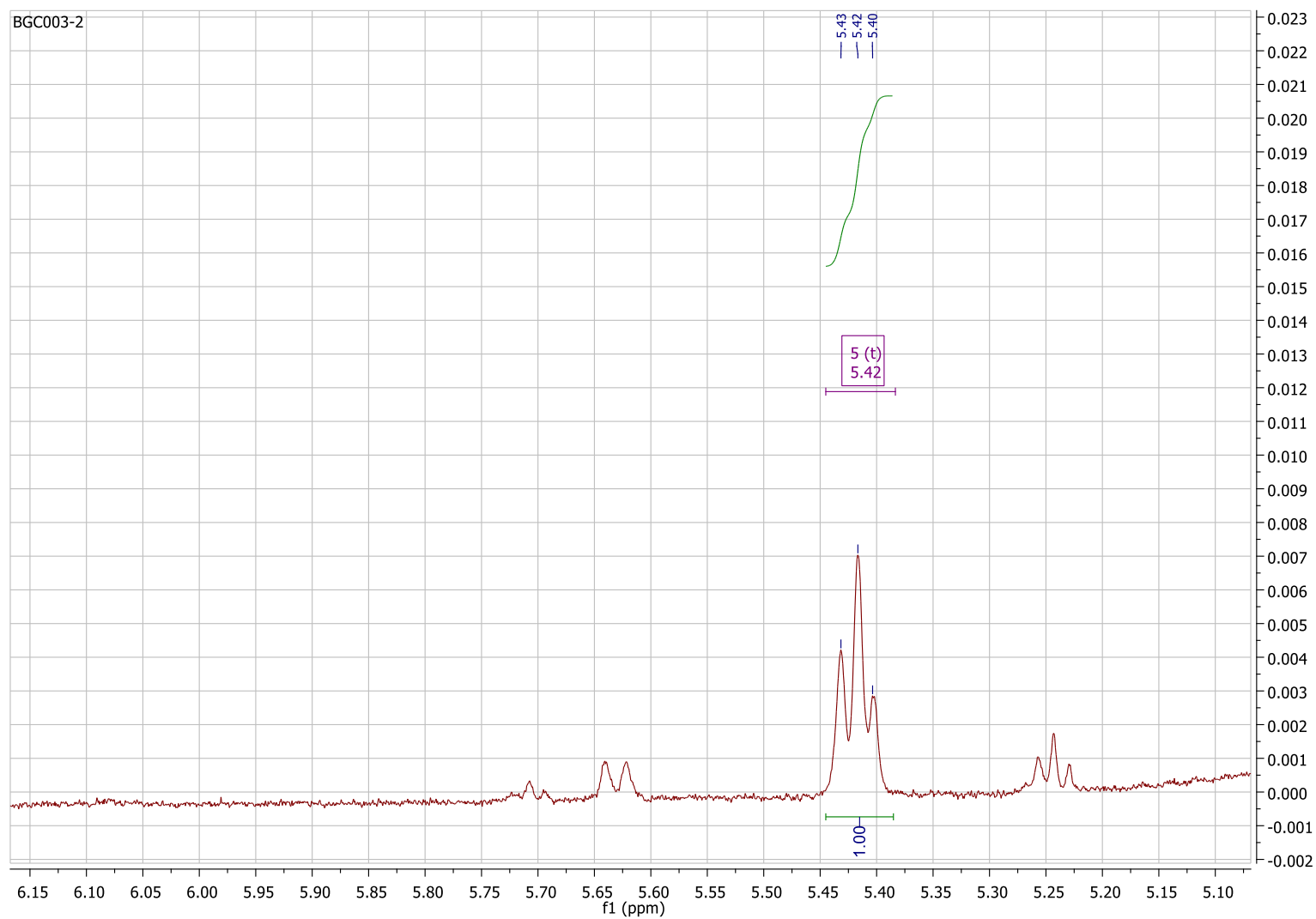
174. Minowa, Y., Araki, M. & Kanehisa, M. Comprehensive Analysis of Distinctive Polyketide and Nonribosomal Peptide Structural Motifs Encoded in Microbial Genomes. *J. Mol. Biol.* **368**, 1500–1517 (2007).
175. Rausch, C. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res.* **33**, 5799–5808 (2005).
176. Röttig, M. *et al.* NRPSpredictor2—a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **39**, W362–W367 (2011).
177. Mahenthiralingam, E. *et al.* Enacyloxins are Products of an Unusual Hybrid Modular Polyketide Synthase Encoded by a Cryptic *Burkholderia ambifaria* Genomic Island. *Chem. Biol.* **18**, 665–677 (2011).
178. Dorival, J. *et al.* Insights into a dual function amide oxidase/macrocyclase from lankacidin biosynthesis. *Nat. Commun.* **9**, 3998 (2018).
179. Kosol, S., Jenner, M., Lewandowski, J. R. & Challis, G. L. Protein–protein interactions in *trans*-AT polyketide synthases. *Nat. Prod. Rep.* **35**, 1097–1109 (2018).
180. Al-Dhelaan, R. *et al.* Condensation-Incompetent Ketosynthase Inhibits *trans*-Acyltransferase Activity. *ACS Chem. Biol.* **14**, 304–312 (2019).
181. Song, L. *et al.* Discovery and Biosynthesis of Gladiolin: A *Burkholderia gladioli* Antibiotic with Promising Activity against *Mycobacterium tuberculosis*. *J. Am. Chem. Soc.* **139**, 7974–7981 (2017).
182. Wu, C.-Y. *et al.* Small molecules targeting severe acute respiratory syndrome human coronavirus. *Proc. Natl. Acad. Sci.* **101**, 10012–10017 (2004).
183. Pettit, G. R. *et al.* Antineoplastic agents. Part 409: Isolation and structure of montanastatin from a terrestrial Actinomycete. *Bioorg. Med. Chem.* **7**, 895–899 (1999).
184. Ryoo, I.-J. *et al.* Selective Cytotoxic Activity of Valinomycin against HT-29 Human Colon Carcinoma Cells via Down-Regulation of GRP78. *Biol. Pharm. Bull.* **29**, 817–820 (2006).
185. Pimentel-Elardo, S. M. *et al.* Anti-Parasitic Compounds from *Streptomyces* sp.

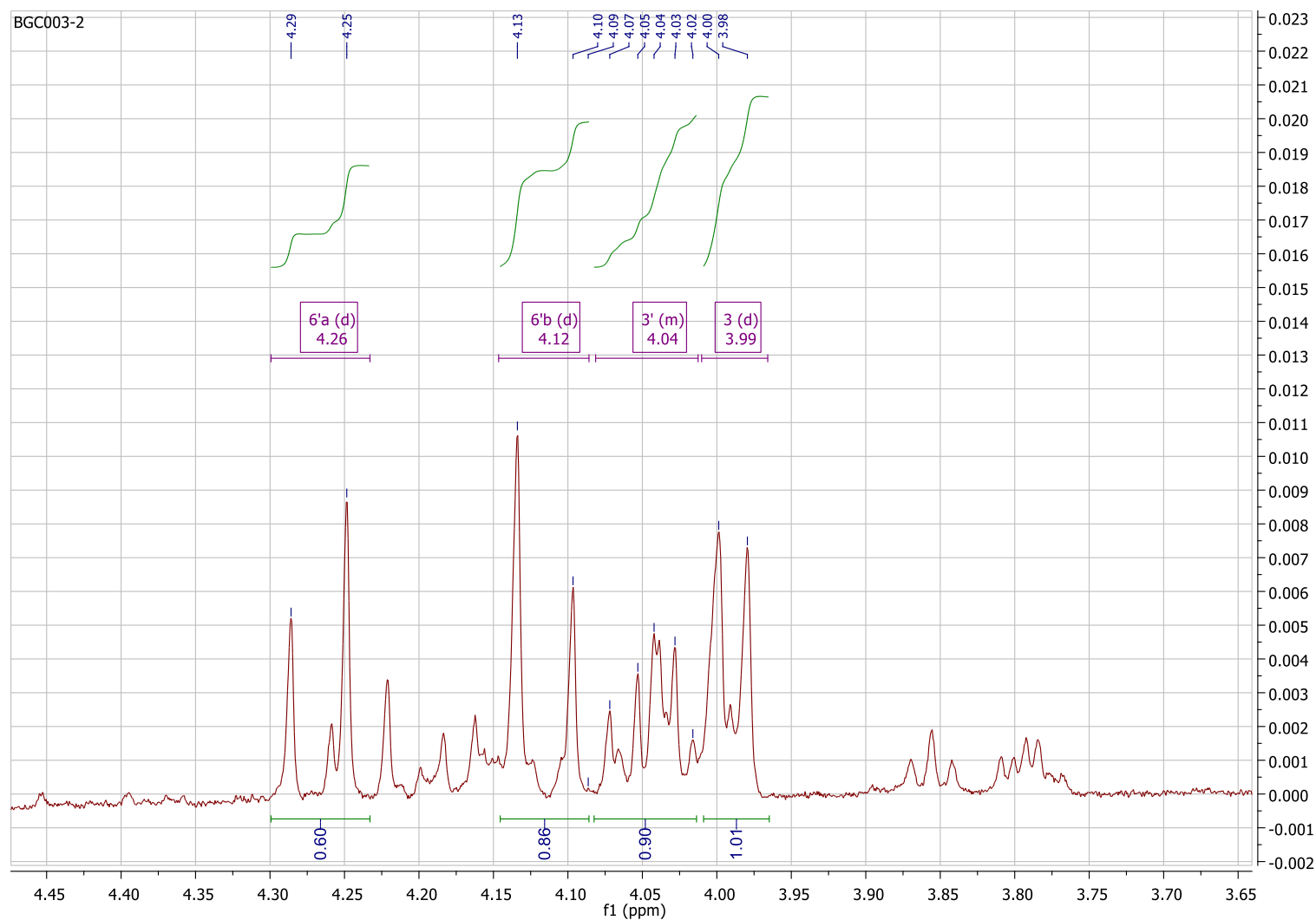
- Strains Isolated from Mediterranean Sponges. *Mar. Drugs* **8**, 373–380 (2010).
186. Ladeuze, S., Lentz, N., Delbrassinne, L., Hu, X. & Mahillon, J. Antifungal Activity Displayed by Cereulide, the Emetic Toxin Produced by *Bacillus cereus*. *Appl. Environ. Microbiol.* **77**, 2555–2558 (2011).
 187. Preto, I. D. Identificação do potencial biotecnológico de microrganismos endofíticos na produção de compostos inseticidas e biorremediação. (Universidade de São Paulo (USP), 2018). doi:10.11606/D.64.2018.tde-13112018-144821
 188. Iacobazzi, R. M. *et al.* Antitumor Potential of Conjugable Valinomycins Bearing Hydroxyl Sites: In Vitro Studies. *ACS Med. Chem. Lett.* **4**, 1189–1192 (2013).
 189. Daoud, S. S. & Juliano, R. L. Reduced Toxicity and Enhanced Antitumor Effects in Mice of the Ionophoric Drug Valinomycin When Incorporated in Liposomes. *Cancer Res.* **46**, 5518–5523 (1986).
 190. Jaitzig, J., Li, J., Süssmuth, R. D. & Neubauer, P. Reconstituted Biosynthesis of the Nonribosomal Macrolactone Antibiotic Valinomycin in *Escherichia coli*. *ACS Synth. Biol.* **3**, 432–438 (2014).
 191. Li, J., Jaitzig, J., Hillig, F., Süssmuth, R. & Neubauer, P. Enhanced production of the nonribosomal peptide antibiotic valinomycin in *Escherichia coli* through small-scale high cell density fed-batch cultivation. *Appl. Microbiol. Biotechnol.* **98**, 591–601 (2014).
 192. Li, J. *et al.* Type II thioesterase improves heterologous biosynthesis of valinomycin in *Escherichia coli*. *J. Biotechnol.* **193**, 16–22 (2015).
 193. Li, J., Jaitzig, J., Lu, P., Süssmuth, R. D. & Neubauer, P. Scale-up bioprocess development for production of the antibiotic valinomycin in *Escherichia coli* based on consistent fed-batch cultivations. *Microb. Cell Fact.* **14**, 83 (2015).
 194. Yang, J. Y. *et al.* Molecular Networking as a Dereplication Strategy. *J. Nat. Prod.* **76**, 1686–1699 (2013).
 195. Rose, L. & Jenkins, A. T. A. The effect of the ionophore valinomycin on biomimetic solid supported lipid DPPE/EPC membranes. *Bioelectrochemistry* **70**, 387–393 (2007).

196. Wang, M. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **34**, 828–837 (2016).
197. Duncan, K. R. *et al.* Molecular Networking and Pattern-Based Genome Mining Improves Discovery of Biosynthetic Gene Clusters and their Products from *Salinispora* Species. *Chem. Biol.* **22**, 460–471 (2015).
198. Reynolds, K. A. *et al.* Isolation and structure elucidation of lipopeptide antibiotic taromycin B from the activated taromycin biosynthetic gene cluster. *J. Antibiot. (Tokyo)*. **71**, 333–338 (2018).
199. Luhavaya, H., Sigrist, R., Chekan, J. R., McKinnie, S. M. K. & Moore, B. S. Biosynthesis of L-4-Chlorokynurenine, an Antidepressant Prodrug and a Non-Proteinogenic Amino Acid Found in Lipopeptide Antibiotics. *Angew. Chemie Int. Ed.* **58**, 8394–8399 (2019).

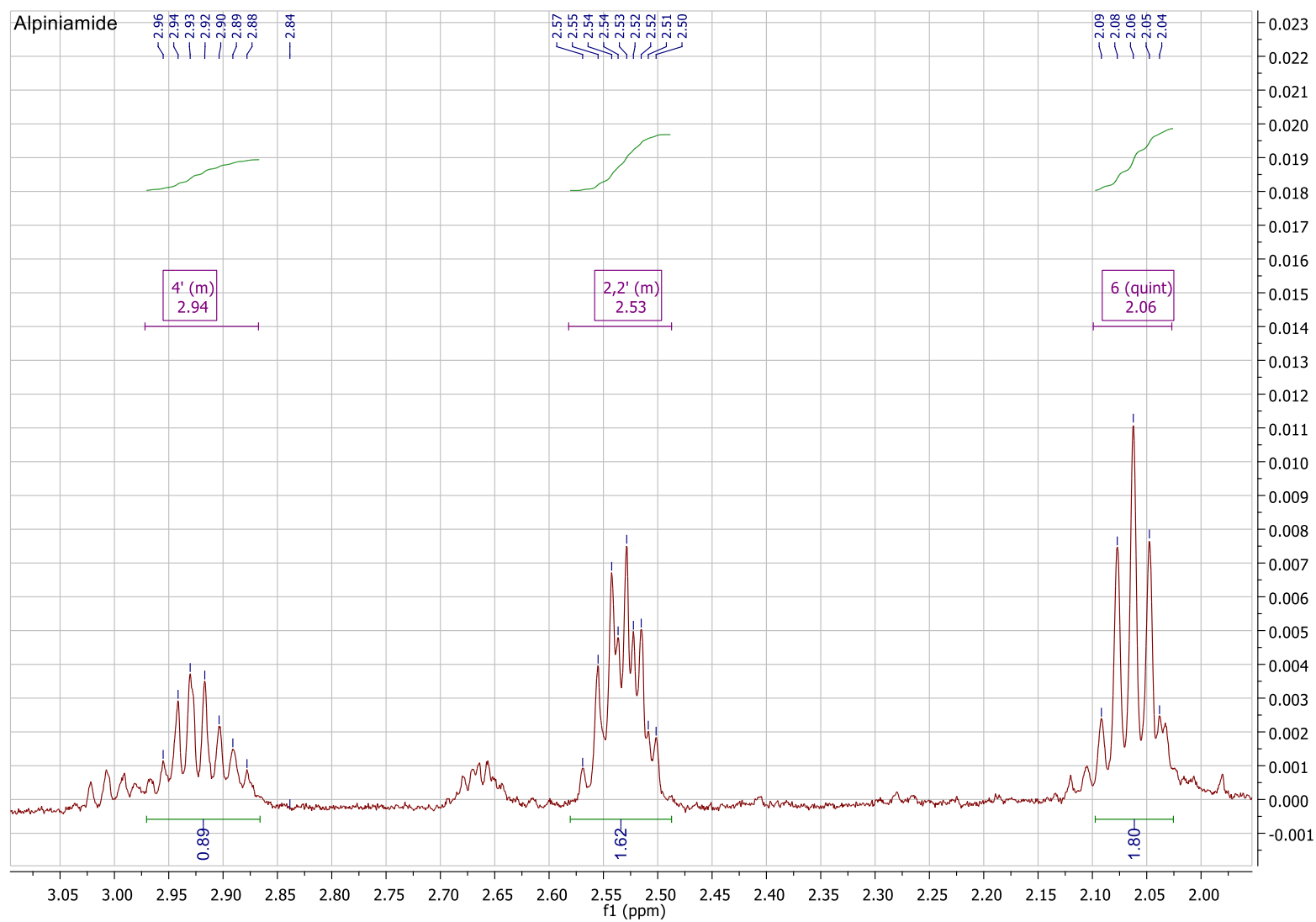
8. APÊNDICE

8.1. Expansão 1 do espectro de RMN de ^1H da alpiniamida (MetOH- d_4 , 500 MHz).

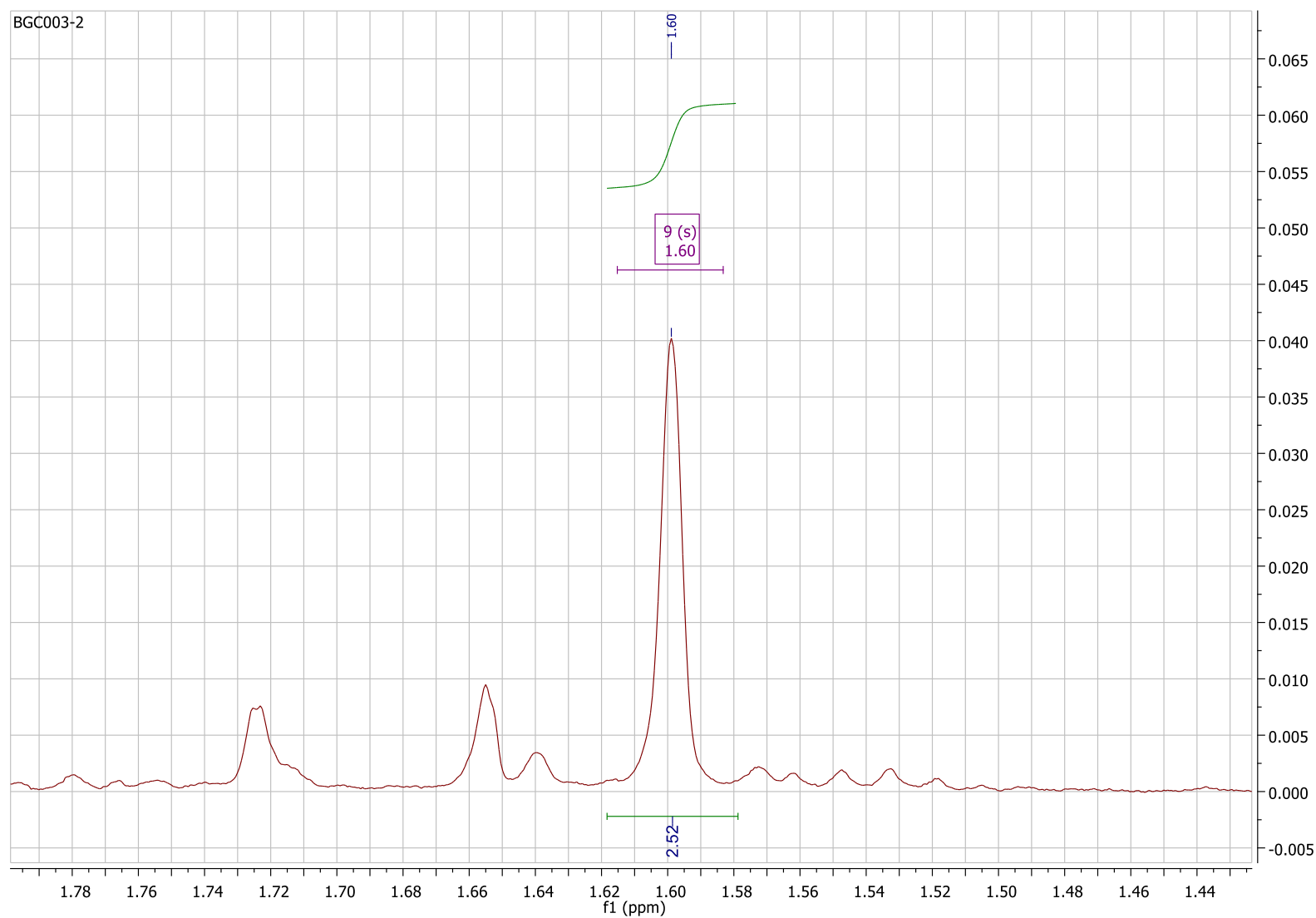


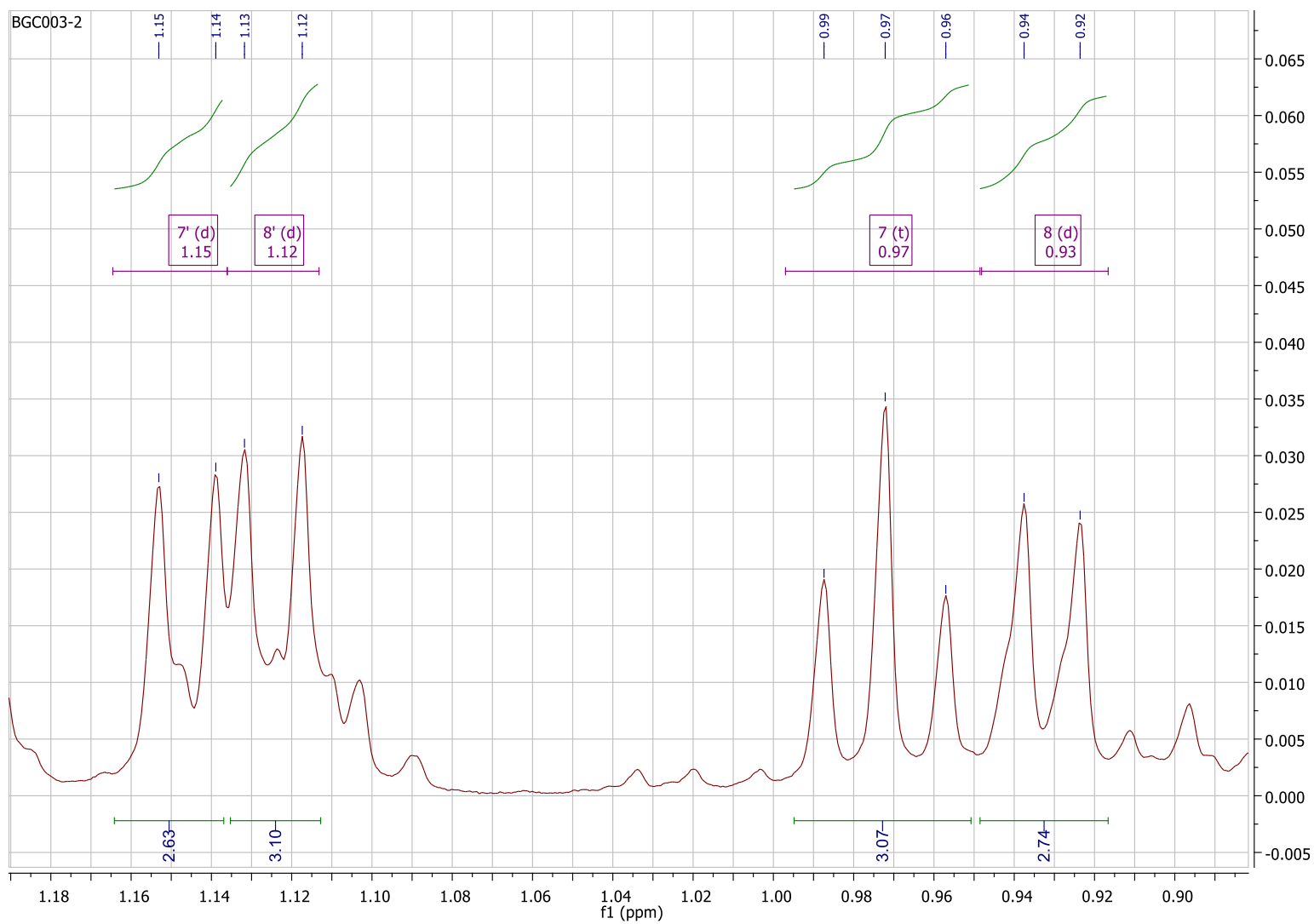
8.2. Expansão 2 do espectro de RMN de ^1H da alpiniamida (MetOH- d_4 , 500 MHz).

8.3. Expansão 3 do espectro de RMN de ^1H da alpiniamida (MetOH- d_4 , 500 MHz).

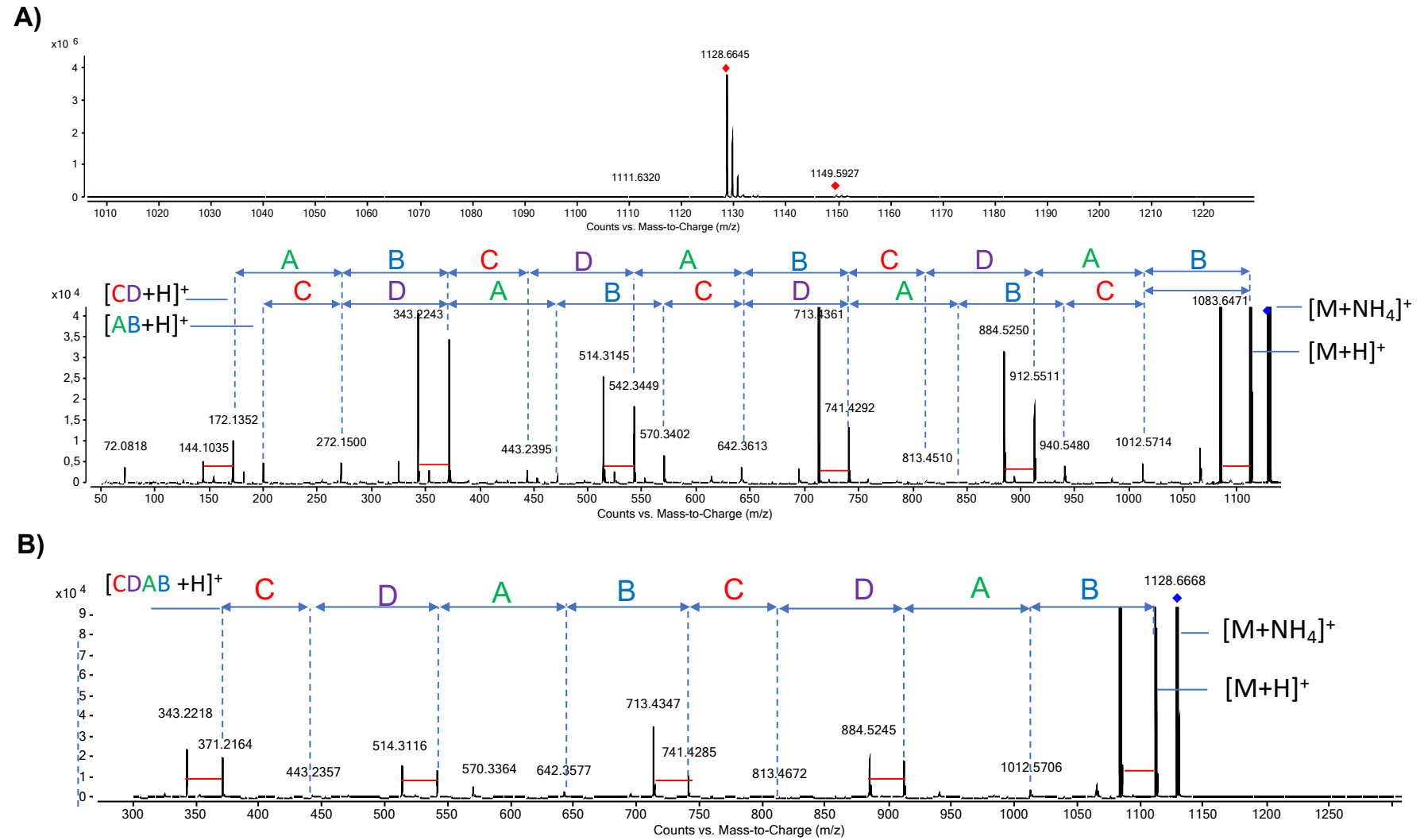


8.4. Expansão 4 do espectro de RMN de ^1H da alpiniamida (MetOH- d_4 , 500 MHz).

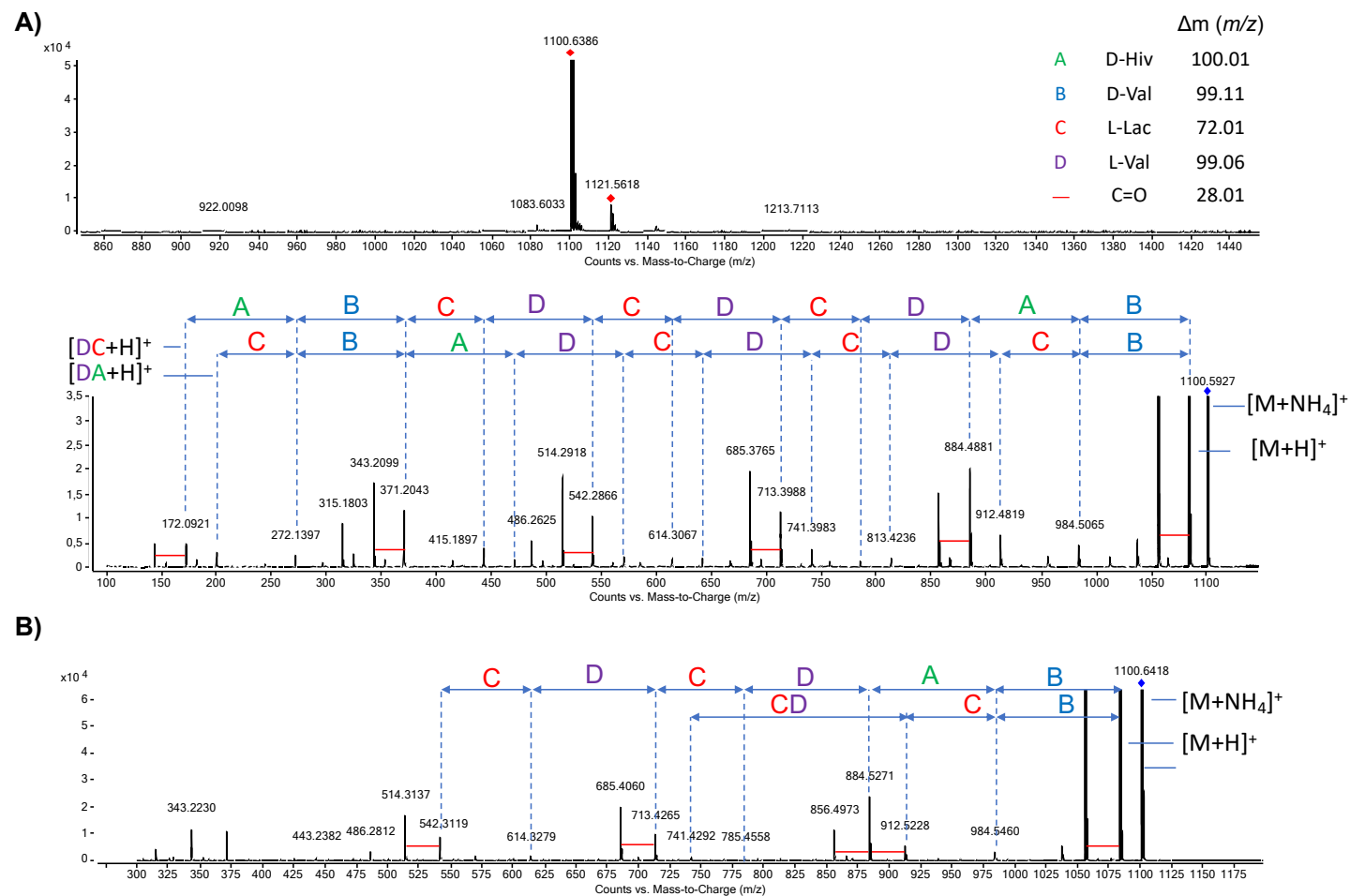


8.5. Expansão 5 do espectro de RMN de ^1H da alpiniamida (MetOH- d_4 , 500 MHz).

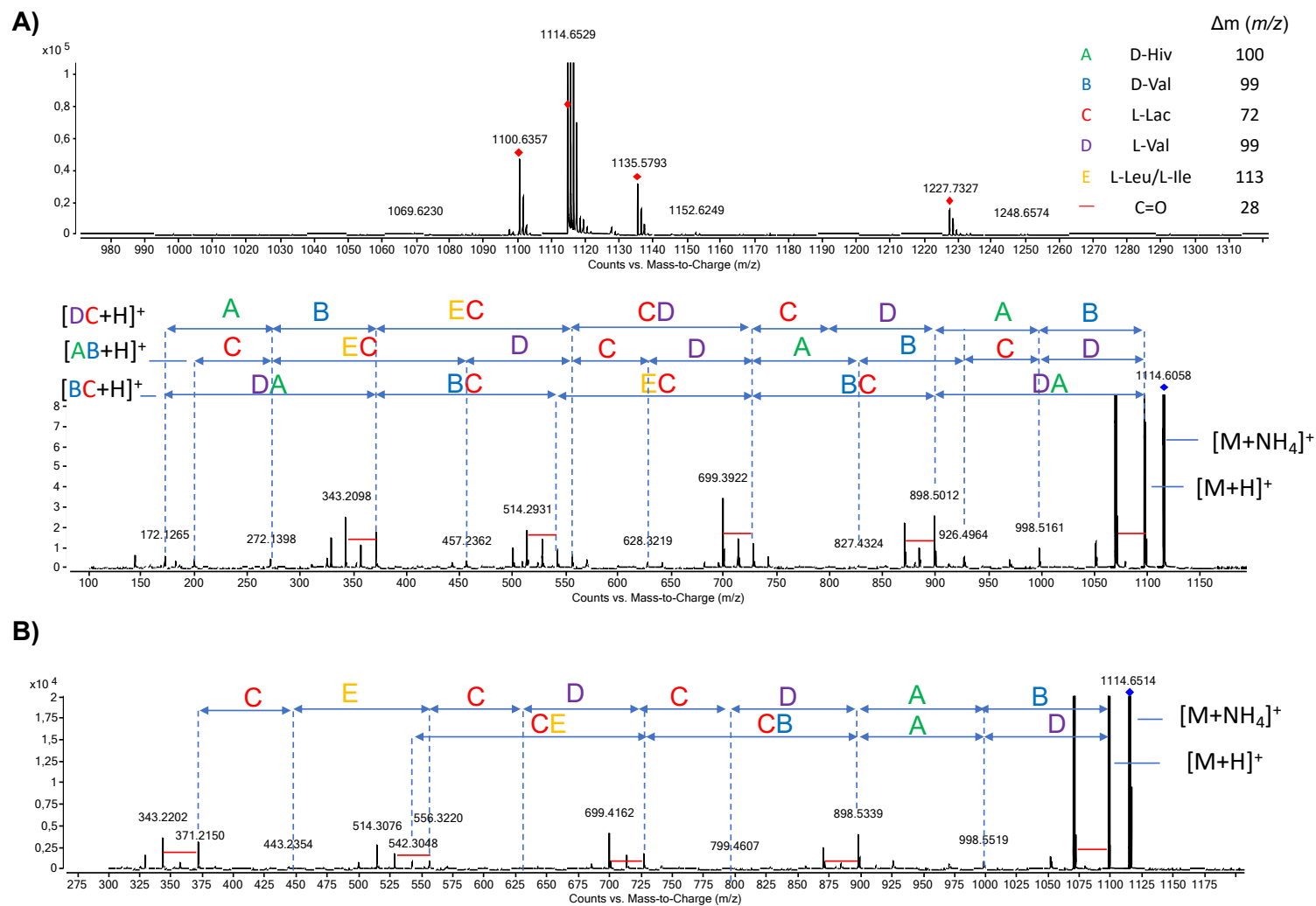
8.6. A. MS e MS/MS em baixa resolução da valinomicina $[M+NH_4]^+$. **B.** MS/MS em alta resolução.



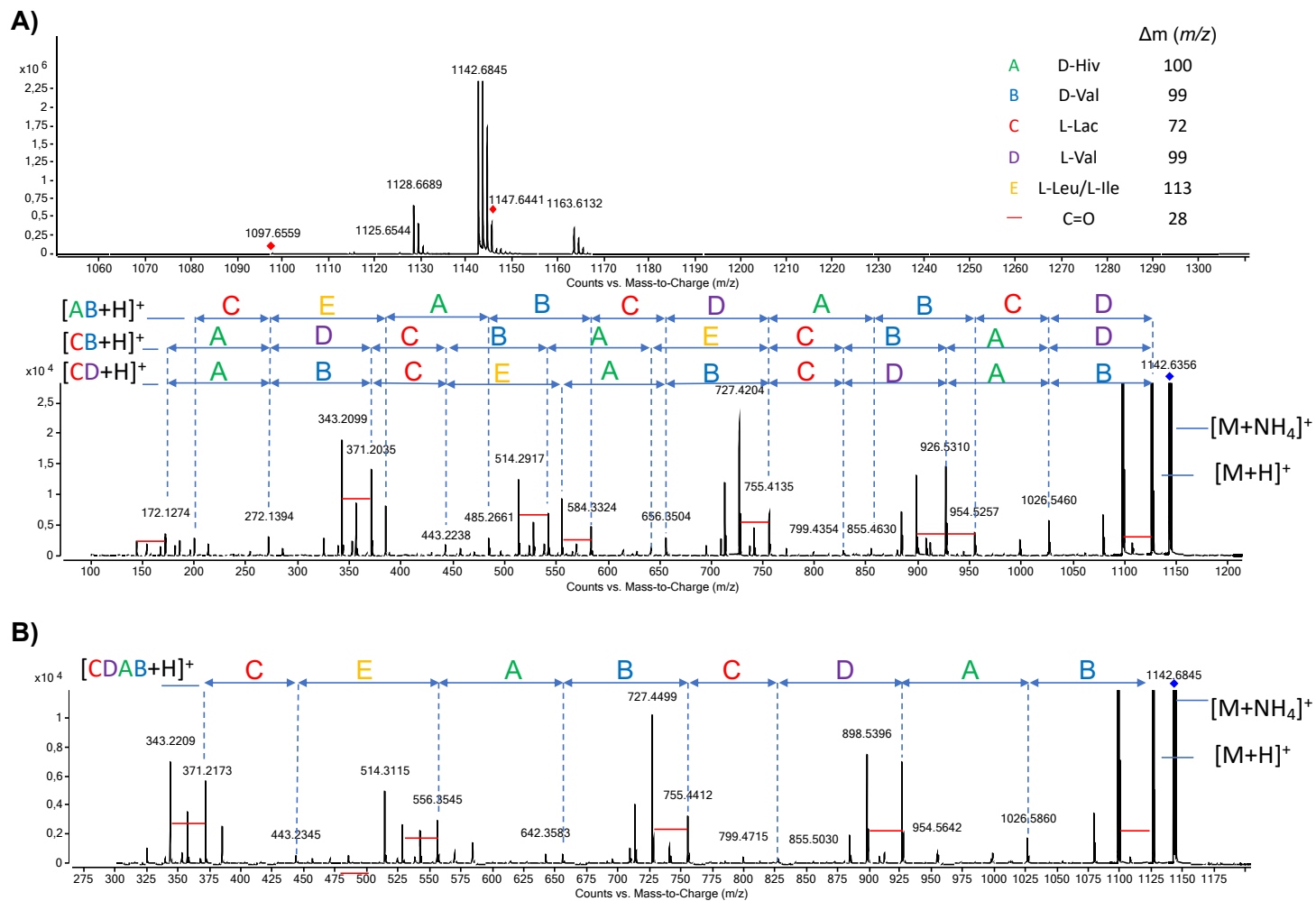
8.7. A. MS e MS/MS em baixa resolução do análogo de m/z 1100,6418 $[M+NH_4]^+$. **B.** MS/MS alta resolução.



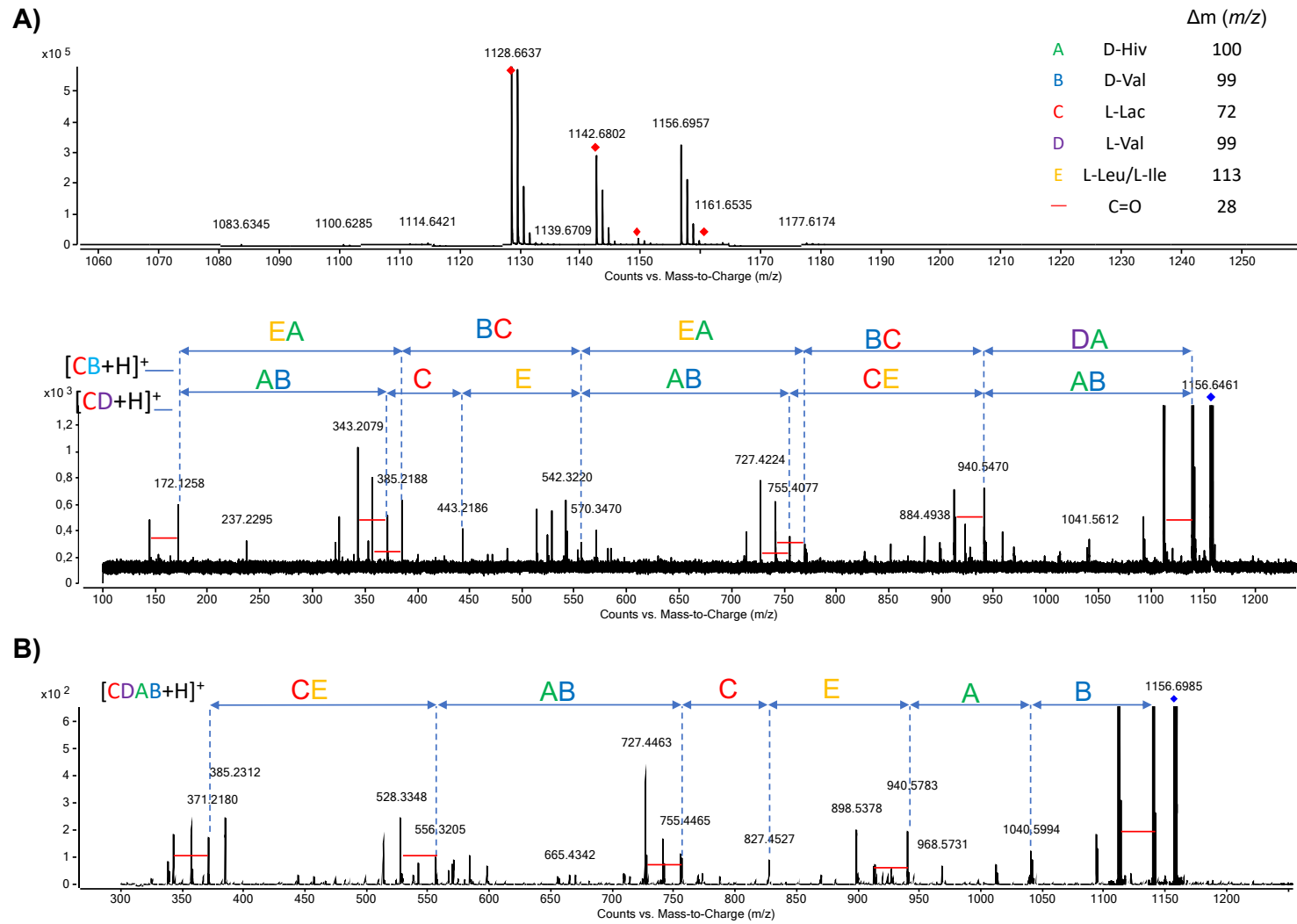
8.8. A. MS e MS/MS em baixa resolução do análogo de m/z 1114,6514 $[M+NH_4]^+$. **B.** MS/MS alta resolução.



8.9. A. MS e MS/MS em baixa resolução do análogo de m/z 1142,6845 $[M+NH_4]^+$. **B.** MS/MS alta resolução.



8.10. A. MS e MS/MS em baixa resolução do análogo de m/z 1142,6845 $[M+NH_4]^+$. **B.** MS/MS alta resolução.



9. ANEXOS

9.1. Oligonucleotídeos sintéticos (*primers*)

Nome	Sequência (5' → 3')
Fragmentos <i>cluster</i> t1PKS-<i>trans</i>-AT PKS-NRPS	
Primer_Inicio_F_1	GGTACACGGAGTACGGCATT
Primer_Inicio_R_1	CCGAAGAACTTCGACCGCTC
Contig51_Final_F3	GGTCCAGTTCCGAGAGGTC
Contig51_Final_R3	ATCTGTAGTCCCGGATCGAA
Contig51_Final_F	CCGGTGGTAGGTGACCGGGCCGATCAGCGCGGCAACGGTCGCC
Contig51_Inicio_R	GCTTCTCCTCCGCTTGCGGGGCGCAGCCGGTCAGGGCGG
pSET_contig51_Inicio_R	CGCGGGTCGACGCCCTGCCGCTCGGATCCTCTAGAGTCGACCTGC
pSET_contig51_Inicio_F	GCAGGTGCACTCTAGAGGATCCGAGCGGCAGGGCGTCGACCCGCG
Contig51_pSET_Final_R	CGATATCGCGCGCGGCCGCGGATCCGCGGGCCTCACCGGAGTC
Contig51_3B_R	GGTAGATCCGCAGGTAGGTGTGACACCCGCGAGG
Contig51_3_F	GGAGAGTTCGACATGGACAACACCACC
Contig51_3_R	GGAGACGGAGCTGAGCAGAGAACTGG
Contig51_4_F	CCTCGGTGACTCCTTCGACCTTCTTGG
Contig51_4_R	CGTTTCGGGCTCGTCGGACCGACCG
Contig51_5_F	GGATGTTCCCGAGGTCCCGTCGTGG
Contig51_5_R	GGACATCGAAGCGTTCTACGCACGTGC
Contig51_6_F	CGACGACGACCTGCAGTGCCGCGTCG
Contig51_6_R	GGAGGGATTCCGCTCGATCGCCTTCC
Contig51_7_F	CGACGTGGGCACGCAGCCGGTCTCC
Contig51_7_R	CGATGAACTGCGCACGCTCGGAGTGC
Contig51_8_F	CGTACTCGTGGACGTGGTGACGCTGG
Contig51_8_R	CCTTCGACCGCATCCTGCACTGCCTCG
Contig51_9_F	GGAAGTACCGGTCCACCGCGAGGACC
Contig51_9_R	GGTCGTTGGTACGACGGGGTGAAGC
Contig51_10B_F	CCTGGAGTTCGAGACTCCAGTGACGAAGGTGCGCGGCGACGATGC
Triagem da biblioteca genômica	
Primer_Inicio_F_1	GGTACACGGAGTACGGCATT
Primer_Inicio_R_1	CCGAAGAACTTCGACCGCTC
Primer_Final_F_4	CTGCTCCTCGTCTACGTGC
Primer_Final_R_4	CACCAGGACGGAGTGCATG
Contig51_Final_F3	GGTCCAGTTCCGAGAGGTC
Contig51_Final_R3	ATCTGTAGTCCCGGATCGAA
Valinomycin_Left_F	GCTCGGCAGACCCGCCGACGAGGTC
Valinomycin_Left_R	GGAGCCCCAGGCCAGGGGCGAGAGC
Valinomycin_Center_F	CGCCGTCGACCGTCTCGACCACGGC
Valinomycin_Center_R	GGAGCGGTACGGGCTGCCCGGGCAG
Valinomycin_Right_F	GAAGCGGCGGTGCGCCGTATCCGCG
Valinomycin_Right_R	GCCGCACCCCCGCCGCCCTCATT
Ctg51_PAC_Left_R	ACCTTCAGCAGCTTCTCCAC
Ctg51_PAC_Left_F	CGTAACCGCCTCAGACTGT
Ctg51_PAC_Center_F	CGACGTGGGCACGCAGCCGGTCTCC
Ctg51_PAC_Center_R	GGAGGGATTCCGCTCGATCGCCTTCC
Ctg51_PAC_Right_F	CTGTTTCTCGTTCTACCAGGG

Ctg51_PAC_Right_R	GTTCGTGATGTCCAGGAAGAC
TAR-cloning	
TAR_BGC3_01_F	GTAGGATGAGCAATCTAGTCGAGCAACGGAGGTACGGACGAATTCC GAGGAGGAGGCGGCATGAGCGAAGACCCC
TAR_BGC3_01_R	GCAGCGTGTTGATGAAGCAGCCGACGATCGGGTCTGAAGTGCTCCG TCGGG
TAR_BGC3_02_F	GGAGCTTCTTCCTGGCGGCGTTCCTGAC
TAR_BGC3_02_R	CGTGCGACCGGACAGGAACGTCTCGA
TAR_BGC3_03_F	GGTACGGGTGAGCCTGCCGACGTATCC
TAR_BGC3_03_R	GTACCCTGCGGCCAGGAGTGCTTTTAC
TAR_BGC3_04_F	CGGAGCTCGACCATCTCCTCTGGGTGG
TAR_BGC3_04_R	GCATCAGTCCCAGTCGACGGACGTGC
TAR_BGC3_05_F	GTTGGCTCGACGGAGATCCGGACACGC
TAR_BGC3_05_R	GCAGGCCATGGACACGATCACGACGGG
TAR_BGC3_06_F	GCAAGCGATAGCGACCCTGCTGCTGAC
TAR_BGC3_06_R	CCTCCACCAAGCCGTCGTCGAGCACAC
TAR_BGC3_07_F	GTTCGTGCGGGAGCTCGAAGCCGAGGGCGCCGAGCGGG
TAR_BGC3_07_R	GGATGAGTAGCAGCACGTTCCCTTATATGTAGCTTTCGACATATGCGT CGACGGCGATGAAGGTCTGTGCGAGCTT
BGC03_117_176bp_F	GCTGCGGACGCTTCAGGTGT
BGC03_117_176bp_R	GCAGGACGCGCTCAGGTTGT
BGC03_118_153bp_F	TCCGGGCGCCATGGACTACA
BGC03_118_153bp_R	GTGTGCCCGATGCAGGGCTT
BGC03_119_168bp_F	ATCACCGGCGGGCAGGTTTG
BGC03_119_168bp_R	CAGCACCTTGACGCCCTCGG
Deleção domínio de adenilação	
B1_cluster14_FI	TGATCAAGGCGAATACTTCATATGGAAACTGCTCAACTGGCAGATG GT
B1_cluster14_RI	GCTGATCGTCGAGTGGTACTTCCAGTTCACCAGGAACC
B1_cluster14_FII	CTGGAAGTACCACTCGACGATCAGCTACGTCTGAAGCCCACGGCAC G
B1_cluster14_RII	CCGCGCGGTGCGATCCCCGCATATGCCAGCATGTGACCGTGTCGA ACC
Del_triagem_F	GACGTGCTGGAGCACGGAAACTACC
Del_triagem_R	GAAGAACGGCGAATCGTCCCAGTCG

9.2. Certificado da Comissão Interna de Biossegurança – CIBio



UNICAMP

COMISSÃO INTERNA DE BIOSSEGURANÇA
UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA

Cx. Postal 6154 - Campinas - 13083-970 SP - Brasil
<http://www.iqm.unicamp.br>



Solicitação/Protocolo: LGO2019-1

Nome do(a) aluno(a): Renata Sigris

Nível: Doutorado

Período: 01/08/2015 a 31/07/2019

Nome do(a) orientador(a): Profa. Dra. Luciana Gonzaga de Oliveira

Título: Identificação e caracterização de uma t1PKS - trans-AT PKS - NRPS em *Streptomyces* sp via Genome Mining

Resumo: A análise recente do genoma de diversas linhagens de *Streptomyces* revelou uma grande abundância de clusters de genes que codificam enzimas do metabolismo secundário. Essas informações conferem uma importante evidência de que a capacidade para a produção de metabólitos por *Streptomyces* e outras actinobactérias não é completamente acessada sob as condições de cultivo em laboratório. Como uma alternativa às metodologias tradicionais de busca por novos metabólitos, a aplicação de ferramentas genômicas iniciadas principalmente neste século, tem possibilitado o isolamento de novas moléculas, a elucidação de vias biossintéticas, a caracterização de enzimas e suas funções e também a identificação de intermediários de biossíntese. A linhagem endofítica *Streptomyces* sp. CBMAI 2042, isolada de *Citrus sinensis* e sequenciada pelo nosso grupo de pesquisas foi selecionado como tema do projeto. Análise *in silico* dos genes anotados codificando para o metabolismo secundário permitiu identificar um cluster de genes envolvido na biossíntese de um metabólito, sem precedentes em bases de dados, por ação conjunta de enzimas PKS do tipo 1, trans-AT PKS e NRPS. Através da mesma análise, foram também identificados genes responsáveis pela biossíntese de valinomicina, composto já descrito na literatura. Neste estudo propõe-se utilizar diferentes metodologias para decifrar o metabólito produzido pela ação da t1PKS-trans-AT PKS-NRPS e caracterizar o cluster codificando para valinomicina. Linhagens: 1) Biblioteca em vetor do tipo PAC (P1-derived Artificial Chromosome) de *Streptomyces* sp CBMAI 2042 - Classe de Risco I; 2) recombinação/reconstrução *in vivo* e *in vitro*, a partir de fragmentos de DNA - linhagens recombinantes - Classe de Risco I 3) transferência dos plasmídeos para organismo heterólogo (*S. coelicolor* M1146/M1152/M1154) - linhagens heterólogas Classe de Risco I. Volumes de cultivo: variam de 50 a 500 mL para as diferentes linhagens descritas. Descarte e limpeza: os cultivos são descontaminados por adição de hipoclorito, mantidos por até 24h para que o agente bactericida atue completamente. Em seguida, o descarte é feito em pia com água corrente. Placas de cultivo de meio sólido são autoclavadas e posteriormente descartadas em lixo comum. Equipamentos: autoclave; fluxo laminar; shaker rotativo para culturas líquidas (30°C e 37°C); estufa (30°C e 37°C); máquina de PCR; cuba de eletroforese horizontal (gel de agarose - DNA); termomixer.

O projeto de pesquisa acima descrito, a ser conduzido pelo(a) aluno(a) de doutorado Renata Sigris, recebeu autorização desta Comissão Interna de Biossegurança.

Presidente da CIBio-IQ: Prof. Dr. José Alberto Fracassi da Silva
04/04/2019

cibioiq@unicamp.br
Telefone: 19 3521 3002

9.3. Cadastro no Sistema Nacional de Gestão do Patrimônio Genético



Ministério do Meio Ambiente
CONSELHO DE GESTÃO DO PATRIMÔNIO GENÉTICO
 SISTEMA NACIONAL DE GESTÃO DO PATRIMÔNIO GENÉTICO E DO CONHECIMENTO TRADICIONAL ASSOCIADO

Certidão
Cadastro nº A4943D4

Declaramos, nos termos do art. 41 do Decreto nº 8.772/2016, que o cadastro de acesso ao patrimônio genético ou conhecimento tradicional associado, abaixo identificado e resumido, no Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado foi submetido ao procedimento administrativo de verificação e não foi objeto de requerimentos admitidos de verificação de indícios de irregularidades ou, caso tenha sido, o requerimento de verificação não foi acatado pelo CGen.

Número do cadastro: **A4943D4**
 Usuário: **UNICAMP**
 CPF/CNPJ: **46.068.425/0001-33**
 Objeto do Acesso: **Patrimônio Genético**
 Finalidade do Acesso:
☒ **Pesquisa Científica** ☐ **Bioprospecção** ☐ **Desenvolvimento Tecnológico**

Espécie

Streptomyces sp
Streptomyces sp
 Título da Atividade: **Mineração do Genoma de Streptomyces**

Equipe

Luciana Gonzaga de Oliveira	UNICAMP
Renata Sigrist	IQ-UNICAMP
Douglas Cunha Sachito	IQ-UNICAMP
Bruno Sacchetto Paulo	IQ-UNICAMP
Gabriel Rodrigues Martos	PUC-Campinas
Ana Beatriz Gonçalves	IQ-UNICAMP
Gonzalo García Delgado	IQ-UNICAMP
Markiyan Samborskyy	University of Cambridge

Parceiras no Exterior

Department of Biochemistry - University of Cambridge

Data do Cadastro: **05/11/2018 23:11:19**
 Situação do Cadastro: **Concluído**



Conselho de Gestão do Patrimônio Genético

Situação cadastral conforme consulta ao SisGen em 14:57 de 25/03/2019.



SISTEMA NACIONAL DE GESTÃO
 DO PATRIMÔNIO GENÉTICO
 E DO CONHECIMENTO TRADICIONAL
 ASSOCIADO - **SISGEN**