



**UNIVERSIDADE ESTADUAL PAULISTA**  
**“JÚLIO DE MESQUITA FILHO”**  
Campus de São José do Rio Preto

**Tábata Peres Cardoso**

**Elucidação estrutural da especificidade por substratos de enzimas da  
biossíntese de aminoglicosídeos e marginolactonas**

**São José do Rio Preto**

**2020**

**Tábata Peres Cardoso**

**Elucidação estrutural da especificidade por substratos de enzimas da  
biossíntese de aminoglicosídeos e marginolactonas**

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Microbiologia, junto ao programa de Pós-graduação em Microbiologia, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

Financiadora: CAPES

**Orientador: Prof. Dr. Marcio Vinicius Bertacine  
Dias**

**São José do Rio Preto**

**2020**

C268e Cardoso, Tábata Peres  
Elucidação estrutural da especificidade por substratos de enzimas da biossíntese de aminoglicosídeos e marginolactonas / Tábata Peres Cardoso. -- , 2020  
102 f. : il., tabs.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto

Orientador: Marcio Vinícius Bertacine Dias

1. Enzimas. 2. Estrutura. 3. Aminoglicosídeos. 4. Marginolactonas.  
I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências, Letras e Ciências Exatas, São Jose do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Tábata Peres Cardoso

**Elucidação estrutural da especificidade por substratos de enzimas da  
biossíntese de aminoglicosídeos e marginolactonas**

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Microbiologia, junto ao programa de Pós-graduação em Microbiologia, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Financiadora: CAPES

**Comissão Examinadora**

Prof. Dr. Marcio Vinícius Bertacine Dias  
UNESP – Câmpus de São José do Rio Preto  
Orientador

Prof. Dr. Gustavo Orlando Bonilla Rodriguez  
UNESP – Câmpus de São José do Rio Preto

Prof(a). Dr(a). Marli de Fátima Fiore  
USP – Câmpus de Piracicaba

Prof(a). Dr(a). Luciana Gonzaga de Oliveira  
UNICAMP – Campinas

Prof(a). Dr(a). Fátima Pereira de Souza  
UNESP – Câmpus de São José do Rio Preto

São José do Rio Preto

26 de Junho de 2020

Dedico este trabalho ao meus pais, Angela e Valter, que demonstram amor incondicional e que acreditam em mim!

## **AGRADECIMENTOS**

Ao único Deus criador e mantenedor da minha vida, que me fez crescer pessoalmente ao longo desse período, estando comigo nas horas alegres e tristes e me fazendo entender que mesmo as coisas não saindo do jeito como eu gostaria, ainda assim, posso confiar que Ele está no comando.

Ao meus pais, Valter e Angela, por estarem sempre comigo, oferecendo colo, afeto e atenção. Sempre me aconselhando para o melhor! Esse suporte emocional foi imprescindível.

Ao meu namorado, Gabriel Angelin, por ter sido atencioso, paciente, carinhoso, oferecendo apoio nas horas delicadas e entusiasmo na diversão. Agradeço a dedicação despendida a mim nesse período.

A amizade fortalecida com Priscila Bury que nunca me negou auxílio nas diversas áreas da vida e sempre esteve disposta a repartir comigo o conhecimento adquirido. Aos meus amigos de laboratório Catharina Santos, Marcelo Barreto, Laura Rivas, João Ribeiro, Gerardo Zúñiga, Sair Máximo, Larissa de Sá, João Pimenta, Jademilson Santos, Fernanda Paiva, Cristiane Tambascia, Pâmela Pena, Natalia Cerrone, Bruno Kaik, Andreza Araújo, Gabriel Stefani, Nicolas, Brenda e a técnica do grupo, Renata Paulo. Vocês foram fundamentais nessa caminhada, apoiando, suportando, incentivando, colaborando. Sou grata pela presença de cada um em minha vida.

Ao meu orientador, Prof<sup>o</sup> Dr<sup>o</sup> Marcio Vinícius Bertacine Dias que aceitou me conduzir nesse trabalho, oferecendo guia nas encruzilhas científicas mais difíceis que passei e por sempre oferecer ao grupo suporte científico movendo o que fosse necessário para que a pesquisa ocorresse. A você Marcio, meu respeito e admiração.

Ao grupo do professor Peter Leadlay pelas construções em especial a Dra. Fanglu Huang e a Dra. Hui Hong.

Ao programa de pós-graduação em Microbiologia do IBILCE/UNESP- Rio Preto e toda seção técnica que sempre me atenderam cordialmente.

Ao departamento de microbiologia do Instituto de Ciências Biológicas da USP –SP pela infraestrutura oferecida durante o desenvolvimento do projeto.

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

## Resumo

Os produtos naturais desempenham um papel importante no desenvolvimento de medicamentos e parte desse sucesso está relacionado com a diversidade química e estrutural apresentadas pelos compostos dotados de relevância biológica. Um exemplo de produtos naturais oriundos do metabolismo secundário de microrganismos são os antibióticos aminoglicosídicos e policetídeos. A complexidade estrutural implica em dificuldades para produzir análogos com atividades medicinais melhoradas. Por outro lado, a biossíntese combinatorial explora a promiscuidade de substratos e emprega enzimas e vias modificadas para produzir novos produtos "não naturais". No presente trabalho, buscamos elucidar estruturalmente enzimas envolvidas na biossíntese de aminoglicosídeos (NeoB, NeoQ, ParB, ParQ, LivB e LivQ) e policetídeos marginolactonas (AMH\_821, AMH\_828, Medi\_4948) com aplicação biotecnológicas. Para isso, genes foram clonados e expressos em *Escherichia coli* e as proteínas recombinantes foram purificadas. Para as proteínas NeoQ e ParQ, foram realizados teste de expressão a fim de encontrar uma condição de maior rendimento proteico. Para a proteína NeoB, foi realizado um ensaio de deslocamento térmico para escolha do melhor tampão para purificação. Cristais das proteínas NeoB, AMH\_828, Medi\_4948 foram obtidos em diferentes condições. Aplicamos métodos de resolução estrutural por substituição molecular para todas as proteínas estudadas. Foram resolvidas as estruturas das proteínas NeoB e Medi\_4948. A estrutura da enzima Medi\_4948 possui um enovelamento característicos de  $\alpha/\beta$  hidrolases, compartilhado por ureohidrolases e foi possível observar que ela é hexamérica, composta por dois trímeros. Esta proteína contém uma alça flexível na região do N-terminal o qual interage com a unidade subjacente do monomero vizinho. O sítio ativo apresenta um cluster binuclear para um íon metal, o qual, juntamente com uma molécula de água é importante para a catálise. A enzima NeoB revelou uma estrutura que contém contém 12 fitas- $\beta$  e 11  $\alpha$ -hélices. Sete fitas- $\beta$  formam uma folha- $\beta$  mista que é bem característico dos membros da superfamília das aspartato aminotransferases. Foi possível também observar um motivo grampo  $\beta$ . Nós mostramos as primeiras informações estruturais obtidas sobre a amidinohidrolase Medi4948 e também conseguimos resultados para outra enzima, uma aminotransferase da biossíntese de neomicina, Neo B.

**Palavras-chave:** Enzimas. Estrutura. Aminoglicosídeos. Marginolactonas.

## Abstract

Natural products play an important role in the development of drugs and part of this success is related to the chemical and structural diversity of compounds with biological relevance. Examples of natural products derived from the secondary metabolism of microorganisms are the aminoglycoside and polyketide antibiotics. Structural complexity, in many cases, makes it difficult to produce analogues with improved pharmacological properties. Combinatorial biosynthesis explores the promiscuity of substrates and makes use of enzymes and modified pathways to produce new "unnatural" products, substantially increasing the structural diversity of natural products with potential pharmaceutical value. In this work, we aimed to elucidate structures of enzymes involved in the aminoglycoside and polyketide biosynthesis with biosynthetic applications. NeoB, NeoQ, ParB, ParQ, LivB and LivQ for aminoglycoside and AMH\_821, AMH\_828, and Medi\_4948 for polyketide biosyntheses. Several genes were cloned and expressed in *Escherichia coli* and the recombinant proteins were purified. For the NeoQ and ParQ proteins, expression tests were performed in order to identify a condition with a higher yield protein production. In addition, for NeoB, thermal shift test was performed to choose the best buffer for purification. Crystals from NeoB proteins, AMH\_828, Medi\_4948 were obtained under different conditions. We apply structural resolution methods by molecular substitution to all studied proteins. The structures of the NeoB and Medi\_4948 proteins were resolved. The structure of the enzyme Medi\_4948 has a folding characteristic of  $\alpha/\beta$  hydrolases, shared by ureohydrolases and it was possible to observe that it is hexameric, composed of two trimers. This protein contains a flexible loop in the N-terminal region which interacts with the adjacent monomer. The active site has a binuclear cluster for a metal ion, which, together with a water molecule, is important for catalysis. NeoB enzyme revealed a structure that contains 12  $\beta$ -tapes and 11  $\alpha$ -helices. Seven  $\beta$ -strands form a mixed  $\beta$ -sheet that is very characteristic in the members of the aspartate aminotransferase superfamily. It was also possible to observe a  $\beta$  clamp motif. We show the first structural information obtained about the amidinohydrolase Medi\_4948 and also obtained results for another enzyme, a neomycin biosynthesis aminotransferase, NeoB.

**Keywords:** Enzymes. Aminoglycoside antibiotics. Marginolactone. Structure.

## Lista de Ilustrações

- FIGURA 1. ESTRUTURA QUÍMICA DOS PRINCIPAIS ANTIBIÓTICOS AMINOGLICOSÍDEOS DA CLASSE DAS 2-DESOXI- ESTREPTAMINAS. A SÃO 4,5 DISSUBSTITUÍDOS, ENQUANTO B E C SÃO DO GRUPO DAS 4,6 DISSUBSTITUÍDOS. .... 22
- FIGURA 2. ANTIBIÓTICOS CUJO ANEL DERIVADO CICLITOL NÃO É A 2-DESOXI-ESTREPTAMINA. .... 23
- FIGURA 3. SÍNTESE DE AMICACINA. PRODUÇÃO DE AMICACINA A PARTIR DE CANAMICINA UTILIZANDO PROCESSOS QUÍMICOS INDICADOS NA FIGURA. .... 25
- FIGURA 4. VIA BIOSSÍNÉTICA PARA PRODUÇÃO DE NEOMICINA E LIVIDOMICINA. SÃO MOSTRADOS OS PASSOS A PARTIR DO INTERMEDIÁRIO COMUM PAROMOMAMINA. .... 28
- FIGURA 5. BIOSÍNTESE DE PSEUDOTETRASACARÍDEOS. EM AZUL É MOSTRADO O GRUPO AMINO QUE É TRANSFERIDO PARA O PSEUDOSACARÍDEO, PELA AÇÃO DAS ENZIMAS NEOB/NEOQ, QUE SÃO BIFUNCIONAIS. EM AMARELO É MOSTRADO O GRUPO AMINO ADICIONADO TANTO PELA ENZIMA NEOB/NEOQ, COMO PARA A ENZIMA LIVB/LIVQ (QUE SÃO MONOFUNCIONAIS). ADAPTADO DE PARK ET AL., 2011. .... 31
- FIGURA 6. ALINHAMENTO DAS SEQUÊNCIAS DAS ENZIMAS SIS5, LIVB, PARB E NEOB, AMINOTRANSFERASES RELACIONADAS COM A BIOSÍNTESE DE ANTIBIÓTICOS AMINOGLICOSÍDEOS. NO ALINHAMENTO, É POSSÍVEL OBSERVAR ALTA IDENTIDADE SEQUENCIAL ENTRE AS ENZIMAS EVIDENCIANDO A SEQUÊNCIA. .... 32
- FIGURA 7. ALINHAMENTO DAS SEQUÊNCIAS DAS ENZIMAS NEOQ, LIVQ, PARQ, OXIDORREDUTASES RELACIONADAS COM A BIOSÍNTESE DE ANTIBIÓTICOS AMINOGLICOSÍDEOS. NO ALINHAMENTO, É POSSÍVEL OBSERVAR A ALTA IDENTIDADE SEQUENCIAL ENTRE AS ENZIMAS. .... 33
- FIGURA 8. EXEMPLOS DE MARGINOLACTONAS. ESTRUTURAS DAS MARGINOLACTONAS ANTIFÚNGICAS DESERTOMICINA (1), PRIMICINA (2), KANCHANAMICINA (3) E AZOLOMICINA (4). .... 36
- FIGURA 9. AMIDINOHIDROLASE DA DESERTOMICINA. AMIDINOHIDROLASE DSTH ATUANDO NO GRUPO GUANIDIOL DA DESERTOMICINA B PARA FORMAÇÃO DA DESERTOMICINA A. ... 37
- FIGURA 10. SEQUÊNCIA DE ALINHAMENTO DAS UREOHIDROLASES. AS SEQUÊNCIAS DE ONZE UREO-HIDROLASES [AMH\_A828: AMIDINO-HIDROLASE DE *STREPTOMYCES OLIVACEUS TÛ4018* (A828); AMH\_A821: AMIDINO-HIDROLASE DE *SACCHAROMONOSPORA CAESIA* (A821); GBUA\_PA: GUANIDINOBUTIRASE DE *PSEUDOMONAS AERUGINOSA*; GPUA\_PA: GUANIDINOPROPIONASE DE *PSEUDOMONAS AERUGINOSA*; PAH\_SC: AMIDINO HIDROLASE DE IDO PROCLAVAMICO (PAH) DE *STREPTOMYCES CLAVULIGERUS*; AGM\_BT: AGMATINASE DE *BURKHOLDERIA THAILANDENSIS*; AGM\_DR: AGMATINASE DE *DEINOCOCCUS RADIODURANS*; AGM\_CD: AGMATINASE DE *CLOSTRIDIUM. DIFFICILE*; AGM\_TV: AGMATINASE

DE <i>THERMOPLASMA. VULCÃO</i> ; ARG_BC: ARGINASE DE <i>BACILLUS. CALDOVELOX</i> ; ARG_TT: ARGINASE DE <i>THERMUS. THERMOPHILUS</i> ] FORAM ALINHADAS USANDO O MULTALIN. TRÊS SEQUÊNCIAS BEM CONSERVADAS (XGGDH, DAHXD E SXDXDXDPXXXXP) NA MAIORIA DAS UREOHIDROLASES SÃO INDICADAS POR CAIXAS PRETAS. OS SÍTIOS DE LIGAÇÃO DE METAL SÃO INDICADOS COM ASTERISCOS, LIGANTES DE GUANIDINO COM TRIÂNGULOS PRETOS	38
FIGURA 11. EXEMPLO DE POLIENOS LINEARES GIGANTES. 1ª CLETRAMICINA, 1B DESULFOCLETRAMICINA, 1C MEDIMICINA A, 1D MEDIOMICINA B, 2TETRAFIBRICINA E 3 ECO-02301	39
FIGURA 12. PARALELO NA ROTA DE BIOSÍNTESE DE MEDIOMICINA A	40
FIGURA 13. FIGURA DA CAPA DO ARTIGO CONTENDO INFORMAÇÕES SOBRE A AMINOTRANSFERASE NEOB PUBLICADO NO INÍCIO DO ANO DE 2018.	41
FIGURA 14. SEQUÊNCIA CODIFICADORA DA ORF PARQ ( <i>STREPTOMYCES RIMOSUS SUBSP. PAROMOMYCINUS</i> – NÚMERO DE ACESSO Q2MFN7). A SEQUÊNCIA POSSUI 1641 PARES DE BASE QUE CODIFICAM PARA A DESIDROGENASE PARQ. AS SEQUÊNCIAS EM NEGRITO REPRESENTAM OS OLIGONUCLEOTÍDEOS FORWARD (EM AZUL) E REVERSE (EM VERMELHO) (5' E 3') UTILIZADOS COMO INICIADORES (PRIMERS)	43
FIGURA 15. INTERAÇÃO ENTRE RESÍDUOS VIZINHOS DA CAUDA HIS-TAG E A MATRIZ NI-NTA. O QUELANTE EM VERMELHO, O ÁCIDO NITRILOTRIACÉTICO, COORDENA-SE COM QUATRO DOS SEIS SÍTIOS LIGANTES DO NÍQUEL (II). OS DOIS SÍTIOS LIVRES DA ESFERA DE COORDENAÇÃO INTERAGEM COM OS ANÉIS IMIDAZOL DA CAUDA DE 6X HISTIDINAS DAS PROTEÍNAS RECOMBINANTES (EM AZUL)	52
FIGURA 16. ESTRUTURA MOLECULAR DO COMPOSTO IMIDAZOL E DO AMINOÁCIDO HISTIDINA.	52
FIGURA 17. ILUSTRAÇÃO DAS TÉCNICAS DE GOTA SUSPensa E GOTA SENTADA UTILIZADAS NA CRISTALOGRAFIA. OS MÉTODOS CONTÊM A GOTA DE PROTEÍNA E A SOLUÇÃO DE CRISTALIZAÇÃO EM UM SISTEMA VEDADO. (A) GOTA SUSPensa: A GOTA FICA DE CABEÇA PARA BAIXO EM UMA LAMÍNULA. (B) GOTA SENTADA: A GOTA É COLOCADA SOBRE UMA PEQUENA CAVIDADE. EM AMBOS OS MÉTODOS A CRISTALIZAÇÃO ACONTECE POR DIFUSÃO DE VAPOR.	55
FIGURA 18. AMPLIFICAÇÃO DOS GENES CODIFICADORES DAS PROTEÍNAS PARB E PARQ. A COLUNA 1 É O MARCADO DE PESO MOLECULAR 1KB PLUS MASS RULER DNA LADDER DA THERMO SCIENTIFIC. A COLUNA 2 É PARB 1.245PB AMPLIFICADA E A COLUNA 3 É PARQ 1.641PB AMPLIFICADA. O MEIO DE SEPARAÇÃO FOI GEL DE AGAROSE 1% EM TAMPÃO 0.5XTAE. A VOLTAGEM APLICADA NA SEPARAÇÃO FOI 70V POR 50MIN. OS FRAGMENTOS DE DNA NO GEL FORAM VISUALIZADOS EM TRANSILUMINADOR UV A 356NM.	58
FIGURA 19. SEPARAÇÃO DOS PLASMÍDEOS DIGERIDOS ZEROBLUNT CONTENDO GENES PARB E PARQ. A COLUNA 1 É O PADRÃO DE PESO MOLECULAR 1KB PLUS MASS RULER DNA LADDER	

DA THERMOSCIENTIFIC. AS COLUNAS 2 E 3 SÃO OS CLONES POSITIVOS SELECIONADOS A SEREM UTILIZADOS NA ETAPA DE SUBCLONAGEM NOS VETORES DE EXPRESSÃO. O MEIO DE SEPARAÇÃO FOI GEL DE AGAROSE 1 % EM TAMPÃO 0.5X TAE, PH 8,3. A VOLTAGEM APLICADA NA SEPARAÇÃO FOI 70 V POR 50 MIN. OS FRAGMENTOS DE DNA NO GEL FORAM VISUALIZADOS EM TRANSILUMINADOR UV A 365 NM. AS BANDAS ESPECÍFICAS DOS GENES *PARB* *PARQ* E DOS PLASMÍDEOS DIGERIDOS ESTÃO REPRESENTADOS POR SETAS CUJOS TAMANHOS SÃO APROXIMADAMENTE 3500 (VETOR), 1.245PB (*PARB*) E 1.641PB (*PARQ*) RESPECTIVAMENTE. .... 59

FIGURA 20. SEPARAÇÃO DOS PLASMÍDEOS DIGERIDOS PET28A CONTENDO BANDAS REFERENTES A *PARB* E *PARQ*. A COLUNA 1 É O PADRÃO DE PESO MOLECULAR 1KB PLUS MASS RULER DNA LADDER DA THERMOSCIENTIFIC. AS COLUNAS 2 A 7 SÃO OS CLONES REFERENTES ÀS *PARB*, AS COLUNAS 8 A 13 SÃO OS CLONES REFERENTES AS *PARQ* SENDO SOMENTE VISTO UM CLONE POSITIVO NA COLUNA 9. O MEIO DE SEPARAÇÃO FOI GEL DE AGAROSE 1 % EM TAMPÃO 0.5X TAE, PH 8,3. A VOLTAGEM APLICADA NA SEPARAÇÃO FOI 70 V POR 50 MIN. OS FRAGMENTOS DE DNA NO GEL FORAM VISUALIZADOS EM TRANSILUMINADOR UV A 365 NM. A BANDA ESPECÍFICA DO GENE *PARQ* ESTÁ REPRESENTADA POR SETA E CÍRCULO. .... 60

FIGURA 21. AMPLIFICAÇÃO DO GENE CODIFICADOR DA PROTEÍNA LIVB. A COLUNA 1 É O MARCADO DE PESO MOLECULAR 1KB PLUS MASS RULER DNA LADDER DA THERMOSCIENTIFIC. A COLUNA 2 É *LIVB* 1.245PB AMPLIFICADA. O MEIO DE SEPARAÇÃO FOI GEL DE AGAROSE 1% EM TAMPÃO 0.5XTAE. A VOLTAGEM APLICADA NA SEPARAÇÃO FOI 70V POR 50MIN. OS FRAGMENTOS DE DNA NO GEL FORAM VISUALIZADOS EM TRANSILUMINADOR UV A 356NM. .... 61

FIGURA 22. GEIS SDS-PAGE 10% DO TESTE DE EXPRESSÃO PARA NEOQ. ENSAIO DE EXPRESSÃO NAS TEMPERATURAS DE 37°C, 30°C, 25°C E 18°C PARA DIFERENTES CONCENTRAÇÕES DE IPTG (0.2MM; 0.4MM;0.6MM E 0.8MM) POR DISTINTAS HORAS DE INDUÇÃO (2H, 4H, 8H, 16H). A PROTEÍNAS NEOQ TEM PESO MOLECULAR DE 58KDA ALÉM DA CAUDA DE HISTIDINA (~3,2KDA). .... 63

FIGURA 23. GEIS SDS-PAGE 10% DO TESTE DE EXPRESSÃO PARA PARQ. ENSAIO DE EXPRESSÃO NAS TEMPERATURAS DE 37°C, 30°C, 25°C E 18°C PARA DIFERENTES CONCENTRAÇÕES DE IPTG (0.2 MM; 0.4 MM; 0.6 MM E 0.8 MM) POR DISTINTAS HORAS DE INDUÇÃO (2H, 4H, 8H, 16H). A PROTEÍNAS PARQ TEM MASSA MOLECULAR DE 58KDA ALÉM DA CAUDA DE HISTIDINA (~3,2KDA). .... 64

FIGURA 24. PURIFICAÇÃO DA NEOB POR CROMATOGRAFIA DE AFINIDADE. A – CROMATOGRAMA DE PURIFICAÇÃO POR AFINIDADE CONTENDO UM GRADIENTE DE IMIDAZOL (5% ATÉ 100%). B- ANÁLISE DAS FRAÇÕES ELUÍDAS DA COLUNA NI-NTA FOI REALIZADA EM SDS-PAGE A 10%. A COLUNA 1 É O MARCADOR DE MASSA MOLECULAR. A COLUNA 2 REPRESENTA A FRAÇÃO DO *PELLET*. A COLUNA 3 A FRAÇÃO DO EXTRATO CELULAR LISADO. A FRAÇÃO 4

CONTÉM O QUE NÃO FOI LIGADA NA COLUNA DE AFINIDADE. A COLUNA 5 REPRESENTAM A FRAÇÃO DE LAVAGEM COM 25MM DE IMIDAZOL. AS COLUNAS 6 A 12, AS FRAÇÕES DE LAVAGEM COM NO GRADIENTE DE IMIDAZOL REFERENTES AO PICO OBTIDO.....	66
FIGURA 25. CROMATOGRAMA DE EXCLUSÃO MOLECULAR. PROTEÍNA NEOB ELUIDA EM TORNO DE 70 ML ATINGINDO NO PICO MÁXIMO DE 400MAU.....	67
FIGURA 26. FRAÇÕES DA PROTEÍNA NEOB. FRAÇÕES PROTEICAS PROVENIENTES APÓS CROMATOGRAFIA POR EXCLUSÃO MOLECULAR.....	68
FIGURA 27. THERMAL SHIFT DA NEOB RECOMBINANTE. NO GRÁFICO, AS CURVAS DE MELTING DE NEOB NA PRESENÇA DOS TAMPÕES QUE MAIS AUMENTARAM A ESTABILIDADE DA PROTEÍNA QUANDO COMPARADOS COM A CONDIÇÃO CONTROLE.....	70
FIGURA 28. CROMATOGRAMA DE PURIFICAÇÃO POR EXCLUSÃO MOLECULAR DAS ENZIMAS RELACIONADAS COM A BIOSÍNTESE DE NEOMICINA. PICOS MAIORES REFERENTES A ELUIÇÃO DA I-NEOB E II-NEOQ.....	71
FIGURA 29. CRISTAL DA PROTEÍNA NEOB NA FORMA APO. CRISTAL COM FORMATO DE LOSANGO NA CONDIÇÃO 30% PEG 8000, 0.2M SULFATO DE AMÔNIA, 0.1M CACODILATO DE SÓDIO PH 6.5 APÓS 5 DIAS.....	73
FIGURA 30. ESTRUTURA DA NEOB. ESTRUTURA REPRESENTADA EM CARTOON APRESENTADA EM DÍMERO SEPARADAMENTE COLORIDOS DE ROSA CLARO E AZUL CÉU. AS FITAS-B DE CADA MONÔMERO ESTÃO DESTACADAS EM COLORAÇÃO CEREJA E AZUL MARINHO..	77
FIGURA 31. ESTRUTURA TRIDIMENSIONAL DO MONÔMERO DA NEOB. ESTRUTURA REPRESENTADA EM CARTOON DE UM DOS MONÔMEROS DA PROTEÍNA. A-HÉLICE COLORIDAS EM ROSA CLARO E B-FITAS COLORIDAS EM CEREJA. AS TRÊS FITAS-B ANTIPARALELAS DA INSERÇÃO DA FOLHA-B ESTÃO DESTACADAS EM MAGENTA. A IMAGEM FOI GERADA PELO PROGRAMA PYMOL (DELANO, 2002). .....	77
FIGURA 32. SOBREPOSIÇÃO DOS MONÔMEROS DAS ESTRUTURAS DA NEOB. EM ROSA ESTÁ A ESTRUTURA DO MONÔMERO DA NEOB OBTIDA NESTE TRABALHO E EM AZUL A ESTRUTURA DA MESMA PROTEÍNA DEMONSTRADA POR UM GRUPO ESTADUNIDENSE (CÓDIGO DO PDB: 6CBK CADEIA A). O RMSD ENTRE AS CADEIAS FICOU EM 1.81Å.....	78
FIGURA 33. PURIFICAÇÃO DA PROTEÍNA AMH_828 POR CROMATOGRAFIA DE AFINIDADE EM COLUNA DE NÍQUEL. (A) – CROMATOGRAMA OBTIDO NA CROMATOGRAFIA DE AFINIDADE NO QUAL A LINHA VERDE REPRESENTA A APLICAÇÃO DO TAMPÃO DE ELUIÇÃO NUM GRADIENTE DE IMIDAZOL. (B) – SDS-PAGE 10% CORADO COM COOMASSIE BRILLIANT BLUE R250, NO QUAL PODE-SE VISUALIZAR: M – MARCADOR DE MASSA MOLECULAR EM KDA (PRECISION PLUS PROTEIN DUAL COLOR STANDARTS -BIORAD), 1- PELLETT 2- FRAÇÃO LISADA 3- FRAÇÃO REFERENTES AO FLOWTHROUGH, 4- FRAÇÃO REFERENTE A LAVAGEM COM 5% DO TAMBÉM B, 5 A 12– FRAÇÕES CORRESPONDENTES AO PICO DE ELUIÇÃO DA PROTEÍNA AMH 828.....	80

- FIGURA 34. PURIFICAÇÃO DA PROTEÍNA MEDI\_4948 POR CROMATOGRAFIA DE AFINIDADE EM COLUNA DE NÍQUEL. (A) – CROMATOGRAMA OBTIDO NA CROMATOGRAFIA DE AFINIDADE NO QUAL A LINHA VERDE REPRESENTA A APLICAÇÃO DO TAMPÃO DE ELUIÇÃO NUM GRADIENTE DE IMIDAZOL. (B) – SDS-PAGE 10% CORADO COM *COOMASSIE BRILLIANT BLUE R250*, NO QUAL PODE-SE VISUALIZAR: M – MARCADOR DE MASSA MOLECULAR EM KDA (PRECISION PLUS PROTEIN DUAL COLOR STANDARTS -BIORAD), 1- *PELLET* 2- FRAÇÃO LISADA 3- FRAÇÃO REFERENTES AO *FLOWTHROUGH*, 4- FRAÇÃO REFERENTE A LAVAGEM COM 5% DO TAMBÉM B, 5- FRAÇÃO CORRESPONDENTES AO PICO DE ELUIÇÃO DA PROTEÍNA MEDI 4948..... 81
- FIGURA 35. PURIFICAÇÃO DA PROTEÍNA AMH\_828 POR CROMATOGRAFIA DE EXCLUSÃO MOLECULAR. (A) – CROMATOGRAMA OBTIDO NA CROMATOGRAFIA DE EXCLUSÃO MOLECULAR. (B) – SDS-PAGE 10% CORADO COM *COOMASSIE BRILLIANT BLUE R250* NO QUAL PODE-SE VISUALIZAR: M – MARCADOR DE MASSA MOLECULAR EM KDA (PRECISION PLUS PROTEIN DUAL COLOR STANDARTS -BIORAD), 1 A 6–FRAÇÕES ELUÍDAS NO VOLUME DE CORRIDA DE APROXIMADAMENTE 60ML. .... 82
- FIGURA 36. PURIFICAÇÃO DA PROTEÍNA MEDI\_4948 POR CROMATOGRAFIA DE EXCLUSÃO MOLECULAR. (A) – CROMATOGRAMA OBTIDO NA CROMATOGRAFIA DE EXCLUSÃO MOLECULAR. (B) – SDS-PAGE 10% CORADO COM *COOMASSIE BRILLIANT BLUE R250* NO QUAL PODE-SE VISUALIZAR: M – MARCADOR DE MASSA MOLECULAR EM KDA (PRECISION PLUS PROTEIN DUAL COLOR STANDARTS -BIORAD), 1A 4–FRAÇÕES ELUÍDAS DURANTE A CORRIDA DE ACORDO COM OS PICOS CORRESPONDENTES. .... 83
- FIGURA 37. ESTRUTURA TRIDIMENSIONAL DO MONÔMERO DA MEDI\_4948. ESTRUTURA REPRESENTADA EM CARTOON DE UM DOS MONÔMEROS DA PROTEÍNA. A-HÉLICES COLORIDAS EM CIANO E B-FITAS COLORIDAS EM MAGENTA. A IMAGEM FOI GERADA PELO PROGRAMA PYMOL (DELANO, 2002)..... 89
- FIGURA 38. ESTRUTURA TRIDIMENSIONAL DO HEXÂMERO DA MEDI\_4948. AS SEIS SUBUNIDADE (A – F) DA MEDI\_4948 ESTÃO COLORIDAS EM AZUL, AMARELO, ROSA CLARO, LARANJA, MAGENTA E VERDE CLARO, RESPECTIVAMENTE. TRÊS SUBUNIDADES DA FRENTE ESTÃO EM CORES FORTES ENQUANTO AS OUTRAS TRÊS SUBUNIDADES POSICIONADAS ATRÁS ESTÃO EM CORES MAIS CLARAS. .... 90
- FIGURA 39. INTERAÇÃO MONÔMERO-MONÔMERO EM TRÍMEROS DO HEXÂMERO. A ALÇA ESTENDIDA DE CADA MONÔMERO INTERAGINDO COM A SUBUNIDADE ADJACENTE (DESTACADOS NAS CORES CIANO, VERMELHO E MAGENTA) EM DIREÇÃO DA ENTRADA DO SÍTIO ATIVO NO MONÔMERO CIRCUNVIZINHO. ESFERAS VERDES INDICAM DOIS ÍONS METÁLICOS NO SÍTIO ATIVO. .... 91
- FIGURA 40. LIGAÇÕES NO SÍTIO ATIVO DA MEDI\_4948. A) CARBONO, NITROGÊNIO, OXIGÊNIO E ÍONS METÁLICOS ESTÃO COLORIDOS EM BRANCO, AZUL, VERMELHO E VERDE, RESPECTIVAMENTE. MOLÉCULAS DE ÁGUA REPRESENTADAS POR ESFERAS EM

VERMELHO. POSSÍVEIS INTERAÇÕES ENTRE ÁTOMOS ESTÃO INDICADAS EM LINHA TRACEJADA AMARELA. B) SUPERFÍCIE DE POTENCIAL ELETROSTÁTICO DO HEXÂMERO COM SETA INDICANDO A ENTRADA DO SÍTIO ATIVO. CARGAS POSITIVAS DO POTENCIAL ELETROSTÁTICO ESTÃO COLORIDAS EM AZUL ENQUANTO AS NEGATIVAS SÃO MOSTRADAS EM VERMELHO. O MAPA DA SUPERFÍCIE FOI OBTIDO UTILIZANDO A FERRAMENTA APBS NO PYMOL. AS IMAGENS FORAM GERADAS NO PYMOL (DELANO, 2007). .....	93
FIGURA 41. RESÍDUOS COM <i>CIS</i> -LIGAÇÃO. RESÍDUOS DE GLICINA POSICIONANDO RESÍDUO DE HISTIDINA NO SÍTIO ATIVO PARA INTERAÇÃO COM ÍON METÁLICO. ....	94
FIGURA 42. SOBREPOSIÇÃO DO MONÔMERO DE MEDI_4948 (AZUL) E FORMIMINOGLUTAMASE DE <i>V. CHOLERAE</i> (VERMELHO). .....	95
FIGURA 43. ALINHAMENTO DAS SEQUÊNCIAS DAS ENZIMAS DA SUPERFAMÍLIA UREOHIDROLASES. ELEMENTOS DE ESTRUTURA SECUNDÁRIA DA MEDI_4948 SÃO MOSTRADOS EM PRETO NA PARTE SUPERIOR. RESÍDUOS CONSERVADOS SÃO DESTACADOS EM GRIFADO DE VERMELHO ENQUANTO QUE OS RESÍDUOS SEMICONSERVADOS SÃO MOSTRADOS NA COR DE LETRA VERMELHA. A SEQUÊNCIA DE NOVE UREOHIDROLASES ( <i>S. MEDIOCIDICUS</i> : MEDI_4948; <i>P. AERUGINOSA</i> : GBUA -Q9I3S3; <i>P. AERUGINOSA</i> : GPUA- Q9I6K2; <i>S. CLAVULIGERUS</i> : PAH - P0DJQ3; <i>B. THAILANDENSIS</i> : BUR-Q2T3W4 [4DZ4 CÓDIGO DO PDB]; <i>D.RADIODURANS</i> : DR_AGM- Q9RZ04; <i>B. CALDOVELOX</i> : BC_ARG - P53608; <i>RATTUS NORVEGICUS</i> : ARG11_RAT - P07824; <i>V. CHOLERAE</i> : HUTG - Q9KSQ2 [1XFK CÓDIGO DO PDB]) FORAM ALINHADAS USANDO O PROGRAMA MULTIALING (CORPET, 1988). TRÊS SEQUÊNCIAS CONSERVADAS (XGGDH, DAHXD E SXDXDXDPXXP) NA MAIORIA DAS UREOHIDROLASES ESTÃO INDICADAS POR CAIXAS VERDES. ESTA FIGURA FOI GERADA COM O PROGRAMA ESPRIPT (GOUET <i>ET AL.</i> , 1999). .....	96

## Lista de Tabelas

TABELA 1. OLIGONUCLEOTÍDEOS INICIADORES UTILIZADOS NA AMPLIFICAÇÃO POR PCR. AS BASES EM AZUL REPRESENTAM O SÍTIO DE RESTRIÇÃO PARA A ENZIMA NDEI E EM VERMELHO O SÍTIO DE RESTRIÇÃO DA ENZIMA XHOI PRESENTES NA SEQUÊNCIA DAS ORFS DE INTERESSE NAS QUAIS CODIFICAM PARA A METIONINA INICIAL.....	45
TABELA 2. CONJUNTO DE DADOS E ESTATÍSTICAS DO REFINAMENTO.....	73
TABELA 3. CRISTAIS DA PROTEÍNA NEOB EM DIFERENTES CONDIÇÕES DE CRISTALIZAÇÃO	78
TABELA 4. CRISTAIS DAS PROTEÍNAS MEDI_4948 E AMH_828 EM DIFERENTES CONDIÇÕES DE CRISTALIZAÇÃO.....	84
TABELA 5. ESTATÍSTICAS DE COLETA DE DADOS E PROCESSAMENTO DO CRISTAL NATIVO AMH_828. OS DADOS ENTRE PARÊNTESES INDICAM OS VALORES DA MENOR E MAIOR RESOLUÇÃO, RESPECTIVAMENTE.....	85
TABELA 6: ESTATÍSTICAS DE COLETA DE DADOS E PROCESSAMENTO DO CRISTAL NATIVO MEDI_4948. OS DADOS ENTRE PARÊNTESES INDICAM OS VALORES DA MENOR E MAIOR RESOLUÇÃO, RESPECTIVAMENTE.....	87

## Lista de Abreviaturas e Siglas

Å	Angstrom
ACP	Proteína carregadora de Acil
AT	Aciltransferase
°C	Graus Celsius
µg	Micrograma
µL	Microlitro
µM	Micromolar
BLAST	Basic Local Alignment Search Tool
cm	centímetro
CV	volume de coluna
DH	Deidratase
EDTA	Ácido tilenodiaminotetracético
ER	Enoil Redutase
FAD	flavina adenina dinucleotídeo
FAS	Ácido Graxo Sintase
FPLC	Fast Pressure Liquid Chromatography
<i>g</i>	aceleração de gravidade
g	Grama
GMC	glicose-metanol-colina
IPTG	isopropil-β-D-tiogalactopiranosídeo
KDa	QuiloDalton
KR	Ceto-redutase
KS	Ceto-sintase
M	Molar (mol/L)
mg	miligrama
mL	mililitro
ng	nanograma
PLP	Pirodoxal Fosfato
PKS	Sintase de policetídeo

## Sumário

1. Introdução.....	17
1.1 Produtos Naturais.....	17
1.2 Antibióticos e suas origens .....	18
1.3 Aminoglicosídeos.....	20
1.4 Biossíntese combinatorial e sua aplicação em aminoglicosídeos .....	23
1.5 Policetídeos .....	33
1.6 Adversidades no percurso .....	40
2. Objetivos .....	42
2.1 Objetivos gerais .....	42
2.2 Objetivos específicos .....	42
3. Materiais e Métodos.....	43
3.1 Desenho dos <i>primers</i> e Amplificação dos Genes .....	43
3.2 Purificação dos Produtos Amplificados por PCR. ....	46
3.3 Clonagem em Vetor de Clonagem Zero Blunt™ .....	46
3.4 Transformação da Bactéria <i>Escherichia coli</i> DH5α com os Plasmídeos Recombinantes pCR®-Blunt .....	47
3.5 Análise de Restrição de Recombinantes pCR®-Blunt de <i>E. coli</i> DH5α .....	47
3.6 Subclonagem dos Genes no Plasmídeo de Expressão pET28a.....	48
3.7 Análise de Restrição dos Plasmídeos Recombinantes pET28a de <i>E. coli</i> DH5α.....	48
3.8 Transformação da Bactéria <i>E. coli</i> BL21(DE3) com os Plasmídeos Recombinantes pET28a	49
3.9 Expressão Heteróloga das Proteínas Recombinantes.....	49
3.10 Purificação das Proteínas Recombina ntes .....	50
3.11 Cromatografia de Afinidade em Resina de Níquel.....	51
3.12 Cromatografia de Exclusão por Tamanho.....	53
3.13 Protocolo de <i>Thermal Shift</i> (HUYNH; PARTCH, 2015).....	54

3.14	Experimentos de cristalização.....	55
3.15	Coleta de dados de difração de raios X e processamento da Medi 4948 e NeoB .....	57
4.	Resultados e Discussão.....	58
4.1	Clonagem.....	58
4.2	Subclonagem no Vetor de Expressão pET28a(+). .....	59
4.3	Amplificação das enzimas correspondentes em <i>Streptomyces lividus</i> .....	60
4.4	Teste de Expressão e Expressão Heteróloga das Proteína NeoB, NeoQ, ParB e ParQ a partir dos Vetores de Expressão.....	61
4.5	Purificação da Proteína NeoB por Cromatografia de Afinidade e Exclusão por Tamanho 65	
4.6	Ensaio de Deslocamento Térmico (Thermal Shift) para NeoB em presença de diferentes tampões .....	68
4.7	Purificação das enzimas NeoB e NeoQ da biossíntese de Neomicina em condição obtida pelo <i>Thermal Shift</i> . .....	70
4.8	Cristalização da NeoB, coleta e processamento de dados .....	72
5.	Resultados obtidos para as enzimas Amidinohidrolases.....	79
5.1	Produção e purificação das proteínas recombinantes Medi_4928 e AMH_828 .....	79
5.2	Cristalização da AMH_828 e Medi_4948.....	83
5.2	Estrutura do Monômero.....	88
5.3	Oligomerização da Medi_4948.....	89
5.4	Sítio de ligação ao metal .....	92
6.	Conclusão.....	97
	Referências .....	98

## **1. Introdução**

### **1.1 Produtos Naturais**

A natureza, rica e diversa, sempre presente e necessária à vida do homem, despertou uma grande atração não somente por ser fonte de nutrientes e manutenção, mas também por ser fonte fundamental de inspiração e aprendizado para o desenvolvimento de tecnologias: seja em química, biologia, medicina, agricultura, no ramo de transportes e até na arquitetura. A utilização longínqua de produtos naturais pela humanidade, a fim de alívio e cura de doenças, bem como na utilização para controle de pragas, nos relata que durante o desenvolvimento das civilizações, a utilização de recursos foi fortemente usada, com destaque para a Medicina Tradicional Chinesa, civilizações Egípcias, Greco-romano, as quais têm episódios documentados da utilização de plantas medicinais, produtos animais e minerais. Durante a idade média, os remédios fitoterápicos foram utilizados na Europa no início do século XIX, iniciando com o isolamento da morfina e outros alcaloides do ópio. Porém, com tempo, houve então, uma transição de remédios naturais de conteúdo desconhecido para a medicina ocidental moderna (DEWICK, 1997).

Os produtos naturais já desempenham um papel importante no desenvolvimento de medicamentos. Hoje, a maioria das principais classes de antibióticos são baseadas em produtos naturais isolados a partir de microrganismos, como são duas das três principais classes atualmente utilizadas de antifúngicos: polienos e equinocandinas. Além disso, muitos outros compostos usados para tratar várias doenças são produtos naturais ou derivados desses. Estruturalmente, mais da metade de todos os medicamentos aprovadas entre 1981 e 2014 se assemelham a produtos naturais. Desses, 6% são produtos naturais inalterados, 26% são derivados de produtos naturais e 32% mimetizam um produto natural e/ou contêm um núcleo farmacofórico presente em produtos naturais (NEWMAN; CRAGG, 2016). Os benefícios dos produtos naturais como pontos de partida para a descoberta de drogas podem ser explicados pela relevância biológica de suas estruturas. Produtos naturais e fármacos derivados deles têm maior diversidade química e estrutural do que os compostos sintéticos (HARVEY; EDRADA-EBEL; QUINN, 2015). No entanto, os produtos naturais derivam da adaptação de um organismo ao seu ambiente e, portanto, têm um propósito biológico específico baseado na evolução. Sendo

assim, faz sentido que os produtos naturais geralmente contenham estruturas biologicamente relevantes. Subestruturas ou *scaffold* de produtos naturais são frequentemente consideradas estruturas privilegiadas, significando que elas podem oferecer bioatividade melhorada contra alvos diversos (Cragg, 2005) (PRACHAYASITTIKUL; WORACHARTCHEEWAN; SHOOMBUEATONG; PRACHAYASITTIKUL *et al.*, 2015) (RODRIGUES; REKER; SCHNEIDER; SCHNEIDER, 2016). Um estudo realizado por Jérôme Hert e colaboradores em 2009 (HERT; IRWIN; LAGNER; KEISER *et al.*, 2009) identificou cerca de 1300 estruturas que possuem anéis (com até 11 átomos de C, N, O e F na molécula) em produtos naturais e cerca de 83% deles não são encontrados em bibliotecas de compostos comerciais.

Nesse segmento, é possível expandir a diversidade gerada pela natureza utilizando ferramentas químicas, tais como a alteração de grupos funcionais, estereoquímica e os esqueletos naturais por aproximação sintética e também por ferramentas biológicas, como a manipulação genética de enzimas participantes da biossíntese, apresentando assim múltiplas oportunidades para obtenção de análogos estruturais e, com isso, aperfeiçoar a efetividade de um fármaco (KIM; THIESSEN; BOLTON; CHEN *et al.*, 2016).

## **1.2 Antibióticos e suas origens**

Segundo a definição, antibióticos são compostos naturais ou sintéticos capazes de inibir o crescimento ou causar a morte de fungos ou bactérias. Podem ser classificados como bactericidas, quando causam a morte da bactéria, ou bacteriostáticos, quando promovem a inibição do crescimento microbiano (Walsh, 2003). Com a descoberta dos antibióticos houve uma redução significativa nas taxas de mortalidade e morbidade de doenças sociais e epidemiológicas, como por exemplo, tuberculose, sífilis, pneumonia, gonorreia e doenças transmissíveis comuns na infância. A revolução ocasionada pela antibioticoterapia foi um marco no século 20, conseqüentemente a eficácia e eficiência notável dos antibióticos causou um senso de admiração pela nova opção medicamentosa (ALANIS, 2005).

A era dos antibióticos pode ser dividida em quatro grandes momentos. No primeiro período, antes de 1940, as pesquisas eram conduzidas na busca por agentes químicos

que apresentassem atividade antibiótica. O agente etiológico de várias doenças, como exemplo: tuberculose, cólera, febre tifoide, havia sido identificado pela metade do século 19. No entanto, esforços ainda deveriam continuar na busca por substâncias capazes de conter ou mesmo interferir na proliferação de tais microrganismos. Em meio a esta busca, o pesquisador alemão Paul Ehrlich, em 1910, desenvolveu o salvarsan, primeiro antimicrobiano de origem sintética usado contra sífilis (Walsh, 2005). Já em 1934, chega ao mercado a proflavina, um antibiótico utilizado principalmente contra infecções de feridas profundas o qual foi extensivamente empregado na Segunda Guerra Mundial (Patrick, 1995). Entretanto, devido à toxicidade apresentada frente à utilização do medicamento em infecções sistêmicas, evidenciou a necessidade de novos agentes mais eficazes (Patrick, 1995). Em 1935, surge o protosil, o qual apresentava atividade *in vivo* contra infecções causadas por espécies de *Streptococcus*. Surge então a primeira classe de agentes sintéticos efetivos contra infecções sistêmicas, as sulfas ou sulfonamidas (Nicolaou, 2008). Introduzidas no início dos anos 1940, as sulfas atualmente apresentam espectro de ação limitado e são pouco utilizadas.

O segundo período, de 1940 a 1960, é caracterizado por triagens de produtos naturais e obtenção de novos antibióticos. Descoberta por Alexander Fleming em 1928, a penicilina G (benzilpenicilina) somente foi introduzida como agente terapêutico a partir de 1940. O grande marco dos anos 40 foi a descoberta da penicilina, a qual possuía atividade superior à das sulfas. A segunda Guerra Mundial foi um fator importante para o processo de industrialização da penicilina. A partir deste momento, esforços significativos foram empregados na descoberta e desenvolvimento de novos antibióticos (PROJAN; SHLAES, 2004). Nesse período, muitos produtos naturais foram introduzidos, principalmente para o tratamento de infecções causadas por bactérias Gram positivas:  $\beta$ -lactâmicos (cefalosporina), aminoglicosídeos (estreptomicina), tetraciclina (clorotetraciclina), macrolídeos (eritromicina), peptídeos (vancomicina) e outros (cloranfenicol, rifamicina B, clindamicina e polimixina B) (FERNANDES, 2006). Já o terceiro período compreende entre os anos de 1960 a 1980. Neste período, os antibióticos semi-sintéticos, ou seja, análogos aos antibióticos naturais, tomam conta do mercado. Sendo eficazes no tratamento de patógenos Gram positivos e Gram negativos, a maior parte foi obtida a partir de protótipos naturais microbianos, como por exemplo, os derivados  $\beta$ -lactâmicos (análogos de penicilina e cefalosporina, ácido clavulânico,

aztreonam), análogos da tetraciclina, aminoglicosídeos (gentamicina, tobramicina, amicacina) (FERNANDES, 2006).

A partir de 1980 até os anos 2000, tem-se o quarto período: caracterizado por buscas de novos antibióticos, utilizou-se ferramentas como genômica e triagens de coleções de compostos. No entanto, os resultados não vieram de acordo com as buscas. Houve uma redução dramática na identificação de novos protótipos antibióticos e paralelamente, uma elevação na incidência de resistência bacteriana. O fato marcante deste período é a introdução da classe das fluoroquinolonas sintéticas, desenvolvidas a partir do ácido nalidíxico (HAPPI; GBOTOSHO; FOLARIN; AKINBOYE *et al.*, 2005; KUMAR; SCHWEIZER, 2005; MOELLERING, 1998). A partir dos anos 2000 poucos antibióticos foram introduzidos para a terapêutica antimicrobiana. Mas um fato interessante é que de 2000 a 2010, o consumo de antibióticos aumentou em 76% em países dos BRICS (Brasil, Rússia, Índia, China e África do Sul) e concomitantemente a indicação sem prescrição médica também cresceu de forma acentuada (VAN BOECKEL; GANDRA; ASHOK; CAUDRON *et al.*, 2014).

### 1.3 Aminoglicosídeos

Dentre os antibióticos existentes, os aminoglicosídeos pertencem a uma classe que foi revelada décadas atrás. O primeiro aminoglicosídeo a ser descoberto foi a estreptomicina, isolada de *Streptomyces griseus*, em 1944 pelo pesquisador Selman A. Waksman, o que lhe rendeu o prêmio Nobel de medicina em 1952. A partir de então, várias outras substâncias similares foram descobertas, como a neomicina (1949, *S. fradiae*), paromomicina (1950, *S. krestomuceticus*), canamicina (1957, *S. kanamyceticus*), gentamicina (1963, *Micromonospora purpurea*) e tobramicina (1967, *S. tenebrarius*). Com o objetivo de reorientar os problemas obtidos pela resistência bacteriana, foram introduzidos na terapêutica a partir de 1970 os antibióticos semi-sintéticos amicacina, dibecacina e netilmicina (WOODRUFF, 2014).

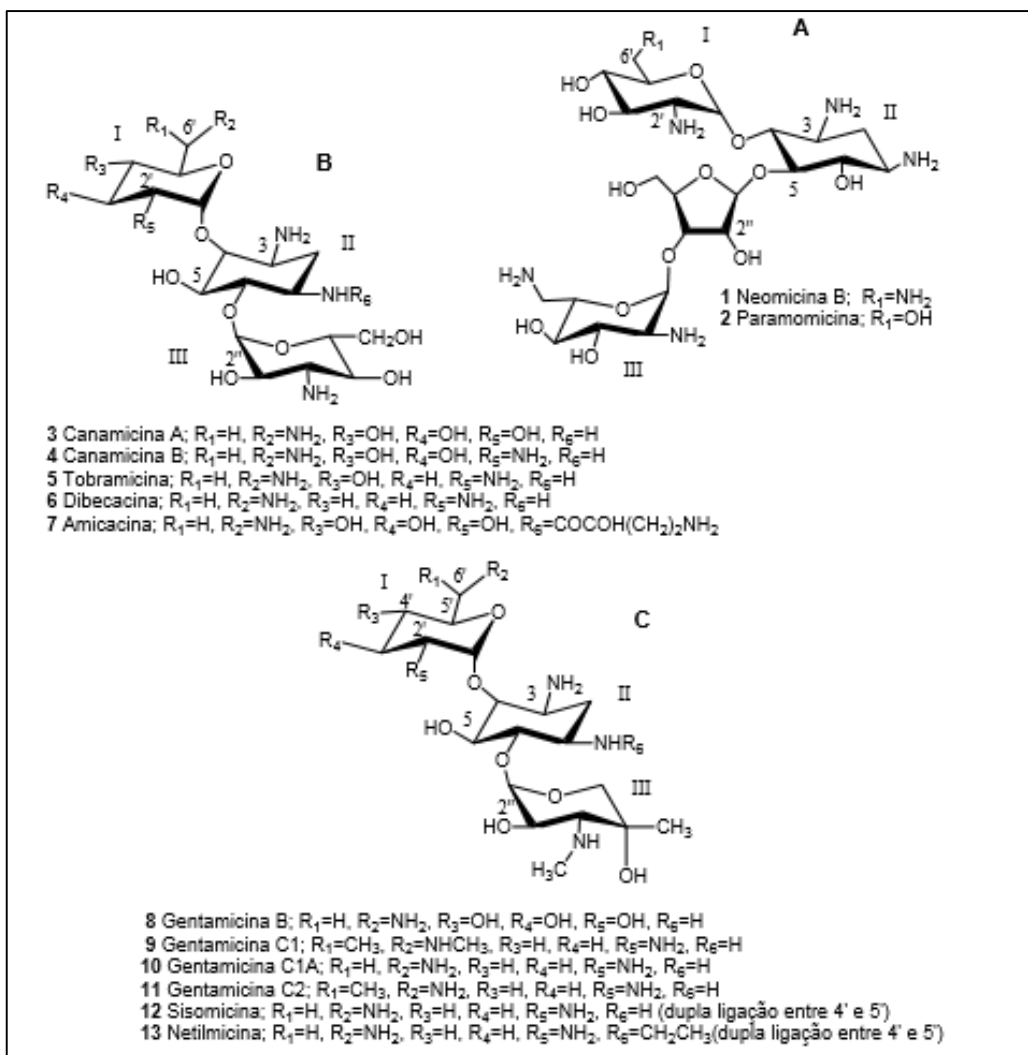
Os antibióticos aminoglicosídeos são obtidos de *Streptomyces* spp. (estreptomicina, neomicina e tobramicina), ou *Micromonospora* spp. (gentamicina) ou obtidos por semi-síntese (netilmicina, amicacina, arbecacina e isepamicina). Eles exibem atividade antimicrobiana contra um amplo espectro de microrganismos, incluindo

bactérias Gram-negativas e Gram-positivas, micobactérias e protozoários (DURANTE-MANGONI; GRAMMATIKOS; UTILI; FALAGAS, 2009).

A descrição estrutural dos aminoglicosídeos revela que estes são moléculas hidrofílicas formadas por um anel aminociclitol ligado a um ou mais amino açúcar através de ligação glicosídica. O grupo aminociclitol geralmente é a 2-desoxi-estreptamina, que pode ser bissubstituída na posição 4 e 5, ou 4 e 6 (Figura 1) (MAGNET; BLANCHARD, 2005). A classificação mais usada refere-se ao anel I como sendo o amino açúcar que se liga na posição 4 da 2-desoxi-estreptamina. Alterações neste anel revelaram que dependendo de quais fossem realizadas acarretariam uma atividade modificada. Como por exemplo, a aminação da posição 6' e 2' presente na canamicina B torna-a mais ativa em comparação com a canamicina A, que possui apenas um grupo amino e uma hidroxila na posição 6' e na posição 2', respectivamente. Além disso, a canamicina C também é detentora de um grupo hidroxila na posição 6' e de um amino na posição 2' (MAGNET; BLANCHARD, 2005). A potência bacteriana não é diminuída nas canamicinas caso seja retirado a hidroxila em 3' ou 4' (SILVA; CARVALHO, 2007). Netilmicina, sisomicina e gentamicinas não possuem função hidroxila nestas posições. Mas, por outro lado, a enzima aminoglicosídeo fosfotransferase (APHs) não é capaz de fosforilar a hidroxila na posição 3', que é um mecanismo clássico de resistências a aminoglicosídeos (Martin, 1998).

O anel II é o grupo aminociclitol central (2-desoxi-estreptamina), enquanto o anel III é o amino açúcar que se liga à posição 5 ou 6 da desoxi-estreptamina. A modificações no anel II são um tanto restritas visto que este é o anel envolvido na ligação do antibiótico com o receptor. Entretanto, na canamicina A, o grupo amino na posição 1' pode ser acilado (amicacina) sem perder a atividade. O anel III, quando comparado com o anel I ou II parece possuir grupos funcionais menos sensíveis as modificações estruturais. No entanto, as 2''-desoxi-gentamicinas são menos ativas que as 2''-hidroxi. Os derivados 2''-amino são mais ativos. O anel IV corresponde a qualquer anel adicional ligado ao anel III (MAGNET; BLANCHARD, 2005; SILVA; CARVALHO, 2007).

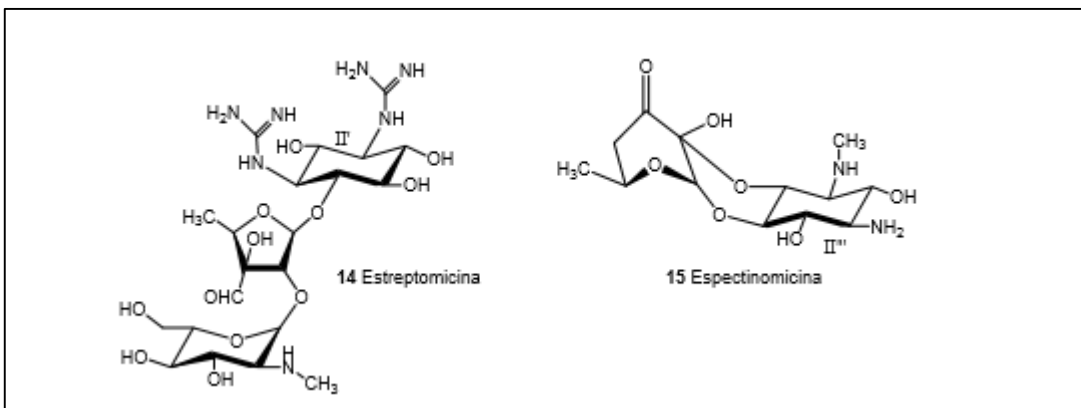
Figura 1. Estrutura química dos principais antibióticos aminoglicosídicos da classe das 2-desoxi-estreptaminas. A são 4,5 dissustituídos, enquanto B e C são do grupo das 4,6 dissustituídos.



Fonte: MAGNET; BLANCHARD, 2005

Contudo existem aminoglicosídeos atípicos (Figura 2), que não se encaixam nesta classificação como, por exemplo, a estreptomina que possui como aminociclitol estreptidina (II'), e a espectinomicina que consiste em três anéis fundidos, onde o anel aminociclitol é a espectinamina (II'') (MAGNET; BLANCHARD, 2005).

Figura 2. Antibióticos cujo anel derivado ciclitol não é a 2-desoxi-estreptamina.



Fonte: MAGNET; BLANCHARD, 2005

O mecanismo de ação dos aminoglicosídeos baseia-se na inibição da síntese proteica através da ligação de alta afinidade com o sítio A da região decodificadora na porção 16S da subunidade 30S do RNA ribossômico bacteriano. O ribossomo é um complexo de ribonucleoproteína responsável pela síntese de cadeias polipeptídicas, baseadas em informações contidas no RNA mensageiro. Esta ligação causa uma alteração conformacional no sítio A do ribossomo e como consequência, a tradução acarreta em uma seleção incorreta de aminoácidos na formação de proteínas. (KOTRA; HADDAD; MOBASHERY, 2000).

O fato de os aminoglicosídeos serem todos estruturalmente próximos, sugere muitas possibilidades para produção de novas moléculas. As ferramentas disponíveis para tal manipulação são variadas.

#### 1.4 Biossíntese combinatorial e sua aplicação em aminoglicosídeos

A biossíntese combinatoria envolve a engenharia genética das vias biossintéticas naturais para a produção de novos compostos usando a maquinaria enzimática existente em cada organismo, especialmente para obtenção de antibióticos derivados de policetídeos e peptídeos não ribossomais (WRIGHT; SUTHERLAND, 2007).

A complexidade estrutural dos produtos naturais, em muitos casos, faz com que seja difícil produzir análogos com atividades medicinais melhoradas. Em contraste, a biossíntese combinatorial explora a promiscuidade de substratos e emprega enzimas e

vias modificadas para produzir novos produtos "não naturais", ampliando substancialmente a diversidade estrutural de produtos naturais com potencial valor farmacêutico. Existem três vantagens da biossíntese combinatorial para a descoberta de medicamentos. Em primeiro lugar, ela ajuda a enriquecer a originalidade e a diversidade de produtos naturais, o que potencialmente melhora as suas características biológicas. Em segundo lugar, oferece uma forma ambientalmente amigável para produzir análogos de produtos naturais, enquanto a síntese química geralmente envolve múltiplos passos de reações de proteção/desproteção, condições adversas, solventes orgânicos tóxicos e resíduos de subprodutos. Em terceiro lugar, através da biossíntese combinatorial, a expressão heteróloga pode se tornar mais eficiente em alguns hospedeiros geneticamente maleáveis levando a um aumento de quantidade do composto, eventualmente resultando em drogas com valor de mercado mais acessível (SUN; LIU; ZHAO; ANG, 2015).

Em 1985, Hopwood clonou os genes biossintéticos do antibiótico actinorodina de *Streptomyces coelicolor* nos microrganismos produtores de medermicina e di-hidrogranaticina, respectivamente. O actinomiceto *Streptomyces sp.* AM-7161 modificado geneticamente produziu, além de medermicina, grandes quantidades de uma nova substância mederorrodina, a qual contém um grupo hidroxila adicional característico da actinorodina. Já o *Streptomyces violaceoruber* Tü 22, modificado para produzir di-hidrogranaticina, produziu uma nova substância, di-hidrogranatirodina, preservando as configurações dos precursores nos estereocentros da nova molécula. Este trabalho representa a primeira utilização do conceito de biossíntese combinatorial de produtos naturais e abriu uma nova perspectiva para a diversificação estrutural de produtos naturais de interesse biotecnológico (HOPWOOD; MALPARTIDA; KIESER; IKEDA *et al.*, 1985).

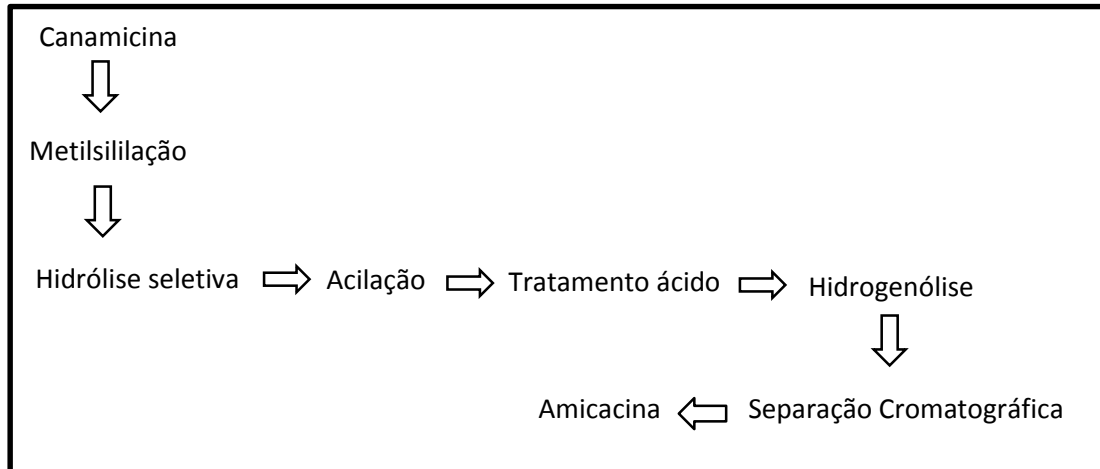
Outro exemplo de sucesso para a biossíntese combinatorial foi o trabalho elaborado por Baltz e colaboradores, cuja manipulação genética das vias biossintéticas de *Streptomyces roseosporus*, produtor da daptomicina, originou uma plataforma robusta com mais de 60 novos antibióticos lipopetídicos, os quais seriam dificilmente obtidos por síntese química (BALTZ; BRIAN; MIAO; WRIGLEY, 2006).

Os antibióticos aminoglicosídeos são alvos de trabalhos utilizando biossíntese combinatorial. No trabalho de Ni, 2016, duas enzimas da biossíntese de canamicina,

KanJ e KanK foram integradas no genoma, através de substituição de gene, da gentamicina B. Foi construída uma nova via biossintética de alto nível de produção de gentamicina B, o produto principal, com um rendimento máximo de 880 ug/ml. Outro fator de destaque foi que, os subprodutos da via foram removidos. A biossíntese combinatória proporcionou novos *insights* sobre a biossíntese de metabólitos secundários da biossíntese de gentamicina B (Xiapung, 2016).

Esta abordagem visando a produção de produtos economicamente mais viáveis é extremamente favorável, visto que atualmente, modificação química específica de aminoglicosídeos é cara e muitas vezes sinteticamente difícil. A especificidade de conversão microbiológica é uma grande vantagem. A amicacina é um dos mais eficazes aminoglicosídeos e para produção desta molécula quimicamente (Figura 3), a partir da canamicina, a torna muito cara com conseqüente valor elevado para o consumidor (Thompson, 1984).

Figura 3. Síntese de Amicacina. Produção de Amicacina a partir de Canamicina utilizando processos químicos indicados na figura.



Fonte: Modificada de Thompson, 1984.

O emprego da técnica de biossíntese combinatorial pode gerar uma boa expectativa na descoberta de novos antibióticos. Porém uma das limitações desse método é que o hospedeiro talvez não expresse o fragmento de DNA incorporado, ou então, as vias metabólicas concorrentes podem desviar preciosos precursores. Outras dificuldades estão relacionadas à imprevisível especificidade do substrato, eficiência

funcional das enzimas de produção, desconhecida especificidade do sistema de efluxo e o potencial de geração de compostos tóxicos incapazes de serem evitados (COATES; HU, 2007). Grandes esforços têm sido feitos no campo da descoberta de fármacos, no entanto, a produção de novos compostos ainda é um grande desafio. A biossíntese combinatória é, sem dúvida, uma importante ferramenta a ser explorada para obtenção de produtos naturais.

No entanto, para o uso da engenharia biossintética ser plausível é necessário que o conhecimento da biossíntese do aminoglicosídeos seja um pré-requisito (LLEWELLYN; SPENCER, 2006).

A maioria dos aminoglicosídeos possuem um aminocicitol como aglicona, o qual passa a ser decorado por deoxiaminoacúcares gerando uma diversidade estrutural na família destes antibióticos. Grande parte dos antibióticos aminoglicosídeos contém a 2-desoxistreptamina (2DOS) como aglicona, do qual deriva o 2-desoxi-cilo-inosamina (2DOIA). Uma outra família de aminociclítóis é derivada de mio-inositol 1-fosfato (MIP), o qual é biossintetizado a partir de D-glicose 6-fosfato (G6P) pelas enzima mio-inositol-1-fosfato sintase e convertido através de uma combinação de fosfatases, desidrogenases e aminotransferases (KUDO; EGUCHI, 2016). Comparado com açúcares polimerizados, o grupo 2-DOS, também conhecido como 1,3-diamino-4,5,6-trihidroxiciclohexeno, é quimicamente estável pois não apresenta função acetal. Esta estabilidade pode ser uma das razões da maioria dos antibióticos aminoglicosídeos possuírem 2-DOS como aglicona central (KUDO; TAMEGAI; FUJIWARA; TAGAMI *et al.*, 1999).

A butirosina, estreptomicina, canamicina e neomicina são os aminoglicosídeos que apresentam suas rotas biossintéticas mais estudadas do ponto de vista funcional (PARK; PARK; BAN; SOHNG *et al.*, 2013), porém outras moléculas apresentam suas vias de síntese praticamente inexploradas. Entretanto, do ponto de vista de biologia estrutural, apenas algumas enzimas relacionada com a biossíntese de aminoglicosídeos apresentam suas estruturas tridimensionais resolvidas (POPOVIC; TANG; CHIRGADZE; HUANG *et al.*, 2006) (Parthier *et al.*, 2012), mostrando um enorme gargalho no entendimento das enzimas e reações que ocorrem na biossíntese desses compostos. Os estudos do Prof. Peter Leadlay da Universidade de Cambridge, adicionaram um avanço significativo no entendimento da função das enzimas da biossíntese de aminoglicosídeos,

particularmente a gentamicina (GUO; HUANG; HUANG; DUAN *et al.*, 2014; HUANG; HUANG; MOISON; GUO *et al.*, 2015).

Para a aplicação de bioengenharia na produção de novos derivados de antibióticos através do uso de enzimas relacionadas com biossíntese, muitas vezes, é necessária a obtenção de estruturas tridimensionais para a identificação dos resíduos chave e que são parte do sítio ativo, delimitação da cavidade, mecanismos de catálise e resíduos essenciais para a atividade enzimática (THIBODEAUX; MELANCON; LIU, 2007).

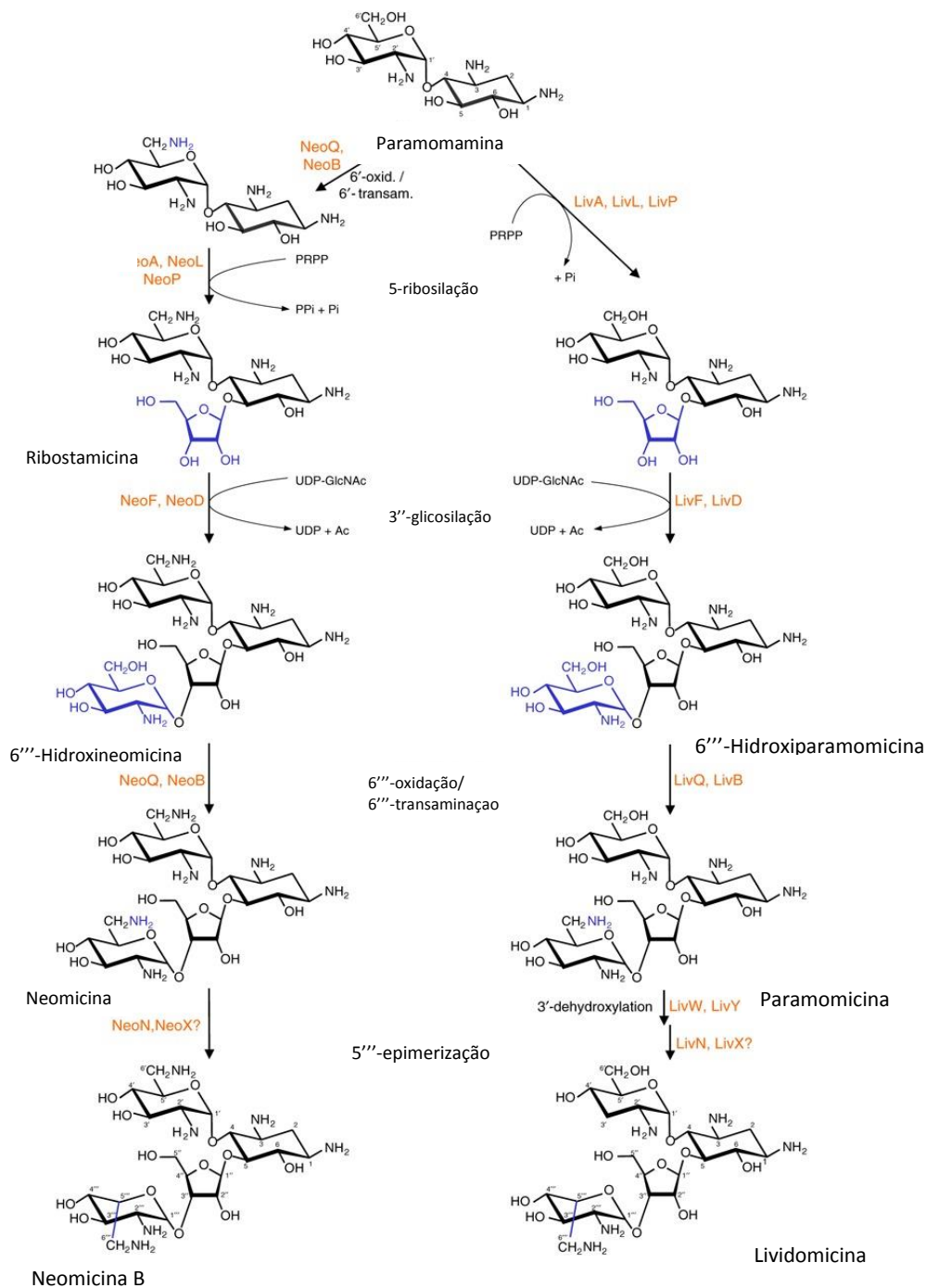
O aminoglicosídeo contendo ciclitol, como o 2-DOS, são amplamente utilizados na terapêutica médica e são foco de muitas pesquisas biológicas, bioquímicas e farmacêuticas. O conhecimento dos clusters gênicos estudados para quase todos os aminoglicosídeos contendo ciclitol revelou que, as vias biossintéticas são amplamente baseadas em enzimas monofuncionais que catalisam passos únicos nas rotas biológicas.

O aminoglicosídeo neomicina, isolado de *S. fradiae* e *S. albogriseus*, foi o primeiro antibiótico contendo 2-DOS a ser conhecido (WAKSMAN; LECHEVALIER, 1949). Os clusters gênicos para compostos relacionados com neomicina são encontrados em bactérias taxonomicamente mais distantes, como em membros das famílias *Streptomycetaceae*, *Micromonosporaceae* (actinobactérias) e *Bacillaceae* (firmicutes), de modo que os aminoglicosídeos, juntamente com outras moléculas do tipo antibiótico, possam ser arranjados, neste sentido, como fósseis moleculares de uma antiga via metabólica. As vias biossintéticas para clusters gênicos correspondentes a neomicinas podem representar uma antiga relação metabólica (Davies et al 1993; Piepersberg et al., 2002, Piepersberg et al., 2007). O cluster da biossíntese de neomicina (grupo neo-gene; código de acesso: AJ629247) tem sido bem conservado durante a evolução. No entanto, foi observado um desvio considerável no conteúdo genético, organizacional e de identidade de sequência para a produção de vários aminoglicosídeos neo-relacionados, tais como butirosina, paromomicina, ribostamicina e lividomicina (Piepersberg et al 2007).

O intermediário paromamina é um precursor comum em quase todos os caminhos 2-DOS de aminoglicosídeos. Por outro lado, a sua posterior transaminação para formar neamina (6'-amino-6'-deoxiparomamina) parece ocorrer em apenas um passo biossintético em vias relacionadas com a neomicina. A variação do neo-tipo 6'-desamino-6'-hidroxi, como na lividomicina (produzido por *Streptomyces lividus*; liv-gene código de

acesso: AJ748832) e paramomicina (produzido por *S. rimosus* ssp *paramomicinus*; par-gene código de acesso: AJ628955), também pode ocorrer (Figura 4).

Figura 4. Via biossintética para produção de Neomicina e Lividomicina. São mostrados os passos a partir do intermediário comum paramomamina.



Fonte: Modificada de Piepersberg et al 2007.

Um conjunto de enzimas relacionadas com a biossíntese aminoglicosídeos e que podem apresentar um grande interesse na aplicação de biossíntese combinatorial são as aminotransferases. Essas enzimas modificam os diferentes pseudosacarídeos que formam os aminoglicosídeos. Dentre as aminações que ocorrem nos aminoglicosídeos, duas que se mostram bastante interessantes são aquela que ocorrem na posição 6' e 6''' da 6'-oxoparomamina para formar neamina na biossíntese de neomicina. O grupo 6'''-NH<sub>2</sub> na neomicina, lividomicina e paramomicina são introduzido por mecanismo de dois passos, pela atuação de uma oxidorreductase (NeoQ, LivQ e ParQ, respectivamente) e uma aminotransferase (NeoB, LivB e ParB, respectivamente) (HUANG; HAYDOCK; MIRONENKO; SPITELLER *et al.*, 2005) (Figura 5). Por outro lado, somente a neomicina apresenta a aminação na posição 6', que é introduzido pelo mesmo par de enzimas.

Desta maneira as enzimas NeoQ e NeoB (relacionadas a produção de neomicina em *Streptomyces fradiae*) são consideradas como sendo bifuncionais por realizar ambas as aminações, enquanto o par de enzimas LivQ/LivB (relacionadas a produção de lividomicina em *Streptomyces lividus*) e ParQ/ParB (relacionadas a produção de paramomicina em *Streptomyces rimosus*) são consideradas monofuncionais.

Clausnitzer et al. (2011) mostrou que, embora as oxidorreductases NeoQ, LivQ e ParQ tenham uma identidade sequencial de mais de 72%, somente a enzima NeoQ tem a capacidade de realizar a oxidorredução da posição 6' da oxoparomamina devido a sua especificidade para o substrato. Por outro lado, as transaminases NeoB, LivB e ParB, que também são altamente conservadas (71% de identidade) são capazes de transaminar a posição 6' da oxoparomamina quando realizado um ensaio acoplado com a neoQ, mas não com a LivQ ou ParQ. O fato é que alguns pseudo-tetrassacarídeos da família neo dos aminoglicosídeos contém o grupo 6'-hidroxil (lividomicina e paramomicina) e outros contém o grupo 6'-NH<sub>2</sub> (neomicina) e que ambos os grupos possuem NH<sub>2</sub> na posição 6''', isto tem levado a uma discussão com diferentes caminhos. Tem sido especulado que a 6'-transaminação é o último passo para a biossíntese de aminoglicosídeos neo-tipos e que alguns microrganismos produtores utilizam o precursor 6'-dihidroxilado de modo a torná-lo mais eficiente ao longo da evolução. A preferência por produzir aminoglicosídeos 6'-hidroxilados (lividomicina e paramomicina) indicaria que o produto 6'-transaminado provavelmente surja como um componente menor no meio de

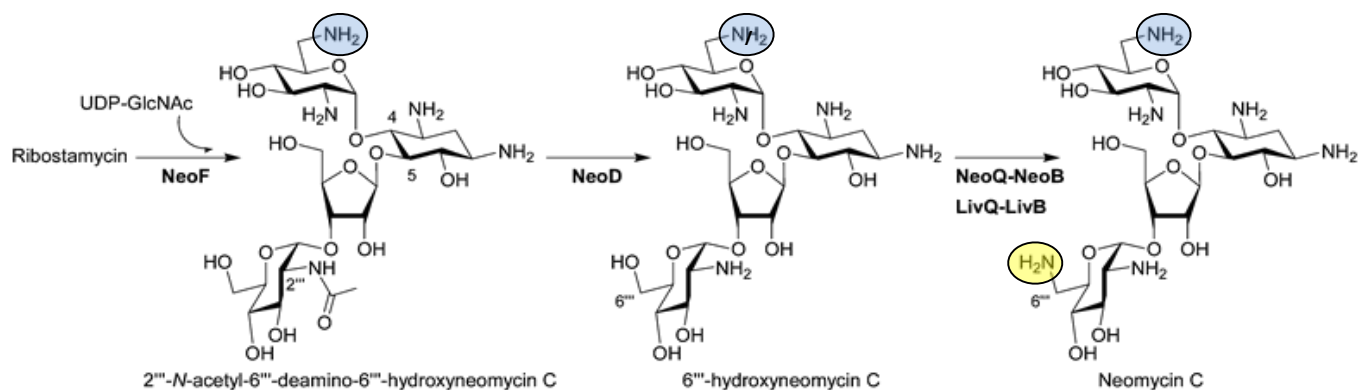
cultura. Uma segunda hipótese sugeriu que um sutil mecanismo de controle para diferenciação do substrato possa distinguir os dois sistemas biossintéticos (LLEWELLYN; SPENCER, 2006).

Estudos realizados por Clausnitzer em 2011 (CLAUSNITZER; PIEPERSBERG; WEHMEIER, 2011), demonstraram que o substrato altera a especificidade das oxidoredutases envolvidas na transaminações das respectivas posições responsáveis pelos diferentes padrões de aminação encontrados em neomicina, lividomicina e paramomicina. Embora as oxidoredutases (NeoQ, LivQ e ParQ) sejam altamente relacionadas e exibam a mesma conservação com as transaminases (NeoB, LivB e ParB), a especificidade do substrato das oxidoredutases se diferem, enquanto o substrato das aminotransferases são funcionalmente idênticos. Esta observação mostra claramente que mesmo as enzimas altamente conservadas podem diferir-se em sua capacidade bioquímica e que não se pode prever propriedade bioquímicas idênticas apenas a partir de alinhamentos de proteína. Do ponto de vista evolutivo, é bastante evidente a observação de que as oxidoredutases apresentem especificidade alterada ao substrato e que as aminotransferases apontem para certa ausência de tal especificidade.

Ainda não é claro como as oxidoredutases, transaminases bifuncionais as quais são encontradas na via biossintéticas de aminoglicosídeos conseguem lidar com diferentes substratos. As enzimas NeoQ/B reconhecem o pseudo-dissacarídeo (paramomina) e um pseudo-tetrassacarídeo (6''-hidroxineomicina C). Ao comparar as sequências proteicas, observa-se um trecho de seis aminoácidos correspondentes aos aminoácidos 172-178 da sequência de LivQ e que está presente também na sequência de ParQ, mas que não é conservado na sequência da enzima NeoQ. Não se sabe se esta sequência específica pode determinar que a LivQ e ParQ sejam enzimas monofuncionais e que a NeoQ seja uma enzima bifuncional. Estudos adicionais ainda precisam ser realizados a fim de verificar tal hipótese. Não é claro se as interações das enzimas com os diferentes substratos são baseadas em resíduos idênticos no sítio ativo ou se estas enzimas contêm dois locais de ligação ao substrato. Enquanto as estruturas cristalográficas não estiverem disponíveis, é especulação inferir como os substratos estruturalmente diferentes estão interagindo com o sítio ativo das enzimas, e que parte do substrato entra em contato com a enzima.

Isso sugere promiscuidade desta enzima e ainda indica que a simples troca de uma enzima desta reação, que necessita de dois passos, pode introduzir uma nova alteração em um aminoglicosídeo, o que apresenta alto valor biotecnológico na produção de novos derivados de aminoglicosídeos (CLAUSNITZER; PIEPERSBERG; WEHMEIER, 2011). Porém, a base estrutural destas especificidades e funcionalidades ainda não são compreendidas e somente com a determinação das estruturas tridimensionais destas enzimas é que esse mecanismo pode ser desvendado.

Figura 5. Biossíntese de pseudotetrasacarídeos. Em azul é mostrado o grupo amino que é transferido para o pseudosacarídeo, pela ação das enzimas NeoB/NeoQ, que são bifuncionais. Em amarelo é mostrado o grupo amino adicionado tanto pela enzima NeoB/NeoQ, como para a enzima LivB/LivQ (que são monofuncionais). Adaptado de Park et al., 2011.



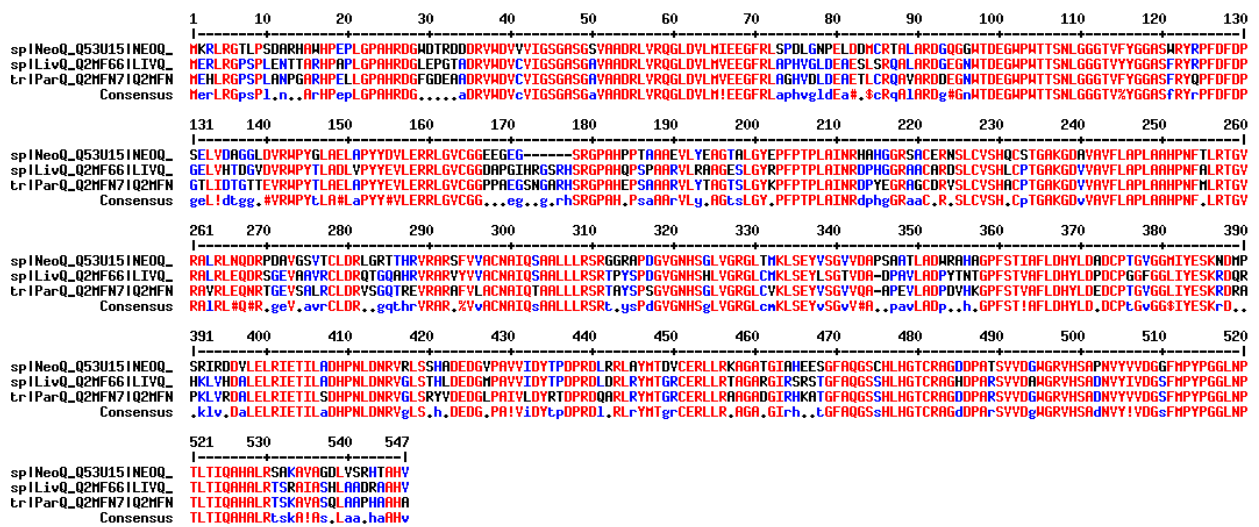
Fonte: Adaptado de Park et al., 2011.

As aminotransferases NeoB, LivB e ParB são consideradas como aminotransferases dependentes de piridoxal fosfato (PLP), um derivado da vitamina B<sub>6</sub>, que está covalentemente ligada a um grupo amino de um resíduo de lisina no sítio ativo da enzima. O grupo amino de um aminoácido é transferido para a porção piridoxal da coenzima, gerando piridoxamina-fosfato. A forma piridoxal da coenzima reage com um α-cetoácido para formar um aminoácido, ao mesmo tempo regenerando a forma aldeído original da enzima. São classificadas como sendo do tipo III. Geralmente as aminotransferases desta família são homodiméricas com dois sítios ativos independentes, mas próximo da interface dos dois protômeros. Estas enzimas



diferentes propriedades enzimáticas que podem explicar as variações de produtos que não são evidentes a partir da análise das sequências proteicas. Do ponto de vista biotecnológico, a simples troca dos genes que codificam as oxidoredutases de vias estreitamente relacionadas podem ser utilizados para a geração *in vivo* de produtos modificados com padrões alterados de aminação.

Figura 7. Alinhamento das sequências das enzimas NeoQ, LivQ, ParQ, oxidoredutases relacionadas com a biossíntese de antibióticos aminoglicosídeos. No alinhamento, é possível observar a alta identidade sequencial entre as enzimas.



Fonte: Elaborado pelo autor.

## 1.5 Policetídeos

Os policetídeos (PK), oriundos do metabolismo secundário, são uma classe de produtos naturais que apresentam maior variedade em suas estruturas química e que envolvem uma porção significativa dos metabólitos microbianos com atividade fisiológica. Várias destas moléculas possuem aplicações comerciais como antibacterianos (a eritromicina e a tetraciclina), imunossuppressores (a rapamicina e o ácido micofenólico), agentes anticancerígenos (a doxorubicina e adaunorubicina), agentes antifúngicos (a anfotericina e a griseofulvina), produtos veterinários (a ivermectina) e anticolésterolêmicos (estatinas) (CANE, 1997; COX, 2007).

A biossíntese de policetídeos se dá por grandes blocos enzimáticos multimodulares, denominados policetídeo sintase ou PKS (do inglês *polyketide synthase*), em que a sua síntese é governada por um conjunto de genes envolvidos desde sua formação e modificação até a sua regulação e própria resistência (MENZELLA; REID; CARNEY; CHANDRAN *et al.*, 2005). Este complexo proteico consiste em vários sítios ativos específicos, composto por três domínios básicos: ceto-sintase (KS), proteína carregadora de acil (ACP) e o domínio acil transferase (AT). Estes domínios podem estar acompanhados de outros domínios ou enzimas que modificam, posteriormente, a molécula policetídica (SCHUMANN; HERTWECK, 2006).

As PKSs estão presentes em bactérias, fungos e plantas e são classificadas em três tipos de acordo com sua estrutura e funcionalidade: PKSs do tipo I, normalmente encontradas em bactérias e fungos; são amplamente estudadas e apresentam o maquinário biossintético dividido em módulos, os quais se subdividem em vários domínios. Podem atuar de modo interativo (domínios se repetem durante a biossíntese) ou não iterativo (canônico), onde cada ciclo de extensão da cadeia policetídica é realizado por um módulo específico e os domínios não se repetem. As PKSs do tipo II, se apresentam como um complexo de proteínas dissociável e possuem um único conjunto de domínios com atividade interativa, sendo essas muito comuns para a formação de compostos aromáticos em bactérias. Por último, as PKSs do tipo III possuem estrutura homodiméricas, ação interativa e não são dependentes de ACP para carregar a cadeia extensora. Estão presentes principalmente em plantas, atuando na biossíntese de flavonoides e chalconas. Algumas bactérias e fungos também possuem essas organizações do tipo III atuando na produção de policetídeos como a flaviolina (CORTES; VELASCO; FOSTER; BLACKABY *et al.*, 2002; SHEN, 2003) (Dewick 2009).

A síntese dos PK assemelha com a síntese dos ácidos graxos, embora as duas vias metabólicas tenham divergido numa fase inicial durante a evolução (HERTWECK, 2009). A síntese dos PK inicia-se pela ligação de uma unidade acila (acetil-CoA ou malonil-CoA) ao domínio KS, conhecida como unidade iniciadora, enquanto o domínio AT transfere outras moléculas acilas ao grupo prostético do ACP. O domínio KS continua com as condensações descarboxilativas (Claisen) entre as moléculas acilas para gerar a  $\beta$ -ceto-acil-ACP. Assim, este processo, repetidamente, origina cadeias com múltiplos grupos ceto dentro da sua estrutura, molécula conhecida como os policetídeo.

(MENZELLA; REID; CARNEY; CHANDRAN *et al.*, 2005; WEISSMAN, 2007).

Outros grupos podem participar no processamento da molécula policetílica, especialmente, na redução do grupo  $\beta$ -cetona: a cetoreductase (KR) reduz o grupo  $\beta$ -cetona estereoseletivamente, a desidratase (DH) que remove a hidroxila (-OH) produzindo uma ligação dupla entre os carbonos adjacentes (C=C) e a enoil redutase (ER) que reduz o C=C para formar uma ligação simples (C-C) (REEVES, 2003).

As PKSs possuem similaridade (tanto em sequência quanto em estrutura) com as enzimas responsáveis pela produção de ácido graxos, denominadas Ácido Graxo Sintase (FAS). A PKS tipo I, análoga com a FAS tipo I de fungos e vertebrados, se caracteriza por ser uma grande proteína multifuncional onde seus sítios ativos estão ligados ordenada e covalentemente. A PKS tipo II, análoga com a FAS tipo II de bactérias, se caracteriza por ser um conjunto de proteínas monofuncionais, cada uma com uma atividade própria e específica (CASTOE; STEPHENS; NOONAN; CALESTANI, 2007; SHEN, 2003).

Mesmo que as duas vias metabólicas compartilhem similaridades na sua enzimologia e na propagação das cadeias crescentes, constituem uma clara diferença no ramo metabólico entre metabolismo primário e secundário. As PKSs se diferenciam das FASs, não somente pela sua habilidade de utilizar uma maior gama de unidades na construção das moléculas, mas também pela formação de vários comprimentos da cadeia policetílica. Mais importante, enquanto FAS catalisa um ciclo redutivo completo depois do processo de alongação, em PKS os passos redutivos são opcionais gerando um padrão de funcionalidade mais complexo (HERTWECK, 2009).

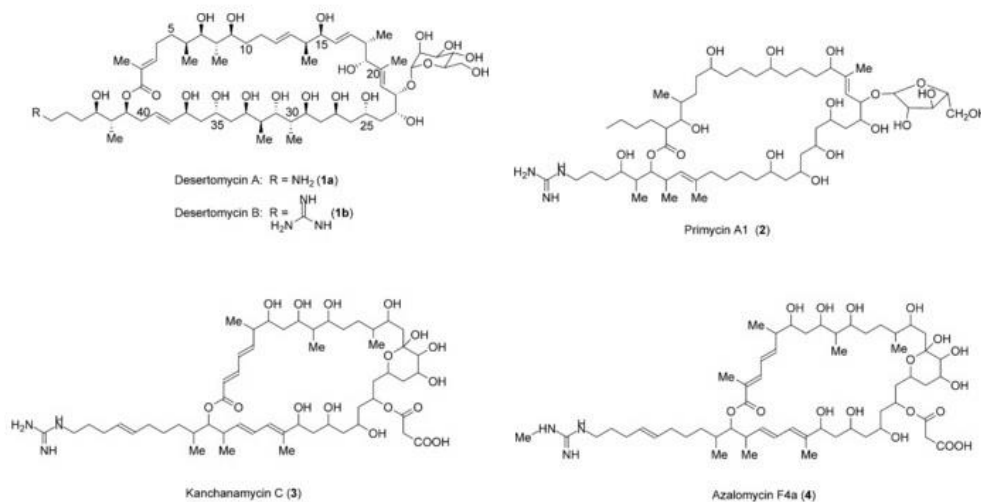
Conforme a atuação do conjunto das enzimas (ou dos domínios) PKS durante a biossíntese, os policetídeos podem ser compostos aromáticos poliidroxilados (como a maioria dos pigmentos fúngicos), compostos alifáticos pouco oxigenados (ou policetídeos parcialmente reduzidos, como a lovastatina) e alifáticos altamente reduzidos (ácidos graxos) (Pastre et al. 2007).

Apesar de que poucas enzimas PKS têm sido isoladas e caracterizadas, é possível, mediante o uso de ferramentas da engenharia genética, sequenciar, clonar, isolar e expressar enzimas desta classe (DONADIO; STAVER; MCALPINE; SWANSON *et al.*, 1991).

Dentre os policetídeos encontram-se os macrolídeos, os quais apresentam importantes atividades biológicas e muitos deles, ou seus derivados, são usados como fármacos (eritromicina, ciclosporina, vancomicina e anfotericina B). As atividades biológicas dos macrolídeos, substâncias com enorme variedade estrutural, são atribuídas, em parte, a uma característica comum: a restrição conformacional em relação a compostos de cadeia aberta. A maior parte destas moléculas naturais cíclicas possui núcleos que compreendem entre 12 e 20 membros. Os macrolídeos são classificados de acordo com o número de membros que compõem a lactona, geralmente sendo de 12, 14, 16 ou 20 membros e a maioria contém ao menos um aminoaçúcar ligado em sua estrutura.

Entre os macrolídeos que são agentes bioativos encontra-se a Desertomicina, um membro das marginolactonas (macrociclos gigante, com mais de 30 átomos de carbono no anel) (HONG; SAMBORSKY; LINDNER; LEADLAY, 2016) (Figura 8). Possui atividade antifúngica e está substituída com um grupo amino ou guanidino.

Figura 8. Exemplos de marginolactonas. Estruturas das marginolactonas antifúngicas desertomicina (1), Primicina (2), Kanchanamicina (3) e Azolomicina (4).

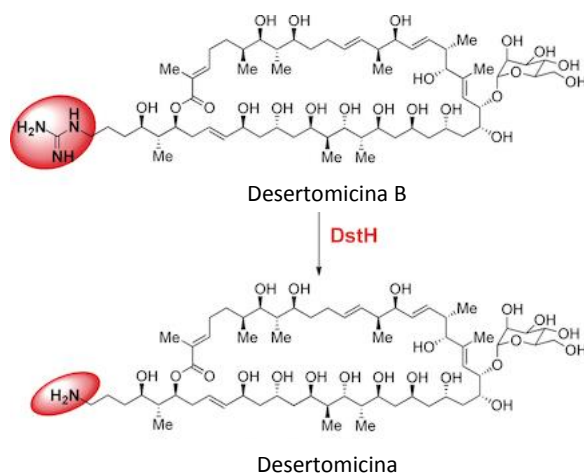


Fonte: Modificado de (HONG; SAMBORSKY; LINDNER; LEADLAY, 2016).

Estudos anteriores revelaram que a 4-guanidinobutanoil-CoA derivada da L-arginina fornece a unidade inicial para a biossíntese da azalomicina F (Figura 8).

Previamente, acreditava-se que a unidade iniciadora da desertomiciana e compostos similares (marginolactonas) fossem biossintetizados pela PKS tendo como unidade iniciadora  $\gamma$ -aminobutanoil-CoA, mas o que foi observado é que, a fim de se evitar a ciclização na unidade inicial, a síntese se dava pela unidade iniciadora 4-guanidinobutanoil-CoA. Para tal, é necessário que haja uma desaminação final do produto natural substituído por guanidino correspondente. Essa etapa é catalisada por uma amidinohidrolase, como um passo tardio na biossíntese (Figura 9).

Figura 9. Amidinohidrolase da desertomiciana. Amidinohidrolase DstH atuando no grupo guanidiol da desertomicina B para formação da desertomicina A.



**Fonte:** Adaptado de (HONG; SAMBORSKY; LINDNER; LEADLAY, 2016).

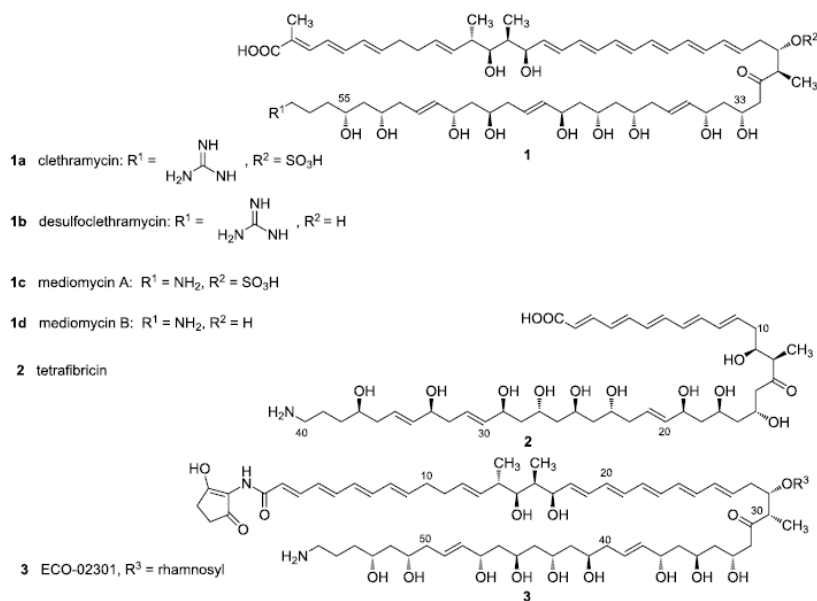
A enzima amidinohidrolase encontrada na biossíntese de desertomicina, denominada DstH, é uma enzima da superfamília da ureohidrolase pois apresenta as mesmas sequências conservadas entre a maioria das ureohidrolases destacadas em caixa preta no alinhamento de algumas enzimas desta superfamília (Figura 10). A superfamília da ureohidrolase abrange diversas agmatinases, arginases, guanidinobutirases, formiminoglutamases e proclavaminato hidrolases (DOWLING; DI COSTANZO; GENNADIOS; CHRISTIANSON, 2008; ELKINS; CLIFTON; HERNANDEZ; DOAN *et al.*, 2002; LEE; KIM; KIM; LEE *et al.*, 2011). O mecanismo dessas enzimas envolve ataque nucleofílico ao carbono amidino por um íon hidróxido ligando dois íons metálicos bivalentes.



Fonte: Hong, 2016.

Curiosamente, Hong e colaboradores, demonstraram em 2017, que os polienos lineares gigantes (Figura 11) Cletramicina e Mediomicina (derivados de *Streptomyces malaysiensis* DSM4137 e *Streptomyces mediocidicus*, respectivamente), aparentemente construídos a partir da unidade de partida 4-guanidinobutanoato e 4-aminobutanoato, com 27 unidades extensora de policetídeo e contendo uma modificação O-sulfonado no grupo hidroxila do C-29, possuem diferenças quanto a amidinohidrolases dos polienos cíclicos (HONG; SAMBORSKY; USACHOVA; SCHNATZ *et al.*, 2017).

Figura 11. Exemplo de polienos lineares gigantes. 1<sup>a</sup> Cletramicina, 1b desulfocletramicina, 1c medicicina A, 1d mediomicina B, 2tetrafabricina e 3 Eco-02301

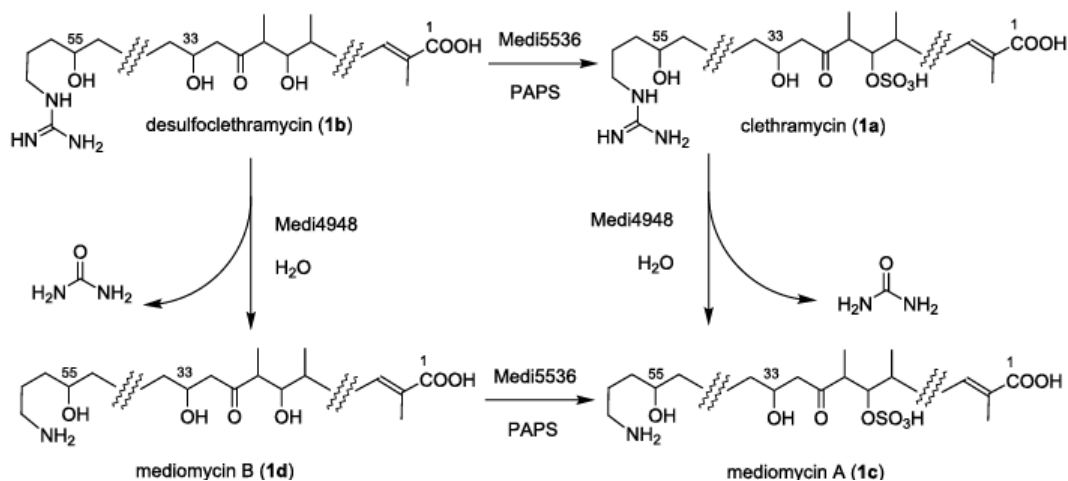


Fonte: Hong, 2017.

O cluster gênico da mediomicina não codifica uma amidinohidrolase, mas quando três genes candidatos a amidinohidrolases de outras partes do genoma do *S. Mediocidicus* foram individualmente expressos em *E.coli* e analisados, apenas um deles (medi\_4948), localizado a 670 Kpb do cluster gênico da mediomicina, catalisou a remoção do grupo amidino na dessulfocletramicina, demonstrando assim um caso raro

em que um gene biossintético essencial não é co-localizado com outros genes da via. Tanto a dessulfocletramicina como a cletramicina são substratos para a amidinohidrolase Medi\_4948 (Figura 12).

Figura 12. Paralelo na rota de biossintese de mediomicina A



Fonte: Hong, 2017.

## 1.6 Adversidades no percurso

Inicialmente, este projeto estava focado em responder questões pertinentes a aminotransferases da biossíntese de neomicina, lividomicina e paramomicina. Entretanto, a competitividade na produção científica nos revelou o trabalho de pesquisadores do departamento de bioquímica, da Universidade de Wisconsin, o qual traz exatamente revelações sobre uma das enzimas alvo do projeto inicial (Figura 13). Sendo assim, nos vimos sujeitos a incluir outras distintas enzimas a fim de buscar elucidação estrutural.

Figura 13. Figura da capa do artigo contendo informações sobre a aminotransferase NeoB publicado no início do ano de 2018.



## **The three-dimensional structure of NeoB: An aminotransferase involved in the biosynthesis of neomycin**

**Garrett T. Dow, James B. Thoden, and Hazel M. Holden\***

Department of Biochemistry, University of Wisconsin, Madison, WI 53706

Received 12 February 2018; Accepted 5 March 2018

DOI: 10.1002/pro.3400

Published online 8 March 2018 proteinscience.org

Fonte: Elaborado pelo autor

## **2. Objetivos**

### **2.1 Objetivos gerais**

Este trabalho tem como objetivo geral obter estruturas tridimensionais de enzimas que representam pontos chaves na biossíntese de alguns aminoglicosídeos, e que possam apresentar interesse biotecnológico na aplicação de biossíntese combinatorial.

### **2.2 Objetivos específicos**

Clonagem dos genes, expressão e purificação das enzimas NeoB e NeoQ da biossíntese de neomicina, LivB e LivQ da biossíntese de lividomicina e ParB e ParQ da biossíntese da paromomicina; e expressão e purificação das enzimas adicionais AMH\_828, 821 e Medi\_4948.

Obtenção das estruturas cristalográficas das enzimas propostas, essencialmente, com a intenção de delinear as suas bases estruturais pelos seus substratos.

### 3. Materiais e Métodos

#### 3.1 Desenho dos *primers* e Amplificação dos Genes

As sequências nucleotídicas das ORFs que codificam as proteínas putativas desidrogenases ParQ produzida por *Streptomyces rimosus subsp. paromomycinus* e aminotransferases ParB estão presentes no genoma do mesmo organismo; e as putativas desidrogenases LivQ e aminotransferase LivB de *Streptomyces lividus*, foram amplificadas através da reação em cadeia da polimerase (PCR). As construções contendo os ORFs referentes as proteínas NeoB, NeoQ e as amidinohidrolases foram gentilmente cedidas pelo Prof. Peter Leadlay da Universidade de Cambridge, UK.

A Figura 14 demonstra a sequência de nucleotídeos para a ORF referente a proteína ParQ como um exemplo da estratégia de amplificação utilizada para obtenção dos genes de interesse neste estudo.

Figura 14. Sequência codificadora da ORF ParQ (*Streptomyces rimosus subsp. paromomycinus* – número de acesso Q2MFN7). A sequência possui 1641 pares de base que codificam para a desidrogenase ParQ. As sequências em negrito representam os oligonucleotídeos *forward* (em azul) e *reverse* (em vermelho) (5' e 3') utilizados como iniciadores (*primers*).

```

ATGGAGCACCTACGGGACCCTCGCCGCTGGCGAATCCCGGCCCGCCGGCACC CGGA ACTCCTCGG
GCCCGCGCACCGCGACGGTTTCGGCGACGAGGCCGCGGACCGCGTGTGGGACGTCTGCGTCATCG
GCAGCGGGGCGAGCGGGGCCGTCGCCGCGGACCGGCTGGTCCGCCAGGGGCTCGACGTCCTGAT
GGTCGAGGAGGGCTTCCGGCTCGCCGGACACGTGGACCTGGACGAGGCCGAGACCCTGTGCCGG
CAGGCCGTCGCCCGCGACGACGAAGGCAACTGGACCGACGAGGGCTGGCCGTGGACCACCTCCAA
CCTCGGCCGGCGGCACGGTCTTCTACGGCGGCGCCTCGTTCCGCTACCAGCCCTTCGACTTCGACCC
GGGCACCCTCATCGACACCGGCACCACCGAGGTGCGCTGGCCGTACACCCTCGCAGAGCTGGCGC
CCTACTACGAGGTGCTGGAGCGGCGTCTGGGCGTGTGCGGGCGGCCCGCCGGCCGAGGGCTCGAA
CGGCGCGCGGCACAGCCGCGGCCCGCCACGAGCCGTCCGCCCGCCGCGGGGTGCTGTACACG
GCCGGCACGTCGCTGGGCTACAAGCCGTTCCCCACGCCGCTGGCCATCAACCGTGACCCGTACGA
GGGGCGCGCCGGGTGCGACCGGGTGTGCGTGTGCGTCAGCCACGCGTGCCCCACCGGCGCCAAG
GGCGATGTCGTCGCGGTCTTCTGGCGCCGCTGGCCGCCACCCCAACTTCATGCTGCGCACCGG

```

CGTCCGCGCCGTACGCCTGGAACAGAACCGCACCGGCGAGGTCTCCGCGCTGCGCTGCCTGGACC  
GGGTGAGCGGGCAGACCAGGGAGGTGCGCGCCCGCGCGTTTCGTCCTGGCGTGCAACGCGATCCA  
GACGGCGGCCCTGCTGCTGCGCTCCCGCACGGCGTACTCCCCCTCCGGCGTCGGCAACCACTCCG  
GCCTGGTGGGGCGGGGCCTGTGCGTCAAGCTGAGCGAGTACGTCAGCGGCGTCGTGCAGGCCGC  
GCCGGAGGTCCTCGCGGACCCGGACGTCCACAAGGGCCCGTTCTCCACCGTGGCGTTCCTGGACC  
ACTACCTGGACGAGGACTGCCCCACGGGCGTGGGCGGCCTGATCTACGAGTCCAAGCGGGACCGG  
GCGCCCAAGCTCGTCCGGGACGCGCTGGAAGTGC GGATCGAGACGATCCTGTCCGACCACCCCAA  
CCTGGACAACCGCGTCGGGCTGTCCCGCTACGTCGACGAGGACGGCCTGCCCGCCATCGTCCTCG  
ACTACCGCACCGACCCGCGCGACCAGGCCCGGCTGCGCTACATGACCGGACGCTGCGAGCGGCTG  
CTGCGCGCCGCGGAGCGGACGGCATCCGCCACAAGGCCACCGGTTTCGCGCAGGGCAGCTCCCA  
CCTGCACGGCACCTGCCGGGCCGGTGACGACCCGGCCCGTTTCGGTGGTGGACGGCTGGGGCCGG  
GTGCACTCGGCGGACAACGTCTACGTCGTCGACGGCAGCTTCATGCCGTACCCCGGCGGTCTCAAC  
CCCACTCTGACCATCCAGGCCACGCCCTGCGCACCAAGGCCGTCGCGTCCCAGCTGGCGGC  
CCCG**CACGCCGCTCATGCGTGA**

Fonte: Elaborado pelo autor.

O artifício empregado se baseou em utilizar dois extensores (*primers*) (oligonucleotídeos iniciadores) complementares às extremidades 3' e 5' do gene de interesse os quais hibridizavam com esta região e servem como ponto de partida para a amplificação pela *Taq* DNA polimerase. O DNA genômico foi utilizado como molde para a amplificação das ORFs (*full-length*) *parB*, *parQ*, *livB* e *livQ*. Os oligonucleotídeos *forward* e *reverse* (listados na Tabela 1) com sítios de restrição para as enzimas *NdeI* e *XhoI*, respectivamente. Foram sintetizados através da empresa *Life Technologies* e utilizados na amplificação a fim de facilitar a clonagem dentro do vetor de expressão escolhido.

Tabela 1. Oligonucleotídeos iniciadores utilizados na amplificação por PCR. As bases em azul representam o sítio de restrição para a enzima *NdeI* e em vermelho o sítio de restrição da enzima *XhoI* presentes na sequência das ORFs de interesse nas quais codificam para a metionina inicial.

PROTEÍNA	OLIGOS	SEQUÊNCIA
ParB	ParB_Fw	5' CGT <b>CATATG</b> ACTCAAATTTTCGCGACCCTG'3
	ParB_RV	5' ATT <b>CTCGAG</b> TCAGTCGTCCAGGAACCG'3
ParQ	ParQ_FW	5' AAT <b>CATATG</b> GAGCACCTACGGGGACCCTC'3
	ParQ_RW	5' ATT <b>CTCGAG</b> TCACGCATGAGCGGCGTG'3
LivB	LivB_FW	5' CGG <b>CATATG</b> ACGCGAAATTCCTCCACG'3
	LivB_RV	5' ATT <b>CTCGAG</b> TCAGTCGTCCAGGAGACG'3
LivQ	LivQ_FW	5' CAT <b>CATATG</b> GAGCGCCTGCGGGGACC'3
	LivQ_RW	5' ATT <b>CTCGAG</b> TCACACATGGGCGGCGCG'3

Fonte: Elaborado pelo autor

A amplificação via PCR foi realizada em um termociclador T100 ThermalCycler (Biorad) empregando-se um volume final de 50  $\mu$ L constituído de 100 pmol de cada *primer*, 100ng de DNA genômico, 10 mM de dNTPs, 10% de tampão 5X phusion buffer, 1 U de polimerase *TaqPhusion High Fidelity* 10% de DMSO.

As condições de amplificação compreenderam uma desnaturação inicial a 95°C por 5 minutos, seguida de 30 ciclos constituídos de desnaturação a 95°C por 45 s, hibridização (*annealing*) variando de 55 e 58°C (de acordo com o gene empregado) por 50 s e extensão dos *primers* a 72 °C por 60 s; com extensão final a 72 °C por 7 min. Ao final dos ciclos as amostras foram mantidas a 4°C.

Os produtos das amplificações foram analisados aplicando-se 5 $\mu$ L das misturas de reação em gel de agarose a 1% contendo 0.005% de *Envirosafe* (Helixx Technologies), em tampão 0.5X TAE (80 mM Tris; 40 mM ácido acético glacial; 2,5 mM EDTA, pH 8,3) e em seguida foi submetido a uma tensão de 70 V por 45 min. Em seguida, as bandas foram visualizados sob luz UV após migração.

### 3.2 Purificação dos Produtos Amplificados por PCR.

Os produtos de PCR foram extraídos do gel e purificados utilizando-se o *GeneJET Gel Extraction and DNA Cleanup Micro Kit*<sup>®</sup> (ThermoScientific), de acordo com as instruções do fabricante. O resultado da extração foi analisado por eletroforese em gel de agarose 1% e quantificado por espectrofotometria no UV-visível.

### 3.3 Clonagem em Vetor de Clonagem Zero Blunt™

O sistema Zero Blunt™ PCR cloning Kit (ThermoFisherScientific) é uma tecnologia que contém plasmídeo projetado para clonar fragmentos de PCR com extremidades abruptas. O vetor pCR<sup>®</sup>-Blunt contém um gene letal para *E. coli* *ccdB*, fusionado com a região C-terminal do gene *lacZ* $\alpha$ . A ligação do fragmento a ser clonado interrompe a expressão do gene de fusão *lacZ* $\alpha$ -*ccdB* permitindo, assim, apenas o crescimento de recombinantes positivos após transformação. Nas células com vetores não-recombinantes, há expressão do gene *ccdB*, o que impossibilita o crescimento destas células.

As sequências codificantes dos genes de interesse amplificadas, foram inseridas no vetor pCR<sup>®</sup>-Blunt através de uma reação de ligação. Cada mistura de ligação foi preparada empregando-se um volume final de 10  $\mu$ L constituídos de 1-5 $\mu$ L produtor de PCR que garanta razão molar 10:1 (inserto:vetor), 1  $\mu$ L da enzima T4 DNA ligase (5U), 2 $\mu$ L do tampão de ligação opara a enzima ExpressLink™ T4 DNA ligase (5X) e 1 $\mu$ L da solução contendo o vetor pCR<sup>®</sup>-Blunt a 25 ng. Água tipo I foi utilizada para completar o volume total. Esta reação foi mantida a temperatura ambiente por 4 horas. As misturas de ligação foram utilizadas na transformação por choque térmico de células competentes de *E. coli* DH5 $\alpha$  (Promega), competentes por tratamento com cloreto de rubídio.

### **3.4 Transformação da Bactéria *Escherichia coli* DH5 $\alpha$ com os Plasmídeos Recombinantes pCR $\text{\textcircled{R}}$ -Blunt**

As suspensões de células competentes DH5 $\alpha$ , produzidas no momento de uso ou estocadas a -80 °C, foram mantidas durante pelo menos 10 min em gelo. A cada 50  $\mu\text{L}$  de células DH5 $\alpha$  foram adicionados 2  $\mu\text{L}$  das reações de ligação. As misturas resultantes foram homogeneizadas gentilmente e mantidas em gelo por 10 min. Em seguida, estas misturas foram submetidas à incubação em banho de água a 42 °C por 40 segundos e subsequente em banho de gelo por mais 5 min (choque térmico). Após este período, 100  $\mu\text{L}$  de meio LB foram adicionados a cada um dos tubos contendo as células transformadas com as reações de ligação. As novas misturas foram mantidas a 37 °C por 1 hora sob agitação de 200 rpm. Estas foram plaqueadas em placas de Petri contendo 20 mL de meio seletivo Luria-Bertani sólido (LB; 15 g extrato de levedura, 10 g de triptona, 10 g de NaCl e 15 g de ágar por litro de solução a pH 7,5) suplementado com 10  $\mu\text{L}$  do antibiótico de seleção canamicina a 100 mg mL<sup>-1</sup> e incubadas *overnight* a 37 °C em estufa.

### **3.5 Análise de Restrição de Recombinantes pCR $\text{\textcircled{R}}$ -Blunt de *E. coli* DH5 $\alpha$**

Após o período de incubação, algumas colônias resistentes (possíveis recombinantes ou clones positivos) foram selecionadas a fim de realizar a extração dos seus DNAs plasmidiais.

As colônias previamente selecionadas foram repicadas, individualmente, em tubos estéreis para cultivo de bactérias contendo 10 mL de meio LB líquido contendo 10  $\mu\text{L}$  de canamicina a 50 mg mL<sup>-1</sup> (pré-inóculo). Os tubos foram incubados a 37 °C por 16 horas, sob agitação de 200 rpm. A seguir, as células tiveram seus DNAs plasmidiais extraídos através do *kit GeneJ ET Plasmid Miniprep*<sup>®</sup> (ThermoScientific) para obtenção dos genes a serem posteriormente subclonados no vetor de expressão.

A primeira análise de restrição foi realizada empregando-se um volume final de 10  $\mu\text{L}$  constituídos de 1-2 $\mu\text{g}$  dos DNAs plasmidiais, 1  $\mu\text{L}$  da enzima *NdeI* (10 U), 0,5 $\mu\text{L}$  da enzima *XhoI* (10 U) e 1  $\mu\text{L}$  do tampão de enzima R (10X), estes últimos fornecidos pelo fabricante. Água tipo 1 foi utilizada para completar o volume total. Esta reação foi mantida

a 37 °C em estufa por 4 horas. Os produtos de digestão foram analisados em gel de agarose preparativo 1% e visualizados sob luz UV.

### **3.6 Subclonagem dos Genes no Plasmídeo de Expressão pET28a**

O plasmídeo do sistema pET utilizado possui um gene que confere resistência à canamicina e podem apresentar uma sequência nucleotídica na região N- ou C- terminal do sítio de múltipla clonagem que codifica seis histidinas denominada *His- tag*. Assim, as ORFs de interesse foram subclonados em fusão com esta sequência facilitando a purificação das proteínas recombinantes através de cromatografia de afinidade em metal imobilizado.

As sequências codificantes de interesse foram inseridas no vetor pET28a através de uma nova reação de ligação. Cada mistura de ligação foi preparada empregando-se um volume final de 10 µL constituídos de 1,5 µL do inserto a 100 ng/µL, 1,7 µL do vetor pET28a a 90 ng/µL, 5 µL do tampão de ligação para a enzima T4 DNA ligase (2X) e 1 µL da enzima T4 DNA ligase (6U). Água tipo 1 foi utilizada para completar o volume total. Esta reação foi mantida a 20 °C por 24 horas.

As novas misturas de ligação foram utilizadas na transformação por choque térmico de células competentes de *E. coli* DH5α para nova propagação plasmidial.

### **3.7 Análise de Restrição dos Plasmídeos Recombinantes pET28a de *E. coli* DH5α**

Após o período de incubação, algumas colônias resistentes foram selecionadas para realizar a extração dos seus DNAs plasmidiais. A seleção dos recombinantes foi realizada através da análise de restrição a fim de verificar a eficiência na transformação e propagação e obter quantidade suficiente dos plasmídeos para posterior transformação em linhagens de expressão.

As colônias selecionadas foram repicadas, individualmente, em 10 mL de meio LB líquido contendo 10 µL de canamicina a 50 mg mL<sup>-1</sup>. Os tubos foram incubados *overnight* a 37 °C em estufa sob agitação por 16 horas. As células tiveram seus DNAs plasmidiais

extraídos através do *kit Gene JET Plasmid Miniprep*<sup>®</sup> (ThermoScientific) quantificados e submetido ao sequenciamento automático.

A análise de restrição foi realizada empregando-se um volume final de 10 µL de acordo ao procedimento descrito na seção 3.5. Uma vez confirmada a presença dos DNAs plasmidiais nestas células, eles foram utilizados na transformação por choque térmico de células competentes de *E. coli* BL21(DE3) para expressão heteróloga.

O sequenciamento de DNA foi realizado utilizando-se o método do dideoxinucleotídeo marcado (*dideoxy chain termination method*) com auxílio de um sequenciador automático *ABI 3730 DNA Analyser*(ThermoScientific) (Centro de Pesquisa sobre Genoma Humano e Células Tronco- Instituto de Biologia/USP-SP), seguindo o protocolo do fabricante. Após o sequenciamento, para verificar a integridade da sequência obtida, esta foi submetida a uma busca em bancos de dados públicos como NCBI através da ferramenta BLAST (<http://www.ncbi.nlm.nih.gov/>).

Os genes *neoB* e *neoQ* de *S. fradiae* NCIMB 8233 foram cortesmente repassados por nossos colaboradores da Universidade de Cambridge, os quais vieram clonados em pET28a(+) nos sítios *NdeI/HindIII*, *BamHI/XhoI*, respectivamente.

### **3.8 Transformação da Bactéria *E. coli* BL21(DE3) com os Plasmídeos Recombinantes pET28a**

O procedimento empregado na transformação de células competentes de *E. coli* BL21(DE3) foi o mesmo descrito na transformação de células DH5α (seção 3.4). Porém, 400ng de cada solução de DNA plasmidial foram utilizados para transformar por choque térmico células competentes de *E. coli* BL21(DE3). A linhagem empregada nesta etapa do trabalho também passou pelo tratamento com cloreto de rubídio.

### **3.9 Expressão Heteróloga das Proteínas Recombinantes**

Para realizar os testes de expressão foram feitos pré-inóculos das colônias resultantes da transformação de células BL21(DE3) selecionadas em meio sólido com antibiótico. Colônias isoladas contendo cada uma destas construções foram escolhidas

aleatoriamente. Estas colônias foram inoculadas, individualmente, em 50 mL de meio LB líquido contendo canamicina (50 µL a 50 mg mL<sup>-1</sup>) e cultivadas *overnight* a 37 °C sob 200 rpm de agitação.

As culturas foram então diluídas na proporção 1:100 com LB líquido contendo canamicina e crescidas, novamente, a 37 °C até alcançar a OD<sub>600</sub> de aproximadamente 0.6. Antes da adição de IPTG no restante da cultura, alíquotas de 250 µL de cada frasco, representando os controles não induzidos foram centrifugadas a 2355 g por 5 min em microcentrífuga, ressuspensas em 20 µL de água tipo 1 e congeladas a - 20°C.

O restante das culturas de células foi induzido com IPTG nas concentrações de 0,2 mM; 0,4 mM; 0,6 mM; 0,8 mM e o crescimento foi mantido por um período de 2, 4, 8 e 16 horas nas temperaturas de 18 °C, 25 °C, 30 °C e 37 °C. Alíquotas de 100 µL da cultura induzida foram retiradas conforme hora demarcada para avaliação do nível de expressão das proteínas recombinantes. Todos os precipitados (*pellets*) provenientes das alíquotas retiradas durante a etapa de expressão proteica foram adequadamente ressuspensos e preparados de acordo com as subseqüentes análises em gel SDS-PAGE 10%.

De acordo com o resultado do teste de expressão, a melhor condição encontrada foi utilizada para superexpressão das enzimas a partir de 4-10 litros de cultura celular.

### **3.10 Purificação das Proteínas Recombinantes**

Os métodos cromatográficos mais empregados para isolamentos de proteínas são cromatografias por afinidade, troca iônica e de exclusão por peso molecular. No presente trabalho, os métodos aplicados para isolar as proteínas recombinantes foram cromatografias de afinidade em resina de níquel e exclusão por tamanho em coluna Superdex™ 200-HL 16/600 (General Eletrics-GE, USA).

A solução tampão utilizada na purificação das proteínas recombinantes foi inicialmente, tampão de lise Tris-HCl constituído de 20 mM de Tris base, 150 mM de NaCl pH 8,0 ajustado com HCl concentrado e posteriormente adaptado para 50mM Tris-HCl pH 7.5, 100mM KCl e 10% Glicerol e o tampão de eluição, semelhante ao de lise mas acrescido de 500mM de imidazol.

Os *pellets* bacterianos descongelados obtidos a partir de 6-10L de cultura induzida foram ressuspensos em 50-80mL de tampão de lise Tris-HCl, acrescentando lisozima (10ng/mL), DNase (1mg/mL) e PMSF (100mM) e mantidos em gelo por 30 min.

As suspensões celulares mantidas em gelo foram submetidas à lise celular utilizando-se o sonicador Sonics Vibra-Cell VCX 750™ com sistema de controle de amplitude Optotune™ (Vibra-Cell Processors) acoplado a uma sonda *microtipe* uma fonte de alta voltagem. O processo de sonicação foi realizado empregando-se 10 a 20 ciclos de 30 segundos com intervalos de 30 segundos. Os extratos proteicos brutos provenientes das células lisadas foram clarificados por centrifugação a 33250 g por 60 min a 4°C. Os sobrenadantes foram empregados na purificação por cromatografias de afinidade e exclusão molecular.

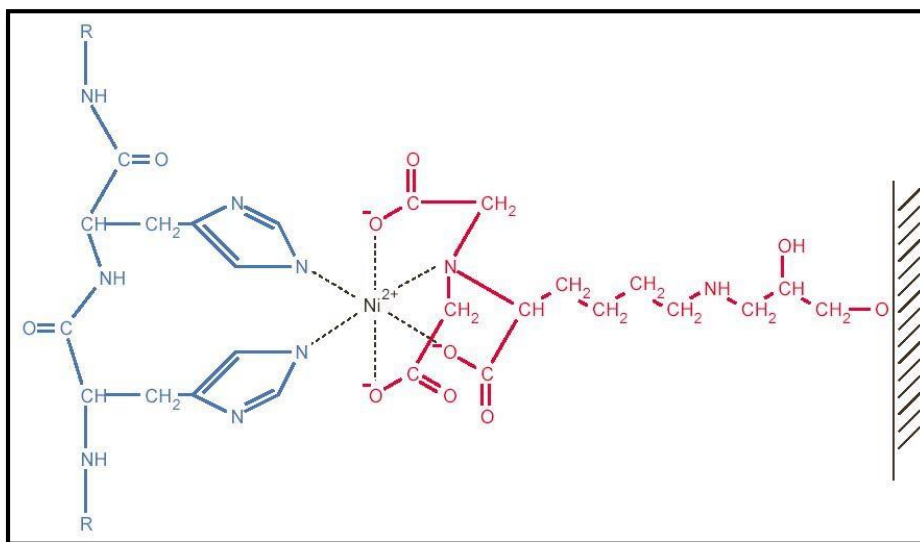
Para as proteínas recombinantes amidinohidrolases, os tampões utilizados foram os seguintes:

- AMH\_828, AMH\_821 e Medi\_4948, tampão A composto por 50mM Tris-HCl pH 8.0, 150 mM NaCl; 1 mM DTT, 10% Glicerol e 10 mM Imidazol; Tampão B idem ao tampão A acrescido de 500 mM de imidazol. Tampão C composto por 20 mM Hepes pH 7.5, 150 mM NaCl, 10% Glicerol e 1 mM DTT.

### **3.11 Cromatografia de Afinidade em Resina de Níquel**

A técnica empregada para a purificação das proteínas recombinantes contendo cauda de hexahistidinas utiliza a prerrogativa da cromatografia de afinidade ao metal. Neste tipo de cromatografia, íons metálicos são imobilizados pelo uso de um agente quelante que é capaz de propiciar o metal para ligação à proteína. Certos aminoácidos, principalmente histidina, apresentam alta especificidade de ligação pelo metal imobilizado. Assim, proteínas ricas em histidinas ou com um grupo destes aminoácidos podem ser seletivamente eluídas da resina carregada com íons metálicos e, conseqüentemente, isoladas por este método. Uma representação da interação dos resíduos de histidinas com a matriz Ni-NTA é apresentada na Figura 15.

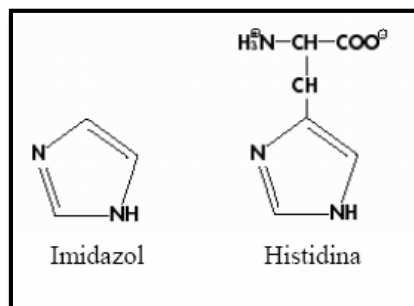
Figura 15. Interação entre resíduos vizinhos da cauda His-Tag e a matriz Ni-NTA. O quelante em vermelho, o ácido nitrilotriacético, coordena-se com quatro dos seis sítios ligantes do Níquel (II). Os dois sítios livres da esfera de coordenação interagem com os anéis imidazol da cauda de 6X histidinas das proteínas recombinantes (em azul)



Fonte: Qiagen, 2003.

A eluição das proteínas de interesse é realizada utilizando-se uma solução contendo o composto imidazol, que também possui o anel imidazólico, mas tem maior afinidade ao Níquel (II) do que a histidina. A Figura 16 apresenta a estrutura molecular do imidazol e da histidina que se ligam aos íons níquel que estão imobilizados pelos grupos NTA na matriz.

Figura 16. Estrutura molecular do composto imidazol e do aminoácido histidina.



Fonte: Qiagen, 2003

Os sobrenadantes foram aplicados nas colunas contendo 5 mL de resina de níquel *Ni-NTA HiTrap™* (GE Healthcare BioSciences) pré-equilibrada com tampão de lise. A resina foi então lavada com um gradiente inicial de imidazol variando de 0 a 100 mM em tampão de lise Tris-HCl. As etapas de lavagem consistiram de 20 mL de tampão de lise sem imidazol, seguido de 50 mL de tampão de lise contendo 5 mM de imidazol, e conseqüente, um gradiente iniciando em 5 mM até 500 mM do tampão de lise contendo imidazol num volume de 70 mL e por fim, 10 mL de tampão contendo 500 mM de imidazol. Para cada proteína recombinante obteve-se um perfil cromatográfico específico eluindo em diferentes concentrações de imidazol.

Alíquotas de 20 µL de todas as frações de lavagem e eluição foram avaliadas e o grau de pureza das proteínas recombinantes avaliado por eletroforese em gel de poliacrilamida 10%.

As frações provenientes das eluições contendo as proteínas recombinantes purificadas foram reunidas e concentradas através de concentradores Centriprep Ultra 10 da Amicon (Millipore Corporation, Bedford, MA, EUA). Cada proteína recombinante possuía um concentrador individual para evitar eventuais contaminações.

As concentrações das proteínas recombinantes purificadas foram determinadas medindo-se suas absorvâncias em espectrofotômetro UV-Vís a 280 nm. As absorvâncias foram relacionadas aos seus respectivos coeficientes de extinção molar obtidos através do cálculo do conteúdo teórico de resíduos de triptofanos, tirosinas e fenilalaninas (GASTEIGER *et al.*, 2005). Todos os estudos subsequentes foram realizados com as proteínas recombinantes contendo a cauda de 6 histidinas.

### **3.12 Cromatografia de Exclusão por Tamanho**

Como alternativa para a retirada do imidazol das frações proteicas provenientes da purificação em coluna de níquel e para prever o grau de oligomerização das proteínas recombinantes, a cromatografia de exclusão por tamanho foi realizada a temperatura ambiente empregando-se a coluna Superdex 200 HL 16/600, acoplada a um sistema HPLC preparativo (ÄKTA™-purifier 10, GE Healthcare BioSciences, Uppsala, Sweden). As membranas foram previamente lavadas com água tipo 1 e pré-equilibradas com

tampão de lise Tris-HCl. As frações eluídas foram monitoradas por absorvância no UV a 280 nm e coletados em frações de 0.8 mL empregando-se uma velocidade de fluxo de  $0,5 \text{ mL min}^{-1}$ . As frações contendo as proteínas recombinantes purificadas foram reunidas e concentradas através dos concentradores Centriprep Ultra 10 da Amicon (Millipore Corporation, Bedford, MA, EUA)

As concentrações das proteínas recombinantes purificadas foram determinadas medindo-se suas absorvâncias por espectrofotômetro UV a 280 nm. Todos os estudos subsequentes foram realizados com as proteínas recombinantes contendo a cauda de 6 histidinas.

### **3.13 Protocolo de *Thermal Shift* (HUYNH; PARTCH, 2015)**

Primeiramente uma placa de 96 poços com capacidade para 25  $\mu\text{L}$  em cada poço foi preparada de modo que cada poço tivesse um tampão específico 5 X mais concentrado. Em seguida, a proteína recombinante purificada foi diluída em tampão de diluição (10 mM fosfato de sódio pH 7.0, 50 mM NaCl) até completar o volume final de 5 mL (suficiente para uma placa de 96 poços). Adicionou-se 4  $\mu\text{L}$  da sonda fluorescente (concentração de estoque 5000 vezes) SYPRO Orange (ThermoFisher®, EUA), de modo que a concentração final de proteína e de sonda seja 5  $\mu\text{M}$  e 2 vezes respectivamente. A solução foi homogeneizada e 40  $\mu\text{L}$  dessa mistura foi distribuída em cada poço de uma placa de 96 poços para PCR com uso de uma pipeta multicanal. A placa foi brevemente centrifugada para remover bolhas e em seguida, 10  $\mu\text{L}$  de cada tampão a ser testado (os tampões estão na placa de 96 poços com capacidade para 2 mL) foram adicionados a cada poço da placa teste com uso de pipeta multicanal (volume final da reação 50  $\mu\text{L}$ ). A placa de teste foi selada com adesivo e centrifugada a 800  $\times g$  por 2 min a temperatura ambiente. Por fim a placa contendo as reações a serem testadas foi colocada em termociclador para PCR em tempo real (CFX Connect™ - BioRad) com a seguinte configuração: 1) Desnaturação inicial por 10 segundos e variação de temperatura de 25 °C com aumento gradativo de 0,5°C até a temperatura final de 95 °C.

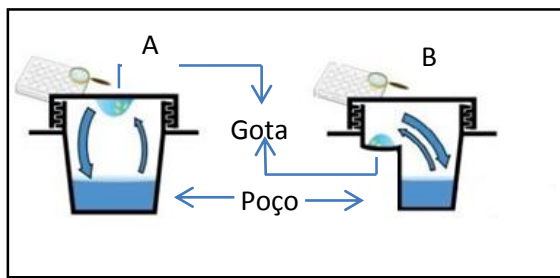
Curvas desnovelamento térmico foram determinadas a partir dos dados de PCR utilizando o software Origin (OriginPro9.0). Os dados foram calculados através de um ajuste não linear de parâmetros.

### 3.14 Experimentos de cristalização

Após a cromatografia de exclusão molecular, a proteína foi submetida a testes de cristalização pelas técnicas da gota suspensa e da gota depositada (Figura 17), sendo essa última realizada em robô Oryx4 (Douglas Instruments) nas dependências do Laboratório de Biologia Estrutural Aplicada (USP-São Paulo). Em ambas as técnicas, a cristalização ocorre por difusão de vapor, na qual um poço contendo a solução de cristalização e uma gota formada pela mesma solução e pela proteína encontram-se em um sistema vedado. Ocorre, então, através da difusão de vapor, uma redução do volume da gota que perde água, possibilitando a supersaturação da concentração da proteína e consequente formação dos cristais (Li; Chang, 2009).

A diferença entre as duas técnicas citadas acima está no fato de que, na gota suspensa, a gota é depositada de cabeça para baixo em uma lamínula, como mostrado na Figura 17-A. Já na gota depositada a gota é depositada sobre uma cavidade (Figura 17-B). Essa última é bastante utilizada em condições de baixa tensão superficial da gota.

**Figura 17. Ilustração das técnicas de gota suspensa e gota sentada utilizadas na cristalografia.** Os métodos contêm a gota de proteína e a solução de cristalização em um sistema vedado. (A) Gota suspensa: a gota fica dependurada em uma lamínula. (B) Gota sentada: a gota é colocada sobre uma pequena cavidade. Em ambos os métodos a cristalização acontece por difusão de vapor.



Fonte: Modificada de Wikiwand,2010.

Em uma triagem inicial, foram avaliadas mais de 500 condições de cristalização constituintes dos kits comerciais *Crystal Screen 1 e 2*, PEG-ION 1-2 (Hampton Research), Classic 1-4, Classic 5-8, JCSG, Pi-Minimal e Basic 1-4 (JenaBioscience) em placas MRC2 de 96 poços (Swissci) através da técnica da gota sentada. Essas soluções foram preparadas no laboratório de acordo com as formulações do kit. Cada condição de cristalização do kit é constituída de um tampão em pH específico, um sal e um precipitante, mas em algumas soluções o tampão e/ou sal estão ausentes.

Nesses testes, 400  $\mu\text{L}$  da solução de cristalização foi adicionado a cada poço e nas lamínulas foram adicionadas gotas com 1  $\mu\text{L}$  da proteína e 1  $\mu\text{L}$  da solução do poço. Já nos testes de gota sentada, foram utilizados 1  $\mu\text{L}$  de proteína e 1  $\mu\text{L}$  da solução de poço em uma placa de 96 poços. Cada poço foi preenchido com 30  $\mu\text{L}$  da solução de cristalização. A concentração da proteína utilizada foi aproximadamente 7-10 mg/mL, tanto nos testes da gota suspensa quanto nos testes da gota sentada.

Após o *screening* inicial, algumas condições com crescimento de pequenos cristais foram identificadas e elas foram posteriormente otimizadas em placas VDX de 24 poços (Hampton Research). Para cada uma dessas condições, foram variados (para valores superiores ou inferiores aos da condição inicial) o pH do tampão (na faixa de tamponamento permitida pelo mesmo), a concentração de precipitante e a concentração de sal.

Além do pH e concentrações de sal e precipitante, outra variação foi inserida nos testes de otimização. Foram testados os seguintes volumes e proporções de proteína para solução de cristalização: 1  $\mu\text{L}$  proteína: 1  $\mu\text{L}$  solução de cristalização, 2  $\mu\text{L}$  proteína: 1  $\mu\text{L}$  solução de cristalização e 0.5  $\mu\text{L}$  proteína: 1  $\mu\text{L}$  solução de cristalização.

Todas as placas de cristalização foram incubadas a 18°C e a presença de cristais foi observada através de um microscópio estereoscópio com aumento de até 100 vezes.

### 3.15 Coleta de dados de difração de raios X e processamento da Medi 4948 e NeoB

As coletas de dados de difração de raios X foram realizadas na linha W01B- MX2 (GUIMARAES; SANFELICI; NEUENSCHWANDER; RODRIGUES *et al.*, 2009) no LNLS. Foi coletado um conjunto de dados de um cristal nativo. Para realização da coleta de dados do cristal nativo no LNLS (Laboratório Nacional de Luz Síncrotron) –Campinas, Brasil, o mesmo foi transferido por alguns segundos com o auxílio de um *loop* de *nylon* para uma solução crioprotetora constituída por 30% glicerol em solução de cristalização. O cristal foi, então, congelado em um fluxo de nitrogênio.

Durante a coleta de dados do cristal nativo, este foi exposto por 0,5 segundos para coleta de cada imagem de difração. Ao todo foram coletadas 3600 imagens com 0,1 graus por imagem, totalizando 360 graus de rotação.

O processamento dos dados de difração de Raios X foi realizado pelo programa XDS (KABSCH, 2010) e as reflexões foram escalonadas através do Aimless (EVANS; MURSHUDOV, 2013) presente no pacote de programa CCP4 (Bailey, 1994). As fases foram obtidas por substituição molecular através do Phaser (MCCOY; GROSSE-KUNSTLEVE; ADAMS; WINN *et al.*, 2007) utilizando como modelos iniciais as entradas do PDB:1GQ6\_A O limite de resolução foi determinado baseado em  $I/\sigma \geq 2,0$ . O refinamento e as inspeções visuais dos modelos foram realizados com REFMAC V. 5.0.32 (MURSHUDOV; VAGIN; DODSON, 1997) e COOT (EMSLEY; LOHKAMP; SCOTT; COWTAN, 2010) usando mapas de densidade eletrônica 2Fo-Fc, respectivamente. As figuras foram preparadas com Pymol (Delano, 2002).

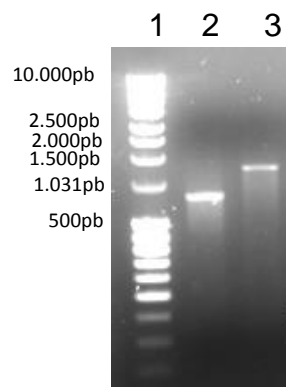
## 4. Resultados e Discussão

### 4.1 Clonagem

Com base na temperatura de hibridização utilizada para o par de oligonucleotídeos desenhado, foi possível à amplificação desejada para *parB*. No entanto, bandas inespecíficas estiveram presentes em ambas as extremidades do gel utilizado para avaliar amostra referente *parQ*. Estas foram atribuídas a prováveis excessos de dNTPs, DNA genômico e eventuais subprodutos da reação de amplificação. Diante deste primeiro resultado, fez-se necessário um processo de otimização da PCR a fim de garantir o máximo de eficiência na obtenção do produto desejado. A amplificação desejada para *parQ* foi possível quando a temperatura de hibridização dos *primers* foi alterada para 58°C.

A Figura 18 ilustra o resultado das amplificações as quais geraram um produto de 1.245 e 1.641 pares de base. O produto de amplificação obtido pela PCR teve o seu tamanho correspondente com o previsto para *parB* e *parQ*, respectivamente, não apresentou bandas inespecíficas, indicando o sucesso da otimização e resultando em quantidades suficientes para a etapa seguinte da clonagem.

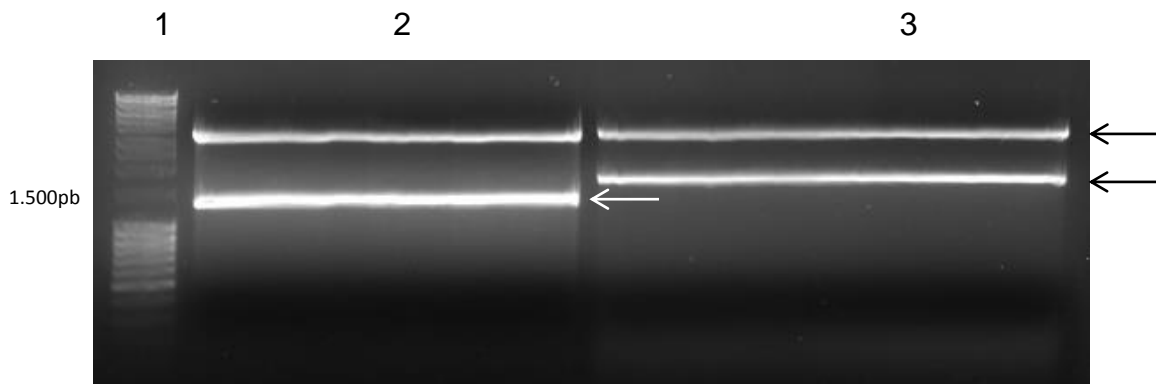
Figura 18. Amplificação dos genes codificadores das proteínas ParB e ParQ. A coluna 1 é o marcado de peso molecular 1kb plus Mass Ruler DNA ladder da Thermo Scientific. A coluna 2 é *parB* 1.245pb amplificada e a coluna 3 é *parQ* 1.641pb amplificada. O meio de separação foi gel de agarose 1% em tampão 0.5xTAE. A voltagem aplicada na separação foi 7cm/V. Os fragmentos de DNA no gel foram visualizados em transiluminador UV a 356nm.



Fonte: Elaborado pelo autor

A análise de restrição empregando as enzimas NdeI e XhoI foi realizada a fim de confirmar a presença do inserto nos clones selecionados. Nesta análise, o tamanho do inserto foi avaliado pela clivagem do plasmídeo em sítios de restrição adjacentes ao sítio de clonagem, com posterior eletroforese em gel de agarose. A Figura 19 apresenta o perfil de restrição dos clones selecionados.

Figura 19. Separação dos plasmídeos digeridos ZeroBlunt contendo genes *parB* e *parQ*. A coluna 1 é o padrão de peso molecular 1kb plus Mass Ruler DNA ladder da ThermoScientific. As colunas 2 e 3 são os clones positivos selecionados a serem utilizados na etapa de subclonagem nos vetores de expressão. O meio de separação foi gel de agarose 1 % em tampão 0.5X TAE, pH 8,3. A voltagem aplicada na separação foi 7cm/V . Os fragmentos de DNA no gel foram visualizados em transiluminador UV a 365 nm. As bandas específicas dos genes *parB* e *parQ* e dos plasmídeos digeridos estão representados por setas cujos tamanhos são aproximadamente 3500 (vetor), 1.245pb (*parB*) e 1.641pb (*parQ*) respectivamente.



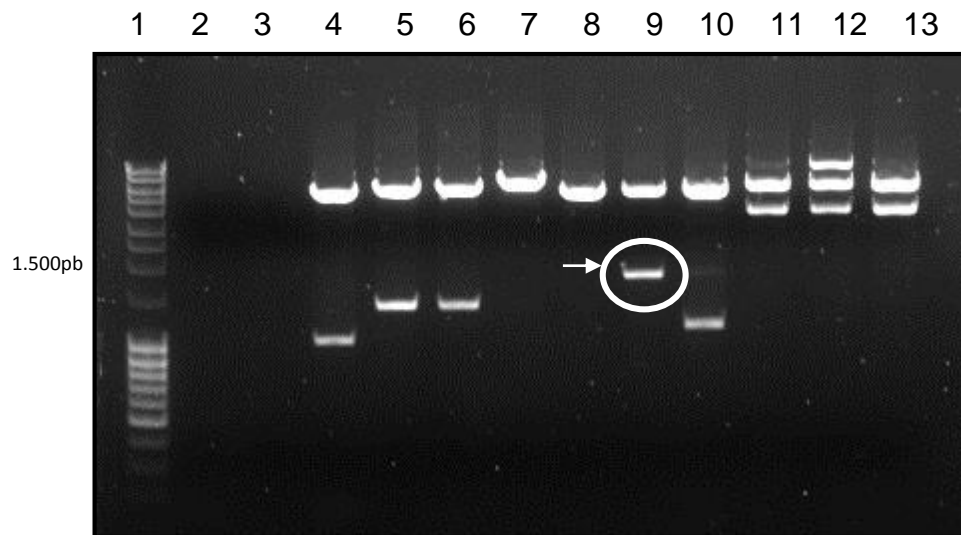
Fonte: Elaborado pelo autor

A presença e o tamanho dos insertos confirmaram o sucesso das etapas de clonagem e transformação das bactérias. Os produtos da digestão de 1.254 pb e 1.641 pb correspondente a *parB* e *parQ* foram retirados do gel, purificados utilizando-se novamente o *kit GeneJET Gel Extraction and DNA Cleanup*, quantificados e utilizados na etapa de subclonagem no vetor de expressão. Estas amostras de DNA apresentaram em geral concentrações variando de 100 a 150 ng $\mu$ L<sup>-1</sup>

## 4.2 Subclonagem no Vetor de Expressão pET28a(+)

Os fragmentos de DNA de 1.245pb que codificam para a proteína ParB e 1.641pb que codificam para a proteína ParQ foram subclonados nos vetores pET28a. A Figura 20 ilustra o novo perfil de restrição destes doze clones selecionados, seis referentes a *parB* e outros seis referentes a *parQ*.

Figura 20. Separação dos plasmídeos digeridos pET28a contendo bandas referentes a *parB* e *parQ*. A coluna 1 é o padrão de peso molecular 1kb plus Mass Ruler DNA ladder da ThermoScientific. As colunas 2 a 7 são os clones referentes às *parB*, As colunas 8 a 13 são os clones referentes as *parQ* sendo somente visto um clone positivo na coluna 9. O meio de separação foi gel de agarose 1 % em tampão 0.5X TAE, pH 8,3. A voltagem aplicada na separação foi 70 V por 50 min. Os fragmentos de DNA no gel foram visualizados em transiluminador UV a 365 nm. A banda específica do gene *parQ* está representada por seta e círculo.



Fonte: Elaborado pelo autor

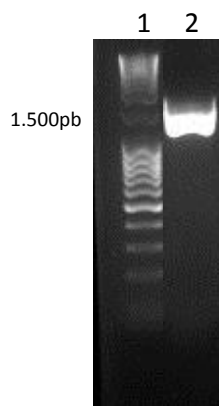
A presença do inserto em clone selecionado confirmou o sucesso das etapas de subclonagem e transformação das bactérias somente para a proteína ParQ. Uma vez confirmada a presença do inserto, o clone 2 (equivalente a 9ª canaletas do gel) o qual

apresenta maior intensidade de banda, foi utilizado para transformar as células competentes de expressão de *E. coli* BL21(DE3). A seleção do clone contendo a proteína de interesse foi subclonado em vetor de expressão e foi realizado após o plaqueamento das células em meio adequado, este clone foi empregado na produção da proteína expressa de forma heteróloga.

#### 4.3 Amplificação das enzimas correspondentes em *Streptomyces lividus*

Obtivemos a amplificação a partir do DNA genômico do *livB* (Figura 21) após ajuste dos parâmetros da PCR. Para o *livQ*, houve dificuldade quanto à amplificação da sequência e novos parâmetros devem ser estabelecidos em trabalhos futuros.

Figura 21. Amplificação do gene codificador da proteína LivB. A coluna 1 é o marcado de peso molecular 1kb plus Mass Ruler DNA ladder da ThermoScientific. A coluna 2 é *livB* 1.245pb amplificada. O meio de separação foi gel de agarose 1% em tampão 0.5xTAE. A voltagem aplicada na separação foi 70V por 50min. Os fragmentos de DNA no gel foram visualizados em transiluminador UV a 356nm.

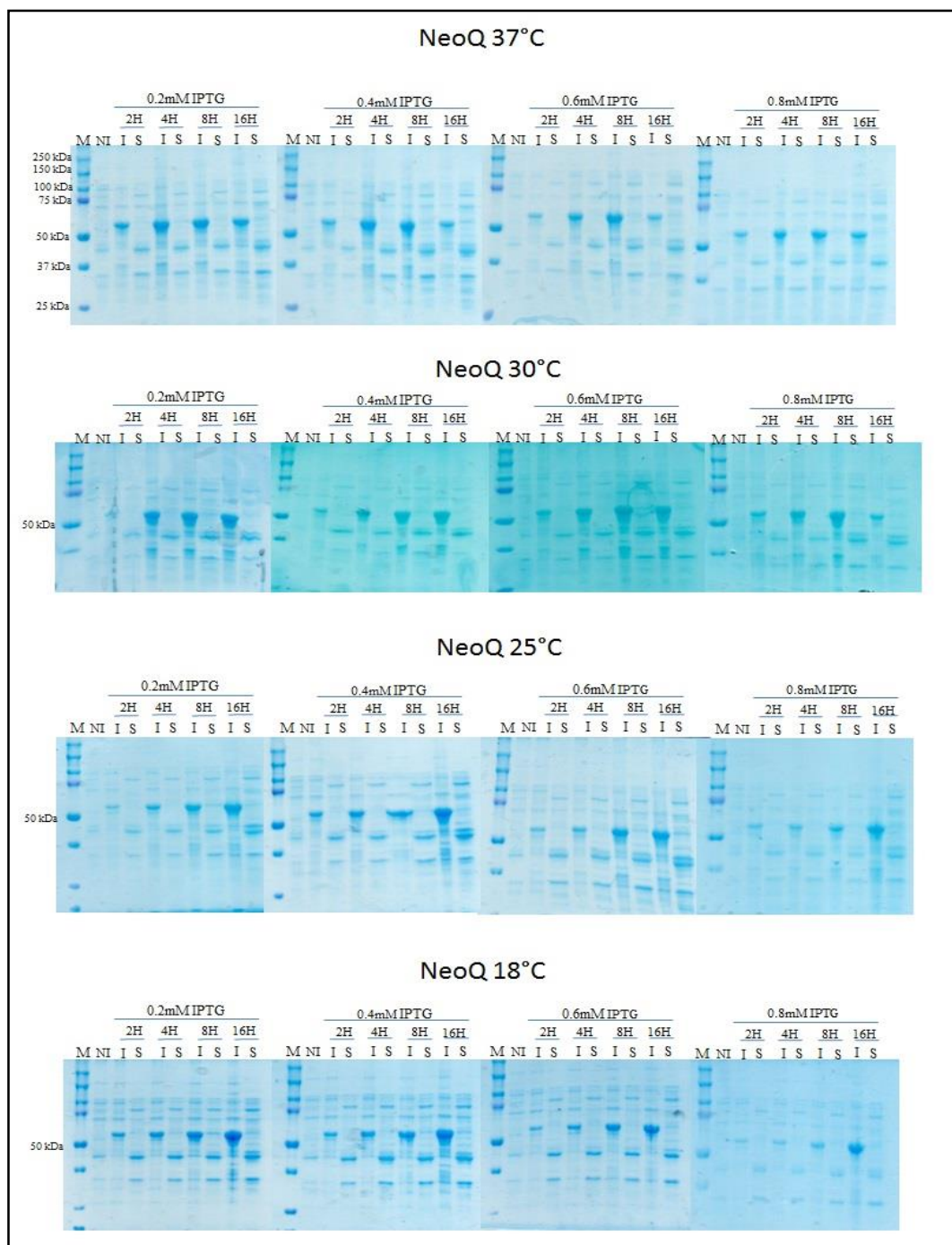


Fonte: Elaborado pelo autor

#### **4.4 Teste de Expressão e Expressão Heteróloga das Proteína NeoB, NeoQ, ParB e ParQ a partir dos Vetores de Expressão**

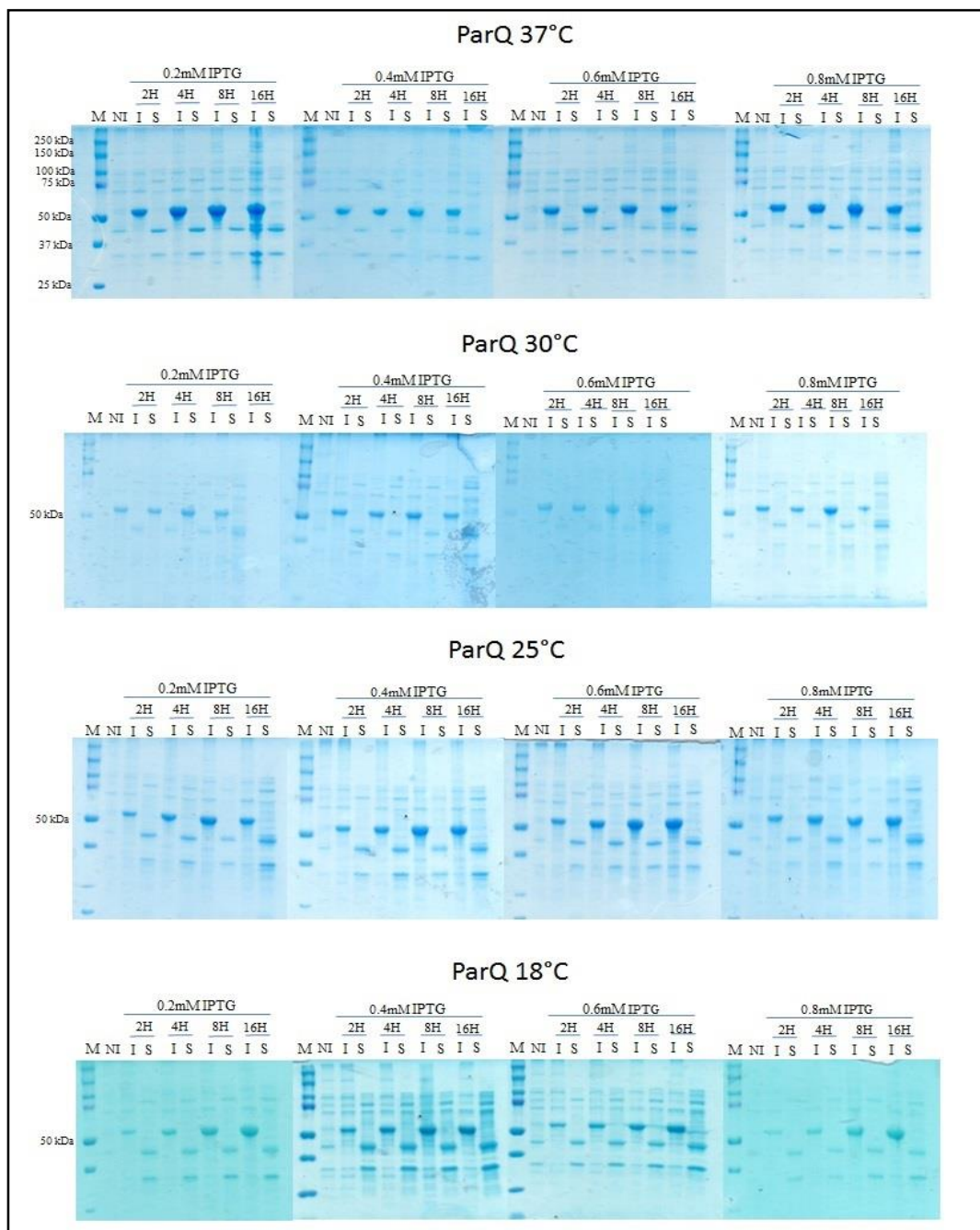
Devido a um baixo rendimento observado, no final da purificação, referente a enzima NeoQ, foi realizado um teste de expressão a fim de melhorar níveis de expressão (Figura 22). Para a proteína ParQ, também foi realizado teste de expressão em célula de BL21(DE3), variando concentração de IPTG e tempo de indução (Figura 23).

Figura 22. Geis SDS-PAGE 10% do teste de expressão para NeoQ. Ensaio de expressão nas temperaturas de 37°C, 30°C, 25°C e 18°C para diferentes concentrações de IPTG (0.2mM; 0.4mM; 0.6mM e 0.8mM) por distintas horas de indução (2h, 4h, 8h, 16h). A proteína NeoQ tem peso molecular de 58kDa além da cauda de histidina (~3,2kDa).



Fonte: Elaborado pelo autor

Figura 23. Geis SDS-PAGE 10% do teste de expressão para ParQ. Ensaio de expressão nas temperaturas de 37°C, 30°C, 25°C e 18°C para diferentes concentrações de IPTG (0.2 mM; 0.4 mM; 0.6 mM e 0.8 mM) por distintas horas de indução (2h, 4h, 8h, 16h). A proteína ParQ tem massa molecular de 58kDa além da cauda de histidina (~3,2kDa).



Fonte: Elaborado pelo autor

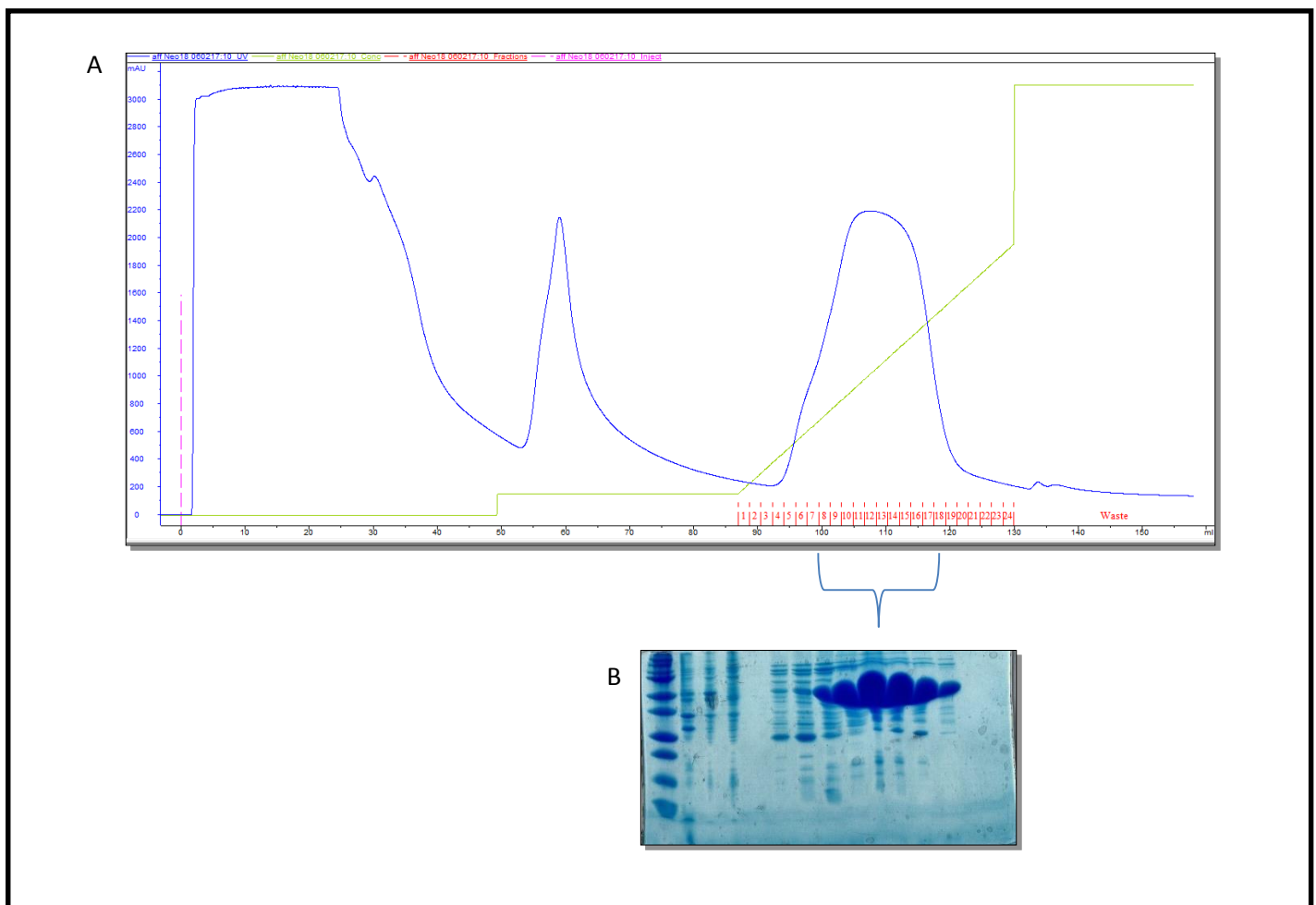
Os resultados indicam que a expressão da ParQ não foi identificada na fração solúvel, sendo visivelmente presente na fração insolúvel. Todas as condições testadas não foram favoráveis para solubilização da enzima. Para a enzima NeoQ, é possível ver nos géis que houve uma pequena expressão fração solúvel quando induzida por 0.2mM de IPTG por 8h e 16h. No entanto, tal rendimento é relativamente baixo quando comparado à banda presente na fração insolúvel, indicando que a maior parte que compõem a proteína se encontra majoritariamente insolúvel.

Portanto, baseado nesses resultados, novos testes de expressão variando outras condições, como célula de expressão, por exemplo, devem ser realizados a fim de encontrar uma condição que possa tornar solúveis as proteínas em questão.

#### **4.5 Purificação da Proteína NeoB por Cromatografia de Afinidade e Exclusão por Tamanho**

As frações solúveis das culturas induzidas para produção de NeoB e NeoQ provenientes da lise celular foram submetidas à cromatografia de afinidade. A Figura 24 mostra o cromatograma referente NeoB a separação eletroforética em gel de poliacrilamida SDS-PAGE a 10%, dos eluídos da coluna de Ni-NTA.

Figura 24. Purificação da NeoB por cromatografia de afinidade. A – Cromatograma de purificação por afinidade contendo um gradiente de imidazol (5% até 100%). B- Análise das frações eluídas da coluna Ni-NTA foi realizada em SDS-PAGE a 10%. A coluna 1 é o marcador de massa molecular. A coluna 2 representa a fração do *pellet*. A coluna 3 a fração do extrato celular lisado. A fração 4 contém o que não foi ligada na coluna de afinidade. A coluna 5 representam a fração de lavagem com 25mM de imidazol. As colunas 6 a 12, as frações de lavagem com no gradiente de imidazol referentes ao pico obtido.



Fonte: Elaborado pelo autor

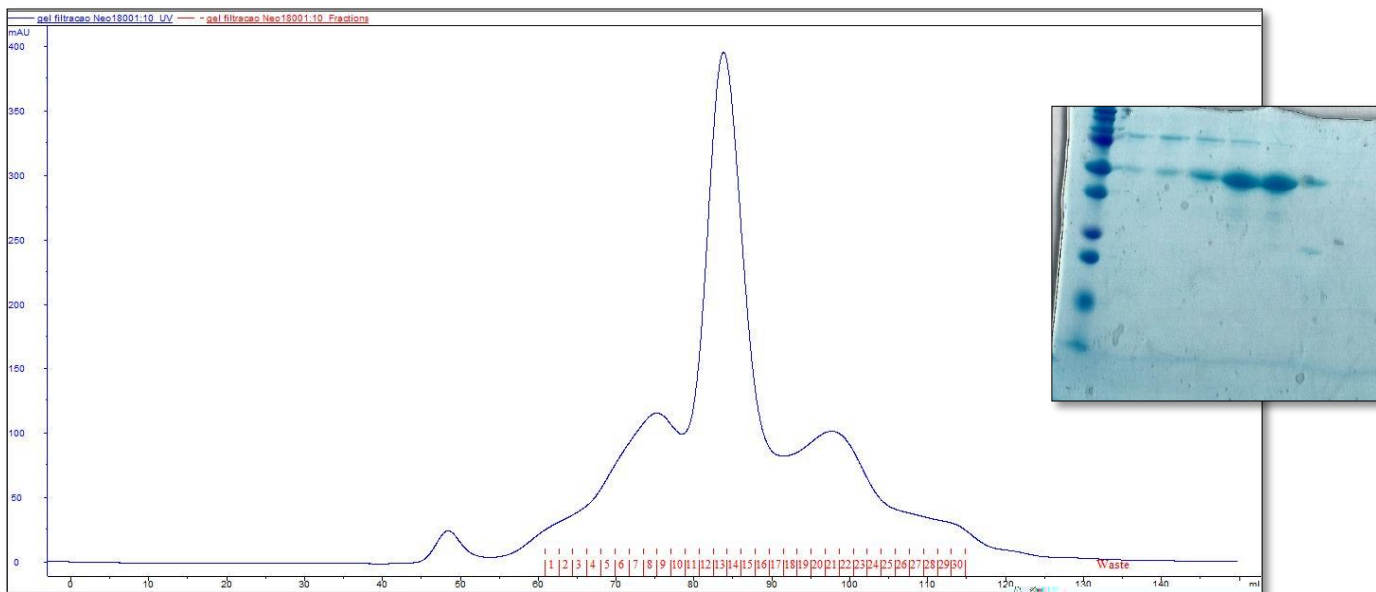
Em algumas das purificações realizadas neste estudo, eventuais proteínas contaminantes foram observadas nas frações de eluição. Para tentar contornar a presença de contaminantes após a purificação em resina de afinidade, a estratégia de

re-cromatografia de afinidade e cromatografia por exclusão molecular foram empregadas quando necessário.

A proteína NeoB submetida à cromatografia de exclusão por tamanho foi eluída constituinte pelo pico representado pelo volume de eluição próximo a 70mL (Figura 25). O emprego desta técnica mostrou-se capaz de separar a fração correspondente a proteína NeoB em fração distinta daquelas dos contaminantes e do imidazol. A eliminação do excesso de imidazol comumente presente após cromatografia de afinidade permitiu que a proteína de interesse fosse utilizada diretamente nas etapas de cristalização sem a necessidade de empregar a diálise para sua remoção, o que foi muito benéfico em termos de estabilidade proteica.

A purificação da proteína recombinante NeoB foi alcançada em duas etapas cromatográficas. A pureza foi confirmada pelas frações eluídas com tampão de lise 50mM trisHCl pH 7,5, 100mM NaCl contendo 500 mM de imidazol nas quais apresentaram uma única banda de aproximadamente 44 kDa, respectivamente, sob condições desnaturantes no SDS-PAGE.

Figura 25. Cromatograma de Exclusão molecular. Proteína NeoB eluída em torno de 70 mL.



Fonte: Elaborado pelo autor

As frações da NeoB correspondentes ao pico de maior UV possuíam coloração rosa (Figura 26). Estas, depois de confirmado o grau de pureza satisfatório, foram concentradas a fim de iniciar os testes de cristalização. No entanto, foi observado que grande parte da proteína estava precipitando durante a concentração em concentrador Amicon. O rendimento da purificação foi elevado, todavia, as perdas causadas pela precipitação da proteína foram inesperadas, levando-nos a buscar um recurso para superar tal situação. De igual modo, ocorreu perda expressiva para purificação da enzima NeoQ, não sendo possível realizar etapa de re-cromatografia por exclusão de peso molecular. Entretanto, para melhorar a estabilidade da proteína purificada NeoB, foi realizado um ensaio de deslocamento térmico.

Figura 26. Frações da proteína NeoB. Frações proteicas provenientes após cromatografia por exclusão molecular.



Fonte: Elaborado pelo autor

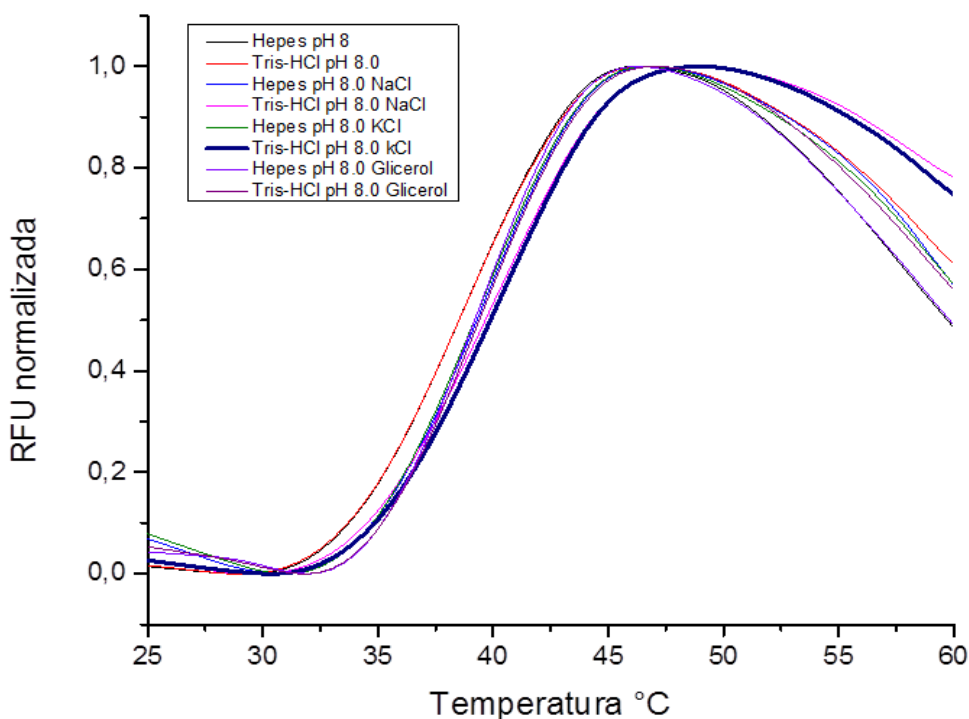
#### **4.6 Ensaio de Deslocamento Térmico (Thermal Shift) para NeoB em presença de diferentes tampões**

Testes iniciais foram realizados utilizando tampão Tris-HCL 50mM pH8.0 e 100mM de NaCl. No entanto, a proteína NeoB se apresentou muito instável em tal condição resultando em quantidade significativa de proteína precipitada. Frente à dificuldade de estabilização encontrada perante proteína em tampão inicial testado, foi realizado um

ensaio de deslocamento térmico na presença de diferentes tampões a fim de encontrar uma condição que se mostrasse mais favorável para estabilizar a proteína. *Thermal Shift* é uma técnica que permite verificar tampões e ligantes que maximizam a estabilidade de uma dada proteína. Neste ensaio, NeoB foi incubada com a sonda fluorescente SYPRO Orange e diferentes tampões, em seguida submetida a um aumento gradativo de temperatura. À medida que a proteína desnatura, esta expõe resíduos hidrofóbicos que se ligam a sonda emitindo uma fluorescência. Curvas de desenovelamento térmico foram então construídas de modo a permitir determinar a temperatura de *melting* de NeoB em cada condição testada. Mudanças na estabilidade da proteína podem ser correlacionadas com mudanças na temperatura de *melting*. Os resultados destes ensaios estão demonstrados na Figura 27.

A estabilidade de proteína recombinante NeoB aumenta quando incubada em tampão contendo como sal o KCl, glicerol e pH da solução em torno de 8.0. É possível ver no gráfico que a proteína apresentou maior estabilidade quando em presença de tampão Tris-HCl pH8.0 KCl, pois houve um deslocamento de temperatura da condição inicial em torno de 5 °C, evidenciando que as condições apresentadas pelo tampão favoreceram a proteína a se manter mais estável na variação de temperatura. Portanto, a fim de eleger um tampão que contivesse componentes favoráveis a estabilização, foi elaborado a solução tampão contendo 50mM Tris-HCl pH 7.5, 100mM KCl e 10% de glicerol. Esta condição foi utilizada para realizar nova purificação.

Figura 27. Thermal Shift da NeoB recombinante. No gráfico, as curvas de melting de NeoB na presença dos tampões que mais aumentaram a estabilidade da proteína quando comparados com a condição controle. Em vermelho condição inicial que a enzima se apresentava e em azul escuro a condição que melhor se mostrou favorável deslocando a temperatura para direita.

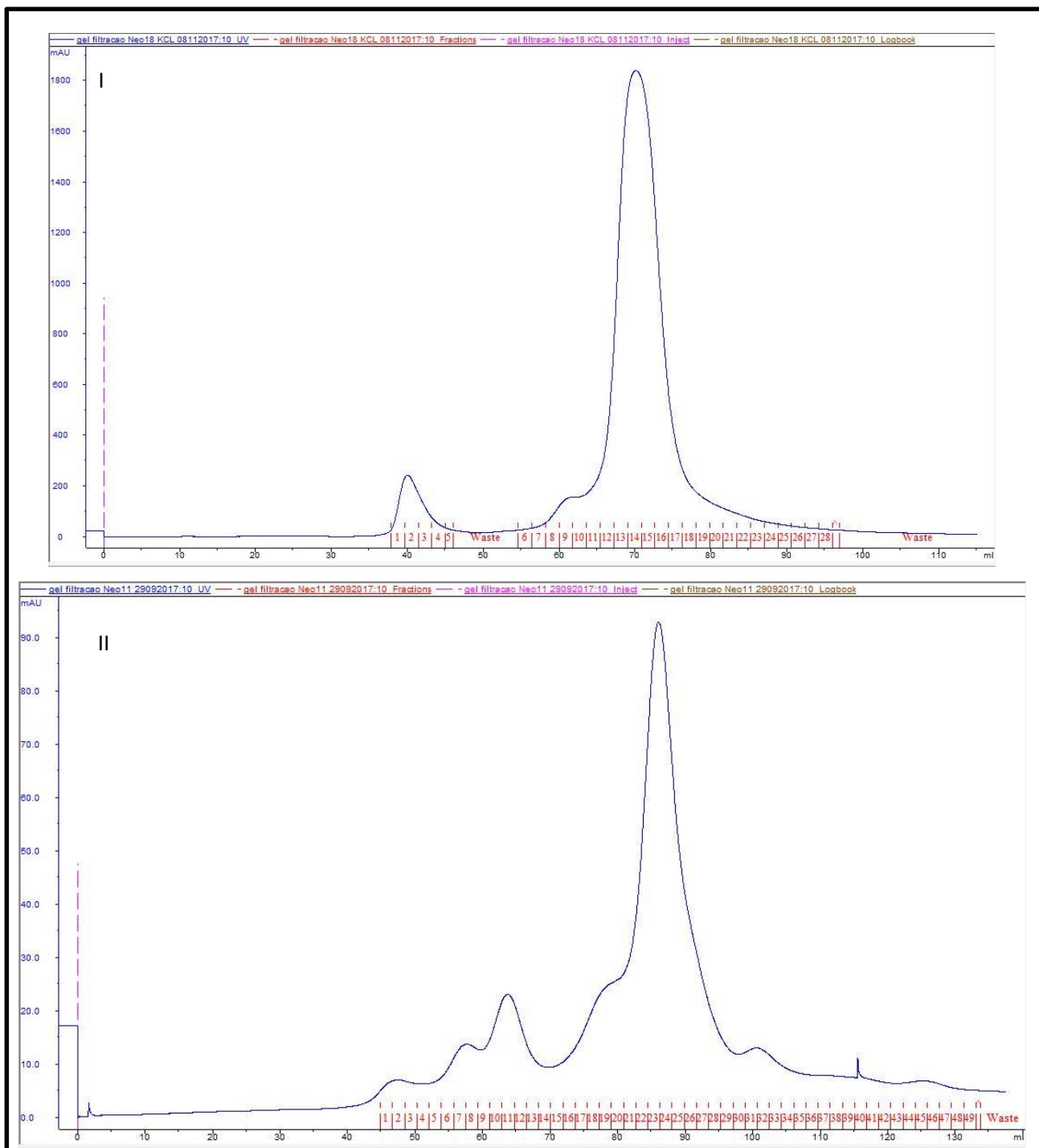


Fonte: Elaborado pelo autor

#### 4.7 Purificação das enzimas NeoB e NeoQ da biossíntese de Neomicina em condição obtida pelo *Thermal Shift*.

Com base no resultado de descolamento térmico, foi realizado uma nova purificação por afinidade em coluna de Ni-NTA seguida de purificação por exclusão peso molecular tanto para proteína NeoB quanto para NeoQ com o tampão 50mM Tris-HCl pH 7.5, 100mM KCl e 10% de glicerol. O perfil cromatográfico resultante da purificação por exclusão molecular pode ser observado na Figura 28. É possível observar que a proteína NeoB apresentou um pico em torno de 1800m mAU referente a enzima purificada, que foi concentrada e armazenada para passos seguintes de investigação cristalográfica.

Figura 28. Cromatograma de purificação por exclusão molecular das enzimas relacionadas com a biossíntese de neomicina. Picos maiores referentes a eluição da I-NeoB e II-NeoQ.



Fonte: Elaborado pelo autor

No entanto, como é possível observar, o pico referente a enzima NeoQ teve seu máximo em torno de 95mAU, sendo obtida pura, porém em baixas quantidades. O rendimento não foi satisfatório e estratégias de expressão foram revisadas a fim de aumentar o rendimento.

Não foi observada precipitação proteica quando purificada em tampão contendo Tris-HCl, KCl e glicerol. A solução tampão adotada para purificação foi aquela que contém 50mM Tris-HCl pH 7.5, 100mM KCl e 10% Glicerol, visto que tal condição soluciona o problema inicial de perda por precipitação.

#### **4.8 Cristalização da NeoB, coleta e processamento de dados**

A cristalografia de macromoléculas é, atualmente, uma das principais técnicas experimentais para obtenção de estruturas tridimensionais de proteínas. Assim, com o objetivo de obter a estrutura da proteína NeoB, testes de cristalização foram realizados tanto pelo método da gota suspensa quanto pelo método da gota sentada.

Os ensaios de cristalização foram iniciados com a proteína NeoB purificada por cromatografia de afinidade por níquel e exclusão molecular. Foram testadas aproximadamente 192 condições de cristalização utilizando o robô Oryx4 nas dependências do Laboratório Biologia Estrutural Aplicada (USP-São Paulo) utilizando kits de triagem padronizados e comercializados pela Hampton Research. A partir dos experimentos iniciais realizados, houve o surgimento de cristais após 5 dias em uma condição do Kit CrystalScreen composta por 30% PEG 8000, 0.2M Sulfato de Amônio, 0.1M Cacodilato de Sódio pH 6.5 (Figura 29)

Figura 29. Cristal da proteína NeoB na forma APO. Cristal com formato de losango na condição 30% PEG 8000, 0.2M Sulfato de Amônia, 0.1M Cacodilato de Sódio pH 6.5 após 5 dias.



Fonte: Elaborado pelo autor

Este cristal foi levado para linha MX2 no Laboratório Nacional de Luz Síncrotron (LNLS) (Campinas) para ser submetido a difração por raio X. Dentre os crio-protetores testados o glicerol na concentração 30% v/v foi considerado mais adequado ao experimento. A resolução dos dados de difração de raios X coletados foi aproximadamente de 1.8 Å. Os dados coletados foram, então, processados no programa XDS (KABSCH, 2010) no qual verificou-se que o cristal pertence ao sistema cristalino monoclinico (**Tabela 2**).

Tabela 2. Conjunto de dados e estatísticas do refinamento.

Coleta de Dados	Cristal NeoB
Fonte de raios X	CNPEM- LNLS- MX2-Campinas/Brasil
Comprimento de onda (Å)	1,458
Faixa de resolução (Å)	47,95 – 1,8 (1,864 – 1,8)
Grupo espacial	P 1 21 1
Célula unitária	
a, b, c (Å)	60,07 111,21 70,74
$\alpha$ , $\beta$ , $\gamma$ (°)	90 110,9 90

Reflexões totais	158678 (15837)
Reflexões únicas	79948 (7972)
Multiplicidade	2,0 (2,0)
Completeza (%)	99,63 (99,50)
$I / \bar{I}$ (I)	12,26 (1,77)
B-factor de Wilson	24,47
R-merge	0,03 (0,46)
R-meas	0,047 (0,65)
CC1 / 2	0,999 (0,694)
Reflexões utilizadas no refinamento	79940 (7972)
Reflexões usados para R-free	2006 (195)
R-work	0,2089 (0,3555)
R-free	0,2198 (0,3558)
Número de átomos diferentes de hidrogênio	6605
macromoléculas	6145
Resíduos de aminoácidos	811
RMS das ligações	0,026
RMS dos ângulos	1,66
Ramachandran favoráveis (%)	97,40
Ramachandran permitido (%)	2,11
Ramachandran não permitidos (%)	0,50
Rotâmeros não permitidos (%)	3,0
solvente	31,24

(\*)Os valores em parênteses são referentes a camada de mais alta resolução

<sup>a</sup>  $R_{\text{sym}} = \frac{\sum_h |I_h - \langle I \rangle|}{\sum_h I_h}$ , onde  $I_h$  é a intensidade da reflexão ;  $\langle I \rangle$  é a intensidade média de todas as reflexões relacionados por simetria.

Fonte: Elaborado pelo autor

Outras 480 condições foram testadas no robô Oryx4 utilizando kits de cristalização comercializados pela Hampton Research e Jean Bioscience. Houve surgimento de cristais em 6 seis condições. Três destas condições foram selecionadas para otimização manual através do método de gota pendurada (Tabela 3).

Tabela 3. Cristais da proteína NeoB em diferentes condições de cristalização

<b>Condição</b>	<b>Cristal</b>
<b>1,6M Sulfato de Amônia; 500mM Cloreto de Lítio</b>	
<b>25% PEG 3350; 100mM BisTris pH 5,5; 200mM Cloreto de magnésio</b>	



Fonte: Elaborado pelo autor

Com estes novos cristais, tentativas de co-cristalização e soaking foram realizadas com ligantes, no entanto, ao serem levados para linha do PETRA III, DESY, *Deutsches Elektronen-Synchrotron* – Hamburgo – Alemanha, SLS (Villigan, Suíça) não houve difração.

Utilizando os dados de difração da Tabela 2, a estrutura da NeoB foi resolvida por substituição molecular, refinada a 1,8 Å. O Coeficiente de Matthews aponta para duas moléculas na unidade assimétrica com 47% de solvente. O dímero apresenta A proteína NeoB é composta por 416 resíduos de aminoácidos e demonstrava uma identidade de sequência menor que 30% a qualquer estrutura de aminotransferase depositada no Protein Data Bank (PDB). Durante o curso desta investigação, fomos capazes de resolver a estrutura molecular da NeoB por substituição molecular utilizando a estrutura da Sis5, uma enzima da biossíntese de sisomicina, um aminoglicosídeo similar as gentamicinas e que apresenta uma similaridade sequencial de aproximadamente 50% com a NeoB.

Na Figura 30 é possível observar uma representação da estrutura tridimensional do dímero da NeoB. Cada subunidade do dímero contém 12 fitas- $\beta$  e 11  $\alpha$ -hélices. Sete fitas- $\beta$  formam uma folha- $\beta$  mista que é bem característico nos membros da superfamília da aspartato aminotransferase (Figura 30). A folha beta antiparalela formada por 3 fitas que se iniciam no resíduo Tyr26 até o resíduo Ile41 que está destacado em pink na Figura 31, é uma inserção não observada no enovelamento clássico “aspartato

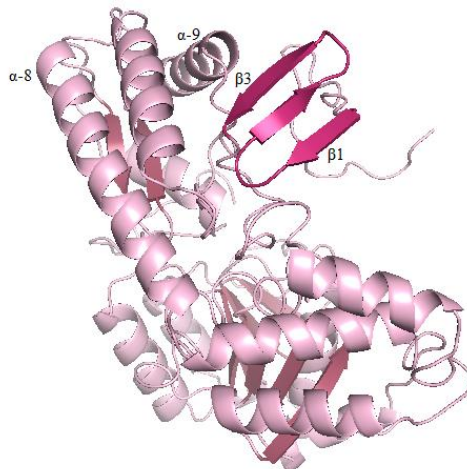
aminotransferase”. As duas fitas restante formam um motivo grampo  $\beta$  ( $\beta$ -hairpin) que conecta as hélices 8 e 9 (Figura 31).

Figura 30. Estrutura da NeoB. Estrutura representada em Cartoon apresentada em dímero separadamente coloridos de rosa claro e azul céu. As fitas- $\beta$  de cada monômero estão destacadas em coloração cereja e azul marinho.



Fonte: Elaborado pelo autor

Figura 31. Estrutura tridimensional do Monômero da NeoB. Estrutura representada em Cartoon de um dos monômeros da proteína.  $\alpha$ -hélice coloridas em rosa claro e  $\beta$ -fitas coloridas em cereja. As três fitas- $\beta$  antiparalelas da inserção da folha- $\beta$  estão destacadas em magenta. A imagem foi gerada pelo programa PyMOL (DeLano, 2002).

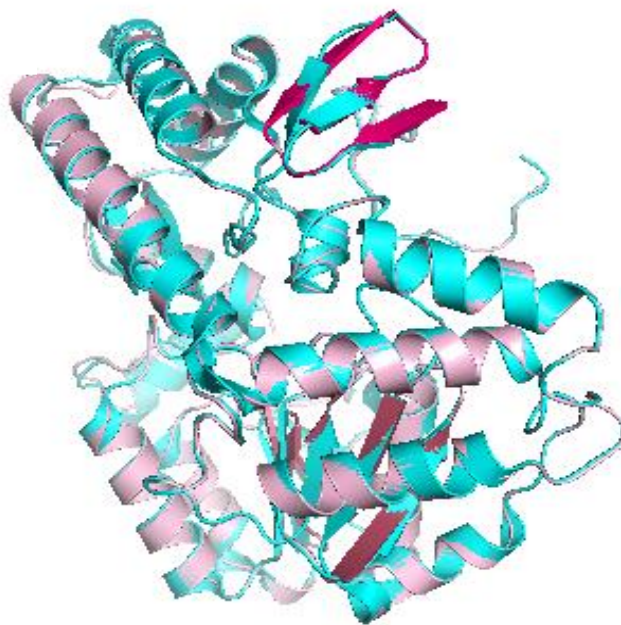


Fonte: Elaborado pelo autor

Essa visão abrangente da NeoB está em total consonância com a estrutura publicada por um grupo estadunidense (DOW, 2018) no ano de 2018, como é possível observar a similaridade das estruturas na sobreposição dos monômeros (Figura 32), antecipando assim a possível novidade a ser revelada pela estrutura da aminotransferase.

Figura 32. Sobreposição dos monômeros das estruturas da NeoB. Em rosa está a estrutura do monômero da NeoB obtida neste trabalho e em azul a estrutura da mesma proteína demonstrada por um grupo estadunidense (Código do PDB: 6CBK cadeia A). O RMSD entre as cadeias ficou em 1.81Å.

RMSD 1.81Å



Fonte: Elaborado pelo autor

## **5. Resultados obtidos para as enzimas Amidinohidrolases**

### **5.1 Produção e purificação das proteínas recombinantes Medi\_4928 e AMH\_828**

Após a superexpressão das proteínas recombinantes, o primeiro passo de purificação foi uma cromatografia de afinidade em coluna de níquel. As figuras 33 e 34 apresenta o cromatograma referente à purificação da proteína AMH\_828 Medi\_4948 por afinidade. Pode-se observar um pico de eluição da proteína recombinante com aproximadamente 12% do tampão de eluição (linha verde no eixo secundário), o que representa 60mM de imidazol.

Figura 33. Purificação da proteína AMH\_828 por cromatografia de afinidade em coluna de níquel. (A) – Cromatograma obtido na cromatografia de afinidade no qual a linha verde representa a aplicação do tampão de eluição num gradiente de imidazol. (B) – SDS-PAGE 10% corado com *Coomassie Brilliant Blue R250*, no qual pode-se visualizar: M – marcador de massa molecular em kDa (Precision Plus Protein Dual Color Standarts -BioRad), 1- *Pellet* 2- fração lisada 3- fração referentes ao *flowthrough*, 4- fração referente a lavagem com 5% do também B, 5 a 12– frações correspondentes ao pico de eluição da proteína AMH 828.

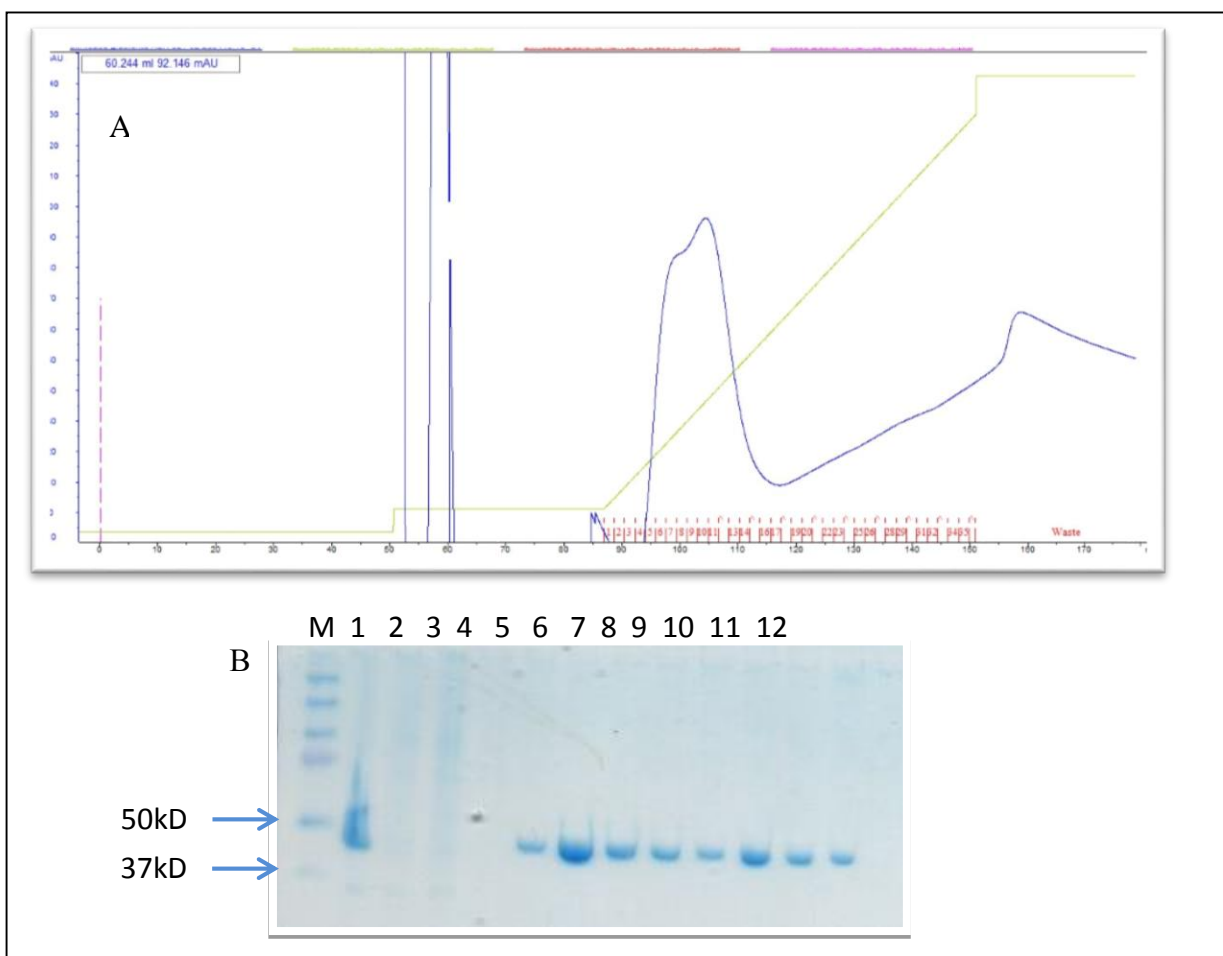
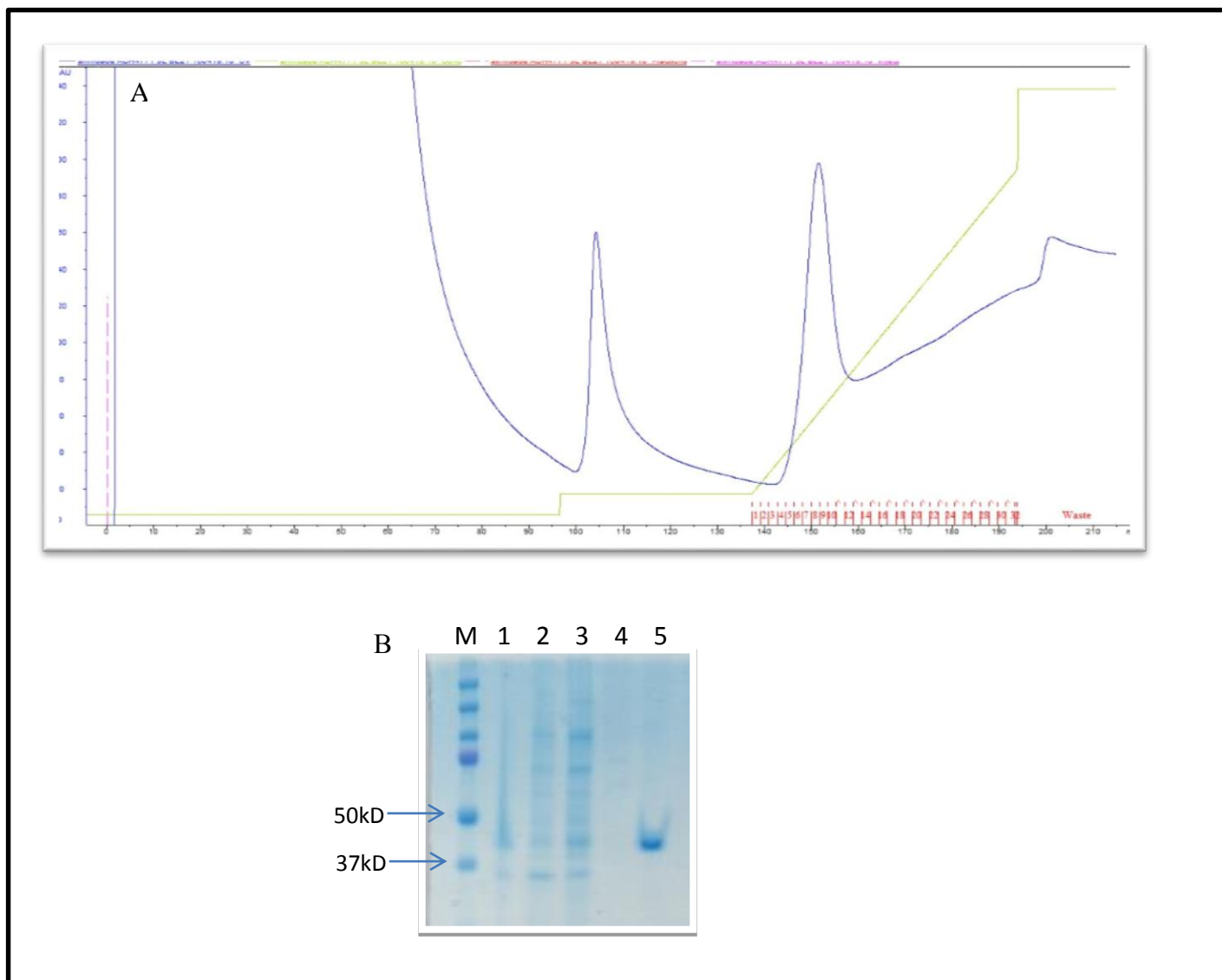


Figura 34. Purificação da proteína Medi\_4948 por cromatografia de afinidade em coluna de níquel. (A) – Cromatograma obtido na cromatografia de afinidade no qual a linha verde representa a aplicação do tampão de eluição num gradiente de imidazol. (B) – SDS-PAGE 10% corado com *Coomassie Brilliant Blue R250*, no qual pode-se visualizar: M – marcador de massa molecular em kDa (Precision Plus Protein Dual Color Standarts - BioRad), 1- *Pellet* 2- fração lisada 3- fração referentes ao *flowthrough*, 4- fração referente a lavagem com 5% do também B, 5- fração correspondentes ao pico de eluição da proteína Medi 4948.

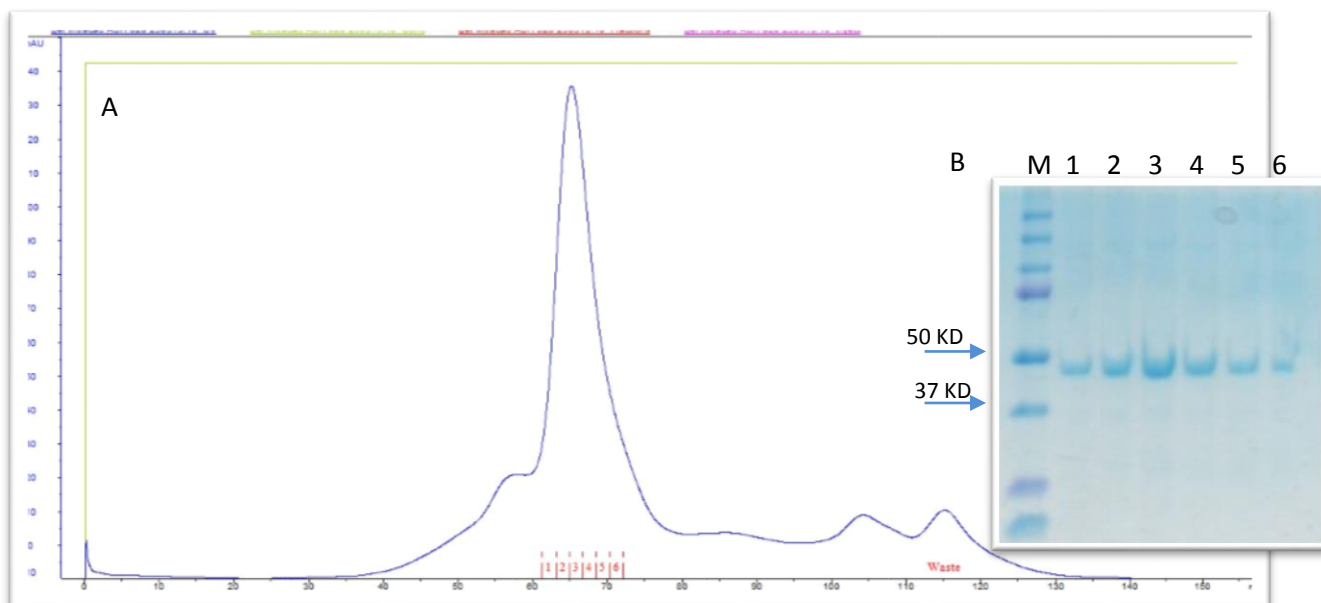


Fonte: Elaborado pelo autor

Após a cromatografia de afinidade, as amostras eluídas foram analisadas por SDS-PAGE (figura 33B e 34B) e verificou-se a necessidade de um segundo passo de purificação e/ou polimento da amostra.

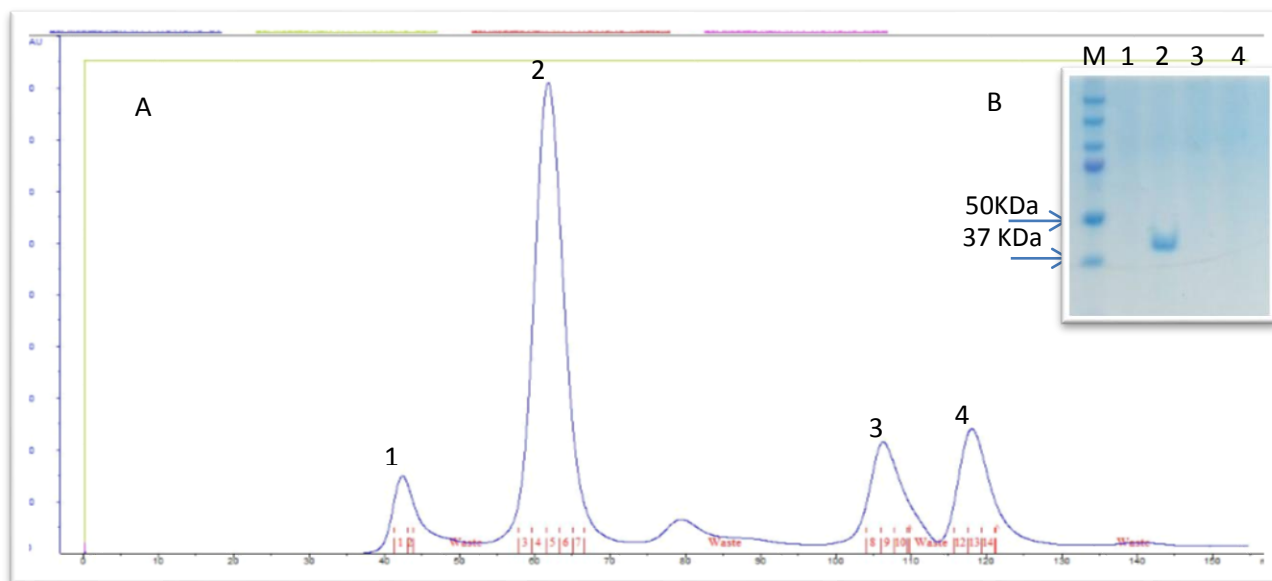
A segunda etapa de purificação consistiu em uma cromatografia de exclusão molecular, que separa as proteínas de acordo com suas massas. A amostra proveniente da cromatografia de afinidade foi aplicada em uma coluna de exclusão molecular, da qual a proteína foi eluída como um pico principal em aproximadamente 62 mL (Figura 35A e Figura 36A).

Figura 35. Purificação da proteína AMH\_828 por cromatografia de exclusão molecular. (A) – Cromatograma obtido na cromatografia de exclusão molecular. (B) – SDS-PAGE 10% corado com *Coomassie Brilliant Blue R250* no qual pode-se visualizar: M – marcador de massa molecular em kDa (Precision Plus Protein Dual Color Standarts -BioRad), 1 a 6 – frações eluídas no volume de corrida de aproximadamente 60mL.



Fonte: Elaborado pelo autor

Figura 36. Purificação da proteína Medi\_4948 por cromatografia de exclusão molecular. (A) – Cromatograma obtido na cromatografia de exclusão molecular. (B) – SDS-PAGE 10% corado com *Coomassie Brilliant Blue R250* no qual pode-se visualizar: M – marcador de massa molecular em kDa (Precision Plus Protein Dual Color Standarts -BioRad), 1 a 4 – frações eluídas durante a corrida de acordo com os picos correspondentes.



Fonte: Elaborado pelo autor

Pela análise de um gel SDS-PAGE 10% (Figura 35B e 36B), foi possível observar a pureza das amostras o que permitiu dar seguimento a etapa de cristalização.

A etapa de expressão e purificação para a enzima AMH\_821 não foi satisfatória quando utilizada células BL21(DE3). Novas alternativas para obtenção foram empregadas na otimização da expressão para que possamos obter quantidade suficiente da proteína no passo de purificação.

## 5.2 Cristalização da AMH\_828 e Medi\_4948

Com finalidade de se obter cristal proteico, foram testadas aproximadamente 475 condições de cristalização utilizando o robô Oryx4 nas dependências do Laboratório Biologia Estrutural Aplicada (USP-São Paulo) utilizando kits de cristalização padronizados e comercializados pela Hampton Research e Jena Bioscience. A partir dos

experimentos iniciais realizados, houve o surgimento de cristais após 2 dias em 3 condições, destas, 1 condição foi selecionada para otimização manual através do método de gota pendurada para a proteína Medi\_4849 e em 1 condição para a proteína AMH\_828 (Tabela 4).

Tabela 4. Cristais das proteínas Medi\_4948 e AMH\_828 em diferentes condições de cristalização

Condição	Cristal Medi 4948	Cristal AMH 828
0.1M HEPES pH 7.5 20% PEG 10.000		
15% MPD 100mM MÉS Sodium Salt pH 6.5 200mM Magnesium Acetate		
200mM Magnesium formate		
30%PEG 400 100mM Hapes pH 7.5 200mM Magnesium Chloride		

Fonte: Elaborado pelo autor

Com intenção de obter cristais maiores, a condição contendo 30% PEG 400, 100mM Hapes pH7,5 e 200mM Cloreto de Magnésio foi otimizada variando a concentração de PEG 400, de precipitante e volume/proporção das gotas, totalizando 24 novas condições testadas.

O cristal inicial da enzima AMH\_828 foi coletado e difratado na linha PETRA III, DESY, *Deutsches Elektronen-Synchrotron* – Hamburgo – Alemanha, apresentando resolução a de 1,8 Å. Os dados do processamento inicial revelaram um grupo espacial C222. O Coeficiente de Matthews sugeriu a possibilidade de um trímero na unidade assimétrica. As estatísticas da coleta de dados, refinamento e parâmetros da qualidade do modelo estão resumidas na Tabela 5. Após a substituição molecular utilizando como modelo a proteína Medi\_4948, os valores finais de Rfree e Rfactor se mantiveram altos tornando assim a inviabilização da estrutura após ciclos de refinamento. Isso pode ter ocorrido devido a um alto grau de twinning no cristal coletado.

Tabela 5. Estatísticas de coleta de dados e processamento do cristal nativo AMH\_828. Os dados entre parênteses indicam os valores da menor e maior resolução, respectivamente.

Parâmetros	AMH_828
Faixa de resolução (Å)	44,78 - 1,861 (1,928 - 1,861)
Grupo espacial	C 2 2 2
Célula unitária a, b, c (Å) $\alpha$ , $\beta$ , $\alpha$ (°)	119,024 136,001 143,938 90 90 90
Reflexões totais	195185 (19310)
Reflexões únicas	97601 (9654)
Multiplicidade	2,0 (2,0)
Completeza (%)	99,67 (98,77)
I / $\delta$ (I)	10,66 (2,51)
B-factor de Wilson	26,40
R-merge	0,02566 (0,2478)
R-meas	0,03629 (0,3505)

R-pim	0,02566 (0,2478)
CC1 / 2	0,998 (0,766)
Reflexões utilizadas no refinamento	1 (0,931)
Reflexões usados para R-free	97344 (9592)
R-work	0,4598 (0,3535)
R-free	0,4904 (0,5085)
Número de átomos diferentes de hidrogênio	4898
Macromoléculas	4898
Resíduos de aminoácidos	642
RMS das ligações	0.011
RMS dos ângulos	1,36
Ramachandran favoráveis (%)	97,34
Ramachandran permitido (%)	2,35
Ramachandran não permitidos (%)	0,31
Rotâmeros não permitidos (%)	1,18

$\$R_{merge} = \frac{\sum hkl \sum i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum hkl \sum i I_i(hkl)}$ , onde  $I_i(hkl)$  é a intensidade observada para  $I(hkl)$  e  $\langle I(hkl) \rangle$  é a média das intensidades. Entre parênteses os valores referentes à camada mais externa de resolução.

Fonte: Elaborado pelo autor

Para a coleta de dados, um cristal nativo da proteína\_Medi 4948 foi submetido a um experimento de difração de raios X. Este cristal apresentou um limite de difração até 1,60 Å de resolução. Os dados coletados foram, então, processados no programa XDS (KABSCH, 2010) no qual verificou-se que o cristal pertence ao sistema cristalino monoclinico (Tabela 6). A tentativa de resolução da estrutura foi determinada por substituição molecular usando como modelo uma proteína aminohidrolase de *Streptomyces clavuligerus* (Protein Data Bank código 1GQ6\_A). Por haver 6 moléculas na unidade assimétrica, simetria não-cristalográfica foi empregada fazendo com que a aumentasse consideravelmente a qualidade do mapa de densidade

eletrônica.

Tabela 6: Estatísticas de coleta de dados e processamento do cristal nativo Medi\_4948. Os dados entre parênteses indicam os valores da menor e maior resolução, respectivamente.

<b>Coleta de Dados</b>	<b>Cristal Medi_4948</b>
Fonte de raios X	CNPEM- LNLS- MX2-Campinas/Brasil
Comprimento de onda (Å)	1,458
Faixa de resolução (Å)	36,54 – 1,597 (1,65 – 1,59)
Grupo espacial	P 1 21 1
Célula unitária a, b, c (Å) $\alpha$ , $\beta$ , $\alpha$ (°)	82,07 143,71 83,23 90 113 90
Reflexões totais	444177 (42988)
Reflexões únicas	229534 (22460)
Multiplicidade	1,9 (1,9)
Completeza (%)	98 (96)
$I / \bar{\sigma} (I)$	14,6 (2,2)
B-factor de Wilson	18,68
R-merge	0,032 (0,389)
R-meas	0,045 (0,550)
CC1 / 2	0,999 (0,777)
Reflexões utilizadas no refinamento	229482 (22446)
Reflexões usados para R-free	11306 (1096)
R-work	0,16 (0,239)
R-free	0,19 (0,269)

Número de átomos diferentes de hidrogênio	16175
macromoléculas	14641
Resíduos de aminoácidos	1922
RMS das ligações	0,007
RMS dos ângulos	0,92
Ramachandran favoráveis (%)	98
Ramachandran permitido (%)	2
Ramachandran não permitidos (%)	0,37
Rotâmeros não permitidos (%)	2,8
solvente	30,68

<sup>a</sup>  $R_{\text{sym}} = \frac{\sum_h |I_h - \langle I \rangle|}{\sum_h I_h}$ , onde  $I_h$  é a intensidade da reflexão;  $\langle I \rangle$  é a intensidade média de todas as reflexões relacionados por simetria.

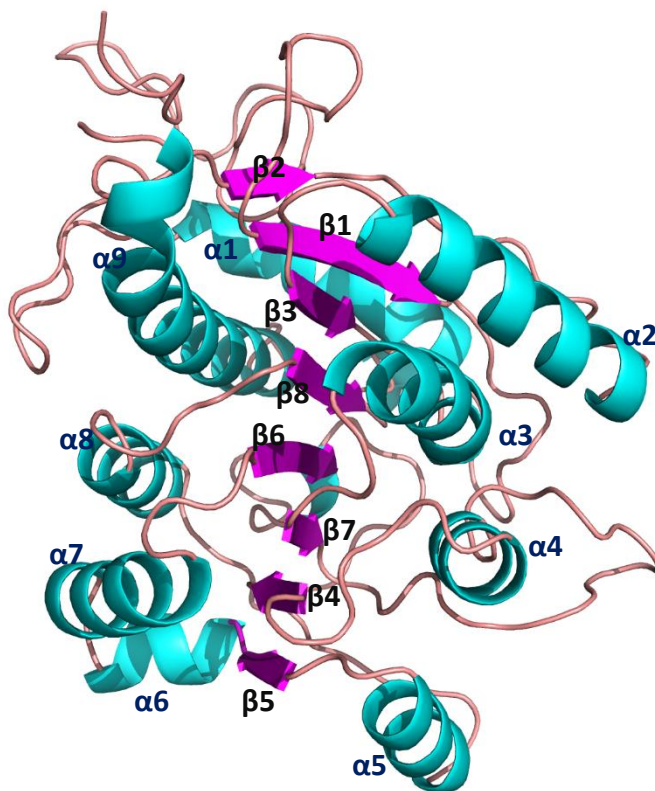
Fonte: Elaborado pelo autor

## 5.2 Estrutura do Monômero

A estrutura do monômero da Medi\_4948 é mostrada na Figura 37. Cada monômero é dobrado em um único  $\alpha/\beta$  domínio com dimensões aproximadas de 40x40x50 Å e é composto por uma folha beta paralela de oito fitas flanqueada por  $\alpha$ -hélices em ambos os lados. O mesmo domínio  $\alpha/\beta$  é compartilhado por outras ureohidrolases como arginase de fígado de rato (KANYO; SCOLNICK; ASH; CHRISTIANSON, 1996), arginase II de fígado humano (CAMA; COLLELUORI; EMIG; SHIN *et al.*, 2003), arginase de *B. cladovelox* (BEWLEY; JEFFREY; PATCHETT; KANYO *et al.*, 1999), amidinohidrolase de *S. clavuligerus* (ELKINS; CLIFTON; HERNANDEZ; DOAN *et al.*, 2002), agmatinase de *Deinococcus radiodurans* (AHN; KIM; LEE; HA *et al.*, 2004) e guanidinobutirase e guanidinopropionase de *P. aeruginosa* (LEE; KIM; KIM; LEE *et al.*, 2011). As fitas centrais da folha- $\beta$  estão organizadas na ordem  $\beta 2$ - $\beta 1$ - $\beta 3$ - $\beta 8$ - $\beta 6$ - $\beta 7$ - $\beta 4$ - $\beta 5$ . Uma ou duas  $\alpha$ -hélices conectam uma fita a próxima de tal forma que um conjunto

de quatro hélices ( $\alpha 2$ ,  $\alpha 3$ ,  $\alpha 4$  e  $\alpha 5$ ) cobrem uma face da folha- $\beta$  e o conjunto de outras cinco hélices ( $\alpha 1$ ,  $\alpha 9$ ,  $\alpha 8$ ,  $\alpha 7$  e  $\alpha 6$ ) cobre a outra face (Figura 37).

Figura 37. Estrutura tridimensional do monômero da Medi\_4948. Estrutura representada em Cartoon de um dos monômeros da proteína.  $\alpha$ -hélices coloridas em ciano e  $\beta$ -fitas coloridas em magenta. A imagem foi gerada pelo programa PyMOL (DeLano, 2002).

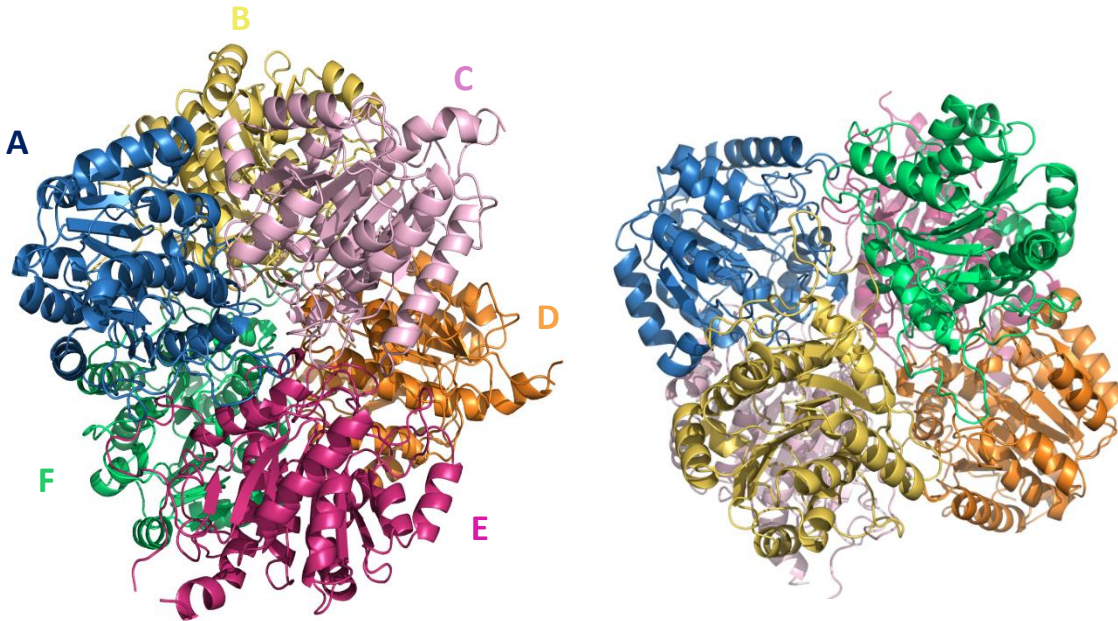


Fonte: Elaborado pelo autor

### 5.3 Oligomerização da Medi\_4948

Arginases de eucariotos são triméricas, entretanto as arginases de bactérias são hexaméricas. Amidinohidrolase de *S. clavuligerus* existe como um hexâmero em solução e sob condição de cristalização, compostas por duas espécies triméricas. Foi possível observar que a Medi\_4948 também é um hexâmero constituído por dois trímeros (figura 38).

Figura 38. Estrutura tridimensional do Hexâmero da Medi\_4948. As seis subunidades (A – F) da Medi\_4948 estão coloridas em azul, amarelo, rosa claro, laranja, magenta e verde claro, respectivamente. Três subunidades da frente estão em cores fortes enquanto as outras três subunidades posicionadas atrás estão em cores mais claras.

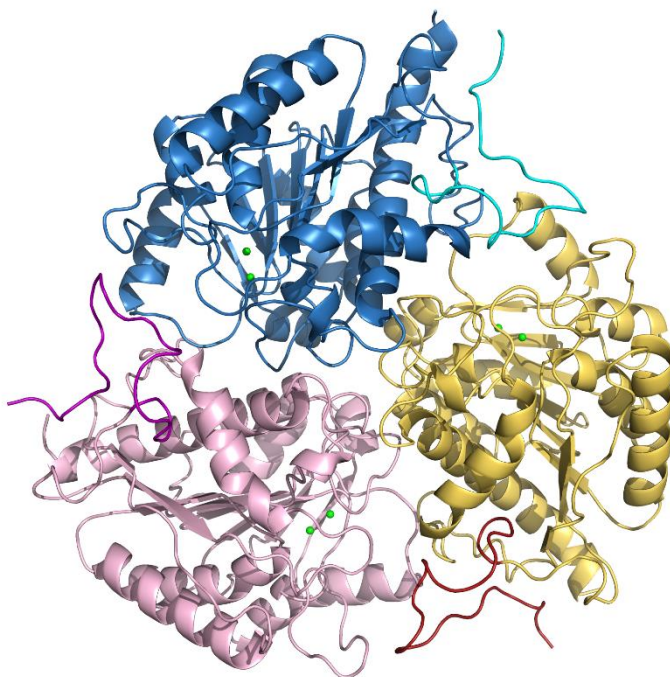


Fonte: Elaborado pelo autor

A estrutura cristalina de arginase de fígado de rato indicou que 19 aminoácidos do C-terminal formaram um motivo de oligomerização em forma de S, que mediou aproximadamente 54% dos contatos intermonômeros entre os monômeros nas unidades triméricas (Kanyo, 1996). Posteriormente, esta cauda no C-terminal foi proposta para estabilizar a natureza oligomérica da arginase através de uma rede ponte salinas que se iniciam a partir da Arg-308. A estrutura cristalina da Medi\_4948, diferentemente da arginase, em vez de uma extensão no C-terminal, possui um loop flexível no seu N-terminal, incluindo os resíduos  $8 \pm 37$  e presumidamente os resíduos  $1 \pm 7$  (que não estão visíveis na estrutura do cristal), estando de acordo com as estruturas de PAH e guanidinobutirase (LEE; KIM; KIM; LEE *et al.*, 2011), as quais apresentam o mesmo

padrão no N-terminal. Assim, dentro de cada unidade trimérica da Medi\_4948, cada monômero (A, B e C) está conectado por ligação do N-terminal de monômero anterior esticado ao monómero adjacente, bloqueando assim a entrada do sítio ativo. Esta alça está envolvido em uma extensa rede de ligações de hidrogênio envolvendo moléculas de água. O modo de associação inter-subunidades responsável para a formação de um dímero de unidades triméricas na Medi 4948 é semelhantes aqueles em PAH *S. clavuligerus* e agmatinase de *D. radiodurans*. Um monomero de uma unidade trimérica da Medi\_4948 interage com dois monômeros vizinhos na outra unidade trimérica (Figura 39).

Figura 39. Interação monômero-monômero em trómeros do hexâmero. A alça estendida de cada monômero interagindo com a subunidade adjacente (destacados nas cores ciano, vermelho e magenta) em direção da entrada do sítio ativo no monômero circunvizinho. Esferas verdes indicam dois íons metálicos no sítio ativo.



Fonte: Elaborado pelo autor

#### 5.4 Sítio de ligação ao metal

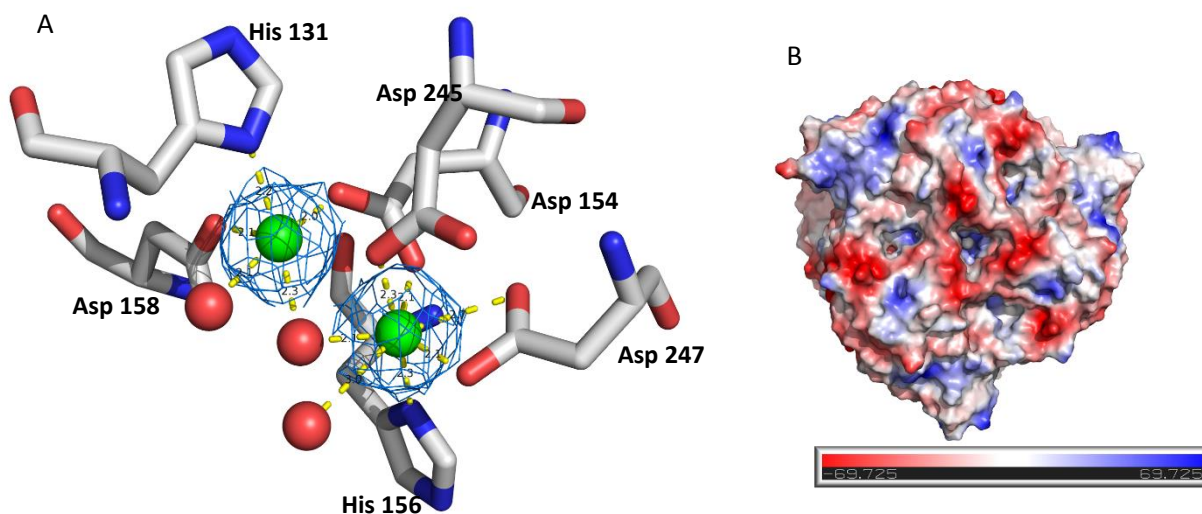
O sítio ativo da Medi\_4948 é formado pela combinação de alças que ligam os elementos estruturais secundários. Está situado em uma fenda profunda e contém dois íons metálicos dispostos assimetricamente separados e ligados por uma molécula de água. Nenhum outro local de ligação a metal é evidente na estrutura. De acordo com a literatura, os íons que mais comumente estimulam o efeito da atividade enzimática são os de  $Mn^{2+}$ ,  $Ni^{2+}$ ,  $Co^{2+}$ ,  $Mg^{2+}$  (Hong, 2016). Para este trabalho, não identificamos o exato íon metálico envolvido na estrutura cristalina, adotamos a nomenclatura de geral (íon metálico), sem especificar qual o íon. Futuras análises serão realizadas a fim de identificar o íon correspondente.

Um íon metálico, designado MA, está posicionado no fundo da cavidade do sítio ativo, enquanto o íon designado MB está mais próximo da superfície. Cada íon metálico é coordenado por um arranjo octaédrico distorcido de ligantes. MA está interagindo com as cadeias laterais de His-131, Asp-158 e Asp-245, a duas moléculas de água, a molécula de água de ligação (como íon hidróxido) (HOH 1600) e uma segunda água (HOH 508). A molécula de água de ligação (HOH 1600) também se liga ao outro íon metálico (MB), assim como um oxigênio de cadeia lateral do Asp154. O resíduo Asp-247 se liga a MB através de ambos os oxigênios carboxilados, e interage com as cadeias laterais de His-156 e Asp-245, completando assim a coordenação octaédrica (Figura 40). A coordenação dos resíduos por metal é conservada entre membros da superfamília ureohidrolases. A ponte entre o íon metálico e a água de ligação, que atua como nucleófilo, é importante para a catálise de ureohidrolases. Uma vez que, embora seja difícil distinguir entre uma molécula de água e um íon hidróxido em estruturas de cristais de raio X, supõe-se que o mecanismo de ação esteja conexo, a hidrólise ocorra por uma molécula de água da ponte metálica ou íon hidróxido que inicia um ataque nucleofílico ao átomo de carbono da guanidina do substrato.

Foi observada uma ligação *cis*-peptídica entre Gly-128 e Gly-129 (Figura 41). Esses resíduos de glicina fazem parte do motivo conservado xGGDH proteínas da superfamília ureohidrolase. Este motivo é importante, pois posiciona His-131 para coordenação com o íon metálico. Embora as ligações *cis*-peptídicas raramente envolvam outros aminoácidos além dos resíduos de prolina (STEWART; SARKAR; WAMPLER,

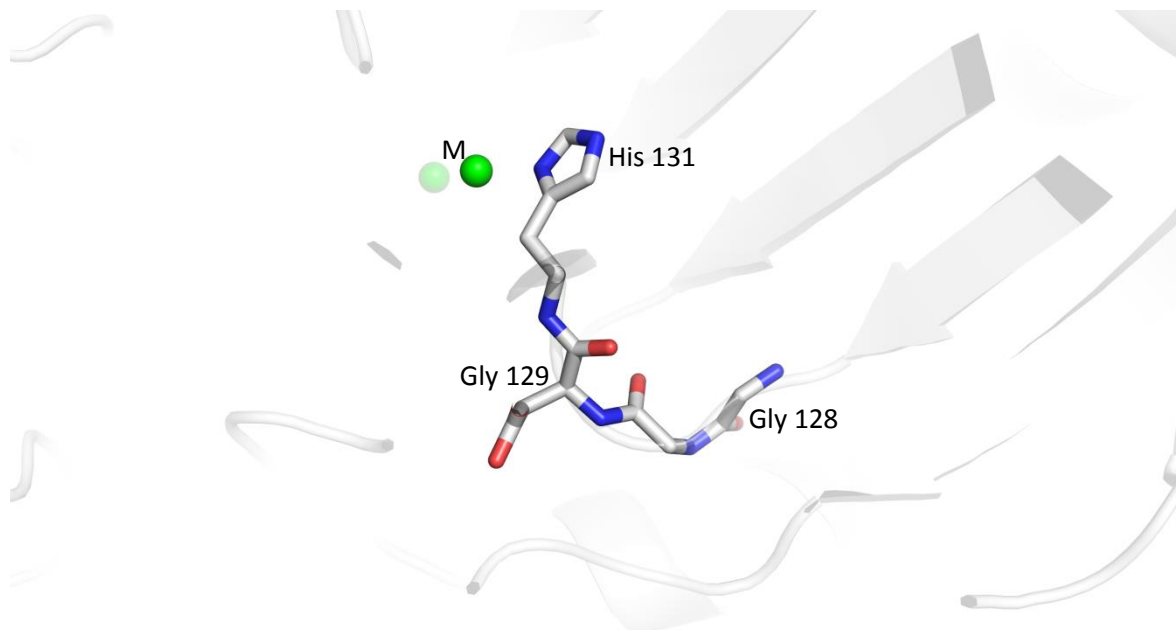
1990), as vezes elas ocorrem para permitir a definição da forma da cavidade do sítio ativo (BOND; CLEMENTS; ASHBY; COLLYER *et al.*, 1997; VARGHESE; GARRETT; COLMAN; CHEN *et al.*, 1994).

Figura 40. Ligações no sítio ativo da Medi\_4948. A) Carbono, nitrogênio, oxigênio e íons metálicos estão coloridos em branco, azul, vermelho e verde, respectivamente. Moléculas de água representadas por esferas em vermelho. Possíveis interações entre átomos estão indicadas em linha tracejada amarela. B) Superfície de potencial eletrostático do hexâmero com seta indicando a entrada do sítio ativo. Cargas positivas do potencial eletrostático estão coloridas em azul enquanto as negativas são mostradas em vermelho. O mapa da superfície foi obtido utilizando a ferramenta APBS no PYMOL. As imagens foram geradas no PYMOL (Delano, 2007).



Fonte: Elaborado pelo autor

Figura 41. Resíduos com *cis*-ligação. Resíduos de glicina posicionando resíduo de histidina no sítio ativo para interação com íon metálico.

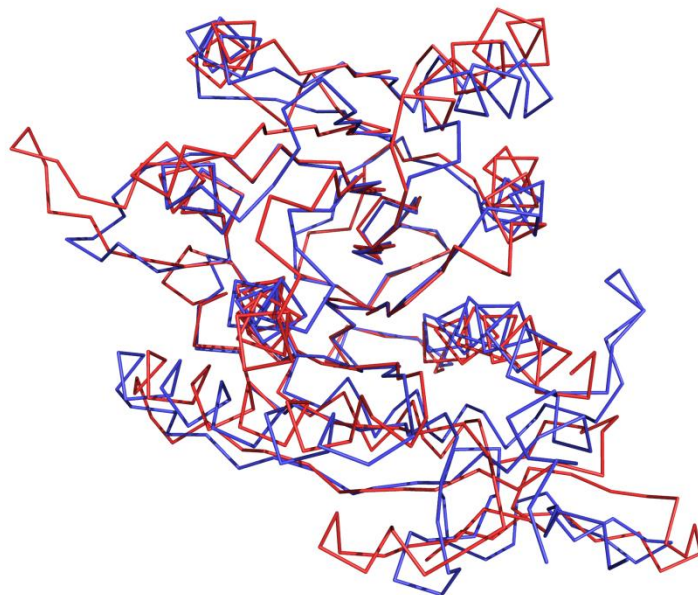


Fonte: Elaborado pelo autor

A estrutura do monômero da Medi\_4948 é similar a outras ureohidrolases. De acordo com uma busca no DALI (HOLM; ROSENSTROM, 2010) uma hipotética agmatinase de *Burkholderia thailandensis* (4DZ4 código do PDB) apresenta alta similaridade estrutural com a Medi4948 (cadeia A), com um r.m.s.d. de 1.9Å, Z-score de 45.3 e 40% de identidade sequencial. A próxima proteína pautada, uma formiminoglutamase de *Vibrio cholerae* (1XFK código do PDB) apresenta um desvio r.m.s.d. de 2.3Å, Z-score de 35.0 e um sequência de identidade de 25%. Interessantemente, essa última estrutura (da formiminoglutamase de *V. cholerae*, PDB 1XFK) não aparenta ser uma metaloenzima, pois não contém íons de metais ligado no sítio ativo. Uma comparação estrutural entre as enzimas amidinohidrolase Medi\_4948 e formiminoglutamase (PDB código 1XFK) é mostrada na figura 42, e podemos observar que elas diferem consideravelmente quanto as regiões de dois loops, os quais podem estar envolvidos com o correto posicionamento do substrato e interação intermonômero.

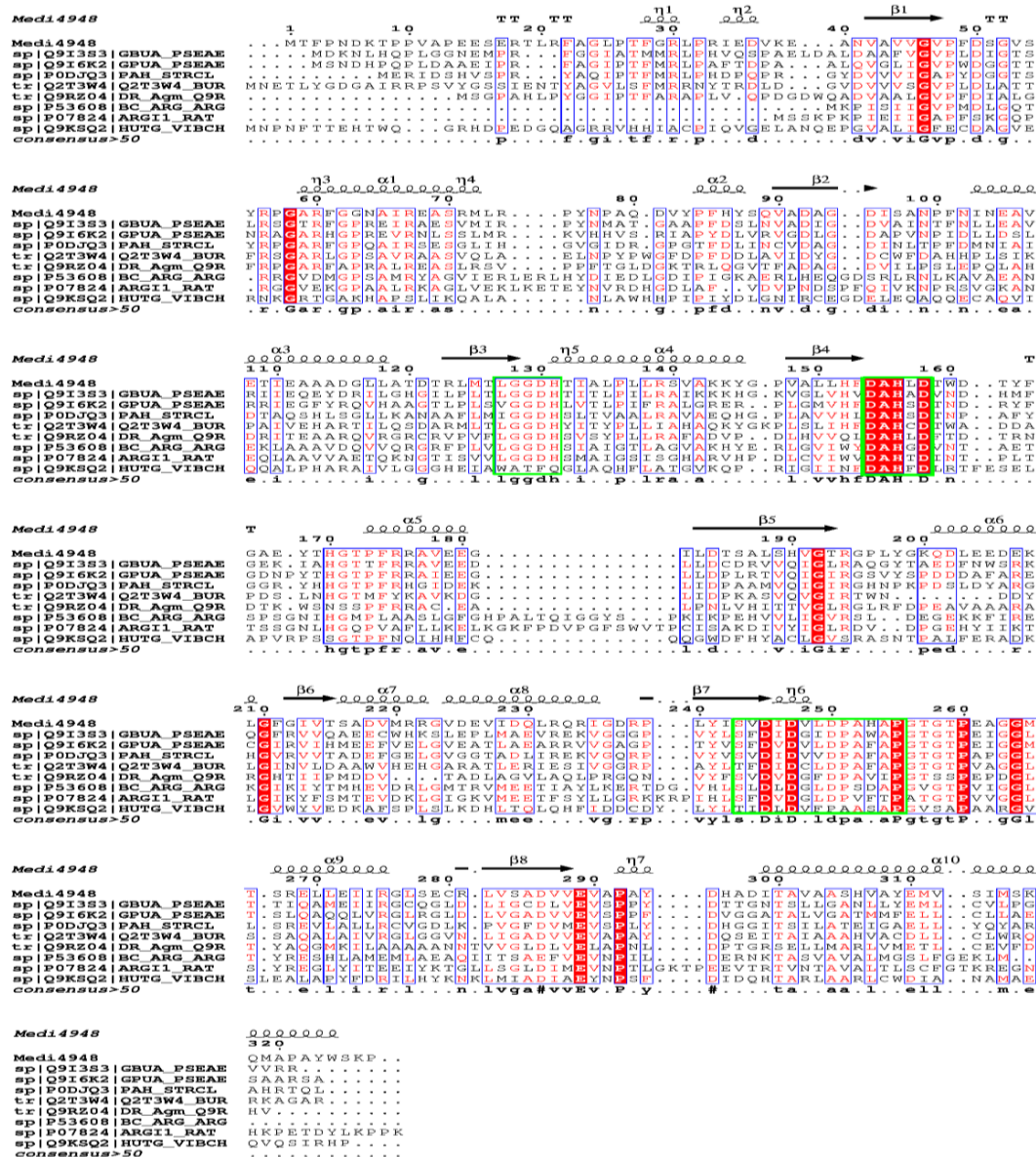
Um aspecto interessante é que os membros da família ureohidrolases contém, tipicamente, dois resíduos de histidinas e quatro resíduos de aspartato envolvidos na ligação com os íons manganês no sítio ativo. Além disso, de acordo com um alinhamento de sequência primária (Figura 43) foi observado que a sequência da proteína formiminoglutamase de *V. cholerae* não possui os resíduos conservados de acordo com família das ureohidrolases. Embora essa tenha sido a segunda proteína elencada no DALI (HOLM; ROSENSTROM, 2010), notamos diferença significativa quanto aos resíduos apresentados nas três sequências (xGGDH, DAHxD e SxDxDxxDPxxxP) muito conservadas entre as ureohidrolases, na sequência de formiminoglutamase essa a primeira sequência é totalmente diferente das outras ureohidrolases, variando de xGGDH para xATFQ, sequência esta que possui uma das histidinas que participa da coordenação do íon metálico no sítio ativo.

Figura 42. Sobreposição do monômero de Medi\_4948 (azul) e Formiminoglutamase de *V. cholerae* (vermelho).



Fonte: Elaborado pelo autor

Figura 43. Alinhamento das seqüências das enzimas da superfamília ureohidrolases. Elementos de estrutura secundária da Medi\_4948 são mostrados em preto na parte superior. Resíduos conservados são destacados em grifado de vermelho enquanto que os resíduos semiconservados são mostrados na cor de letra vermelha. A seqüência de nove ureohidrolases (*S. medicidicus*: Medi\_4948; *P. aeruginosa*: GbuA - Q9I3S3; *P. aeruginosa*: GpuA- Q9I6K2; *S. clavuligerus*: PAH - P0DJQ3; *B. thailandensis*: Bur-Q2T3W4 [4DZ4 código do PDB]; *D. radiodurans*: DR\_agm- Q9RZ04; *B. caldovelox*: BC\_ARG - P53608; *Rattus norvegicus*: ARG11\_RAT - P07824; *V. cholerae*: HUTG - Q9KSQ2 [1XFK código do PDB]) foram alinhadas usando o programa MultiAling (Corpet, 1988). Três seqüências conservadas (xGGDH, DAHxH e SxDxDxxDPxxxP) na maioria das reohidrolases estão indicadas por caixas verdes. Esta figura foi gerada com o programa ESPript (GOUET; COURCELLE; STUART; METOZ, 1999).



Fonte: Elaborado pelo autor

## 6. Conclusão

Neste trabalho foram apresentadas as etapas gerais de clonagem, as quais se expuseram bem-sucedidas para a enzima ParQ. No entanto, para alguns genes, como aqueles que codificam para as enzimas ParB, LivB e LivQ tivemos dificuldade. A expressão heteróloga em *E. coli* foi bem-sucedidas para NeoB, porém para as enzimas ParQ e NeoQ somente foram obtidas na sua forma insolúvel. Também obtivemos sucesso na expressão, purificação e cristalização da Medi\_4948 e AMH\_828.

Tendo em vista que o processo cristalográfico é totalmente empírico, das proteínas propostas conseguimos obter cristais de NeoB, Medi\_4948 e AMH\_828.

Cristais de proteína da NeoB foram obtidos e difratados, possibilitando assim o processamento dos dados e resolução da estrutura. A enzima se apresentou como dímero e nós conseguimos observar o enovelamento característico de aspartato aminotransferase. A NeoB se mostrou em consonâncias com dados previamente publicados.

As etapas para obtenção da estrutura tridimensional avançaram para a proteína Medi\_4948. Realizamos a elucidação estrutural da enzima e foi possível fazer uma análise proteína da estrutura frente ao que se encontrava na literatura. Nós conseguimos observar uma região de *loop* próximo ao N-terminal da enzima que se sobrepõem ao monômero vizinho influenciando na obstrução do sítio ativo. Muito provavelmente, este loop esteja relacionado a especificidade ao substrato, visto que outras enzimas da mesma família possuem um N-terminal menor quando comparada a sequência primária. Frente a isso, podemos inferir que as diferenças de especificidade apresentadas entre as amidinohidrolases estejam relacionadas ao loop que acomoda o substrato.

Para a enzima AMH\_828, os dados obtidos da difração foram processados e posteriormente, ao serem refinados, continuaram a apresentar valores de parâmetros acima do desejável provavelmente causados por defeitos no processo de formação dos cristais.

## Referências

- AHN, H. J.; KIM, K. H.; LEE, J.; HA, J. Y. *et al.* Crystal structure of agmatinase reveals structural conservation and inhibition mechanism of the ureohydrolase superfamily. **J Biol Chem**, 279, n. 48, p. 50505-50513, Nov 26 2004.
- ALANIS, A. J. Resistance to antibiotics: are we in the post-antibiotic era? **Arch Med Res**, 36, n. 6, p. 697-705, Nov-Dec 2005.
- BALTZ, R. H.; BRIAN, P.; MIAO, V.; WRIGLEY, S. K. Combinatorial biosynthesis of lipopeptide antibiotics in *Streptomyces roseosporus*. **J Ind Microbiol Biotechnol**, 33, n. 2, p. 66-74, Feb 2006.
- BEWLEY, M. C.; JEFFREY, P. D.; PATCHETT, M. L.; KANYO, Z. F. *et al.* Crystal structures of *Bacillus caldovelox* arginase in complex with substrate and inhibitors reveal new insights into activation, inhibition and catalysis in the arginase superfamily. **Structure**, 7, n. 4, p. 435-448, Apr 15 1999.
- BOND, C. S.; CLEMENTS, P. R.; ASHBY, S. J.; COLLYER, C. A. *et al.* Structure of a human lysosomal sulfatase. **Structure**, 5, n. 2, p. 277-289, Feb 15 1997.
- CAMA, E.; COLLELUORI, D. M.; EMIG, F. A.; SHIN, H. *et al.* Human arginase II: crystal structure and physiological role in male and female sexual arousal. **Biochemistry**, 42, n. 28, p. 8445-8451, Jul 22 2003.
- CANE, D. E. Introduction: Polyketide and Nonribosomal Polypeptide Biosynthesis. From *Collie* to *Coli*. **Chem Rev**, 97, n. 7, p. 2463-2464, Nov 10 1997.
- CASTOE, T. A.; STEPHENS, T.; NOONAN, B. P.; CALESTANI, C. A novel group of type I polyketide synthases (PKS) in animals and the complex phylogenomics of PKSs. **Gene**, 392, n. 1-2, p. 47-58, May 1 2007.
- CAVENER, D. R. GMC oxidoreductases. A newly defined family of homologous proteins with diverse catalytic activities. **J Mol Biol**, 223, n. 3, p. 811-814, Feb 5 1992.
- CLAUSNITZER, D.; PIEPERSBERG, W.; WEHMEIER, U. F. The oxidoreductases LivQ and NeoQ are responsible for the different 6'-modifications in the aminoglycosides lividomycin and neomycin. **J Appl Microbiol**, 111, n. 3, p. 642-651, Sep 2011.
- COATES, A. R.; HU, Y. Novel approaches to developing new antibiotics for bacterial infections. **Br J Pharmacol**, 152, n. 8, p. 1147-1154, Dec 2007.
- CORTES, J.; VELASCO, J.; FOSTER, G.; BLACKABY, A. P. *et al.* Identification and cloning of a type III polyketide synthase required for diffusible pigment biosynthesis in *Saccharopolyspora erythraea*. **Mol Microbiol**, 44, n. 5, p. 1213-1224, Jun 2002.
- COX, R. J. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. **Org Biomol Chem**, 5, n. 13, p. 2010-2026, Jul 7 2007.
- DEWICK, P. M. The biosynthesis of C5-C25 terpenoid compounds. **Nat Prod Rep**, 14, n. 2, p. 111-144, Apr 1997.

DONADIO, S.; STAVER, M. J.; MCALPINE, J. B.; SWANSON, S. J. *et al.* Modular organization of genes required for complex polyketide biosynthesis. **Science**, 252, n. 5006, p. 675-679, May 3 1991.

DOWLING, D. P.; DI COSTANZO, L.; GENNADIOS, H. A.; CHRISTIANSON, D. W. Evolution of the arginase fold and functional diversity. **Cell Mol Life Sci**, 65, n. 13, p. 2039-2055, Jul 2008.

DURANTE-MANGONI, E.; GRAMMATIKOS, A.; UTILI, R.; FALAGAS, M. E. Do we still need the aminoglycosides? **Int J Antimicrob Agents**, 33, n. 3, p. 201-205, Mar 2009.

ELKINS, J. M.; CLIFTON, I. J.; HERNANDEZ, H.; DOAN, L. X. *et al.* Oligomeric structure of proclavaminic acid amidino hydrolase: evolution of a hydrolytic enzyme in clavulanic acid biosynthesis. **Biochem J**, 366, n. Pt 2, p. 423-434, Sep 1 2002.

EMSLEY, P.; LOHKAMP, B.; SCOTT, W. G.; COWTAN, K. Features and development of Coot. **Acta Crystallogr D Biol Crystallogr**, 66, n. Pt 4, p. 486-501, Apr 2010.

EVANS, P. R.; MURSHUDOV, G. N. How good are my data and what is the resolution? **Acta Crystallogr D Biol Crystallogr**, 69, n. Pt 7, p. 1204-1214, Jul 2013.

FERNANDES, P. Antibacterial discovery and development--the failure of success? **Nat Biotechnol**, 24, n. 12, p. 1497-1503, Dec 2006.

GOUET, P.; COURCELLE, E.; STUART, D. I.; METOZ, F. ESPript: analysis of multiple sequence alignments in PostScript. **Bioinformatics**, 15, n. 4, p. 305-308, Apr 1999.

GUIMARAES, B. G.; SANFELICI, L.; NEUENSCHWANDER, R. T.; RODRIGUES, F. *et al.* The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS. **J Synchrotron Radiat**, 16, n. Pt 1, p. 69-75, Jan 2009.

GUO, J.; HUANG, F.; HUANG, C.; DUAN, X. *et al.* Specificity and promiscuity at the branch point in gentamicin biosynthesis. **Chem Biol**, 21, n. 5, p. 608-618, May 22 2014.

HAPPI, C. T.; GBOTOSHO, G. O.; FOLARIN, O. A.; AKINBOYE, D. O. *et al.* Polymorphisms in Plasmodium falciparum dhfr and dhps genes and age related in vivo sulfadoxine-pyrimethamine resistance in malaria-infected patients from Nigeria. **Acta Trop**, 95, n. 3, p. 183-193, Sep 2005.

HARVEY, A. L.; EDRADA-EBEL, R.; QUINN, R. J. The re-emergence of natural products for drug discovery in the genomics era. **Nat Rev Drug Discov**, 14, n. 2, p. 111-129, Feb 2015.

HERT, J.; IRWIN, J. J.; LAGGNER, C.; KEISER, M. J. *et al.* Quantifying biogenic bias in screening libraries. **Nat Chem Biol**, 5, n. 7, p. 479-483, Jul 2009.

HERTWECK, C. The biosynthetic logic of polyketide diversity. **Angew Chem Int Ed Engl**, 48, n. 26, p. 4688-4716, 2009.

HOLM, L.; ROSENSTROM, P. Dali server: conservation mapping in 3D. **Nucleic Acids Res**, 38, n. Web Server issue, p. W545-549, Jul 2010.

HONG, H.; SAMBORSKY, M.; LINDNER, F.; LEADLAY, P. F. An Amidinohydrolase Provides the Missing Link in the Biosynthesis of Amino Marginolactone Antibiotics. **Angew Chem Int Ed Engl**, 55, n. 3, p. 1118-1123, Jan 18 2016.

HONG, H.; SAMBORSKY, M.; USACHOVA, K.; SCHNATZ, K. *et al.* Sulfation and amidinohydrolysis in the biosynthesis of giant linear polyenes. **Beilstein J Org Chem**, 13, p. 2408-2415, 2017.

HOPWOOD, D. A.; MALPARTIDA, F.; KIESER, H. M.; IKEDA, H. *et al.* Production of 'hybrid' antibiotics by genetic engineering. **Nature**, 314, n. 6012, p. 642-644, Apr 18-24 1985.

HUANG, C.; HUANG, F.; MOISON, E.; GUO, J. *et al.* Delineating the biosynthesis of gentamicin x2, the common precursor of the gentamicin C antibiotic complex. **Chem Biol**, 22, n. 2, p. 251-261, Feb 19 2015.

HUANG, F.; HAYDOCK, S. F.; MIRONENKO, T.; SPITELLER, D. *et al.* The neomycin biosynthetic gene cluster of *Streptomyces fradiae* NCIMB 8233: characterisation of an aminotransferase involved in the formation of 2-deoxystreptamine. **Org Biomol Chem**, 3, n. 8, p. 1410-1418, Apr 21 2005.

HUYNH, K.; PARTCH, C. L. Analysis of protein stability and ligand interactions by thermal shift assay. **Curr Protoc Protein Sci**, 79, p. 28 29 21-14, Feb 2 2015.

KABSCH, W. Xds. **Acta Crystallogr D Biol Crystallogr**, 66, n. Pt 2, p. 125-132, Feb 2010.

KANYO, Z. F.; SCOLNICK, L. R.; ASH, D. E.; CHRISTIANSON, D. W. Structure of a unique binuclear manganese cluster in arginase. **Nature**, 383, n. 6600, p. 554-557, Oct 10 1996.

KIM, S.; THIESSEN, P. A.; BOLTON, E. E.; CHEN, J. *et al.* PubChem Substance and Compound databases. **Nucleic Acids Res**, 44, n. D1, p. D1202-1213, Jan 4 2016.

KOTRA, L. P.; HADDAD, J.; MOBASHERY, S. Aminoglycosides: perspectives on mechanisms of action and resistance and strategies to counter resistance. **Antimicrob Agents Chemother**, 44, n. 12, p. 3249-3256, Dec 2000.

KUDO, F.; EGUCHI, T. Aminoglycoside Antibiotics: New Insights into the Biosynthetic Machinery of Old Drugs. **Chem Rec**, 16, n. 1, p. 4-18, Feb 2016.

KUDO, F.; TAMEGAI, H.; FUJIWARA, T.; TAGAMI, U. *et al.* Molecular cloning of the gene for the key carbocycle-forming enzyme in the biosynthesis of 2-deoxystreptamine-containing aminocyclitol antibiotics and its comparison with dehydroquinase synthase. **J Antibiot (Tokyo)**, 52, n. 6, p. 559-571, Jun 1999.

KUMAR, A.; SCHWEIZER, H. P. Bacterial resistance to antibiotics: active efflux and reduced uptake. **Adv Drug Deliv Rev**, 57, n. 10, p. 1486-1513, Jul 29 2005.

LEE, S. J.; KIM, D. J.; KIM, H. S.; LEE, B. I. *et al.* Crystal structures of *Pseudomonas aeruginosa* guanidinobutyrase and guanidinopropionase, members of the ureohydrolase superfamily. **J Struct Biol**, 175, n. 3, p. 329-338, Sep 2011.

LLEWELLYN, N. M.; SPENCER, J. B. Biosynthesis of 2-deoxystreptamine-containing aminoglycoside antibiotics. **Nat Prod Rep**, 23, n. 6, p. 864-874, Dec 2006.

- MAGNET, S.; BLANCHARD, J. S. Molecular insights into aminoglycoside action and resistance. **Chem Rev**, 105, n. 2, p. 477-498, Feb 2005.
- MCCOY, A. J.; GROSSE-KUNSTLEVE, R. W.; ADAMS, P. D.; WINN, M. D. *et al.* Phaser crystallographic software. **J Appl Crystallogr**, 40, n. Pt 4, p. 658-674, Aug 1 2007.
- MENZELLA, H. G.; REID, R.; CARNEY, J. R.; CHANDRAN, S. S. *et al.* Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. **Nat Biotechnol**, 23, n. 9, p. 1171-1176, Sep 2005.
- MOELLERING, R. C., Jr. Antibiotic resistance: lessons for the future. **Clin Infect Dis**, 27 Suppl 1, p. S135-140; discussion S141-132, Aug 1998.
- MURSHUDOV, G. N.; VAGIN, A. A.; DODSON, E. J. Refinement of macromolecular structures by the maximum-likelihood method. **Acta Crystallogr D Biol Crystallogr**, 53, n. Pt 3, p. 240-255, May 1 1997.
- NEWMAN, D. J.; CRAGG, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. **J Nat Prod**, 79, n. 3, p. 629-661, Mar 25 2016.
- PARK, S. R.; PARK, J. W.; BAN, Y. H.; SOHNG, J. K. *et al.* 2-Deoxystreptamine-containing aminoglycoside antibiotics: recent advances in the characterization and manipulation of their biosynthetic pathways. **Nat Prod Rep**, 30, n. 1, p. 11-20, Jan 2013.
- POPOVIC, B.; TANG, X.; CHIRGADZE, D. Y.; HUANG, F. *et al.* Crystal structures of the PLP- and PMP-bound forms of BtrR, a dual functional aminotransferase involved in butirosin biosynthesis. **Proteins**, 65, n. 1, p. 220-230, Oct 1 2006.
- PRACHAYASITTIKUL, V.; WORACHARTCHEEWAN, A.; SHOOMBUEATONG, W.; PRACHAYASITTIKUL, V. *et al.* Classification of P-glycoprotein-interacting compounds using machine learning methods. **EXCLI J**, 14, p. 958-970, 2015.
- PROJAN, S. J.; SHLAES, D. M. Antibacterial drug discovery: is it all downhill from here? **Clin Microbiol Infect**, 10 Suppl 4, p. 18-22, Nov 2004.
- REEVES, C. D. The enzymology of combinatorial biosynthesis. **Crit Rev Biotechnol**, 23, n. 2, p. 95-147, 2003.
- RODRIGUES, T.; REKER, D.; SCHNEIDER, P.; SCHNEIDER, G. Counting on natural products for drug design. **Nat Chem**, 8, n. 6, p. 531-541, Jun 2016.
- SCHUMANN, J.; HERTWECK, C. Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. **J Biotechnol**, 124, n. 4, p. 690-703, Aug 5 2006.
- SHEN, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. **Curr Opin Chem Biol**, 7, n. 2, p. 285-295, Apr 2003.
- SILVA, J. G.; CARVALHO, I. New insights into aminoglycoside antibiotics and derivatives. **Curr Med Chem**, 14, n. 10, p. 1101-1119, 2007.

STEWART, D. E.; SARKAR, A.; WAMPLER, J. E. Occurrence and role of cis peptide bonds in protein structures. **J Mol Biol**, 214, n. 1, p. 253-260, Jul 5 1990.

SUN, H.; LIU, Z.; ZHAO, H.; ANG, E. L. Recent advances in combinatorial biosynthesis for drug discovery. **Drug Des Devel Ther**, 9, p. 823-833, 2015.

THIBODEAUX, C. J.; MELANCON, C. E.; LIU, H. W. Unusual sugar biosynthesis and natural product glycodiversification. **Nature**, 446, n. 7139, p. 1008-1016, Apr 26 2007.

VAN BOECKEL, T. P.; GANDRA, S.; ASHOK, A.; CAUDRON, Q. *et al.* Global antibiotic consumption 2000 to 2010: an analysis of national pharmaceutical sales data. **Lancet Infect Dis**, 14, n. 8, p. 742-750, Aug 2014.

VARGHESE, J. N.; GARRETT, T. P.; COLMAN, P. M.; CHEN, L. *et al.* Three-dimensional structures of two plant beta-glucan endohydrolases with distinct substrate specificities. **Proc Natl Acad Sci U S A**, 91, n. 7, p. 2785-2789, Mar 29 1994.

WAKSMAN, S. A.; LECHEVALIER, H. A. Neomycin, a New Antibiotic Active against Streptomycin-Resistant Bacteria, including Tuberculosis Organisms. **Science**, 109, n. 2830, p. 305-307, Mar 25 1949.

WEISSMAN, K. J. Mutasyntesis - uniting chemistry and genetics for drug discovery. **Trends Biotechnol**, 25, n. 4, p. 139-142, Apr 2007.

WOODRUFF, H. B. Selman A. Waksman, winner of the 1952 Nobel Prize for physiology or medicine. **Appl Environ Microbiol**, 80, n. 1, p. 2-8, Jan 2014.

WRIGHT, G. D.; SUTHERLAND, A. D. New strategies for combating multidrug-resistant bacteria. **Trends Mol Med**, 13, n. 6, p. 260-267, Jun 2007.

ZAHA, A.; SCHRANK, A.; FERREIRA, H.B.; SCHRANK, I.S.; RODRIGUES, J.J.S.; REGNER, L.P.; PASSAGLIA, L.M.P.; ROSSETTI, M.L.R.; RAUPP, R.M.; SILVA, S.C.; GAIESKY, V.L.V. **Biologia Molecular Básica**. Porto Alegre: Mercado Aberto, p. 336, 1996.