

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA E IMUNOLOGIA

**Clonagem, expressão e purificação das enzimas
NahE e NahK de *Pseudomonas putida* para
determinação de suas estruturas cristalográficas**

SAMUEL LEITE GUIMARÃES

Belo Horizonte

2011

SAMUEL LEITE GUIMARÃES

**Clonagem, expressão e purificação das enzimas
NahE e NahK de *Pseudomonas putida* para
determinação de suas estruturas cristalográficas**

Dissertação de Mestrado apresentada ao Departamento de Bioquímica e Imunologia do Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Bioquímica e Imunologia.

Área de concentração: Estrutura e função de proteínas e peptídeos.

Orientador: Ronaldo Alves Pinto Nagem

Coorientadora: Glória Regina Franco

Belo Horizonte

2011



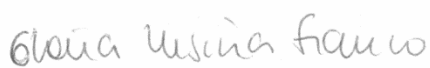
Universidade Federal de Minas Gerais -
 Curso de Pós-Graduação em Bioquímica e Imunologia ICB/UFMG
 Av. Antônio Carlos, 6627 – Pampulha
 31270-901 - Belo Horizonte – MG
 e-mail: pg-biq@icb.ufmg.br (31)3409-2615



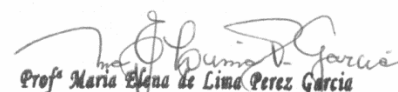
ATA DA DEFESA DA DISSERTAÇÃO DE MESTRADO DE SAMUEL LEITE GUIMARÃES. Aos dezessete dias do mês de fevereiro de 2011 às 10:00 horas, reuniu-se no Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais a Comissão Examinadora da dissertação de Mestrado, indicada *ad referendum* do Colegiado, para julgar, em exame final, o trabalho intitulado “Clonagem, expressão e purificação das enzimas NahE e NahK de *Pseudomonas putida* para determinação de suas estruturas cristalográficas” requisito final para a obtenção do grau de Mestre em Bioquímica e Imunologia. Abrindo a sessão, o prof. Ronaldo Alves Pinto Nagem, Orientador, da Universidade Federal de Minas Gerais, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato para apresentação de seu trabalho. Seguiu-se a argüição pelos examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações: Dr. Shaker Chuck Farah, da Universidade de São Paulo, aprovado; Dr. Carlos Renato Machado, da Universidade Federal de Minas Gerais, aprovado; Dra. Glória Regina Franco, Coorientadora, da Universidade Federal de Minas Gerais, aprovado e Dr. Ronaldo Alves Pinto Nagem, Orientador, da Universidade Federal de Minas Gerais, aprovado. Pelas indicações, o candidato foi considerada **APROVADO**. O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente da Comissão encerrou a reunião e lavrou a presente Ata que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 17 de fevereiro de 2011.


 Dr. Shaker Chuck Farah – USP


 Dr. Carlos Renato Machado – UFMG


 Dra. Glória Regina Franco – Coorientadora - UFMG


 Dr. Ronaldo Alves Pinto Nagem – Orientador – UFMG


 Prof. Maria Elena de Lima Perez Garcia
 Coordenadora do Curso de Pós Graduação
 em Bioquímica e Imunologia
 ICB - UFMG

Dedico este trabalho às duas pessoas que mais me deram amor, educação, apoio e incentivo. Pessoas as quais sempre torceram pelos meus sonhos e batalharam junto comigo para que eu pudesse realizá-los, meus pais: Custódio e Marizete.

AGRADECIMENTOS

Agradeço aos meus pais por todo amor e apoio incondicional, e que mesmo distantes, foram extremamente importantes para essa caminhada. Muito obrigado também pelo imenso incentivo por sempre buscarem pelo o que é melhor para mim.

À minha irmã, sobrinha, familiares e amigos do Triângulo mineiro, pelo carinho, apoio e compreensão por tantos momentos de ausência.

Aos meus padrinhos pelo apoio quando precisei e por sempre torcerem e vibrarem pelas minhas conquistas.

Ao meu orientador Ronaldo pelos inúmeros ensinamentos, oportunidades oferecidas e por ajudar em questões profissionais e pessoais. Sem dúvidas minha permanência em Belo Horizonte para a realização deste trabalho também foi de sua ajuda.

Aos meus amigos em Belo Horizonte Natássia, Rubens e Danilo, pelo apoio, amizade e risadas sem as quais os dias aqui não seriam tão bons. Especialmente ao Rubens por ser como um irmão e por me ajudar em diversos momentos difíceis.

À minha coorientadora Glória e à Marcela, assim como aos demais professores, alunos e técnicos do Laboratório de Genética Bioquímica pelos ensinamentos e sugestões nos experimentos de bancada.

Aos colegas do Laboratório de Biologia Estrutural: Simara, Débora, Juliana, Cíntia, Mariana, Sandro e Bruno, pela paciência e apoio em todos os experimentos, bem como pelos bons momentos compartilhados. Todos me ensinaram inúmeras coisas e sou muito grato.

À Renata do Laboratório Nacional de Luz Sincrotron em Campinas - SP pela ajuda com os experimentos de Dicroísmo Circular e Espalhamento Dinâmico de Luz; e ao André, também do LNLS, pela ajuda na difração de raios-X.

A todos os laboratórios e alunos do Departamento de Bioquímica e Imunologia que contribuíram de alguma forma para a realização deste trabalho.

Aos colegas de Bases pelo convívio, amizade e inúmeros bons momentos no decorrer do curso.

A CAPES e FAPEMIG pelo apoio financeiro.

RESUMO

GUIMARÃES, S. L. **Clonagem, expressão e purificação das enzimas NahE e NahK de *Pseudomonas putida* para determinação de suas estruturas cristalográficas.** Dissertação (Mestrado) - Departamento de Bioquímica e Imunologia, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

Devido ao uso do petróleo e seus derivados, diversos tipos de poluentes são liberados no meio ambiente. Alguns deles são classificados como Hidrocarbonetos Aromáticos Policíclicos (HAPs). A maioria dos HAPs, ou seus metabólitos, são capazes de interagir com bases nitrogenadas causando lesões no DNA, potencializando mutações e, conseqüentemente, o desenvolvimento de câncer. Como resultado disso, houve nas últimas décadas um aumento substancial na atenção dada às contaminações de solos e lençóis freáticos por HAPs. Inúmeros microrganismos possuem a capacidade de utilizar essas moléculas tóxicas como fontes de carbono e energia. O uso de determinadas bactérias (ou de suas enzimas) através de técnicas de biorremediação é uma estratégia em potencial para eliminação dos HAPs do ambiente. O HAP mais comum e tóxico é o naftaleno, e graças à capacidade da bactéria *Pseudomonas putida* em degradá-lo completamente, ela tem sido o foco de inúmeras pesquisas. Neste trabalho foram estudadas duas enzimas de *P. putida* envolvidas na degradação de naftaleno: NahE (uma hidratase-aldolase) e NahK (uma decarboxilase). O objetivo principal do trabalho foi a amplificação e clonagem dos genes *nahE* e *nahK*, expressão e purificação das respectivas enzimas para ensaios bioquímicos e estruturais. Os genes foram clonados inicialmente no vetor de expressão pET28a-TEV e expressos em *Escherichia coli*. A proteína recombinante NahE foi detectada em fração insolúvel e diversas estratégias foram utilizadas, sem sucesso, para obtenção da proteína expressa em fração solúvel. NahE solúvel foi obtida utilizando purificação em condições desnaturantes seguida por reenovelamento. Por outro lado, a proteína recombinante NahK foi detectada em fração solúvel e purificada por cromatografias de afinidade e de exclusão molecular. As proteínas purificadas foram submetidas a ensaios de Espalhamento Dinâmico de Luz e Dicroísmo Circular. Também foram realizados ensaios de cristalização com NahK visando o crescimento de cristais para a elucidação de sua estrutura tridimensional por Cristalografia de Raios-X.

Palavras chave: *Pseudomonas*, naftaleno, biorremediação.

Suporte Financeiro: CAPES & FAPEMIG.

ABSTRACT

GUIMARÃES, S. L. **Cloning, expression and purification of NahE and NahK enzymes from *Pseudomonas putida* for determination of their crystallographic structures.** Dissertação (Mestrado) - Departamento de Bioquímica e Imunologia, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

Many pollutants are released to the environment due the use of petroleum and its derivatives. Some of them are classified as Polycyclic Aromatic Hydrocarbons (PAHs). Most of PAHs, or their metabolites, can interact with nitrogenous bases or cause lesions in DNA, potentiating mutations and consequently inducing the development of cancer. As a result of this, in the last decades, the concerns about contamination of soil and groundwater by PAHs have substantially increased. Various microorganisms have the capacity to use these hazardous molecules as carbon and energy source. The use of these microorganisms and/or their enzymes for the elimination of PAHs to the environment is a potential strategy for bioremediation. One of the most common and hazardous PAH is naphthalene, and because of the bacteria *Pseudomonas putida* ability to completely degrade it, it has been the focus of numerous studies. This work was focused in the study of two enzymes from *P. putida* involved in naphthalene degradation: NahE (a hydratase-aldolase) and NahK (a decarboxylase). The main objectives of this work were the amplification and cloning of *nahE* and *nahK* genes, expression and purification of NahE and NahK enzymes for biochemical and structural assays. The genes were initially cloned into the expression vector pET28a-TEV and expressed in *Escherichia coli*. The recombinant protein NahE was mainly detected in insoluble fraction and different trials were performed to obtain the protein in the soluble fraction, without success. Soluble NahE was obtained only when using denaturizing conditions of purification followed by a refolding protocol. On the other hand, NahK was detected in the soluble fraction and was purified by affinity and size-exclusion chromatographies. The purified proteins were submitted to Dynamic Light Scattering and Circular Dichroism experiments. Crystallization assays were also performed with NahK aiming crystal growth for elucidation of its three-dimensional structure by X-Ray Crystallography.

Key words: *Pseudomonas*, naphthalene, bioremediation.

Financial Support: CAPES & FAPEMIG.

LISTA DE FIGURAS

Figura 1: Estrutura química dos 16 HAPs listados pela EPA como prioritários para monitoramento no meio ambiente.	4
Figura 2: Mecanismos de ativação de HAPs.....	6
Figura 3: Vias superior e inferior de degradação do naftaleno em <i>P. putida</i> G7.	10
Figura 4: Reação catalisada pela enzima trans-o-hidroxibenzilidenopiruvato hidratase-aldolase.	16
Figura 5: Representação em <i>cartoon</i> stereo da estrutura tridimensional da proteína Dihidrodipicolinato sintase.....	17
Figura 6: Reação catalisada pela enzima 4-oxalocrotonato decarboxilase.....	19
Figura 7: Representação em <i>cartoon</i> stereo da estrutura tridimensional da proteína HpcG.	20
Figura 8: Esquema do vetor pET28a-TEV.	32
Figura 9: Esquema do vetor pMCSG7.	32
Figura 10: Esquema do vetor pMCSG19C.....	32
Figura 11: Esquema de tratamento do inserto no sistema LIC.....	39
Figura 12: Esquema de tratamento do vetor no sistema LIC.	40
Figura 13: Amplificação do gene <i>nahE</i> de <i>P. putida</i> G7.....	55
Figura 14: Amplificação do gene <i>nahK</i> de <i>P. putida</i> G7.....	56
Figura 15: Digestão do vetor pGEM-T Easy para liberação do fragmento contendo <i>nahE</i>	57
Figura 16: Dupla digestão do vetor de expressão pET28a-TEV.	58
Figura 17: Sequenciamento dos clones 20 e 22 de <i>nahE</i>	59
Figura 18: Sequenciamento do clone 5 de <i>nahK</i>	60
Figura 19: Expressão da proteína NahE no vetor pET28a-TEV em BL21 a 37°C, IPTG 1 mM.	62

Figura 20: Expressões de NahE no vetor pET28a-TEV em BL21 a 30°C (A), 22°C (B) e 12°C (C); IPTG 0,5 mM.	62
Figura 21: Expressão de NahE no vetor pET28a-TEV em ArticExpress a 12°C, IPTG 0,5 mM.	63
Figura 22: Purificação da proteína NahE por cromatografia de afinidade ao níquel.	65
Figura 23: Digestão do vetor pMCSG19C utilizando a enzima Sspl.	68
Figura 24: Expressão da proteína NahE no vetor pMCSG19C em BL21 a 37°C, IPTG 1 mM.	69
Figura 25: Controle negativo em BL21 induzida a 37°C, IPTG 1 mM.	70
Figura 26: Expressão da proteína NahE no vetor pMCSG19C em BL21 a 30°C, IPTG a 1, 0,5 e 0,25 mM.	71
Figura 27: Expressão da proteína NahE no vetor pMCSG19C em BL21 a 22°C, IPTG a 0,5 mM.	71
Figura 28: Expressão da proteína NahE no vetor pMCSG19C em ArticExpress e BL21 a 12°C, IPTG a 0,5 mM.	72
Figura 29: Purificação da proteína NahE por cromatografia de afinidade ao níquel.	74
Figura 30: Expressão de NahE em pMCSG7 em BL21 a 37°C e tratamento da fração insolúvel.	76
Figura 31: Purificação da proteína NahE por cromatografia de afinidade em condições desnaturantes.	76
Figura 32: Reenovelamento da proteína NahE.	77
Figura 33: Cromatograma obtido durante a purificação por cromatografia de afinidade ao níquel da proteína NahE reenovelada.	78
Figura 34: Cromatograma obtido durante a purificação por cromatografia de exclusão molecular da proteína NahE reenovelada.	78
Figura 35: Sequenciamento de <i>nahE</i> em vetor pMCSG7.	79
Figura 36: Expressão da proteína NahK no vetor pET28a-TEV em BL21 a 30°C, IPTG 0,5 mM.	81
Figura 37: Purificação da proteína NahK por cromatografia de afinidade ao níquel.	82

Figura 38: Purificação da proteína NahK por cromatografia de troca iônica.	84
Figura 39: Purificação da proteína NahK por cromatografia de troca iônica.	85
Figura 40: Cromatografia de afinidade ao níquel de NahK para concentração da amostra.	86
Figura 41: Perfis de massas da clivagem de NahK utilizando 5% de TEV.....	87
Figura 42: Perfil de massas de NahK não clivada (curva azul), clivada (curva verde), e protease TEV (curva vermelha).....	88
Figura 43: Purificação da proteína NahK clivada por cromatografia de afinidade.	90
Figura 44: Purificação da proteína NahK clivada por cromatografia de exclusão molecular.....	91
Figura 45: Espectro de Dicroísmo Circular da proteína NahK em Tris-HCl pH 7,4 5mM, NaCl 5mM a 20°C.....	93
Figura 46: Espectro de Dicroísmo Circular da proteína NahK em Tris-HCl pH 7,4 5mM, NaCl 5mM em diferentes temperaturas.....	94
Figura 47: Deconvolução dos dados do espectro da Figura 46 feito pelo programa CDNN 2.1.....	95
Figura 48: Cristais de NahK em condições promissoras de cristalização.	96
Figura 49: Padrão de difração de cristal de NahK.....	97

LISTA DE TABELAS

Tabela 1: Análise de identidade das proteínas da via superior e inferior de degradação do naftaleno de <i>P. putida</i> G7 e de proteínas com estrutura tridimensional já elucidada depositada no PDB.	13
Tabela 2: Informações sobre organismo de origem e organismo de expressão das proteínas com estruturas já resolvidas e de mesma função das proteínas de <i>P. putida</i> G7.....	13
Tabela 3: Oligonucleotídeos sintéticos utilizados como iniciadores para amplificação dos genes <i>nahE</i> e <i>nahK</i>	31
Tabela 4: Reagentes e quantidades utilizadas nas reações de PCR.....	33
Tabela 5: Programa para amplificação dos genes <i>nahE</i> e <i>nahK</i> por PCR.....	34
Tabela 6: Antibióticos utilizados para os respectivos plasmídeos ou cepas, e concentrações de uso.	35
Tabela 7: Enzimas de restrição utilizadas e porcentagem de atividade em cada tampão.	37
Tabela 8: Reagentes e quantidades utilizados no tratamento com T4 DNA polimerase.....	41
Tabela 9: Materiais e quantidades utilizadas na preparação dos géis para SDS-PAGE.	46
Tabela 10: Análise da proteína NahE por espalhamento de luz dinâmico.	66
Tabela 11: Análise da proteína NahK por espalhamento dinâmico de luz.	93
Tabela 12: Dados coletados do cristal de NahK. Dados de alta resolução estão exibidos entre parênteses.	98

LISTA DE ABREVIATURAS E SIGLAS

m/m	percentual massa por massa
m/v	percentual massa por volume
v/v	percentual volume por volume
µg	micrograma (10^{-6} g)
µL	microlitro (10^{-6} L)
µm	micrometro (10^{-6} m)
µM	micromolar (10^{-6} M)
APS	persulfato de amônio
CD	dicroísmo circular (<i>Circular Dichroism</i>)
D.O. ₆₀₀	densidade óptica em comprimento de onda de 600 nm
Da	Dalton
dCMP	desoxicitidina-monofosfato
dCTP	desoxicitidina-trifosfato
dGMP	desoxiguanosina-monofosfato
dGTP	desoxiguanosina-trifosfato
DLS	espalhamento dinâmico de luz (<i>Dynamic Light Scattering</i>)
DMF	N,N-dimetilformamida
dNTP	desoxirribonucleotídeos trifosfatados
DTT	ditiotreitól
EPA	Agência de Proteção Ambiental dos Estados Unidos (<i>United States Environmental Protection Agency</i>)
EDTA	ácido etilenodiaminotetracético
<i>et al.</i>	e demais colaboradores
g	gramas

G	aceleração da gravidade
h	hora
HAPs	hidrocarbonetos aromáticos policíclicos
IPTG	isopropil- β -D-tiogalactopiranosídeo
IUPAC	União Internacional de Química Pura e Aplicada (<i>International Union of Pure and Applied Chemistry</i>)
kb	quilobase (10^3 base)
kDa	quiloDalton (10^3 Dalton)
L	litro
LB	meio de cultura Luria-Bertani
LIC	clonagem independente de ligação (<i>Ligation-Independent Cloning</i>)
M	molar (Mol/L)
MALDI-TOF	espectrometria de massa por ionização/desorção a laser assistida por matriz / tempo de voo (<i>Matrix Assisted Laser Desorption/Ionization / Time-of-Flight</i>)
mAU	mili unidades de absorbância
MBP	Proteína Ligadora de Maltose (<i>Maltose Biding Protein</i>)
mdeg	milidegrees
mg	miligrama (10^{-3} g)
min	minuto
mL	mililitro (10^{-3} L)
mM	milimolar (10^{-3} M)
mS/cm	milisiemens por centímetro
ng	nanograma (10^{-9} g)
nm	nanômetro (10^{-9} m)
pb	pares de bases
PCR	reação em cadeia da polimerase (Polymerase Chain Reaction)

pH	potencial hidrogeniônico
pI	ponto isoelétrico
q.s.p.	quantidade suficiente para
rpm	rotações por minuto
SDS	dodecil sulfato de sódio
SDS-PAGE	eletroforese em gel de poliacrilamida - SDS
seg	segundo
TAE	tampão Tris - acetato - EDTA
TE	tampão Tris - EDTA
TEMED	N,N,N',N'-tetrametiletilenodiamina
TEV	Tobacco Etch Virus
T _m	temperatura de anelamento
Tris	tris (hidroximetil) aminometano
TVMV	Tobacco Vein Mottling Virus
U	unidades
V	volts
vc	volumes de coluna
X	vezes

SUMÁRIO

1.	<u>INTRODUÇÃO</u>	<u>2</u>
1.1	Petróleo	2
1.2	Hidrocarbonetos aromáticos policíclicos	3
1.3	Biorremediação	6
1.4	Degradação de naftaleno em <i>Pseudomonas sp.</i>	8
1.5	Caracterização estrutural das proteínas envolvidas na degradação de naftaleno	11
1.6	Enzima NahE de <i>P. putida</i> G7	16
1.7	Enzima NahK de <i>P. putida</i> G7	18
2.	<u>JUSTIFICATIVA</u>	<u>21</u>
3.	<u>OBJETIVOS</u>	<u>23</u>
3.1	Objetivo geral	24
3.2	Objetivos específicos	24
4.	<u>MATERIAIS E MÉTODOS</u>	<u>25</u>
4.1	Cultivo de <i>P. putida</i> G7	26
4.2	Extração do plasmídeo NAH7	26
4.3	Análise da quantidade e qualidade de DNA	27
4.4	Eletroforese em gel de agarose	28
4.5	Tampão de amostra e padrão para DNA	28
4.6	Bactérias	29
4.7	Preparo de células eletrocompetentes	30
4.8	Desenho dos oligonucleotídeos iniciadores	30
4.9	Amplificação dos genes	33
4.10	Ligação de fragmentos no vetor de clonagem.	34
4.11	Transformação de <i>E. coli</i> por eletroporação	34
4.12	PCR de colônias	35
4.13	Extração de DNA plasmidial	36
4.14	Digestões enzimáticas e ligação de insertos a vetores de expressão	36
4.15	Clonagem Independente de Ligação	38
4.16	Sequenciamento de clones positivos	42
4.17	Expressão das proteínas recombinantes	42
4.18	Lise bacteriana e teste de solubilidade	43
4.19	Otimização da expressão das proteínas recombinantes	44
4.20	Reenovelamento de NahE insolúvel	44

4.21	Eletroforese em gel de poliacrilamida-SDS	45
4.22	Tampão de amostra e padrão para proteínas	47
4.23	Purificação das proteínas recombinantes	47
4.24	Determinação da concentração proteica	48
4.25	Clivagem com protease TEV	49
4.26	Espectrometria de Massa	49
4.27	Espalhamento Dinâmico de Luz	50
4.28	Dicroísmo Circular	51
4.29	Ensaio de cristalização e difração de raios-X	52
5.	RESULTADOS E DISCUSSÃO	54
5.1	Amplificação dos genes <i>nahE</i> e <i>nahK</i>	55
5.2	Subclonagem em vetor pGEM-T Easy	56
5.3	Digestão dos vetores pGEM-T Easy e pET28a-TEV	56
5.4	Clonagem de <i>nahE</i> e <i>nahK</i> no vetor de expressão pET28a-TEV	58
5.5	Expressão de NahE em vetor pET28a-TEV	61
5.6	Cromatografia de afinidade para NahE em pET28a-TEV	63
5.7	Caracterização do estado de agregação de NahE e Cpn60	66
5.8	Clonagem e expressão do gene <i>nahE</i> utilizando sistema LIC	67
5.9	Purificação de NahE expressa em vetor com sistema LIC	73
5.10	Reenovelamento de NahE	75
5.11	Caracterização da estrutura secundária de NahE reenovelada	80
5.12	Expressão e purificação de NahK em vetor pET28a-TEV	80
5.13	Clivagem de NahK pela protease TEV	86
5.14	Purificação de NahK sem cauda de histidina	88
5.15	Caracterização do estado de agregação de NahK	92
5.16	Caracterização da estrutura secundária de NahK	93
5.17	Ensaio de cristalização de NahK	95
6.	DISCUSSÃO	99
7.	CONCLUSÃO	104
8.	PERSPECTIVAS	106
9.	REFERÊNCIAS BIBLIOGRÁFICAS	108

1. INTRODUÇÃO

1.1 Petróleo

Nos últimos 200 anos houve um rápido crescimento da população mundial, resultando numa grande demanda por combustíveis e outros produtos utilizados para melhorar e manter a qualidade de vida.

Atualmente o petróleo é o principal combustível utilizado como fonte de energia pelo homem. Além de servir como base para a fabricação do diesel, gasolina e querosene, que são os principais combustíveis para veículos automotores, embarcações e aeronaves em todo o mundo, seu refino também resulta em vários outros materiais utilizados como base para fabricação dos mais variados produtos. Sua composição é uma combinação complexa e de proporção variável de moléculas orgânicas, composta na sua maioria de hidrocarbonetos alifáticos (também chamados de alcanos ou parafinas), alicíclicos (cicloalcanos ou naftenos) e aromáticos, podendo conter também quantidades pequenas de nitrogênio, oxigênio, enxofre e íons metálicos, principalmente níquel e vanádio [1, 2]. Esses hidrocarbonetos constituem 93 a 97% da massa total do petróleo e estão comumente presentes nas seguintes porcentagens: alcanos 40% - 70%, cicloalcanos 20% - 50%, aromáticos 5% - 20%. Entretanto, a proporção entre alcanos, cicloalcanos e aromáticos pode variar consideravelmente entre diferentes reservas de petróleo e áreas geográficas. Por exemplo, na reserva de petróleo de Troll - Noruega, a porcentagem de alcanos chega a apenas 13,9%, enquanto que cicloalcanos alcança 75,2% e aromáticos 10%. Já na reserva de Norne, também situada na Noruega, alcanos constituem 27,7%, cicloalcanos 34,8% e hidrocarbonetos aromáticos chegam a 37,5% da massa total do petróleo [3].

Devido à necessidade da exploração, produção, armazenamento, transporte, distribuição e destinação final do petróleo e seus derivados, contaminações do meio ambiente com hidrocarbonetos de petróleo tornaram-se frequentes, sendo reconhecido como um problema mundial a partir da metade dos anos 80 [4]. As áreas contaminadas representam riscos à saúde de vários seres vivos.

1.2 Hidrocarbonetos aromáticos policíclicos

Os hidrocarbonetos de petróleo abrangem um espectro muito grande de substâncias que podem ser agrupadas segundo seu grau de saturação, existências de ligações duplas ou triplas, ou segundo sua estrutura (linear, ramificada ou cíclica). Dentre os hidrocarbonetos totais do petróleo, se destaca uma família de compostos caracterizada por possuírem dois ou mais anéis aromáticos fundidos: os hidrocarbonetos aromáticos policíclicos (HAPs) [5].

Os HAPs são gerados naturalmente e liberados no meio ambiente através da combustão incompleta de matéria orgânica, como em queimadas e incêndios florestais, além de atividades vulcânicas [6]. Apesar disso, são estimados que 90% das emissões desses compostos ocorrem por atividade antropogênica através de queima de madeira para produção de carvão vegetal, uso de incineradores e outros processos de combustão incompleta, e, principalmente, devido à extração, transporte, refino, armazenamento e utilização do petróleo e seus derivados [5, 7].

Devido à possibilidade da fusão de um número variável de anéis e das várias posições em que estes podem ligar entre si, há atualmente mais de 100 HAPs reconhecidos pela IUPAC. Dentre estes, 32 foram listados pela EPA como os principais HAPs poluentes, sendo que 16 (Figura 1) são alvos frequentes para monitoramento e dosagem de amostras em função de suas importâncias físico-químicas, toxicológicas, industriais e ambientais [8].

HAPs são moléculas estáveis quimicamente, e, somando-se a essa característica a sua hidrofobicidade, tendem a ser fortemente adsorvidos na fase sólida do solo persistindo no ambiente por muito tempo, o que aumenta a possibilidade de exposição de humanos e outros seres vivos a esses compostos [8]. O tempo de permanência no ambiente depende de vários fatores, como a estrutura química do HAP, concentração e a biodisponibilidade do contaminante. Adicionalmente, fatores como tipo e estrutura de solo, pH, temperatura e presença de níveis adequados de oxigênio podem limitar a capacidade da biota microbiana local em degradar essas moléculas [9, 10].

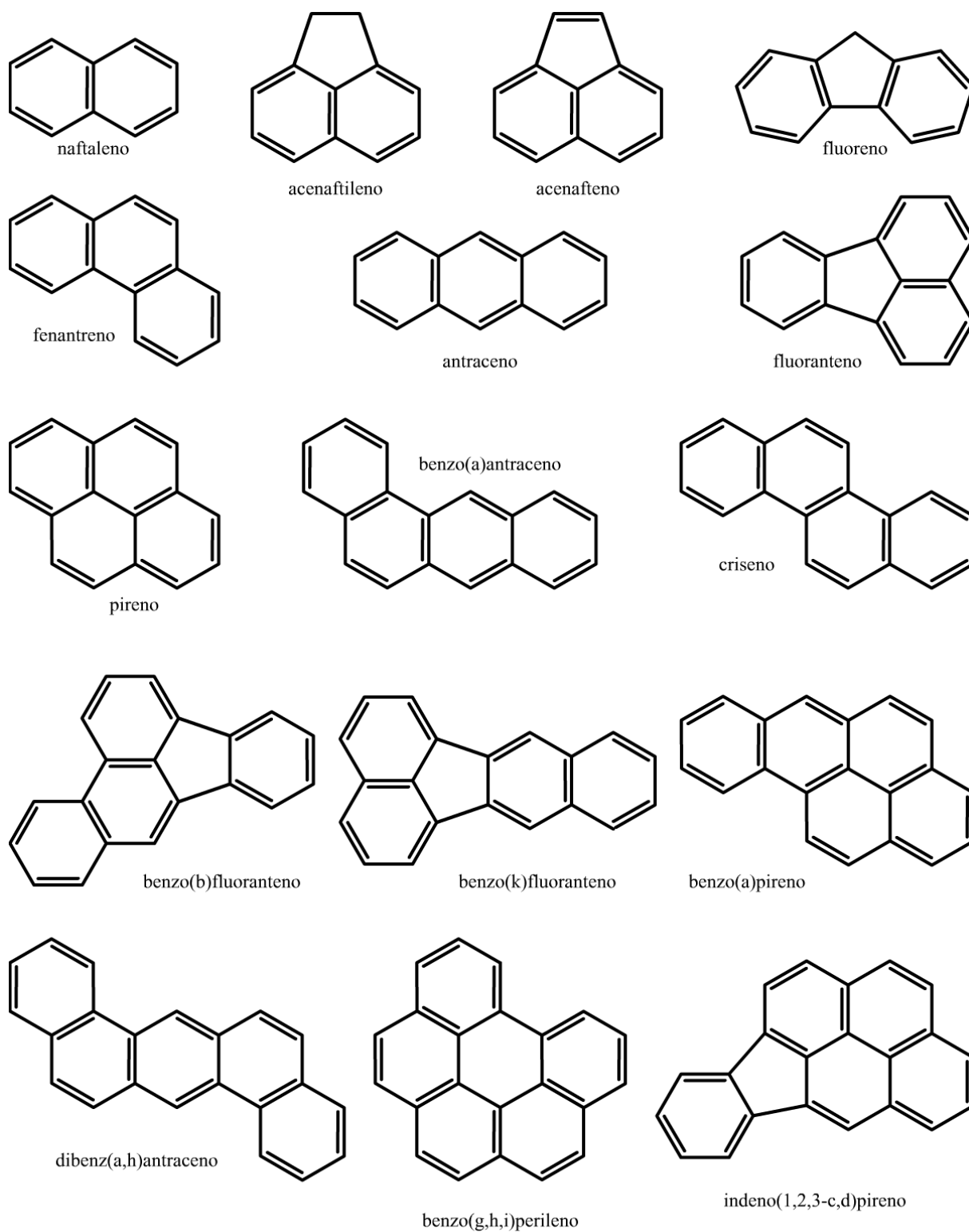


Figura 1: Estrutura química dos 16 HAPs listados pela EPA como prioritários para monitoramento no meio ambiente.

Quando incorporados pela flora e fauna, HAPs atravessam facilmente a membrana celular, ligando-se às moléculas proteicas e, devido ao seu caráter lipofílico, tornam-se parte da reserva lipídica, podendo ser transferidos sem alteração de suas estruturas da presa para o predador e, eventualmente, ao homem. De fato, muitos seres expõem-se à contaminação por hidrocarbonetos derivados do petróleo por ingestão de animais ou plantas que acumularam tais moléculas. Isto traz sérias implicações para a pesca e a saúde pública [11]. HAPs e seus derivados também podem ser absorvidos pela pele ou por inalação, sendo rapidamente distribuídos pelo organismo [5].

Por serem lipofílicos, os HAPs podem permanecer indefinidamente no organismo enquanto associados ao tecido adiposo. Seu metabolismo, no entanto, torna-os mais hidrossolúveis permitindo serem mais facilmente excretados. Apesar disso, ao longo de seu metabolismo, essas moléculas podem ser convertidas a metabólitos ainda mais tóxicos e carcinogênicos.

O metabolismo dos HAPs no organismo envolve uma série de enzimas presentes em todos os tecidos e que catalisam reações de oxidação, redução, hidrólise e conjugação. Devido às diferenças estruturais dos diversos HAPs, várias vias podem ser utilizadas em sua biomodificação, ocorrendo diferenças no metabolismo e na carcinogenicidade de cada molécula.

Alguns mecanismos têm sido propostos para explicar a ativação de HAPs. No mecanismo por formação de diol-epóxido (Figura 2 A), monoxigenases dependentes do citocromo P450 oxidam os HAPs formando compostos epóxidos que podem espontaneamente produzirem fenóis ou, por ação das epóxido hidrolases, produzirem dihidrodióis. Esses últimos podem sofrer nova epoxidação levando à formação de dihidrodióis epóxidos que são capazes de reagir com as bases nucleotídicas do DNA e, eventualmente, iniciar um processo mutagênico. No mecanismo via formação de quinona (Figura 2 B), a desidrogenação enzimática dos dihidrodióis produzem quinonas capazes de reagirem diretamente com o DNA ou capazes de reagirem com O_2 gerando espécies reativas de oxigênio, como radicais hidroxilas ou ânions superóxidos que atacam o DNA. Estes mecanismos não são excludentes, podendo ocorrer simultaneamente [5, 12-15].

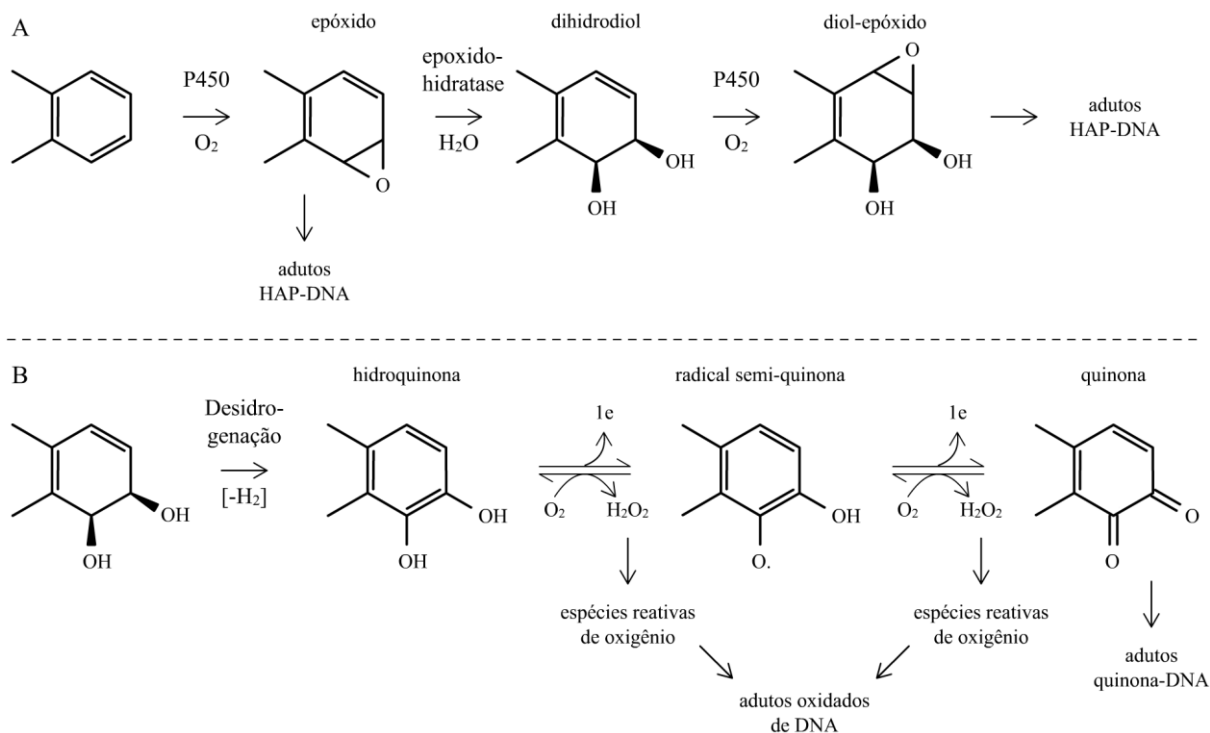


Figura 2: Mecanismos de ativação de HAPs.

1.3 Biorremediação

Processos de remediação físicos e químicos consistem na coleta do material contaminado, seja água ou solo, e remoção dos poluentes por adsorção em carvão ativado, oxidação por meio de aquecimento em temperaturas de até 500°C, ou outros processos demorados e dispendiosos [11].

Devido à presença natural de HAPs no ambiente, diversos microrganismos adquiriram a capacidade de utilizar essas moléculas como fonte de carbono e energia. Desde a década de 1950, diversas bactérias e fungos degradadores destes compostos estão sendo identificados e isolados [16].

Dessa maneira, uma interessante estratégia para eliminação dos HAPs de áreas contaminadas é através da biorremediação. O termo engloba uma série de tecnologias distintas para tratamento de solos e águas contaminadas utilizando consórcios microbianos nativos ou não para a degradação dos contaminantes, liberando para o meio ambiente substâncias inertes como CO_2 e H_2O [17]. Esta

biotecnologia vem sendo utilizada há vários anos em outros países e, em certos casos, apresenta menor custo e maior eficiência na remoção dos contaminantes do que as técnicas físicas e químicas (como incineração e lavagem do solo), sendo atualmente utilizada em escala comercial no tratamento de diversos resíduos e na remediação de áreas contaminadas [10].

A degradação de HAPs por microrganismos é limitada por uma série de fatores, sendo necessários níveis adequados de nutrientes como nitrogênio e fósforo, quantidade de água no solo, pH, temperatura, entre outros [18]. Além disso, geralmente há poucos microrganismos da biota local que são capazes de degradar completamente os poluentes, e a introdução de microrganismos alóctones ou organismos geneticamente modificados resistentes às condições diversas do ambiente e capazes de fazer uma degradação completa dos contaminantes pode implicar em uma competição com a microbiota nativa gerando um desequilíbrio no ecossistema [19].

A degradação dos HAPs dentro dos microrganismos ocorre através de diversas vias contendo séries de reações catalisadas por enzimas. Dessa maneira, uma opção para a biorremediação é a utilização das próprias enzimas ao invés dos microrganismos nos sítios contaminados. Algumas vantagens dessa técnica são [20]:

- as enzimas são específicas, catalisando reações bem definidas sem gerar subprodutos indesejáveis;
- não permanecem no ambiente após o tratamento, pois são degradadas por microrganismos nativos;
- através da tecnologia do DNA recombinante é possível produzir enzimas em grande escala com atividade e estabilidade aumentadas.

Entretanto, alguns fatores devem ser considerados [21]:

- durante o processo devem ser mantidas condições adequadas para atividade das enzimas para que a degradação dos poluentes seja eficiente;
- o custo para produção e uso de todas as enzimas de uma via de degradação pode ser elevado, portanto a seleção e produção de

enzimas específicas que catalisam passos determinantes devem ser analisadas a fim de diminuir os custos;

- ter baixa dependência de cofatores ou coenzimas, como NAD(P)H, que são caros e podem levar a um custo proibitivo de uso em grande escala.

Algumas dessas implicações têm sido contornadas através da resolução da estrutura tridimensional dessas proteínas. A partir do entendimento de sua estrutura-função é possível fazer uso de técnicas de engenharia de proteínas, como mutações sítio-dirigidas, para aumentar a estabilidade e atividade enzimática aprimorando o uso dessas enzimas nos processos de biorremediação [22].

1.4 Degradação de naftaleno em *Pseudomonas sp.*

O desenvolvimento de enzimas para biorremediação envolve etapas cruciais como: selecionar e isolar microrganismos capazes de degradar os compostos de interesse, identificar as proteínas responsáveis pelas vias de degradação, isolar os genes codificadores dessas enzimas, expressar e purificar as enzimas para estudos estruturais, e, finalmente, aprimorar as propriedades enzimáticas em laboratório [21].

As bactérias estão ativamente envolvidas na degradação de poluentes orgânicos de áreas contaminadas. Diversas espécies bacterianas são conhecidas pela capacidade de degradar HAPs, sendo que a maioria das que possuem uma eficiente biodegradação foram isoladas de solos contaminados [23].

O naftaleno é o HAP mais simples, formando por dois anéis benzênicos fundidos. Possui baixa solubilidade em água (30 mg/L), sublima facilmente à temperatura ambiente e tende a ser fortemente absorvido pelo solo. Apesar de sua hidrofobicidade, dentre os HAPs, o naftaleno é o que possui a maior solubilidade em água, sendo um dos primeiros compostos a atingir lençóis freáticos em caso de contaminação do solo, e o primeiro de uma listagem de 16 HAPs a terem seus níveis monitorados no ambiente [24]. Sua ingestão ou inalação pelo homem pode provocar diversos efeitos, dentre eles: anemia hemolítica, catarata, aumento do baço, danos

aos rins e ao cérebro, além de formação de estresse oxidativo nas células, aumentando as chances de mutações no DNA e desenvolvimento de câncer [25].

A degradação de naftaleno por *Pseudomonas putida* G7 tem sido bastante estudada devido à presença de um plasmídeo de 83 kb chamado NAH7 que codifica enzimas capazes de degradar completamente esse contaminante [26]. Os genes do plasmídeo NAH7 envolvidos na via de degradação de HAPs estão organizados em três operons. O primeiro codifica as enzimas da via superior que são responsáveis pela degradação de naftaleno à salicilato. O segundo codifica enzimas da via inferior, que converte salicilato a intermediários do ciclo do ácido cítrico. O terceiro operon codifica uma proteína regulatória denominada NahR, que age como regulador positivo dos outros dois operons, utilizando salicilato como indutor [27]. As enzimas, substratos e produtos da via superior e inferior de degradação do naftaleno estão representados na Figura 3.

A sequência nucleotídica completa e identificação dos genes contidos no plasmídeo NAH7 de *P. putida* G7 está depositada no banco de dados de nucleotídeos do *National Center for Biotechnology Information* (NCBI) [26].

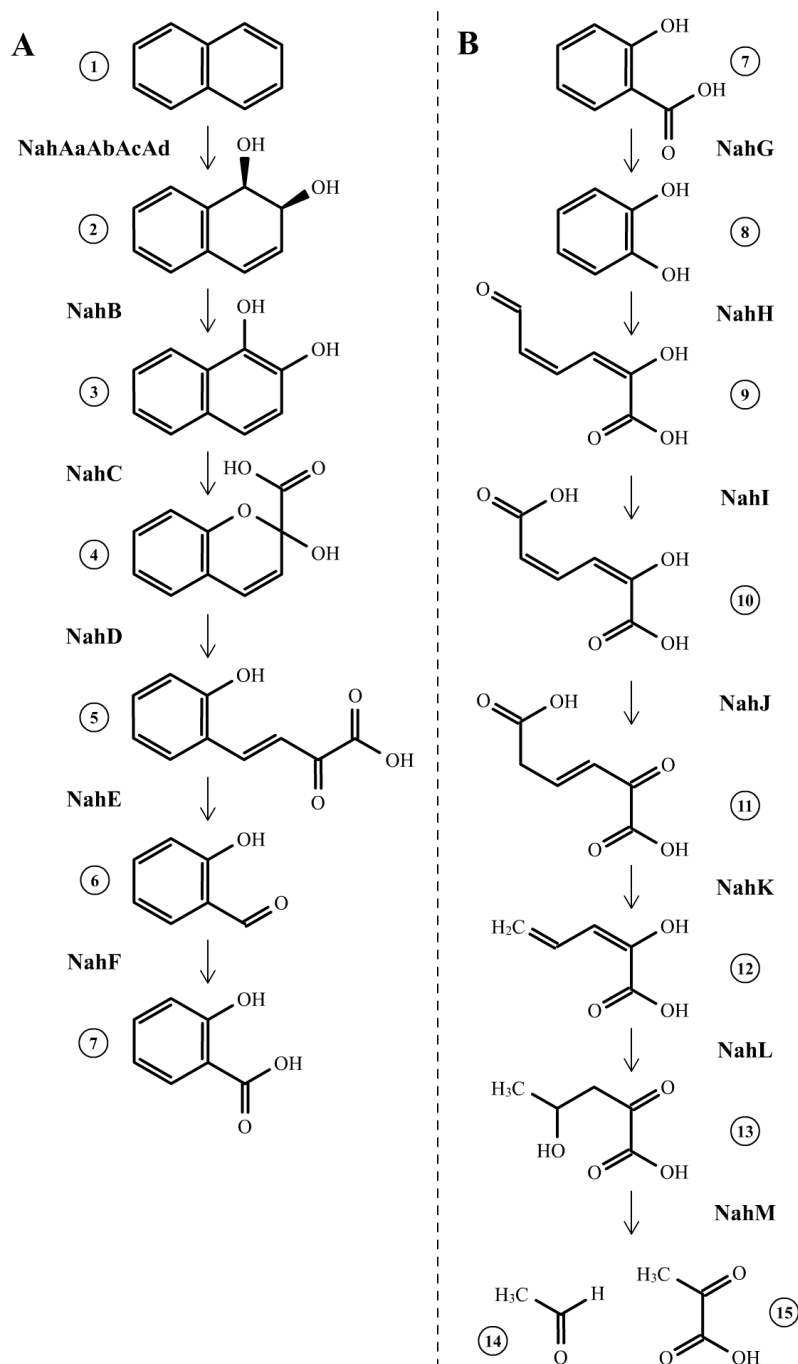


Figura 3: Vias superior e inferior de degradação do naftaleno em *P. putida* G7.

A - via superior; **B** - via inferior; **NahAaAbAcAd** - complexo enzimático da naftaleno dioxigenase; **NahB** - cis-dihidrodiol naftaleno desidrogenase; **NahC** - 1,2-dihidroxi-naftaleno dioxigenase; **NahD** - 2-hidroxi-2H-cromeno-2-carboxilato isomerase; **NahE** - hidroxibenzilidenopiruvato hidratase-aldolase; **NahF** - salicilaldeído desidrogenase; **NahG** - salicilato hidroxilase; **NahH** - catecol 2,3-dioxigenase; **NahI** - aminomuconato-semialdeído desidrogenase; **NahJ** - 4-oxalocrotonato tautomerase; **NahK** - 4-oxalocrotonato decarboxilase; **NahL** - 2-oxopent-4-enoato hidratase; **NahM** - 2-oxo-4-hidroxipentanoato aldolase; **(1)** naftaleno; **(2)** cis-1,2-dihidronaftaleno-1,2-diol; **(3)** 1,2-dihidroxi-naftaleno; **(4)** ácido 2-hidroxi-2H-cromeno-2-carboxílico; **(5)** ácido trans-o-hidroxibenzilidenopiruvato; **(6)** salicilaldeído; **(7)** ácido salicílico; **(8)** catecol; **(9)** 2-hidroxiomuconato semialdeído; **(10)** 2-hidroxiomuconato; **(11)** 4-oxalocrotonato; **(12)** 2-oxopent-4-enoato; **(13)** 4-hidroxi-2-oxopentanoato; **(14)** acetaldeído; **(15)** piruvato.

1.5 Caracterização estrutural das proteínas envolvidas na degradação de naftaleno

O método experimental mais comum para obter um modelo detalhado de uma macromolécula, sendo permitido identificar a posição de átomos individuais, é através da interpretação do padrão de Difração de Raios-X vindo de muitas moléculas idênticas em um arranjo ordenado, como um cristal [28].

Determinar a estrutura tridimensional de uma proteína por difração de raios-X implica em obter cristais de alta qualidade de proteína purificada, medir direções e intensidades de feixes de raios-X difratados pelo cristal. Posteriormente, utilizando programas computacionais, simular por meio de imagens a distribuição das nuvens eletrônicas da macromolécula de forma que a posição e densidade dos átomos no modelo sejam consistentes com os dados obtidos da difração [28].

Em certas circunstâncias, muitos compostos moleculares, incluindo proteínas, podem solidificar na forma de cristais, que consiste num arranjo tridimensional ordenado onde as moléculas adotam orientações idênticas e repetitivas ao longo de todo o cristal. Os cristais crescem através de lenta e controlada precipitação a partir de uma solução aquosa em condições que não desnature a proteína. Para coletar os dados de difração, o cristal é posicionado entre uma fonte e um detector de raios-X. Ao atravessar o cristal, os raios-X são difratados e emergem como se fossem ondas refletidas por centenas de planos cristalinos imaginários, aparecendo como pontos espalhados na imagem capturada pelo detector. Cada onda possui uma amplitude e fase, porém, durante o processo de difração, somente as informações de intensidade e direção de cada onda são coletados. A informação da fase é perdida. Porém, para resolver a estrutura da proteína é necessária essa informação. Um dos métodos utilizado para contornar esse problema é a Substituição Molecular, que consiste em utilizar as fases de uma proteína conhecida como estimativa inicial para as fases da molécula em estudo. A proteína conhecida é referida como modelo de fases e deve apresentar certa identidade com a proteína estudada [28].

O Banco de Dados de Proteínas (PDB - *Protein Data Bank*) é um portal de informação de estruturas macromoleculares biológicas, no qual são depositadas estruturas tridimensionais resolvidas de proteínas e ácidos nucleicos [29]. De um

total de 70.813 estruturas depositadas até o dia 25 de janeiro de 2011, aproximadamente 87% foram resolvidas por Difração de Raios-X e 89% destas proteínas foram obtidas na forma recombinante utilizando *Escherichia coli* como organismo produtor. Nesse portal, utilizando a ferramenta *Sequence Search*, é possível localizar, para uma proteína alvo com sequência de aminoácidos conhecida, proteínas com estrutura tridimensional já resolvida e que possuem certa identidade com a proteína alvo.

Uma análise das sequências de aminoácidos das enzimas da via superior e inferior de degradação do naftaleno de *P. putida* G7 (Tabela 1) indicou que as subunidades Ab, Ac e Ad do complexo enzimático NahAaAbAcAd, bem como as enzimas NahC, NahD, NahH e NahJ apresentam proteínas já depositadas no PDB com mais de 85% de identidade. De fato, essas proteínas já caracterizadas estruturalmente possuem a mesma função que as proteínas de *P. putida* G7. Portanto, essas enzimas não seriam, inicialmente, candidatas à determinação de suas estruturas tridimensionais. Na Tabela 2 é possível observar alguns detalhes referentes ao organismo de origem e organismo no qual essas respectivas enzimas foram expressas para obtenção de amostra para determinação de suas estruturas. Todas as enzimas citadas na Tabela 2 tiveram suas estruturas determinadas por cristalografia de raios-X.

Tabela 1: Análise de identidade das proteínas da via superior e inferior de degradação do naftaleno de *P. putida* G7 e de proteínas com estrutura tridimensional já elucidada depositada no PDB.

Enzima	Código PDB	Identidade	Enzima	Código PDB	Identidade
NahAa	1KRH	30%	NahF	1O9J	36%
NahAb	2QPZ	<u>94%</u>	NahG	2RGJ	26%
NahAc	1EG9	<u>97%</u>	NahH	1MPY	<u>85%</u>
NahAd	1EG9	<u>95%</u>	NahI	2WME	39%
NahB	1BDB	36%	NahJ	1OTF	<u>85%</u>
NahC	2EHZ	<u>97%</u>	NahK	2EB4	41%
NahD	2IMD	<u>87%</u>	NahL	1SV6	39%
NahE	3PUD	30%	NahM	1NVM	55%

Tabela 2: Informações sobre organismo de origem e organismo de expressão das proteínas com estruturas já resolvidas e de mesma função das proteínas de *P. putida* G7.

Enzima	PDB	Organismo de origem	Expressão	Referência
NahAb	2QPZ	<i>Pseudomonas putida</i>	<i>Escherichia coli</i>	[30]
NahAc NahAd	1EG9	<i>Pseudomonas putida</i>	<i>Escherichia coli</i>	[31]
NahC	2EHZ	<i>Pseudomonas sp. C18</i>	<i>Pseudomonas putida</i> Kt2442	a ser publicado
NahD	2IMD	<i>Pseudomonas putida</i>	<i>Escherichia coli</i>	[32]
NahH	1MPY	<i>Pseudomonas putida</i> MT-2	<i>Escherichia coli</i> W3110	[33]
NahJ	1OTF	<i>Pseudomonas sp. CF600</i>	<i>Escherichia coli</i>	[34]

Por outro lado, as proteínas NahB, NahE, NahF, NahG, NahI, NahK e NahL apresentam macromoléculas com estruturas resolvidas com baixa identidade, cerca de 26 a 41%. Logo, a resolução de suas estruturas tridimensionais seria mais interessante por apresentar um maior potencial para a descoberta de novos conhecimentos; principalmente para aqueles envolvidos na relação estrutura-função destas enzimas.

Outras técnicas também são utilizadas para caracterizar a estrutura de proteínas, como Espalhamento Dinâmico de Luz (DLS) e Dicroísmo Circular (CD).

O espalhamento de ondas eletromagnéticas pela matéria, incluindo biopartículas, tem sido usado por décadas como fonte de informação da estrutura e dinâmica de moléculas em solução. As técnicas de espalhamento de luz utilizam ondas eletromagnéticas entre os comprimentos de onda de 400 a 800 nm, sendo amplamente utilizadas para estudo da massa, forma e estado de agregação e oligomerização de proteínas em soluções aquosas [35].

O experimento de DLS mede as flutuações da intensidade da luz espalhada por partículas em movimento browniano em solução. A intensidade de movimentação das partículas é inversamente proporcional ao tamanho das mesmas. Logo, uma molécula proteica monomérica e pequena move-se mais rapidamente do que um dímero ou uma molécula maior. Além disso, proteínas de mesmo peso molecular, mas de diferentes formas são caracterizadas por diferentes coeficientes de difusão. Os coeficientes de difusão das partículas são obtidos a partir das análises das flutuações instantâneas da luz espalhada pelas partículas. O raio hidrodinâmico das partículas é obtido a partir da equação de Stokes-Einstein, na qual se utiliza o coeficiente de difusão da partícula [35, 36].

O DLS é uma técnica bastante útil para se avaliar o potencial de cristalização de proteínas. As interações proteína-proteína estimadas por essa técnica servem para prever condições favoráveis de cristalização, assim como para a solubilidade da proteína. Amostras monodispersas são mais facilmente cristalizadas do que amostras polidispersas [37]. O estado monodisperso é aquele em que a solução apresenta apenas a proteína em seu estado monomérico ou oligomérico natural. Já no estado polidisperso, a solução é formada pela proteína em seu estado natural em baixa concentração e grande quantidade de agregados proteicos [35].

O Dicroísmo Circular é outro excelente método para o estudo de proteínas. Por meio dessa técnica é possível avaliar a estrutura secundária, o enovelamento e as propriedades de ligação de proteínas com substratos ou cofatores. Quando as proteínas interagem com luz circularmente polarizada à direita e à esquerda, elas absorvem diferentemente cada uma dessas ondas, sendo o espectro de CD definido pela diferença dessa absorção [38, 39].

Um feixe de luz possui um campo elétrico e magnético que oscilam perpendicularmente um ao outro e à direção de propagação. Quando a luz é polarizada seu campo elétrico oscila sinusoidalmente num único plano alterando sua amplitude ao longo do eixo de propagação. Luz polarizada circularmente ocorre quando o vetor do campo elétrico rotaciona ao longo do eixo de propagação da onda, mantendo uma amplitude constante. Utilizando filtros é possível separar a onda que circula no sentido horário da que circula no sentido inverso. Quando estas ondas atravessam um meio contendo moléculas quirais, como as proteínas, haverá uma diferença na absorção de uma e outra. Essa diferença é bem pequena, mas possível de ser medida com instrumentação moderna [38]. Os resultados são fornecidos em termos de elipticidade expressa em *milidegrees*. Para proteínas, é mais adequado expressar essa diferença em termos de elipticidade molar em $\text{mdeg.cm}^2.\text{dmol}^{-1}$, que considera a massa molecular da proteína, seu número de resíduos e a concentração proteica utilizada ao realizar o experimento [40].

Os espectros do CD de proteínas são geralmente divididos em três faixas de comprimentos de onda, baseados na energia das transições eletrônicas que dominam cada faixa. Essas são: UV distante (abaixo de 250 nm), onde as contribuições dos peptídeos predominam; UV próximo (250 a 300 nm), onde as contribuições das cadeias laterais dos resíduos aromáticos predominam; e de UV à luz visível (300 a 700 nm), onde as contribuições dos cromóforos extrínsecos predominam [40].

O CD na região do UV distante reflete o conteúdo de estrutura secundária da proteína, uma vez que é dominado pelas contribuições dos cromóforos amidas dos resíduos de aminoácidos, os quais determinam a conformação da cadeia. Já as análises do CD na região de UV próximo fornecem informações sobre as interações terciárias, uma vez que, nessa faixa de comprimento de onda, predominam as contribuições dos resíduos aromáticos e pontes dissulfeto. Na região do UV à luz visível é possível obter informações sobre cromóforos extrínsecos como grupos prostéticos, íons metálicos, inibidores e substratos [40].

Os polipeptídeos formam diferentes estruturas secundárias, baseados no arranjo dos grupos amidas. As principais estruturas secundárias presentes nas proteínas são as hélices- α e as fitas- β . As hélices- α são estabilizadas por interações de hidrogênio de resíduos subsequentes da proteína. As fitas- β ao interagirem umas

com as outras formam as folhas- β que também são estabilizadas por interações de hidrogênio, porém essas se dão entre átomos das diferentes fitas, podendo ser da mesma proteína ou de diferentes subunidades. Essas estruturas são caracterizadas por um conjunto de ângulos φ e ψ . O ângulo da ligação N - C de um aminoácido é denominado φ , e o ângulo da ligação C - C de um aminoácido é ψ . Para as hélices- α o ângulo φ comum é -57° e o ψ é -47° . Para as folhas- β , o ângulo φ comum é -120° e o ψ é $+120^\circ$ [40]. Esses dois tipos de estruturas secundárias têm espectros de CD característicos. O espectro de CD para as hélices- α é caracterizado por duas bandas negativas a 222 e 208 nm e uma banda positiva a 193 nm. Para as folhas- β , o espectro é caracterizado por uma banda negativa a 218 nm e uma positiva a 195 nm, enquanto que para estruturas randômicas e desordenadas, é característica uma banda pouco positiva a 210 nm e uma banda negativa em 195 nm [39].

1.6 Enzima NahE de *P. putida* G7

A proteína NahE nativa de *P. putida* G7, de acordo com análise feita pelo programa ProtParam [41], hospedado no servidor proteômico ExPASy (*Expert Protein Analysis System*) [42] possui 331 resíduos de aminoácidos, peso molecular de 36.636,9 Da e pI teórico de 5,48.

No banco de dados de enzimas BRENDA [43] a enzima é classificada com número EC (*Enzyme Commission*) 4.1.2.45 e a reação catalisada pela mesma está representada na Figura 4.

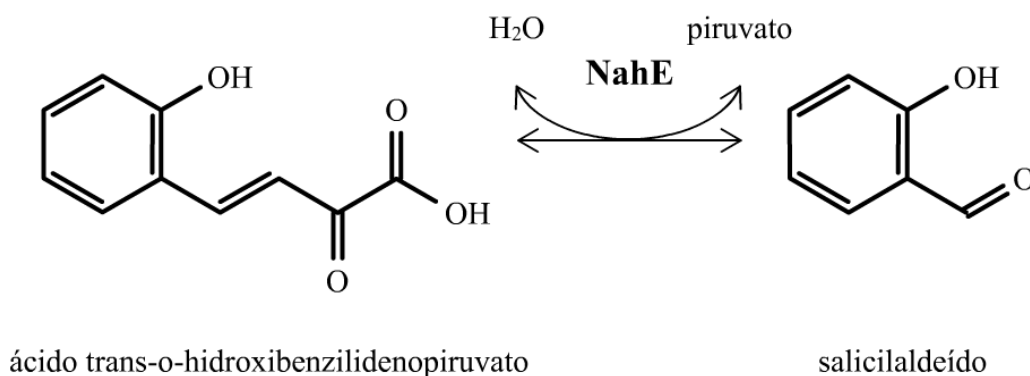


Figura 4: Reação catalisada pela enzima trans-o-hidroxibenzilidenopiruvato hidratase-aldolase.

A avaliação da atividade enzimática de NahE pode ser feita a partir da análise em espectrofotômetro observando a diminuição da absorbância do substrato trans-o-hidroxibelzilidenopiruvato nos comprimentos de onda de 296 e 340 nm, e o aumento da absorbância do produto salicilaldeído a 255 nm [44]. Como a reação é reversível, pode-se também avaliar sua atividade através da junção de salicilaldeído e piruvato (disponíveis comercialmente), gerando o produto trans-o-hidroxibelzilidenopiruvato (indisponível comercialmente).

Através do banco de dados Pfam 24.0 (*Protein Families Database*) [45] foi possível identificar que NahE pertence ao grupo das Dihidro-dipicolinato sintetases, que está dentro da superfamília TIM-fosfatase. Essa superfamília possui proteínas de enovelamento característico em $\alpha\beta$ -Barril bastante conservado [46].

Na Figura 5 é possível observar um enovelamento característico em $\alpha\beta$ -Barril. Este modelo, representado em *cartoon* stereo para evidenciar as estruturas secundárias, corresponde à estrutura tridimensional da proteína Dihidro-dipicolinato sintase de *Acinetobacter baumannii* com código PDB 3PUD (Jithesh, O.; sem publicação) e que compartilha com a enzima NahE 30% de identidade.

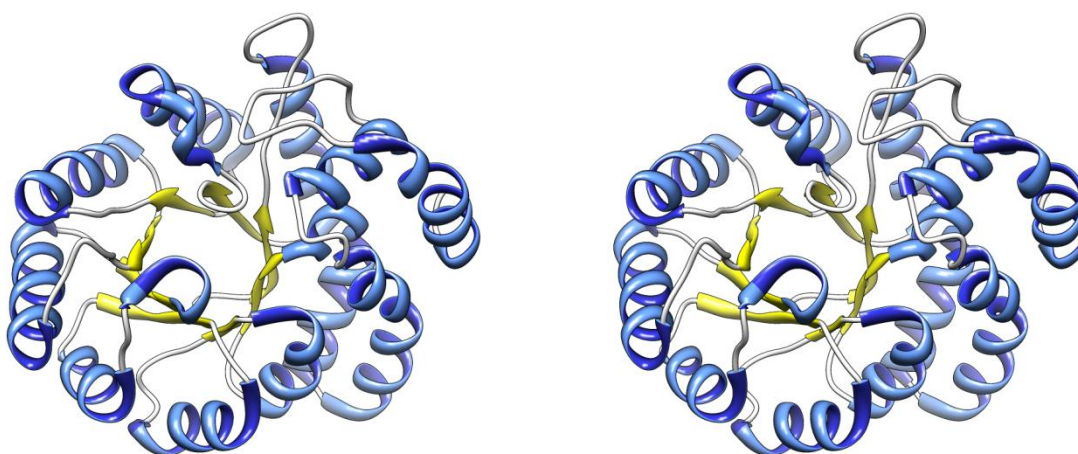


Figura 5: Representação em *cartoon* stereo da estrutura tridimensional da proteína Dihidro-dipicolinato sintase.

Apesar de essa proteína ser a mais semelhante, já resolvida estruturalmente, com NahE, uma análise do alinhamento de suas sequências de resíduos de aminoácidos realizado pela ferramenta BLAST [47], mostra que essa semelhança cobre apenas 50% da sequência total dessas proteínas, sendo que a sequência compreendida entre os resíduos 16 a 183 de NahE (169 de 331 resíduos) e os resíduos 2 a 161 de Dihidrodipicolinato sintase (160 de 291 resíduos) possui 30% de identidade. Não há semelhanças significativas de sequência com outras proteínas de estrutura tridimensional conhecida para os resíduos 1 a 16 e 184 a 331 de NahE.

Além das características da enzima NahE citadas acima, outros parâmetros da mesma podem ser obtidos por programas de bioinformática. O servidor TMHMM Server v2.0 [48], que avalia a estrutura primária de uma proteína e prediz a presença de hélices transmembrana, indica que NahE aparentemente não possui este tipo de motivo estrutural. Já o programa XtalPred Server [49], que faz a predição da 'cristalibilidade' de uma proteína - ranqueando-as de 1 (cristalização ótima) a 5 (cristalização muito difícil), sugere que NahE pertence à classe 2, com boas chances de cristalização. Além disso, o programa Regional Order Neural Network (RONN) [50], que prediz regiões desordenadas na proteína - as quais dificultam o processo de cristalização - indica que NahE pode possuir algumas regiões de desordem estrutural entre os resíduos 1 a 5, 10 a 41, 298 e 299, 301 a 303.

1.7 Enzima NahK de *P. putida* G7

De acordo com o programa ProtParam, a proteína NahK nativa de *P. putida* G7 possui 264 resíduos de aminoácidos, peso molecular de 28.435,6 Da e pI teórico de 5,18. No banco de dados de enzimas BRENDA a enzima é classificada com número EC 4.1.1.77 e a reação catalisada pela mesma está representada na Figura 6.

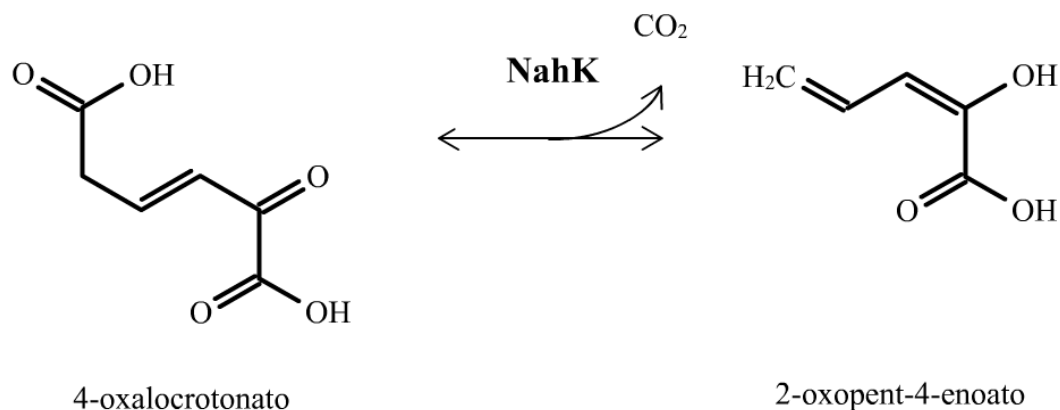


Figura 6: Reação catalisada pela enzima 4-oxalocrotonato decarboxilase.

Em 1984 foram realizados alguns estudos com genes de *P. putida* mt-2. O operon do plasmídeo TOL pWWO dessa cepa foi clonado em *E. coli* K-12 e verificou-se a expressão e atividade de oito enzimas da via de degradação do catecol (correspondentes à via inferior de degradação do naftaleno), dentre elas, NahK [51]. Mais tarde, em 1989, as enzimas NahK e NahL foram copurificadas do extrato de *P. putida* utilizando troca aniônica, cromatografia de interação hidrofóbica e cromatografia de exclusão molecular, sendo verificada a possível formação de um complexo dessas duas proteínas *in vivo* [52]. No mesmo trabalho, NahK foi expressa em *E. coli* K-12 na ausência de NahL, purificada por troca aniônica e por cromatografia de interação hidrofóbica. Após esse último passo de purificação, ocorria perda da atividade de NahK.

Em 2000, a fim de avaliar detalhes da atividade de NahK, os genes *nahK* e *nahL* foram clonados a partir do plasmídeo TOL pWWO de *P. putida* mt-2 e as proteínas coexpressas em *E. coli* BL21(DE3). Mutações sítio dirigidas em NahL baseadas em alinhamentos de proteínas homólogas resultaram na inativação da atividade dessa proteína, enquanto que NahK manteve sua atividade [53]. Em 2004 o mesmo grupo, utilizando alinhamentos de sequências homólogas à NahK, identificaram inúmeros resíduos conservados e provavelmente importantes para sua atividade enzimática. Alguns mutantes de NahK foram construídos e verificou-se a perda da atividade dessa enzima ao substituir alguns resíduos de glutamato para glutamina [54]. Desde então esse grupo anseia pela resolução das estruturas de NahK e NahL para melhor entender seus mecanismos de ação enzimática.

Segundo o banco de dados Pfam, NahK pertence à família das Fumarilacetoacetato hidrolases. Os programas TMHMM, XtalPred e RONN, juntamente com a sequência de NahK, sugerem a inexistência de hélices transmembrana, ótimo potencial de cristalização (classe 1) e a predição de uma única e pequena região de desordem compreendida entre os resíduos 55 a 64.

Utilizando a sequência de resíduos de aminoácidos de NahK e a ferramenta de alinhamento BLAST juntamente com o banco de dados PDB, foi possível identificar que a proteína HpcG de *Escherichia coli*, já resolvida estruturalmente e com código PDB 2EB4 [55], possui 41% de identidade com NahK. A similaridade dessas proteínas cobre 95% de suas sequências, compreendida entre os resíduos 13 a 264 de NahK (252 de 264 resíduos) e os resíduos 10 a 267 de HpcG (258 resíduos de 267).

Na Figura 7 é possível observar a estrutura tridimensional de HpcG representada em *cartoon* stereo para evidenciar suas estruturas secundárias. Devido a grande similaridade entre as sequências de NahK e HpcG, é provável que ambas possuam enovelamento semelhante e que a estrutura de HpcG possa ser utilizada como modelo de fases na Substituição Molecular para a resolução da estrutura tridimensional de NahK.

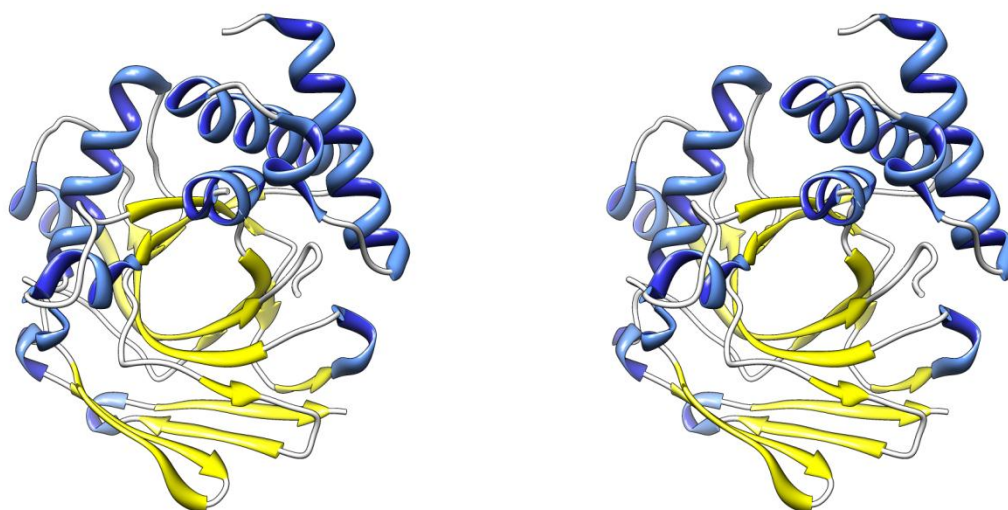


Figura 7: Representação em *cartoon* stereo da estrutura tridimensional da proteína HpcG.

2. JUSTIFICATIVA

Em vista da grande utilização do petróleo e seus derivados, dos quais a atual sociedade é tão dependente, têm ocorrido inúmeros casos de contaminações de solo e água por HAPs. Devido às suas características e a possibilidade de ocasionarem problemas à saúde de humanos e animais, deve ser buscada uma rápida eliminação de ambientes contaminados. A biorremediação é uma alternativa para a descontaminação de ambientes com HAPs, na qual os microrganismos degradadores irão transformá-los em substâncias inertes. Em vista de a maioria dos microrganismos do solo não possuírem a capacidade de degradar estes compostos, há necessidade de se isolar e selecionar os microrganismos degradadores.

O naftaleno é um dos HAPs mais tóxicos e está presente em contaminações por derrames de gasolina, diesel, entre outros produtos derivados do petróleo. Considerando-se as atuais descobertas dos genes de enzimas degradadoras de naftaleno e os problemas gerados por esse composto, torna-se essencial o estudo aprofundado das enzimas que participam de sua degradação.

A determinação das estruturas tridimensionais dessas enzimas é importante, pois, com isso, é possível caracterizar a relação estrutura-função das mesmas. Assim, podem ser propostas modificações nas sequências de aminoácidos das proteínas que promovam aumento da sua estabilidade e atividade catalítica, para que, futuramente, essas macromoléculas possam ser utilizadas como biorremediadoras em condições adversas de pH e temperatura.

Algumas enzimas constituintes da via superior e inferior de degradação do naftaleno apresentam proteínas com alta identidade cujas estruturas tridimensionais já foram elucidadas, por exemplo, algumas proteínas do complexo NahAaAbAcAd, NahC, e NahD, restando outras enzimas para terem suas estruturas resolvidas.

Desta maneira, propõe-se neste trabalho o estudo das enzimas NahE e NahK de *P. putida* G7 integrantes da via de degradação de naftaleno, um composto tomado como modelo na degradação de HAPs e um dos mais comumente encontrados no meio ambiente.

3. OBJETIVOS

3.1 Objetivo geral

Clonar e expressar os genes *nahE* e *nahK* e purificar as enzimas hidroxibenzilidenopiruvato hidratase-aldolase (NahE) e 4-oxalocrotonato decarboxilase (NahK) de *P. putida* G7 a fim de caracterizar a estrutura secundária e terciária de suas formas recombinantes.

3.2 Objetivos específicos

- Amplificar e clonar os genes *nahE* e *nahK* a partir do plasmídeo NAH7 de *P. putida* G7;
- Otimizar a expressão em *E. coli* para obtenção das enzimas recombinantes em fração solúvel;
- Expressar e purificar NahE e NahK em quantidades suficientes para realização dos ensaios bioquímicos e estruturais;
- Determinar o estado de oligomerização das enzimas em solução através de Espalhamento Dinâmico de Luz;
- Caracterizar a estrutura secundária das proteínas através de Dicroísmo Circular;
- Realizar ensaios de cristalização das proteínas visando determinar suas estruturas cristalográficas por Difração de Raios-X.

4. MATERIAIS E MÉTODOS

4.1 Cultivo de *P. putida* G7

A bactéria *P. putida* G7, cedida pelo Dr. Masataka Tsuda do Departamento de Ciências da Vida Ambiental da Faculdade de Ciências da Vida da Universidade de Tohoku, Japão, foi cultivada em meio mineral M9 com 1% m/v de naftaleno. O meio mineral M9 contém por litro: 6 g de Na_2HPO_4 ; 3 g de KH_2PO_4 ; 0,5 g de NaCl; 10 mL de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 0,1 M; e 10 mL de CaCl_2 1M. Foi feita uma solução estoque de naftaleno diluído em DMF na concentração de 300 mg/mL.

O preinóculo de 50 μL do estoque de bactéria foi feito em 10 mL de meio M9 contendo 166,5 μL da solução de naftaleno dissolvido em DMF. A cultura foi incubada a 28°C por 48 a 72 h sob agitação de 150 rpm. Após o turvamento do meio, foi realizado o inóculo de 1 mL da cultura em 19 mL de meio LB líquido sem antibiótico e incubado a 28°C por 48 h sob agitação de 150 rpm.

4.2 Extração do plasmídeo NAH7

Como o plasmídeo NAH7 é muito grande (82.832 pb) [56], não é possível extraí-lo utilizando protocolo padrão de extração DNA plasmidial. Dessa forma a extração do plasmídeo NAH7 foi realizada de acordo com o protocolo de extração de DNA genômico de bactérias [57], com algumas modificações, como descrito a seguir.

Após o cultivo do inóculo, 1 mL de cultura foi transferido para um microtubo e centrifugado a 15.000 *G* por 2 min. O sobrenadante foi descartado e o precipitado lavado com 1 mL de tampão TE (Tris-HCl 10 mM pH 8,0 e EDTA 1 mM) e, em seguida, centrifugado novamente a 15.000 *G* por 2 min. O sobrenadante foi descartado e adicionou-se ao precipitado 230 μL de TE, 15 μL de SDS 10% e homogeneizou-se lentamente. Foram adicionados 245 μL de fenol/clorofórmio 1:1 e misturou-se gentilmente. A amostra foi centrifugada a 15.000 *G* por 10 min. Em seguida, a fase aquosa (fase superior) foi transferida para um novo microtubo e adicionou-se igual volume de clorofórmio. Novamente, centrifugou-se a amostra a

15.000 G por 10 min e a fase aquosa (fase superior) foi transferida para um novo microtubo. Adicionou-se então, 25 µL de NaCl 2 M e 150 µL de isopropanol gelado e homogeneizou-se levemente. O próximo passo foi a incubação por 30 min a -80°C e, após isso, a amostra foi centrifugada a 17.000 G por 10 min. O sobrenadante foi descartado e adicionou-se ao microtubo 250 µL de etanol 70%, sem ressuspender o precipitado. A amostra foi centrifugada a 17.000 G por 10 min e o sobrenadante foi descartado. Novamente foram adicionados 250 µL de etanol 70%, sem ressuspender o precipitado, e centrifugado a 17.000 G por 10 min. O sobrenadante foi descartado e, após o restante do etanol evaporar a temperatura ambiente, o precipitado foi ressuspensionado em 30 µL de tampão TE ou H₂O milli-Q contendo RNase 10 µg/mL. A amostra foi colocada em banho-maria a 37°C por 1 h e depois armazenada a -20°C.

4.3 Análise da quantidade e qualidade de DNA

A quantificação da extração de DNA plasmidial de *P. putida* G7, assim como outras amostras de DNA, foi realizada por eletroforese em gel de agarose, ou mais precisamente por espectrofotometria. Quando utilizada a eletroforese em gel de agarose, a concentração do DNA foi estimada pela comparação da intensidade de fluorescência das bandas, quando expostas a radiação ultravioleta, com um padrão feito com quantidades conhecidas de DNA plasmidial. Já a dosagem mais precisa foi feita por meio da leitura de absorvância a 260 nm em aparelho Nanodrop® ND-1000 Spectrophotometer, sendo o grau de pureza das amostras determinado pela relação das leituras de absorvância a 260 e 280 nm. Conforme especificado pelo fabricante, recomenda-se que essa relação seja de aproximadamente 1,8.

4.4 Eletroforese em gel de agarose

Todas as amostras de DNA analisadas por eletroforese foram corridas em géis de agarose 1% preparados como descrito a seguir.

Um por cento m/v de agarose foi dissolvido por aquecimento em tampão TAE (Tris-acetato 40 mM pH 8,0 e EDTA 1 mM). Para visualização das bandas de DNA foi adicionado SYBR® Safe 0,1 µL/mL (Invitrogen). O tampão utilizado na corrida eletroforética foi o mesmo utilizado na confecção do gel. Após a polimerização do gel, as amostras foram aplicadas nas canaletas e a eletroforese executada a 80 V por cerca de 45 min utilizando o sistema Mini-Sub Cell GT (Bio-Rad) e fonte FB300 (Fisher Scientific). Para visualização do resultado, o gel foi exposto à luz ultravioleta no transiluminador UV 302 nm T26M (Bio Agency). O SYBR® Safe foi utilizado como fluoróforo no lugar do brometo de etídio, pois ele apresenta baixa ou nenhuma atividade mutagênica [58]. Os géis foram fotodocumentados utilizando câmera Vilber Lourmat.

4.5 Tampão de amostra e padrão para DNA

Todas as amostras de DNA aplicadas em gel de agarose foram previamente misturadas com tampão de amostra para DNA 6X preparado com 0,3% m/v azul de bromofenol; 0,3% m/v xileno cianol; 50% v/v glicerol e 1,0 mM EDTA pH 8,0; ou misturados com tampão de amostra 6X Orange DNA Loading Dye (Fermentas), que contém, xileno cianol e orange G. O azul de bromofenol migra na mesma taxa que um fragmento de DNA de aproximadamente 300 pb, o xileno cianol na mesma taxa de um fragmento de 4.000 pb, e o orange G na mesma taxa que um fragmento de 50 pb. Nas eletroforeses foram incluídos os padrões 1 kb plus DNA Ladder (Invitrogen), que possui fragmentos de DNA entre 100 e 12.000 pb; O'GeneRuler™ 1 kb DNA Ladder (Fermentas), que possui fragmentos de DNA entre 250 e 10.000 pb; ou 100 pb EZ Load Molecular Ruler (Bio-Rad), que possui fragmentos de DNA entre 100 pb e 1000 pb.

4.6 Bactérias

As bactérias *E. coli* DH5 α TM (Life Technologies), *E. coli* BL21(DE3) (Novagen) e ArticExpressTM(DE3) (Stratagene) foram utilizadas no preparo de células eletrocompetentes para transformação.

As bactérias da linhagem DH5 α foram utilizadas para replicação das construções plasmidiais. Essa cepa de *E. coli* possui uma mutação no gene *endA*, representada no genótipo por *endA1*. O gene *endA* selvagem codifica a Endonuclease I, que é uma proteína periplasmática de 12 kDa capaz de degradar DNA de dupla fita. Cepas de *E. coli* com essa mutação produzem uma forma inativa da nuclease e são referidas como *EndA*⁻, permitindo assim um significativo aumento no rendimento e qualidade de DNA plasmidial usado para sequenciamentos ou outros ensaios [59]. Essa cepa também carrega outros genes modificados, como a mutação *hsdR17* que elimina a endonuclease de restrição do sistema de restrição-modificação *EcoKI*, permitindo que DNA não metilado por esse sistema não seja degradado [60]. Também está presente o alelo $\Delta(lacZ)M15$, permitindo a coloração azul/branca para identificação de colônias positivas para genes clonados em vetores baseados no sistema *lacZ* [61].

As linhagens BL21(DE3) e ArticExpress(DE3) foram utilizadas para expressão das proteínas recombinantes. São derivadas da cepa *E. coli* B que é deficiente das proteases Lon e OmpT, as quais podem degradar proteínas durante a purificação. Também carrega o gene para a enzima T7 RNA Polimerase, essencial para a superexpressão de genes clonados em vetores que carregam o promotor T7. A cepa ArticExpress é derivada da linhagem BL21-Gold da Stratagene e possui um plasmídeo que codifica as chaperonas Cpn10 e Cpn60 de *Oleispira antartica*. Este plasmídeo apresenta um gene de resistência à gentamicina. A Cpn10 possui massa molecular de aproximadamente 10 kDa e a Cpn60, de 57 kDa. As chaperonas são macromoléculas que facilitam o enovelamento correto de proteínas através da ligação e estabilização de proteínas parcialmente ou incorretamente enoveladas [62]. As chaperonas Cpn10 e Cpn60 de *O. antartica* apresentam alta atividade em temperaturas de 4 a 12°C. Em ArticExpress, essas chaperonas auxiliam no processamento de proteínas expressas em baixas temperaturas, podendo aumentar o rendimento de proteínas recombinantes, tornando-as mais solúveis e ativas.

4.7 Preparo de células eletrocompetentes

Todas as células eletrocompetentes foram preparadas de acordo com o protocolo presente no manual do eletroporador MicroPulser™ (Bio-Rad), como descrito a seguir.

Inicialmente, foi feito um preinóculo da cepa de interesse em 5 mL meio LB líquido que foi incubado 16 h a 37°C sob agitação de 150 rpm. Cepas que contém plasmídeos, como ArticExpress, tiveram os respectivos antibióticos adicionados ao meio de cultura em concentrações exibidas na Tabela 6. Após o crescimento, o preinóculo foi utilizado como inóculo em 500 mL de meio LB líquido. As células do inóculo cresceram a 37°C sob agitação de 200 rpm, sendo o crescimento monitorado em espectrofotômetro até a cultura alcançar uma leitura de D.O.₆₀₀ entre 0,5 a 0,7. Ao alcançar o nível de crescimento desejado, a cultura foi resfriada em gelo por 20 min e, em todos os passos subsequentes, sempre que possível, as células e todos os recipientes foram mantidos em gelo.

As próximas etapas consistem em 5 ciclos de centrifugações a 4.000 G, por 15 min a 4°C, sendo o sobrenadante cuidadosamente descartado, seguidos de ressuspensões das células em volumes cada vez menores de glicerol 10% estéril e gelado. Inicialmente em volume de 500 mL, 250 mL, 20 mL e por fim 2 mL. Após serem ressuspensas no volume final, as células foram imediatamente armazenadas em microtubos contendo alíquotas de 50 µl e incubadas em gelo até serem transportadas e armazenadas a -80°C, estando assim prontas para uso.

4.8 Desenho dos oligonucleotídeos iniciadores

As sequências nucleotídicas dos genes *nahE* e *nahK* foram obtidas a partir das sequências do plasmídeo NAH7 de *P. putida* G7 depositadas no banco de dados nucleotídicos do National Center for Biotechnology Information - NCBI [56]. Os iniciadores foram desenhados no programa Primer Premier 5.0, tendo como base as sequências dos respectivos genes alvo.

De acordo com os métodos para clonagem utilizando o vetor de expressão pET28a-TEV [63] (Figura 8), foram inseridas nos iniciadores pequenas sequências reconhecidas por enzimas de restrição compatíveis com sítio múltiplo de clonagem desse vetor. Para os vetores pMCSG7 e pMCSG19C [64, 65] (Figura 9 e Figura 10), foram inseridas sequências maiores compatíveis com o sistema LIC presente nesses vetores. Os iniciadores direto e reverso tiveram suas T_m calculadas pelo programa citado anteriormente sem considerar as sequências de restrição ou do sistema LIC, e foram ajustadas para não ultrapassar 2°C de diferença. Os iniciadores, sítios das enzimas de restrição e sítios do sistema LIC estão representados na Tabela 3. Os nucleotídeos sublinhados correspondem às sequências reconhecidas pelas enzimas de restrição ou pelos sítios do sistema LIC.

Tabela 3: Oligonucleotídeos sintéticos utilizados como iniciadores para amplificação dos genes *nahE* e *nahK*.

Iniciadores	Sequências nucleotídicas	Enzimas de restrição	T_m (°C)
nahE-fw-NdeI	5' <u>CATATG</u> TTGAATAAAGTTATTTAAAACC 3'	<i>NdeI</i>	50,8
nahE-rev-HindIII	5' <u>AAGCTT</u> CTGCGGAGGTGAACTACTT 3'	<i>HindIII</i>	52,4
nahE-fw-LIC	5' <u>TACTTCCAATCCAATG</u> CGATGTTGAATAAAGTTATTTAAAACC 3'	-	50,8
nahE-rev-LIC	5' <u>TTATCCACTTCCAATG</u> CTGCGGAGGTGAACTACTT 3'	-	52,4
nahK-NheI-F	5' <u>GCTAGCAT</u> GAAACCGTACCCTGACCCGC 3'	<i>NheI</i>	65,3
nahK-HindIII-R	5' <u>AAGCTT</u> TCAGATAAAGCGCGGATACCG 3'	<i>HindIII</i>	66,6

Os iniciadores para *nahE* (SIGMA), e para *nahK* (Invitrogen) foram recebidos liofilizados. O material foi ressuspendido em H₂O milli-Q estéril até uma concentração estoque de 100 µM e armazenado a -20°C em alíquotas de 10 µL. Para uso dos iniciadores, uma alíquota foi diluída em H₂O milli-Q autoclavada até a concentração de 10 µM.

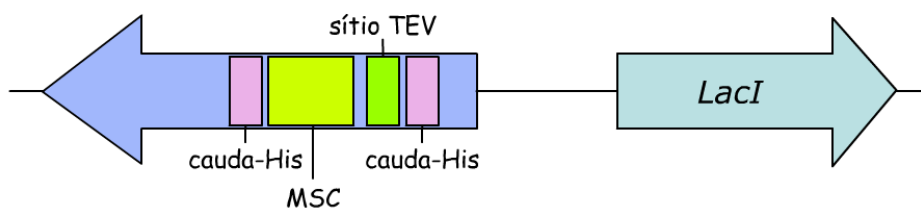


Figura 8: Esquema do vetor pET28a-TEV.

Utilizando determinados sítios de restrição do sítio múltiplo de clonagem (MSC) é possível expressar a proteína com a cauda de histidinas na posição N-terminal seguida do sítio para clivagem pela protease TEV a fim de remover a cauda de histidinas após purificação por cromatografia de afinidade; ou então utilizar uma cauda de histidinas C-terminal, sem a possibilidade de clivagem dessa cauda.

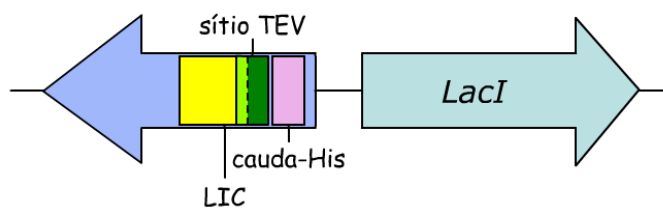


Figura 9: Esquema do vetor pMCSG7.

Utilizando o sistema de Clonagem Independente de Ligaç o no vetor pMCSG7, a prote na   expressa com uma cauda de histidinas na posi o N-terminal que pode ser removida da prote na de interesse pela clivagem da protease TEV.

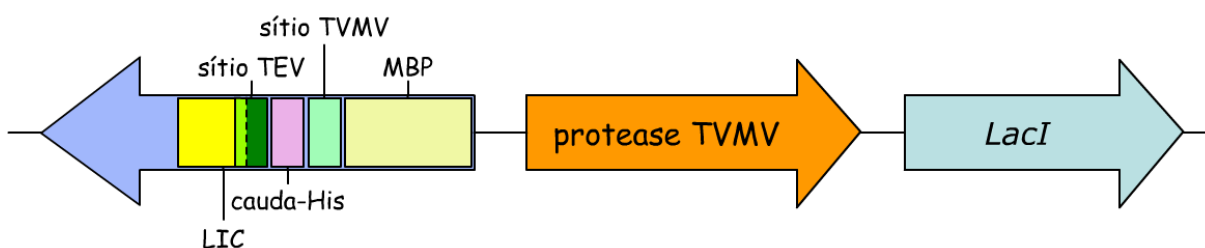


Figura 10: Esquema do vetor pMCSG19C.

Utilizando esse vetor, a prote na   expressa fusionada a prote na MBP, seguida por um s tio de clivagem pela protease TVMV com uma subsequente cauda de histidina e um s tio de clivagem pela protease TEV. Como a protease TVMV   coexpressa com a prote na de interesse, a cauda MBP   clivada *in vivo*, liberando a prote na com uma cauda de histidinas que pode ser removida por clivagem com protease TEV.

4.9 Amplificação dos genes

As amplificações dos genes foram realizadas através de PCR em Termociclador MJ Mini Personal (Bio-Rad), e otimizadas a fim de se obter o mínimo de amplificações inespecíficas. Os reagentes, suas quantidades e suas respectivas concentrações que foram utilizados nas amplificações estão representados na Tabela 4. Os programas utilizados para a amplificação estão representados na Tabela 5. Os fragmentos amplificados dos genes foram visualizados através de eletroforese em gel de agarose.

Tabela 4: Reagentes e quantidades utilizadas nas reações de PCR.

Reagente	Concentração final	Volume
Plasmídeo NAH7 de <i>P. putida</i> G7 a 30 ng/μL	3 ng/μL	5 μL
Tampão 10X	1X	5 μL
50 mM MgSO ₄	2 mM	2 μL
Platinum® Taq DNA Polymerase High Fidelity 5 U/μL (Invitrogen)	1 U	0,2 μL
10 mM dNTPs	0,2 mM	1 μL
10 μM Iniciador direto	0,2 μM	1 μL
10 μM Iniciador reverso	0,2 μM	1 μL
H ₂ O milli-Q autoclavada	-	q.s.p. 50 μL

Tabela 5: Programa para amplificação dos genes *nahE* e *nahK* por PCR.

Etapas	Temperatura (°C)	Tempo
1. Início	94	5 min
2. Desnaturação	94	1 min
3. Anelamento	45 (<i>nahE</i>), 58 (<i>nahK</i>)	1 min
4. Extensão	72	1 min
5. Repetição das etapas 2 a 4 (29X)	-	-
6. Extensão final	72	5 min

4.10 Ligação de fragmentos no vetor de clonagem.

O material amplificado de *nahE*, obtido com os iniciadores *nahE-fw-NdeI* e *nahE-rev-HindIII*, assim como o material amplificado do gene *nahK*, obtido com os iniciadores *nahK-NheI-F* e *nahK-HindIII-R*, após purificação utilizando o kit *illustra GFX™ PCR DNA and Gel Band Purification* (GE Healthcare), foram subclonados no vetor *pGEM®-T Easy* (Promega) utilizando a enzima *T4 DNA Ligase* (Promega) conforme instruções do fabricante. Após a reação de ligação, os plasmídeos foram inseridos por eletroporação em *E. coli* DH5 α eletrocompetentes.

4.11 Transformação de *E. coli* por eletroporação

As transformações das cepas foram realizadas através de eletroporação em aparelho eletroporador *MicroPulser™* (Bio-Rad), utilizando o protocolo descrito em seguida.

Cubetas para eletroporação de 0,2 cm (Bio-Rad) e alíquotas de células eletrocompetentes foram mantidas em gelo por cerca de 5 min. Posteriormente, 3 a 5 μ L de produtos de ligação de fragmentos em vetores, ou plasmídeos purificados,

foram incubados com 50 μ L de células eletrocompetentes e mantidos em gelo por mais 2 min. Em seguida, as células foram transferidas para a cubeta e a eletroporação realizada utilizando o programa EC2 do aparelho. Imediatamente após a eletroporação, foram acrescentados à cubeta 500 μ L de meio LB líquido sem antibiótico e as células eletroporadas foram transferidas para um microtubo de 1,5 mL. A cultura foi incubada a 37°C por 1 h sob agitação de 200 rpm.

Finalizada a incubação, as células foram plaqueadas em meio seletivo LB-ágar 1,5% contendo o antibiótico para o qual o vetor inserido e/ou a cepa conferem resistência. As placas foram incubadas a 37°C por 16 h. Os antibióticos utilizados, bem como suas respectivas concentrações de estoque e de uso são exibidos na Tabela 6.

Tabela 6: Antibióticos utilizados para os respectivos plasmídeos ou cepas, e concentrações de uso.

Antibiótico	Plasmídeo / Cepa	Concentração estoque	Concentração no meio de cultura
Ampicilina	pMCSG7, pMCSG19C, pGEM®-T Easy	100 mg/mL em água	100 μ g/mL
Kanamicina	pET28a-TEV	50 mg/mL em água	50 μ g/mL
Gentamicina	ArticExpress™	40 mg/mL em água	20 μ g/mL

4.12 PCR de colônias

Após a transformação por eletroporação e crescimento de colônias nas placas com meio seletivo, a fim de selecionar clones positivos para os genes de interesse, foram realizadas PCR de colônias utilizando o mesmo protocolo de amplificação dos respectivos genes (Tabela 5). Na PCR de colônia não se coloca amostra de DNA. Em seu lugar é colocada uma pequena amostra da colônia que se deseja analisar. As amostras da PCR de colônia são, posteriormente, analisadas através de eletroforese em gel de agarose. Caso o clone seja positivo, haverá amplificação do gene de interesse; dessa forma a colônia é selecionada para ser crescida e ter o plasmídeo extraído.

4.13 Extração de DNA plasmidial

Clones positivos ou bactérias contendo os plasmídeos de expressão foram cultivados por 16 h a 37°C sob agitação de 200 rpm, em 5~10 mL de meio seletivo LB líquido com antibiótico específico (Tabela 6) para posterior extração dos plasmídeos. As extrações foram realizadas com kits de purificação AxyPrep™ Plasmid Miniprep (Axygen Biosciences), illustra™ plasmidPrep Mini Spin Kit (GE Healthcare), QIAprep® Spin Miniprep Kit (QIAGEN), ou Wizard® Plus Minipreps DNA Purification System (Promega). Em todos os casos foi seguido o protocolo do fabricante.

Todos os kits possuem um procedimento semelhante, que consiste no crescimento das colônias, centrifugação das células e descarte do sobrenadante, ressuspensão das células e posterior lise alcalina. A lise faz extravasar o conteúdo celular juntamente com os plasmídeos. Após a lise, neutraliza-se o pH da solução com uma base fraca, comumente acetato de potássio. A amostra é então centrifugada e a fase aquosa aplicada na coluna do kit, que possui a capacidade de reter ácidos nucleicos. Após uma ou mais etapas de lavagem para remover contaminantes, o DNA é eluído da coluna e sua quantidade e pureza analisada em gel de agarose ou aparelho espectrofotométrico (Nanodrop).

4.14 Digestões enzimáticas e ligação de insertos a vetores de expressão

Os clones positivos para os genes *nahE* e *nahK* ligados no vetor de clonagem pGEM-T Easy, e colônias contendo o vetor pET28a-TEV, foram crescidos e tiveram seus plasmídeos extraídos e purificados. A próxima etapa consiste na extração do gene de seu vetor de clonagem e posterior ligação do gene no vetor de expressão.

Em função da sequência dos iniciadores direto e reverso, os genes ligados ao vetor pGEM-T Easy estão flanqueados por sequências reconhecidas por enzimas de restrição compatíveis com o sítio múltiplo de clonagem do vetor de expressão pET28a-TEV. Os sítios de reconhecimento são para *NdeI* e *HindIII* no gene *nahE*, e

NheI e *HindIII* no gene *nahK*. Para obtenção dos genes e do vetor de expressão com extremidades coesivas e compatíveis para reação de ligação, os vetores de clonagem extraídos dos clones positivos, assim como o vetor de expressão, foram submetidos à reação de dupla digestão com as respectivas enzimas de restrição.

As enzimas utilizadas são do fabricante Fermentas. Os níveis de atividade das enzimas em cada tampão estão listados na Tabela 7. Conforme indicação do fabricante, a reação de dupla digestão utilizando as enzimas *NdeI* e *HindIII* ocorreu utilizando o tampão Tango 2X, no qual ambas enzimas apresentam 50~100% de atividade. Para a reação de dupla digestão utilizando as enzimas *NheI* e *HindIII*, o tampão utilizado foi Tango 1X, no qual a enzima *HindIII* apresenta 50~100% de atividade, e a enzima *NheI* apresenta 100% de atividade.

Tabela 7: Enzimas de restrição utilizadas e porcentagem de atividade em cada tampão.

Enzima	Porcentagem de atividade enzimática em cada tampão			
	Orange	Red	Tango 1X	Tango 2X
<i>NdeI</i>	100%	50-100%	0-20%	50-100%
<i>HindIII</i>	0-20%	100%	50-100%	50-100%
<i>NheI</i>	0-20%	0-20%	100%	0-20%

A digestão dos vetores foi testada com cada enzima separadamente e em reação de dupla digestão, incubadas a 37°C por 16 h. A quantidade DNA em cada teste foi de 500 ng, a quantidade de enzima foi de 5 U/μg de DNA, conforme recomendação do fabricante. Os produtos de digestão foram analisados em gel de agarose. Posteriormente, com a padronização da dupla digestão, 2 a 3 μg de cada vetor foram digeridos para dar sequência no processo de clonagem.

Sendo realizada a digestão dos vetores de clonagem pGEM-T Easy com consequente liberação dos insertos contendo os genes *nahE* e *nahK*, e a linearização do vetor de expressão pET28a-TEV, os produtos de digestão foram submetidos à eletroforese em gel de agarose, e as bandas de interesse foram purificadas do gel utilizando o kit ilustra GFX™ PCR DNA and Gel Band Purification (GE Healthcare). A amostra contendo os insertos e os plasmídeos purificados foram

dosados em espectrofotômetro (Nanodrop). Posteriormente, os insertos foram ligados aos vetores utilizando a enzima T4 DNA ligase (Promega), conforme recomendação do fabricante, e inseridos em *E. coli* DH5 α por eletroporação. Os transformantes obtidos foram submetidos a PCR de colônias e os clones positivos selecionados para extração dos plasmídeos. Os plasmídeos contendo os genes de interesse foram então inseridos por eletroporação em *E. coli* BL21(DE3) para a etapa de expressão das proteínas.

4.15 Clonagem Independente de Ligação

O sistema de clonagem independente de ligação foi desenvolvido para facilitar o processo de clonagem. Ele está presente nos vetores de expressão pMCSG7 e pMCSG19C (Protein Structure Initiative - Material Repository / Harvard). Utilizando esse sistema não há necessidade de uso de vetores intermediários de clonagem (como pGEM ou TOPO), sítios de enzimas de restrição nos produtos de PCR, ou T4 DNA ligase [66]. Eliminam-se então diversas etapas tornando mais rápido o processo de obter clones positivos já nos vetores de expressão.

As extremidades 5' dos iniciadores para esse sistema contêm 15 nucleotídeos adicionais carentes de dGMP (Tabela 3 e Figura 11). Como resultado, os produtos amplificados incluem nas extremidades 3' sequências complementares carentes de dCMP. Tais sequências podem ser removidas pela atividade exonucleásica 3'-5' da T4 DNA polimerase na presença de dCTP. Ao encontrar um nucleotídeo desoxicitosina, a enzima T4 DNA polimerase não removerá mais nucleotídeos, gerando assim, caudas de DNA simples fita nas extremidades 5' com 15 nucleotídeos de tamanho [67] (Figura 11).

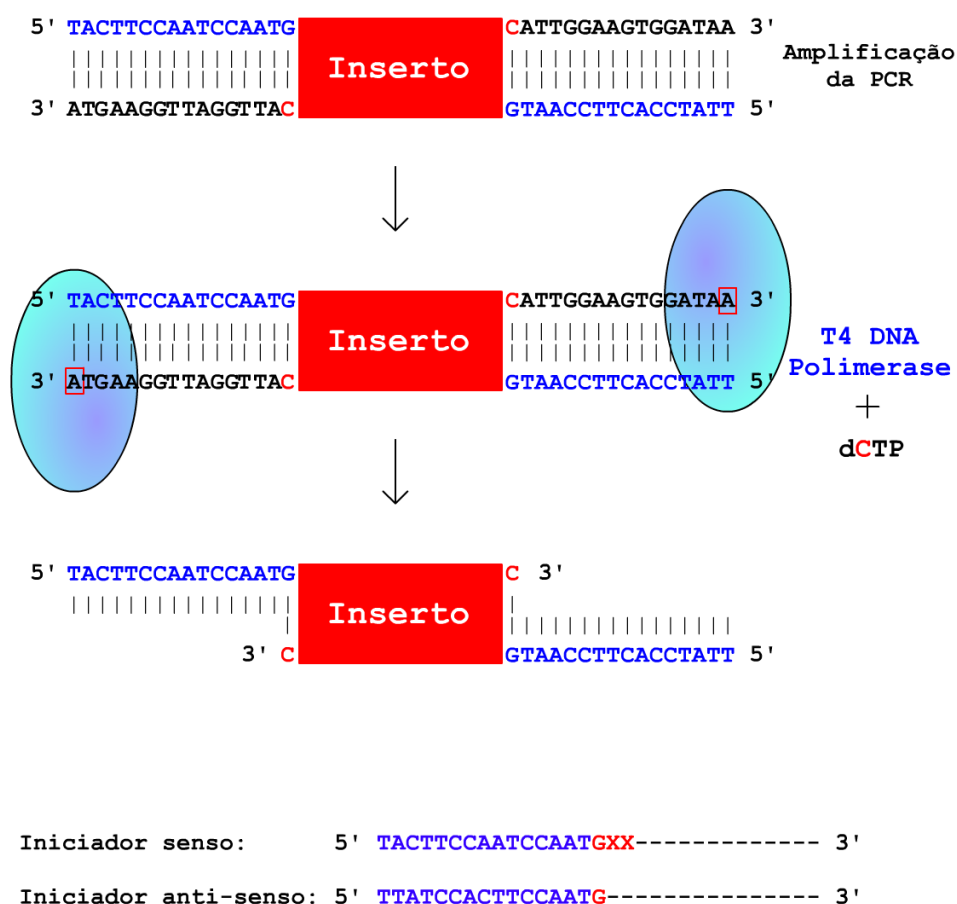


Figura 11: Esquema de tratamento do inserto no sistema LIC.

Quando os vetores de expressão são linearizados pela enzima de restrição SspI, passam a conter em suas extremidades 3', sequências complementares às sequências dos produtos amplificados. Da mesma maneira, ao serem tratados com T4 DNA polimerase na presença de dGTP, a enzima remove nucleotídeos até encontrar um desoxiguanina. Assim os vetores ficam com caudas de DNA simples fita de 15 nucleotídeos de tamanho em suas extremidades 5' (Figura 12).

A recombinação entre os produtos amplificados e os vetores tratados com T4 DNA polimerase ocorre mediada pelas caudas das extremidades 5', já que suas sequências são coesivas. Não ocorre recombinação entre vetores tratados em misturas carentes de produtos amplificados, nem o contrário, já que as sequências não são palíndromos. Não há necessidade do uso de T4 DNA ligase para uma transformação bacteriana eficiente, já que as interações de pontes de hidrogênio das

caudas de 15 nucleotídeos são suficientes para manter os insertos unidos aos vetores [68]. Uma vez inseridos nas bactérias, a própria maquinaria celular é capaz de ligar definitivamente o inserto ao vetor.

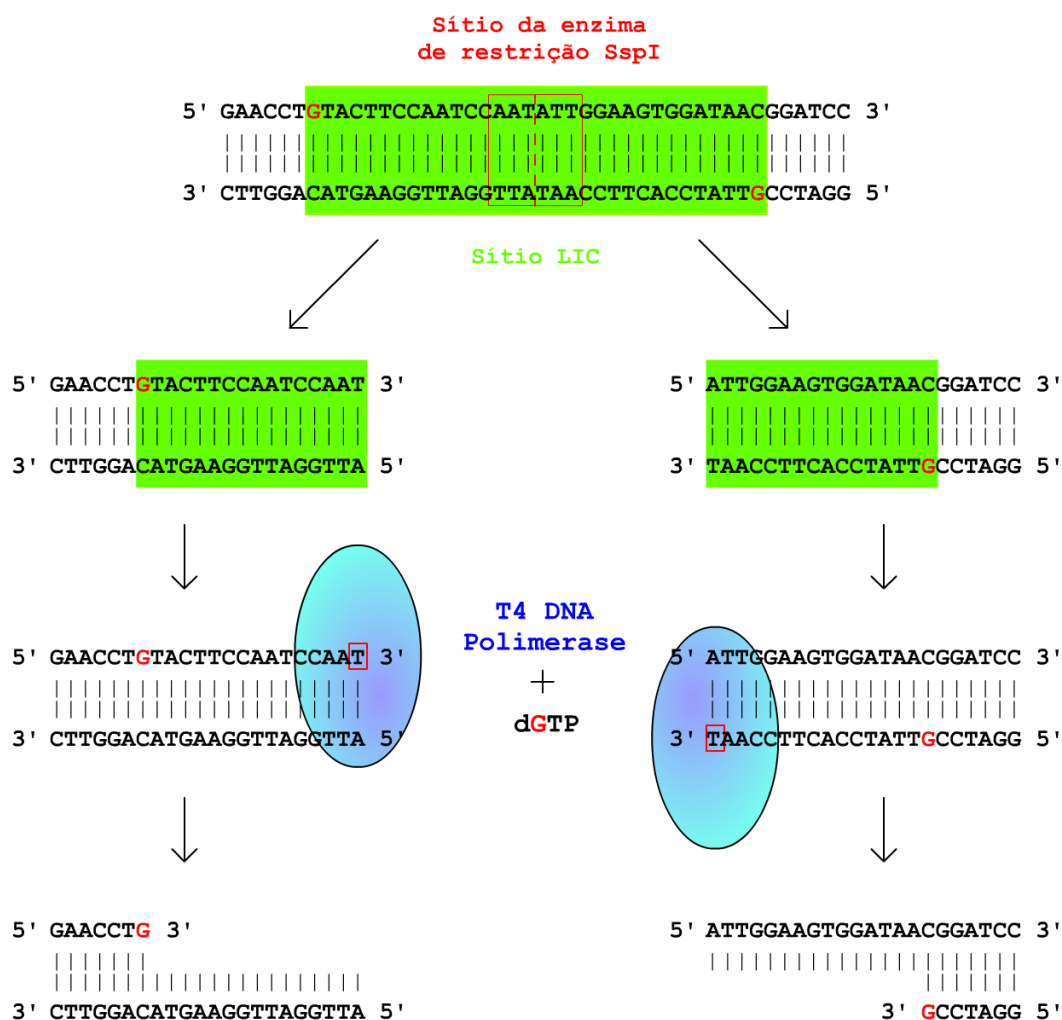


Figura 12: Esquema de tratamento do vetor no sistema LIC.

Três μg dos vetores pMCSG7 e pMCSG19C foram incubados com a enzima SspI (Fermentas) na quantidade de $5 \text{ U}/\mu\text{g}$, a 37°C por 16 h, conforme indicações do fabricante. Após a reação de digestão, a amostra foi analisada em gel de agarose e a banda equivalente ao vetor pMCSG7 linearizado (5286 pb), e ao vetor pMCSG19C linearizado (7468 pb) foi purificada utilizando o Kit illustra GFX™ PCR DNA and Gel Band Purification (GE Healthcare), e a quantidade recuperada de DNA foi dosada em espectrofotômetro Nanodrop. Os vetores linearizados foram submetidos ao tratamento com T4 DNA polimerase na presença de dGTP. O produto amplificado de

nahE obtido com os iniciadores *nahE-fw-LIC* e *nahE-rev-LIC*, após serem purificados utilizando o Kit *illustra GFX™ PCR DNA and Gel Band Purification* (GE Healthcare), e dosados em espectrofotômetro Nanodrop, também foram submetidos ao tratamento com T4 DNA polimerase na presença de dCTP. Os reagentes e as quantidades para a reação com T4 DNA polimerase estão representados na Tabela 8. O tratamento consiste em incubar as amostras a 11°C por 30 min para que ocorra a atividade exonucleásica da T4 DNA polimerase. Posteriormente, a enzima é inativada a 70°C por 10 min.

Após o tratamento, 20 µL da reação do produto amplificado foram misturados com 10 µL da reação de um dos vetores. A mistura foi deixada a temperatura ambiente por 30 min e depois aquecida a 75°C. Após alcançar tal temperatura, a amostra foi resfriada lentamente até temperatura ambiente, e 3 µL foram utilizados na transformação das células eletrocompetentes para expressão.

Tabela 8: Reagentes e quantidades utilizados no tratamento com T4 DNA polimerase.

Reagente	Concentração final	Volume
DNA	500 ng (vetor) 70 ng (inserto)	-
Tampão 5X	1X	8 µL
dGTP 100 mM (para vetor)	2,5 mM	1 µL
dCTP 100 mM (para inserto)	2,5 mM	1 µL
BSA 10 mg/mL	0,1 ng/mL	0,4 µL
T4 DNA polimerase 5 U/µL	0,5 U	0,1 µL
H2O milli-Q autoclavada	-	q.s.p. 40 µL

4.16 Sequenciamento de clones positivos

Os sequenciamentos de alguns dos plasmídeos recombinantes foram feitos pelo método didesoxiterminal [69] no núcleo de Análise do Genoma e Expressão Gênica (NAGE / ICB / UFMG), ou no Laboratório de Genética Bioquímica (LGB / ICB / UFMG), com auxílio do DYEnamic ET Dye Terminator Cycle Sequencing Kit for MegaBACE DNA Analysis (Amersham Biosciences). Para confirmação da integridade das sequências alvo no sequenciamento dos plasmídeos, foram utilizados os iniciadores T7 direto (5' TAA TAC GAC TCA CTA TAG GG 3') e T7 reverso (5' GCT AGT TAT TGC TCA GCG G 3'), T_m 50°C.

4.17 Expressão das proteínas recombinantes

O protocolo inicial das expressões das proteínas NahE nos vetores pET28a-TEV, pMSCG7 e pMCSG19C; e da proteína NahK no vetor pET28a-TEV foi feito como indicado a seguir.

Clones positivos foram preinoculados em meio LB líquido contendo o respectivo antibiótico de cada vetor de expressão, e, quando necessário, da cepa bacteriana (Tabela 6). O preinóculo foi incubado a 37°C por 16 h sob agitação de 200 rpm. Em seguida, a cultura foi utilizada como inóculo na proporção de 1:100 de cultura em meio LB líquido. As células do inóculo foram crescidas a 37°C sob agitação de 200 rpm, sendo seu crescimento monitorado em espectrofotômetro até a cultura alcançar uma leitura de $D.O._{600}$ entre 0,5 e 0,7. Nesse momento, foi adicionado o indutor de expressão IPTG para uma concentração final de 1 mM. A indução foi feita a 37°C por 3 h sob agitação de 200 rpm.

Para verificar a expressão da proteína recombinante, 1 mL do inóculo foi retirado imediatamente antes da indução por IPTG, centrifugado a 10.000 G por 5 minutos e as células ressuspendidas em tampão de amostra de proteína 4X. Após o término da expressão, a $D.O._{600}$ da cultura foi novamente medida, e uma quantidade equivalente de células foi coletada (quando comparada com a $D.O._{600}$ antes da

indução). A amostra foi centrifugada e as células ressuspendidas em tampão de amostra de proteína 4X. Posteriormente as amostras foram submetidas à SDS-PAGE para se verificar a presença de bandas de superexpressão na amostra retirada ao término da expressão.

4.18 Lise bacteriana e teste de solubilidade

Após a etapa de expressão, a cultura foi centrifugada a 10.000 G a 4°C por 10 min, e as células ressuspendidas em tampão de lise (fosfato de sódio 20 mM pH 7,4, NaCl 500 mM, imidazol 30 mM). O volume de tampão utilizado foi na proporção de 20 mL para cada litro de expressão, sendo que em volumes de expressão menores que 250 mL, sempre foram utilizados um mínimo de 5 mL de tampão de lise. Posteriormente, foi adicionada lisozima à suspensão numa concentração final de 100 µg/mL, e mantida em gelo por 30 min. Em seguida, o material foi congelado a -20°C e descongelado a temperatura ambiente por 3 vezes. Após os ciclos de congelamento, o lisado foi sonificado em aparelho sonicador Sonic Dismembrator 500 (Fisher Scientific) em amplitude de 30%, com 4 pulsos de 15 seg ligado e 1 min desligado. Em seguida, o lisado foi centrifugado a 17.000 G a 4°C por 10 min. O sobrenadante foi filtrado utilizando filtro de 0,45 µm.

A fração insolúvel, correspondente ao precipitado após o lisado ser centrifugado, foi ressuspendida no mesmo volume de água e uma amostra retirada para análise por SDS-PAGE. Uma amostra para análise por SDS-PAGE também foi retirada da fração solúvel, correspondente ao sobrenadante do lisado.

4.19 Otimização da expressão das proteínas recombinantes

Para otimizar a obtenção da proteína alvo em fração solúvel, alguns parâmetros foram alterados, como por exemplo, temperaturas menores de expressão (30°C, 22°C e 12°C), menor concentração do agente indutor IPTG (0,5 mM), diferentes cepas bacterianas (ArticExpress), e diferente tampão de lise (Tris-HCl 50 mM pH 7,4, 1% sacarose, 1% Tween-20, 1% glicerol, 30 mM imidazol). Além disso foi testada a clonagem em vetor de expressão com cauda de fusão MBP (vetor pMCSG19C) que também pode auxiliar na solubilidade [70].

Ao utilizar temperaturas mais baixas para a expressão, o preinóculo e inóculo foram crescidos a 37°C. Após o inóculo, ao atingir uma leitura de D.O.₆₀₀ entre 0,5 e 0,7, a cultura foi colocada, por 10 min, na temperatura desejada para a expressão, e em seguida, foi adicionado o indutor IPTG. A 30°C a indução foi feita por 5 h, a 22°C durante 16 h, e a 12°C por 24 h.

4.20 Reenovelamento de NahE insolúvel

Para o caso da insolubilidade de NahE, mesmo utilizando as estratégias para obtenção de expressão solúvel, foi seguido o protocolo de purificação de proteínas de corpos de inclusão [71] e o protocolo de reenovelamento [72] de proteínas, descritos a seguir.

A expressão de NahE insolúvel foi realizada a 37°C por 4 h sob agitação de 200 rpm, utilizando 1 mM de IPTG como indutor. Após a lise utilizando tampão fosfato de sódio 20 mM pH 7,4, NaCl 500 mM; e centrifugação da amostra, conforme descritos em 4.18, o precipitado foi ressuscitado em mesmo volume de tampão de lavagem (Tris-HCl 50 mM pH 8,0, 100 mM NaCl e 0,5% v/v Triton X-100). Após homogeneizar o precipitado, a amostra foi centrifugada a 10.000 G a 4°C por 10 min e o sobrenadante foi descartado. O precipitado foi novamente ressuscitado em tampão de lavagem e centrifugado a 10.000 G a 4°C por 10 min, sendo o sobrenadante descartado. O precipitado foi ressuscitado em H₂O milli-Q e

alíquotas de 100 µL foram coletadas em 8 microtubos. Após a coleta das alíquotas, a amostra maior foi centrifugada a 10.000 G a 4°C por 10 min e o precipitado armazenado a -20°C. As 8 alíquotas de 100 µL foram centrifugadas a 10.000 G a 4°C por 10 min e o sobrenadante descartado. O precipitado de cada microtubo foi ressuspensionado em tampão fosfato de sódio 20 mM pH 7,4, NaCl 100 mM, contendo concentrações crescentes de ureia (de 1 a 8 M). Após homogeneização, as alíquotas foram centrifugadas a 10.000 G a 4°C por 10 min. O sobrenadante foi coletado e tampão de amostra de proteína 4X foi adicionado em cada alíquota, que foram analisadas por SDS-PAGE. Após verificar em qual concentração de ureia NahE torna-se solúvel, a amostra maior foi ressuspensionada em tampão fosfato de sódio 20 mM pH 7,4, NaCl 100 mM, e ureia na concentração determinada.

A purificação de NahE foi realizada por cromatografia de afinidade ao níquel, conforme descrito no item 4.23, porém sob condições desnaturantes, utilizando ureia em todos os tampões na concentração determinada anteriormente. Após purificação, as frações que continham NahE foram unidas e diluídas em tampão de reenovelamento (fosfato de sódio 20 mM pH 7,4, NaCl 100 mM, L-arginina 500 mM) até que a concentração de ureia atingisse 2 M. A solução foi incubada a 4°C por 16 h e posteriormente dialisada contra 100/1 volumes de tampão de reenovelamento sem L-arginina a 4°C por 16h para remover ureia, L-arginina e imidazol. Após diálise, a amostra foi centrifugada a 10.000 G a 4°C por 10 min e o sobrenadante concentrado para ser purificado novamente em cromatografia de afinidade ao níquel e exclusão molecular.

4.21 Eletroforese em gel de poliacrilamida-SDS

As amostras das expressões proteicas, testes de lise, reenovelamento e frações de purificação foram analisadas através de SDS-PAGE.

Os géis foram preparados da mistura de N,N'-metileno-bis-acrilamida 0,8% e acrilamida 30% m/v. O gel de separação (12 ou 15%) foi feito adicionando-se a essa mistura Tris-HCl 1,5 M pH 8,8 e SDS 0,01% m/v. O gel foi polimerizado pela adição de APS a 10% em água destilada (0,5% v/v) e TEMED 0,05% v/v. O gel de

concentração (4%) foi preparado como descrito para o gel de separação, porém utilizando tampão Tris-HCl 0,5 M pH 6,8. As quantidades dos materiais estão exibidas na Tabela 9, e são suficientes para montagem de 2 géis de 0,75 mm de espessura utilizando-se o aparato de montagem para a cuba Mini-PROTEAN Tetra Cell (Bio-Rad).

Tabela 9: Materiais e quantidades utilizadas na preparação dos géis para SDS-PAGE.

Reagente	Gel de separação		Gel de concentração 4%
	12%	15%	
Água destilada	3,35 mL	2,29 mL	1,22 mL
Tris-HCl 1,5 M pH 8,8	2,5 mL	2,5 mL	-
Tris-HCl 0,5 M pH 6,8	-	-	0,5 mL
SDS 10%	100 µL	100 µL	20 µL
Bis-acrilamida 0,8%, acrilamida 30%	4 mL	5 mL	260 µL
APS 10%	100 µL	100 µL	20 µL
TEMED	10 µL	10 µL	2 µL

Para a corrida eletroforética foi utilizado tampão de eletrodo contendo Tris-HCl 25 mM pH 8,3, glicina 192 mM, e SDS 0,1%, em voltagem de 110 V até que as amostras entrassem no gel de concentração e alcançassem o gel de separação. A partir desse momento, a voltagem foi elevada para 150 V até o final da corrida. Foi utilizada a fonte FB300 (Fisher Scientific).

Para visualização das bandas, os géis foram colocados em solução corante por 2 horas (0,5% m/v Comassie Blue R-250, 10% ácido acético, 50% etanol), e depois deixados em solução descorante (25% etanol, 10% ácido acético) até o aparecimento das bandas proteicas e o descoramento das regiões do gel que não tinham proteínas. Os géis foram fotodocumentados utilizando câmera Vilber Lourmat.

4.22 Tampão de amostra e padrão para proteínas

Todas as amostras proteicas a serem analisadas por SDS-PAGE foram misturadas com tampão de amostra 4X (Tris-HCl 250 mM pH 6,8; SDS 0,8% m/v; azul de bromofenol 0,2% m/v; glicerol 45,5% v/v), e incubadas em banho-maria a 100°C por 5 min antes de serem aplicadas nas canaletas do gel.

Os padrões de peso molecular utilizados foram Precision Plus Protein™ All Blue Standards (Bio-Rad), que apresenta bandas proteicas de 10 a 250 kDa; Unstained Protein Molecular Weight Marker (Fermentas), que possui bandas proteicas de 14,4 a 116 kDa; ou BenchMark™ Protein Ladder (Invitrogen), que possui bandas entre 10 e 220 kDa.

4.23 Purificação das proteínas recombinantes

As proteínas recombinantes expressas com os vetores utilizados nesse trabalho possuem em seu amino-terminal uma cauda de 6 histidinas. Dessa forma, as proteínas solúveis foram purificadas inicialmente através de cromatografia de afinidade ao níquel em coluna HisTrap HP de 5 mL (GE Healthcare) utilizando aparelho ÄKTAprime plus (GE Healthcare).

A coluna foi previamente equilibrada com 5 vc com o mesmo tampão utilizado na lise das células (tampão A). Após a aplicação da fração solúvel do lisado, a coluna foi lavada com 5 vc de tampão A, ou até que nenhuma leitura de absorvância a 280 nm fosse observada. A eluição das proteínas que se ligaram à coluna foi feita utilizando um gradiente de concentração crescente com o tampão de lise contendo imidazol 500 mM (tampão B). Amostras recolhidas ao longo de todo o processo foram submetidas à análise por SDS-PAGE.

Para remoção do sal e do imidazol, frações provenientes da coluna de afinidade que continham a proteína NahK foram dialisadas contra 100/1 volumes de tampão Tris-HCl 50 mM pH 7,5, NaCl 50 mM a 4°C por 16 h em membrana de celulose.

Algumas amostras contendo NahK foram injetadas em coluna RESOURCE™ Q de 1 mL (GE Healthcare) previamente equilibrada com 5 vc de tampão fosfato de sódio 20 mM pH 7,4, NaCl 100 mM; ou Tris-HCL 50 mM pH 7,4, NaCl 50 mM. Após a aplicação a coluna foi lavada com 10 vc do mesmo tampão. A eluição foi feita em gradiente crescente de concentração de NaCl utilizando 20 vc de tampão fosfato de sódio 20 mM pH 7,4, NaCl 1 M; ou tampão Tris-HCL 50 mM pH 7,4, NaCl 1 M. Amostras recolhidas ao longo do processo foram analisadas por SDS-PAGE.

Após clivagem com protease TEV, a amostra de NahK foi purificada novamente por cromatografia de afinidade ao níquel utilizando Tris-HCl 50 mM pH 7,5 e NaCl 50 mM como tampão A; e Tris-HCl 50 mM pH 7,5, NaCl 50 mM e imidazol 500 mM como tampão B. Como a protease TEV contém em seu amino-terminal uma cauda de poli-histidina, ela fica retida na coluna juntamente com NahK não clivada e contaminantes provenientes da primeira etapa de purificação. Espera-se que a proteína NahK clivada, por não ter mais a cauda de afinidade, não interaja com a coluna. Amostras recolhidas ao longo de todo o processo foram submetidas à SDS-PAGE.

A amostra de NahK clivada por TEV e purificada pela cromatografia de afinidade foi submetida a uma última etapa de purificação utilizando cromatografia de exclusão molecular em coluna Superdex 200 10/300 GL (GE Healthcare). Foi utilizado para essa purificação o tampão Tris-HCl 50 mM pH 7,5, NaCl 50 mM. A coluna foi equilibrada com 2 vc de tampão e posteriormente a amostra foi injetada. Amostras recolhidas ao longo de todo o processo foram analisadas por SDS-PAGE. Frações dessa etapa de purificação que continham NahK foram unidas, concentradas e tiveram suas concentrações dosadas para posterior análise por DLS, CD e submetidas a ensaios de cristalização.

4.24 Determinação da concentração proteica

A dosagem das proteínas foi realizada através da medida da absorbância a 280 nm em espectrofotômetro modelo UV-160A UV Visible recording spectrophotometer (Shimadzu) ou em Nanodrop®.

Como são conhecidas as sequências de resíduos de aminoácidos das proteínas em estudo, é possível prever o coeficiente de extinção molar de cada uma e aplicar a seguinte fórmula de Lambert-Beer:

$$A = \epsilon * I * C$$

onde **A** é a absorvância a 280 nm; ϵ é o coeficiente de extinção molar em $M^{-1}cm^{-1}$; **I** o caminho óptico utilizado em cm; e **C** a concentração molar da proteína.

Os coeficientes de extinção para as proteínas fusionadas ou não com a cauda de histidina foram obtidos através do programa ProtParam [41], hospedado no servidor proteômico ExPASy - Expert Protein Analysis System [42].

4.25 Clivagem com protease TEV

A cauda de histidina presente na extremidade amino-terminal de NahK recombinante foi clivada utilizando a protease TEV. Essa protease é produzida no Laboratório BIOEST (ICB/UFMG) a partir do vetor pMHT238 Δ . Sua atividade é sítio-específica e reconhece a sequência de resíduos ENLYFQS ou ENLYFQG, presente entre a cauda de afinidade e o início da sequência da proteína recombinante. A clivagem ocorre após o resíduo de glutamina [73].

Testes de clivagem foram feitos a 25°C por 3, 6 e 16 horas, utilizando 5 e 20% m/m de TEV, em tampão Tris-HCl 50 mM pH 7,4 e NaCl 50 mM com a adição de DTT para concentração final de 1 mM. A remoção da cauda de histidina foi confirmada por análises de MALDI-TOF.

4.26 Espectrometria de Massa

O peso molecular da proteína NahK, antes e depois da clivagem pela protease TEV, foi determinado através de MALDI-TOF. Para isso, foram aplicadas em duplicata sobre placa MTP AnchorChip™ 600/384 (Bruker Daltonics) 0,5 μ L de

amostra e 0,5 μ L da matriz ácido sinapínico (10 mg/mL em ACN 30%, TFA 0,3%). As amostras foram analisadas utilizando o programa Flexcontrol 3.0. O equipamento foi calibrado utilizando padrão com proteínas de peso molecular conhecido. Os espectros obtidos foram analisados utilizando o programa Flex analysis 3.0.

4.27 Espalhamento Dinâmico de Luz

A determinação do estado de agregação e oligomerização das proteínas NahK e NahE em solução foi feita por ensaios de DLS. Os experimentos foram realizados no Laboratório de Espectroscopia e Calorimetria do Laboratório Nacional de Luz Sincrotron, utilizando o equipamento DynaPro™ Protein Solutions / Wyatt Technology Corporation.

As amostras proteicas foram mantidas a 0°C ou 4°C até o momento do ensaio e, logo antes da análise, foram centrifugadas a 15.000 G por 10 min para retirar qualquer material suspenso. Os ensaios foram realizados com temperatura constante de 18°C, em cubeta de quartzo com caminho óptico de 1 cm e 80 μ L de amostra. A concentração proteica das amostras analisadas foi de aproximadamente 1 mg/mL. Foram coletadas cerca de 200 a 250 medidas e os dados foram analisados no programa Dynamics V6.

Com essa análise, foram obtidos valores para a polidispersividade das proteínas em solução, o raio médio da molécula em nm e o peso molecular médio da molécula em kDa. Para verificar se havia alguma partícula agregada que não fosse a proteína alvo e que poderia interferir nos resultados, foi feito, antes do experimento com a amostra proteica, uma coleta de 20 medidas de tampão no qual a proteína estava diluída e da cubeta vazia. Nestes casos, os valores do raio médio e do peso molecular médio devem ser próximos de zero.

4.28 Dicroísmo Circular

Para caracterização da estrutura secundária de NahK foram realizados experimentos de CD no Laboratório de Espectroscopia e Calorimetria do Laboratório Nacional de Luz Síncrotron, utilizando o espectropolarímetro Jasco J-810 (Jasco corporation).

A concentração de NahK utilizada para o ensaio foi de 0,2 mg/mL. A proteína foi mantida a 0°C ou 4°C até o momento do ensaio e logo antes da análise a amostra proteica foi centrifugada a 15.000 G a 4°C por 10 min para retirar qualquer material suspenso.

Para a coleta de dados foram utilizados cubetas de quartzo com 0,1 cm de caminho óptico, 130 µL de amostra, comprimentos de onda de 195 a 260 nm e temperaturas de 20 a 90°C em intervalos de 5°C. As medidas foram realizadas de 0,2 em 0,2 nm, com 20 acumulações para cada medida.

Antes dos experimentos com as amostras proteicas, foram feitas medidas com a cubeta vazia para verificar se não havia proteína grudada na superfície da cubeta e, também, foram realizados experimentos com o tampão Tris-HCl 5 mM e NaCl 5 mM, a 20°C, entre os comprimentos de onda de 195 nm a 260 nm, sendo feitas 10 acumulações.

Os espectros gerados correspondem à diferença entre a média das 20 acumulações realizadas para a proteína e as 10 acumulações realizadas para o tampão em que ela estava diluída.

Os dados, obtidos em termos de elipcidade (θ) expressa em mdeg, foram convertidos em elipcidade molar residual $[\theta]_{MR}$ segundo a fórmula:

$$[\theta]_{MR} = \theta / (10 * l * ((C * 1000 * N) / M))$$

onde $[\theta]_{MR}$ é a elipcidade molar residual a ser expressa em degrees.cm².dmol⁻¹, θ é a elipcidade obtida pelo espectropolarímetro em mdeg; l é o caminho óptico da cubeta em cm; C é a concentração da proteína em g/mL; N é o número de resíduos de aminoácidos; e M a massa molecular da proteína em kDa.

Para estimar a porcentagem de cada tipo de estrutura secundária presente nas amostras, a deconvolução dos dados foi feita pelo programa CDNN 2.0,

desenvolvido pelo Dr. Gerald Böhm, Institut für Biotechnologie, Martin-Luther-universität Halle-Wittengerg, Alemanha.

4.29 Ensaios de cristalização e difração de raios-X

Os ensaios de cristalização com a proteína NahK foram realizados segundo a técnica de difusão de vapor em gota suspensa utilizando soluções do Kit Crystal Screen I e II (Hampton Research). Foram utilizadas nos testes placas de cristalização Limbro Box de 24 poços (Hampton Research), lamínulas silanizadas e graxa de silicone.

Um mililitro de cada condição de cristalização dos kits foi colocado em cada poço das placas (96 condições no total). Sobre uma lamínula silanizada foram depositadas duas gotas de NahK clivada a 12 mg/mL, sendo uma de 1 μ L e outra de 2 μ L; e uma gota de 1 μ L de NahK não clivada a 12 mg/mL. Posteriormente, 1 μ L da solução da condição de cristalização do poço foi adicionado e misturado a cada gota de proteína. A lamínula contendo as três gotas foi virada e colocada sobre o poço com a condição de cristalização correspondente, de forma que as gotas ficassem suspensas sobre o poço. O fechamento hermético do poço foi feito utilizando graxa de silicone. As placas de cristalização foram incubadas a 18°C. O monitoramento dos resultados em cada condição de cristalização ocorreu diariamente.

Lentamente, por difusão de vapor, a concentração de precipitante das gotas, inicialmente menor, vai se igualando à concentração do poço. Devido ao aumento de concentração, a gota pode chegar lentamente à condição de supersaturação, na qual a proteína não é mais solúvel e começa a precipitar. Caso ocorra precipitação, ela pode ser de forma desordenada, formando precipitados amorfos ou mudanças para fase oleosa (não desejáveis); ou acontecer de forma sistemática, formando núcleos cristalinos que correspondem as proteínas em uma estrutura organizada. Dependendo da condição de cristalização e das particularidades da macromolécula, após a nucleação, mais moléculas de proteína podem agregar-se aos núcleos de forma ordenada fazendo com que o cristal cresça.

Condições promissoras, nas quais formaram nucleações ou cresceram pequenos cristais, foram amplificadas de forma a prover uma maior gama de condições otimizadas para cristalização.

Os cristais de NaHCO₃ foram transportados até o Laboratório Nacional de Luz Sincrotron para serem submetidos à Difração de Raios-X. Para a coleta de dados, o cristal é capturado da gota de cristalização através de um *loop* de nylon e incubado alguns segundos em uma gota de 1 µL de solução crio-protetora formada pela mistura da solução da condição de cristalização respectiva e etileno-glicol na concentração final de 10% v/v. Posteriormente, o cristal é montado na cabeça de um goniômetro e centralizado. O cristal, a partir desse momento, fica resfriado à -173°C sob constante fluxo de nitrogênio. Inicialmente o detector é posicionado a uma distância de 100 mm do cristal e a qualidade de difração é testada recolhendo duas imagens a 0 e 90° de oscilação.

Para a aquisição de cada imagem, o cristal é submetido ao feixe de raios-X durante 30 a 60 segundos, sendo rotacionado 1° ao longo de cada exposição. As imagens são analisadas no programa iMosflm [74], que localiza as reflexões do padrão de difração e fornece estimativas prévias sobre parâmetros de célula unitária, grupo espacial do cristal e mosaicidade. Além disso, o programa também apresenta estratégias para a coleta de dados sugerindo ângulos preferenciais para o início e o término da coleta.

5. RESULTADOS E DISCUSSÃO

5.1 Amplificação dos genes *nahE* e *nahK*

A amplificação do gene *nahE* (aproximadamente 1000 pb) utilizando ambos pares de iniciadores (sistema pET e LIC) ocorreu mais eficientemente com uma T_m de 45°C, apesar do programa utilizado para o desenho dos iniciadores indicar uma temperatura média de 51,6°C. A concentração final de DNA na reação de PCR que atingiu melhores resultados na amplificação de *nahE* foi cerca de 3 ng/μL, como exibido na Figura 13.

A amplificação do gene *nahK* (aproximadamente 800 pb) ocorreu de forma mais eficiente e com menor quantidade de amplificações inespecíficas utilizando uma T_m de 58°C. Assim como em *nahE*, a concentração de DNA para amplificação na amostra que melhor obteve resultado foi de aproximadamente 3 ng/μL, exibida na Figura 14.

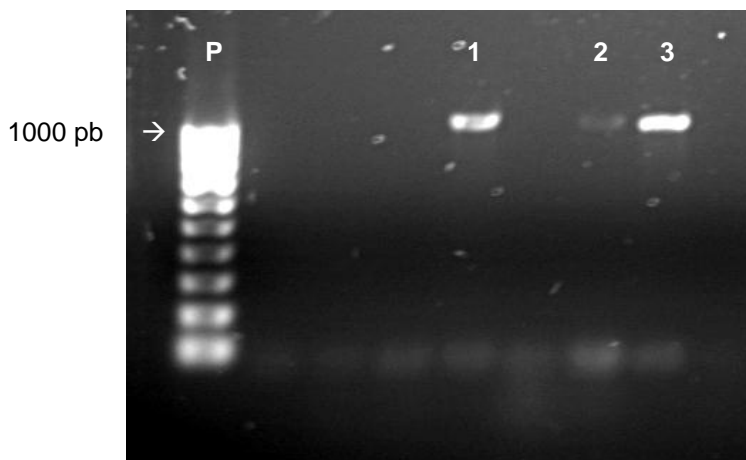


Figura 13: Amplificação do gene *nahE* de *P. putida* G7.

P - padrão 100 pb EZ Load Molecular Ruler; **1**, **2** e **3** - amplificação do gene contendo, respectivamente, DNA para amplificação numa concentração de 4, 15 e 3 ng/μL.

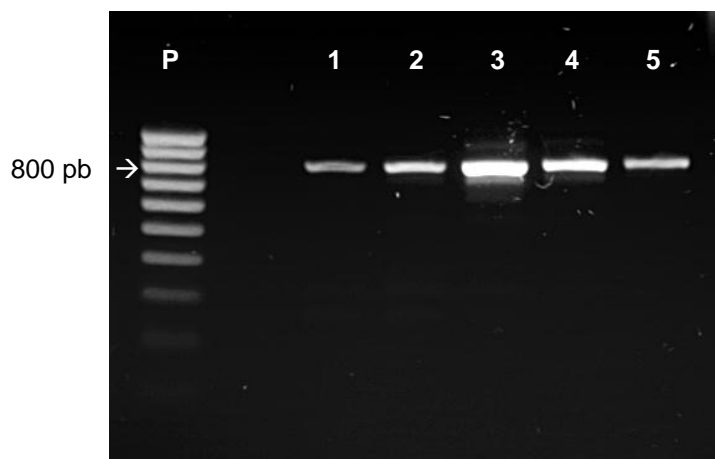


Figura 14: Amplificação do gene *nahK* de *P. putida* G7.

P - padrão 100 pb EZ Load Molecular Ruler; **1 a 5** - amplificações do gene contendo, respectivamente, DNA para amplificação numa concentração de 30, 15, 3, 0,6 e 0,3 ng/μL.

5.2 Subclonagem em vetor pGEM-T Easy

Os produtos das ligações dos fragmentos amplificados com o vetor pGEM-T Easy foram inseridos em *E. coli* DH5α por eletroporação. Através de PCR de colônia foi identificado um clone positivo de *nahE*, e diversos clones positivos foram obtidos para *nahK*. Um clone de cada construção foi selecionado para extração do plasmídeo para a etapa de digestão.

5.3 Digestão dos vetores pGEM-T Easy e pET28a-TEV

As amostras do vetor de clonagem contendo os insertos foram completamente linearizadas utilizando as quantidades de enzima recomendadas pelo fabricante (5 U/μg, canaletas 3 e 4, Figura 15). Entretanto, essa quantidade de enzima não foi capaz de digerir completamente o vetor em reação de dupla digestão (canaleta 5, Figura 15). Nesse caso, o dobro da quantidade recomendada de enzima (10 U/μg)

foi necessária para completa digestão do vetor e liberação do inserto (canaleta 6, Figura 15).

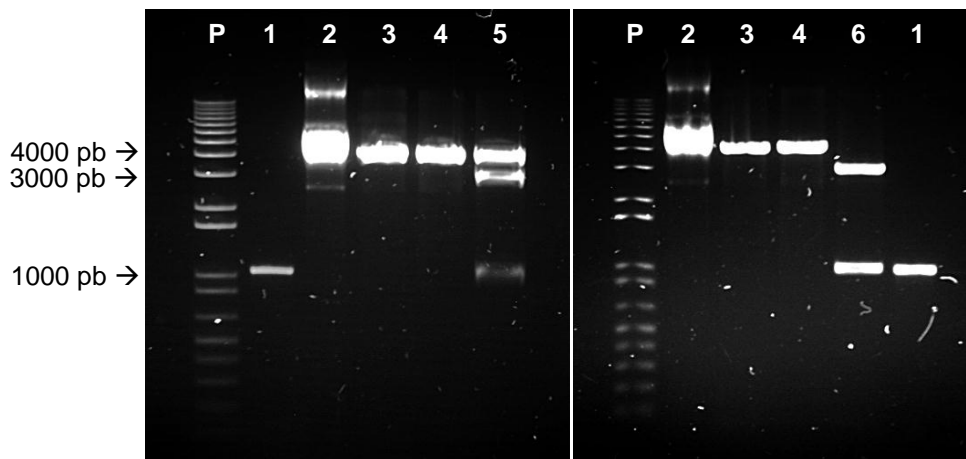


Figura 15: Digestão do vetor pGEM-T Easy para liberação do fragmento contendo *nahE*.

P - 1 kb plus DNA Ladder; **1** - amplificação do gene *nahE* por PCR; **2** - vetor pGEM-T Easy fechado contendo o inserto de *nahE*; **3** - digestão do vetor utilizando NdeI 5 U/μg; **4** - digestão do vetor utilizando HindIII 5 U/μg; **5** - dupla digestão utilizando 5 U/μg de cada enzima; **6** - dupla digestão utilizando 10 U/μg de cada enzima.

Com base nos resultados da dupla digestão utilizando as enzimas NdeI e HindIII para liberação de *nahE*, a dupla digestão do vetor pGEM-T Easy contendo o inserto de *nahK* foi realizada eficientemente utilizando a quantidade de 10 U/μg das enzimas NheI e HindIII (dado não mostrado).

A dupla digestão do vetor pET28a-TEV (aproximadamente 5300 pb) para receber o inserto de *nahE* não foi eficiente utilizando a quantidade das enzimas NdeI e HindIII recomendada pelo fabricante (canaleta 2, Figura 16). O dobro da quantidade também não foi capaz de digerir toda a amostra (canaleta 3, Figura 16), sendo necessária uma quantidade de 20 U/μg das enzimas para completa digestão desse vetor (canaleta 4, Figura 16). A dupla digestão do vetor para receber o inserto de *nahK* foi eficiente utilizando 20 U/μg das enzimas NheI e HindIII.

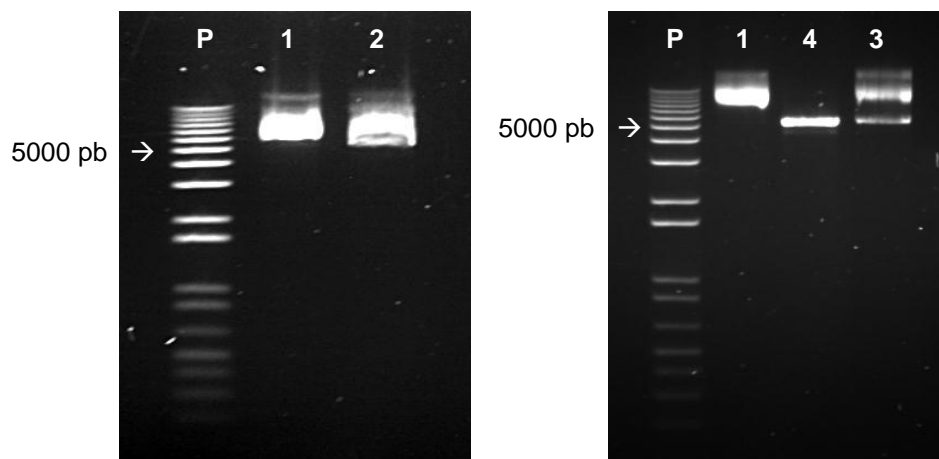


Figura 16: Dupla digestão do vetor de expressão pET28a-TEV.

P - 1 kb plus DNA Ladder; **1** - vetor fechado; **2** - dupla digestão utilizando 5 U/μg das enzimas NdeI e HindIII; **3** - dupla digestão utilizando 10 U/μg das enzimas; **4** - dupla digestão utilizando 20 U/μg.

Os insertos purificados do gel foram ligados aos respectivos vetores pET28a-TEV digeridos e purificados, e inseridos em *E. coli* DH5α por eletroporação.

5.4 Clonagem de *nahE* e *nahK* no vetor de expressão pET28a-TEV

Através de PCR de colônia foram identificados dois clones positivos para *nahE* no vetor pet28a-TEV, denominados colônia 20 e 22, cujo sequenciamento, concluído durante as etapas de expressão, está representado na Figura 17. O clone 20 apresenta uma mutação de desoxiadenina para desoxitimidina na base 318, fazendo com que o resíduo de aminoácido 107 seja trocado de isoleucina para fenilalanina. O clone 22 apresenta uma mutação de desoxiguanidina para desoxiadenina na base 200, trocando o resíduo de aminoácido nativo de glicina para aspartato. O trabalho continuou com o clone 20. Uma busca utilizando a ferramenta BLAST [47] identificou uma proteína chamada Dihidrodipicolinato Sintetase de *Burkholderia* sp., pertencente ao mesmo grupo de proteínas que NahE, a família das dihidrodipicolinato sintetases. Sua identidade com NahE é de 66%, e outros 12% dos resíduos possuem substituições equivalentes, indicando a grande similaridade na sequência de aminoácidos entre essas proteínas. Em Dihidrodipicolinato Sintetase, o resíduo

correspondente à posição isoleucina 107 de NahE é uma fenilalanina. Ambas proteínas são pertencentes à superfamília TIM-fosfatase, que possui proteínas de enovelamento característico e bastante conservado [46].

```

nahE      ATGTTGAATAAAGTTATTTAAAACCACGCGTCTTACCCTGAAGATATCAACGGTGCCTGGACTATAATGCCACACCCGTC 80
Colônia 20 ATGTTGAATAAAGTTATTTAAAACCACGCGTCTTACCCTGAAGATATCAACGGTGCCTGGACTATAATGCCACACCCGTC 80
Colônia 22 ATGTTGAATAAAGTTATTTAAAACCACGCGTCTTACCCTGAAGATATCAACGGTGCCTGGACTATAATGCCACACCCGTC 80
*****

nahE      GACGCTGATGCTTCTGATTGGCGCAGCACTAACACTGTGGACTTAGACGAGACTGCCCGCATAGTTGAAGAGCTGATTG 160
Colônia 20 GACGCTGATGCTTCTGATTGGCGCAGCACTAACACTGTGGACTTAGACGAGACTGCCCGCATAGTTGAAGAGCTGATTG 160
Colônia 22 GACGCTGATGCTTCTGATTGGCGCAGCACTAACACTGTGGACTTAGACGAGACTGCCCGCATAGTTGAAGAGCTGATTG 160
*****

nahE      CAGCTGGTGTCAACGGTATTTAAGTATGGGTACTTTTGGTGTGAGTGCACCACGTTGACCTGGGAGGAGAAACGTGATTAT 240
Colônia 20 CAGCTGGTGTCAACGGTATTTAAGTATGGGTACTTTTGGTGTGAGTGCACCACGTTGACCTGGGAGGAGAAACGTGATTAT 240
Colônia 22 CAGCTGGTGTCAACGGTATTTAAGTATGGGTACTTTTGGTGTGAGTGCACCACGTTGACCTGGGAGGAGAAACGTGATTAT 240
*****

nahE      GTTTCGACGGTTGTGCAACCATTTCGTGGTTCGTGTGCCTTATTCTGTGGCAGCAGCGCCCTGAATACCCGAGAAGTCATC 320
Colônia 20 GTTTCGACGGTTGTGCAACCATTTCGTGGTTCGTGTGCCTTATTCTGTGGCAGCAGCGCCCTGAATACCCGAGAAGTCATC 320
Colônia 22 GTTTCGACGGTTGTGCAACCATTTCGTGGTTCGTGTGCCTTATTCTGTGGCAGCAGCGCCCTGAATACCCGAGAAGTCATC 320
*****

nahE      CGCCAGACCCGAGAGCTTATCGATATTGGCGCTAACGGCACCATGCTAGGCGTGCCTGATGTTGGTGAAGATGGACCTGCC 400
Colônia 20 CGCCAGACCCGAGAGCTTATCGATATTGGCGCTAACGGCACCATGCTAGGCGTGCCTGATGTTGGTGAAGATGGACCTGCC 400
Colônia 22 CGCCAGACCCGAGAGCTTATCGATATTGGCGCTAACGGCACCATGCTAGGCGTGCCTGATGTTGGTGAAGATGGACCTGCC 400
*****

nahE      CACAGCGGTCCAGTTCTATCGTGTGTTGCAGGCGCGGTACCCGAGGCTGCCATTGCGATTTACGCCAACCCCGAAGCAT 480
Colônia 20 CACAGCGGTCCAGTTCTATCGTGTGTTGCAGGCGCGGTACCCGAGGCTGCCATTGCGATTTACGCCAACCCCGAAGCAT 480
Colônia 22 CACAGCGGTCCAGTTCTATCGTGTGTTGCAGGCGCGGTACCCGAGGCTGCCATTGCGATTTACGCCAACCCCGAAGCAT 480
*****

nahE      TCAAATTCGACTTCCCTCGCCCATTTTGGGCAGAGATGTCCAAAATTCCTCAGGTAGTGACTGCCAAGTATCTAGGCATC 560
Colônia 20 TCAAATTCGACTTCCCTCGCCCATTTTGGGCAGAGATGTCCAAAATTCCTCAGGTAGTGACTGCCAAGTATCTAGGCATC 560
Colônia 22 TCAAATTCGACTTCCCTCGCCCATTTTGGGCAGAGATGTCCAAAATTCCTCAGGTAGTGACTGCCAAGTATCTAGGCATC 560
*****

nahE      GGAATGCTTGACTTGGACCTGAAATTGGCGCCTAACATCCGCTTCCCTCCACACGAGGACGACTATTACGCGGCCGACG 640
Colônia 20 GGAATGCTTGACTTGGACCTGAAATTGGCGCCTAACATCCGCTTCCCTCCACACGAGGACGACTATTACGCGGCCGACG 640
Colônia 22 GGAATGCTTGACTTGGACCTGAAATTGGCGCCTAACATCCGCTTCCCTCCACACGAGGACGACTATTACGCGGCCGACG 640
*****

nahE      CATCAATCCCAGCGCATAACCGCGTTCCTGGTCAAGCGGGCCATGTGCGGCCGGCTACCCTATCATGTTGCGTGATG 720
Colônia 20 CATCAATCCCAGCGCATAACCGCGTTCCTGGTCAAGCGGGCCATGTGCGGCCGGCTACCCTATCATGTTGCGTGATG 720
Colônia 22 CATCAATCCCAGCGCATAACCGCGTTCCTGGTCAAGCGGGCCATGTGCGGCCGGCTACCCTATCATGTTGCGTGATG 720
*****

nahE      AAGTGGAGCGGGCCAAGAGTACCAGTGGATCAAGGCCAAAGCCATCTCCGATGATATGCGTGACGCCGATTCGACA 800
Colônia 20 AAGTGGAGCGGGCCAAGAGTACCAGTGGATCAAGGCCAAAGCCATCTCCGATGATATGCGTGACGCCGATTCGACA 800
Colônia 22 AAGTGGAGCGGGCCAAGAGTACCAGTGGATCAAGGCCAAAGCCATCTCCGATGATATGCGTGACGCCGATTCGACA 800
*****

nahE      TTGTTTCCGCGTGGCGACTTTTCGGAGTCTCGAAGTATAACATCGGGCTTGAAAAGGCACGGATGGACCGGGCTGGTTG 880
Colônia 20 TTGTTTCCGCGTGGCGACTTTTCGGAGTCTCGAAGTATAACATCGGGCTTGAAAAGGCACGGATGGACCGGGCTGGTTG 880
Colônia 22 TTGTTTCCGCGTGGCGACTTTTCGGAGTCTCGAAGTATAACATCGGGCTTGAAAAGGCACGGATGGACCGGGCTGGTTG 880
*****

nahE      GTCCTCCCTACAATCTTGTTCAGAAGATTACCTCGTGGTGCACAGAAATCAGGCAAGGCGTGGGCCGCGCTGCACGCT 960
Colônia 20 GTCCTCCCTACAATCTTGTTCAGAAGATTACCTCGTGGTGCACAGAAATCAGGCAAGGCGTGGGCCGCGCTGCACGCT 960
Colônia 22 GTCCTCCCTACAATCTTGTTCAGAAGATTACCTCGTGGTGCACAGAAATCAGGCAAGGCGTGGGCCGCGCTGCACGCT 960
*****

nahE      AAATACAGTAAATAA 995
Colônia 20 AAATACAGTAAATAA 995
Colônia 22 AAATACAGTAAATAA 995
*****

```

Figura 17: Sequenciamento dos clones 20 e 22 de *nahE*.

Vários clones para *nahK* em pET28a-TEV foram identificados através de PCR de colônia. Foram tentados diversos sequenciamentos do clone selecionado para expressão, e mesmo utilizando o iniciador reverso nahK-HindIII-R e iniciador reverso T7, não foi possível concluir o sequenciamento da região gênica correspondente ao carboxi-terminal da proteína. O sequenciamento, que cobriu cerca de 59% do gene *nahK* a partir da região correspondente ao amino-terminal de NahK, está representado na Figura 18.

```

nahK      ATGGGCCATCATCATCATCACGAAAACCTGTATTTTCAGGGCCATATGGCTAGCATGAACCGTACCCTGACCCGCGA 80
Colônia 5 ATGGGCCATCATCATCATCATCACGAAAACCTGTATTTTCAGGGCCATATGGCTAGCATGAACCGTACCCTGACCCGCGA 80
*****

NahK      ACAGGTGCTGGCCCTGGCCGAGCACATTGAAAACGCCGAGCTGAATGTCCACGACATCGGCAAGGTGACCAACGACTTCC 160
Colônia 5 ACAGGTGCTGGCCCTGGCCGAGCACATTGAAAACGCCGAGCTGAATGTCCACGACATCGGCAAGGTGACCAACGACTTCC 160
*****

NahK      CCGAAATGACCTTCGCCGACGCCTACGACGTGCAGTGGGAAATCCGTCCGCGCAAGGAGGCTCGCGCAACAAGATCGTC 240
Colônia 5 CCGAAATGACCTTCGCCGACGCCTACGACGTGCAGTGGGAAATCCGTCCGCGCAAGGAGGCTCGCGCAACAAGATCGTC 240
*****

NahK      GGCCTGAAGATGGGCCTGACCTCCTGGGCGAAGATGGCGCAGATGGGCGTGGAACGCCGATCTACGGCTTCCTCGCCGA 320
Colônia 5 GGCCTGAAGATGGGCCTGACCTCCTGGGCGAAGATGGCGCAGATGGGCGTGGAACGCCGATCTACGGCTTCCTCGCCGA 320
*****

NahK      CTACTTCAGCGTACCGGACGGCGCGTGGTGGATTGTTCCAAGCTGATCCACCCGAAGATCGAGGGCGAAATATCGGTGG 400
Colônia 5 CTACTTCAGCGTACCGGACGGCGCGTGGTGGATTGTTCCAAGCTGATCCACCCGAAGATCGAGGGCGAAATATCGGTGG 400
*****

NahK      TGACCAAGGCGCCCTGCACGGCCCGGTTGCCATCTCGGCGACGTGATTGCGGCGATCGATTACGTGATTCCGACCGTC 480
Colônia 5 TGACCAAGGCGCCCTGCACGGCCCGGTTGCCATCTCGGCGACGTGATTGCGGCGATCGATTACGTGATTCCGACCGTC 480
*****

NahK      GAGGTAATCGACTCGGCTATGAGAACTTCAAGTTCGACCTGATCAGCGTGGTTGCCGACACGCCTCCTCGACCCGCTT 560
Colônia 5 GAGGTAATCGACTCGGCTAT----- 500
*****

NahK      CATCACC GGCGGCATGGCCAGCCTGGAGGAAGTGCACCTGCGCACCCCTTGGCGTGGTTCATGGAGAAGAATGGCGAAG 640
Colônia 5 -----

NahK      TGGTGGAACTCGGCGCTGGCGCGGCGTACTCGGCCATCCGCTGTCCAGCGTGGCGATGCTGGCCAACCTGCTGGCCGAG 720
Colônia 5 -----

NahK      CGCGGCGAACAACATCCCTGCCGGCACCTTCATCATGACCGCGGCATCACC GCGCGGTGCCGCGCGGCGACAA 800
Colônia 5 -----

NahK      CATCACC GTGCGTTACCAGGGTCTCGGCTCGGTATCCGCGCGCTTATCTGA 852
Colônia 5 -----

```

Figura 18: Sequenciamento do clone 5 de *nahK*.

5.5 Expressão de NahE em vetor pET28a-TEV

O vetor pET28a-TEV é um plasmídeo que foi modificado, a partir do vetor pET28a da Novagen, no Centro de Biologia Molecular e Estrutural (CeBiME) pelo grupo do Prof. Dr. Nilson Zanchin do Laboratório Nacional de Luz Síncrotron (LNLS), Campinas - SP. Proteínas expressas por esse vetor apresentam em seu amino-terminal uma cauda de histidina que auxilia na purificação da proteína por cromatografia de afinidade. Uma vez purificada, a cauda de histidina não é mais necessária, e para removê-la, há um sítio para clivagem reconhecido pela protease TEV entre a sequência dos resíduos da proteína e a cauda de histidina (no vetor pET28a, o sítio original de clivagem é específico para Trombina).

A proteína NahE expressa nesse vetor apresenta a cauda de histidina e o sítio para clivagem TEV. Sua massa estimada com a cauda de fusão amino-terminal passa de 36,63 kDa (proteína nativa), para 38,63 kDa. Utilizando a protease TEV para clivagem da cauda, sua massa estimada vai para 36,83 kDa, restando em seu amino-terminal dois resíduos de aminoácidos: uma glicina e uma histidina.

A proteína recombinante foi expressa em *E. coli* BL21. A indução foi feita pela adição de IPTG (concentração final de 1 mM) em meio seletivo LB incubado a 37°C por 4 h. Amostras foram coletadas a cada hora de expressão (Figura 19, A). Após o término da expressão, foi feita a lise em tampão Tris-HCl 50 mM pH 7,4, 1% sacarose, 1% Tween-20, 1% glicerol. Observa-se, após a indução, o aparecimento de uma banda com peso molecular aproximado do valor esperado para NahE. Análises das frações solúvel e insolúvel indicaram que a proteína ficou presente somente na fração insolúvel (Figura 19, B).

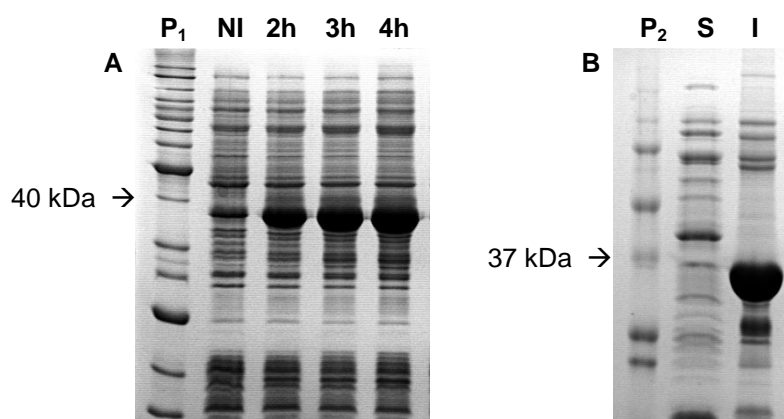


Figura 19: Expressão da proteína NahE no vetor pET28a-TEV em BL21 a 37°C, IPTG 1 mM.

P₁ - padrão de proteínas Invitrogen; **NI** - expressão não induzida; **P₂** - padrão de proteínas Bio-Rad; **S** - fração solúvel; **I** - fração insolúvel.

Como a proteína apresentou-se insolúvel em expressão a 37°C, foram feitas, utilizando o mesmo clone, sucessivas expressões em temperaturas mais baixas, de 30°C, 22°C e 12°C, utilizando IPTG na concentração de 0,5 mM (Figura 20). Em todos esses testes de expressão a proteína NahE apresentou-se somente em fração insolúvel.

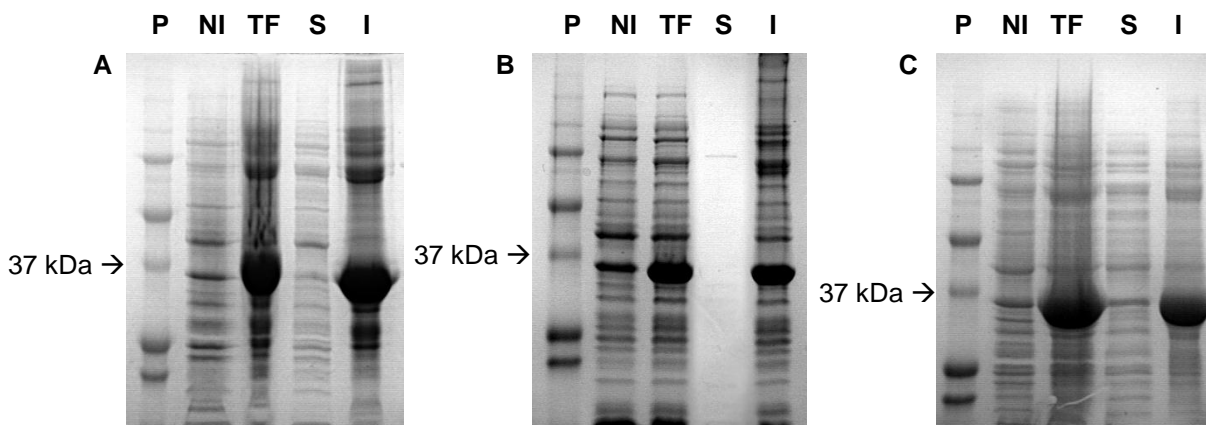


Figura 20: Expressões de NahE no vetor pET28a-TEV em BL21 a 30°C (A), 22°C (B) e 12°C (C); IPTG 0,5 mM.

P - padrão de proteínas Bio-Rad; **NI** - expressão não induzida; **TF** - tempo final de expressão; **S** - fração solúvel; **I** - fração insolúvel.

Devido à insolubilidade de NahE em expressões sob diversas temperaturas em *E. coli* BL21, o vetor pET28a-TEV contendo o gene *nahE* foi inserido em *E. coli* ArticExpress. A proteína foi expressa a 12°C por 24h utilizando IPTG na concentração de 0,5 mM. Após a lise utilizando tampão fosfato de sódio 20 mM pH 7,4, NaCl 500 mM, foi identificada uma banda de expressão de NahE na fração solúvel (Figura 21). Também foi identificada a banda de expressão da chaperona Cpn60 na fração solúvel. Tendo obtida a proteína NahE em fração solúvel, foi feita uma expressão em volume maior para prosseguir para a etapa de purificação.

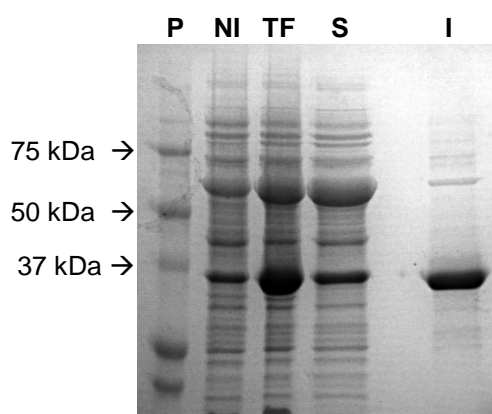


Figura 21: Expressão de NahE no vetor pET28a-TEV em ArticExpress a 12°C, IPTG 0,5 mM.

P - padrão de proteínas Bio-Rad; **NI** - expressão não induzida; **TF** - tempo final de expressão; **S** - fração solúvel; **I** - fração insolúvel.

5.6 Cromatografia de afinidade para NahE em pET28a-TEV

A fração solúvel da proteína NahE, expressa no vetor pET28a-TEV em ArticExpress a 12°C, foi submetida à cromatografia de afinidade em coluna de níquel-sepharose no sistema ÄKTAprime plus.

A coluna utilizada HisTrap HP consiste de uma matriz sólida de agarose contendo grupos quelantes imobilizados carregados com ions Ni^{2+} , os quais se ligarão às histidinas expostas na superfície de proteínas, e com maior afinidade, à cauda de histidina da proteína NahE. Após a amostra ter sido injetada na coluna, foi utilizado tampão contendo imidazol para a eluição das proteínas que se ligaram na coluna. O imidazol é um composto que é semelhante ao radical da histidina e atua

como competidor nas interações com os ions Ni^{2+} , favorecendo o desligamento das proteínas da coluna. O perfil cromatografico está representado na Figura 22 A. Algumas frações da purificação foram analisadas por SDS-PAGE (Figura 22 B).

Através da curva azul do cromatograma é possível observar que grande parte da amostra protéica não interagiu com a coluna, sendo coletadas nas frações entre 0 e 30 mL. A análise dessas frações por SDS-PAGE evidenciou que a maior parte das proteínas nativas de *E. coli* atravessaram a coluna sem se ligar a ela, podendo ser visualizadas bandas de diversos pesos moleculares na canaleta 1 da Figura 22 B.

Como a quantidade de NahE em fração solúvel foi pouca (Figura 21), foi utilizada na eluição um *step* de imidazol ao invés do gradiente, no qual o tampão muda rapidamente de 0 a 100% de imidazol 500 mM. Isso possibilita um aumento na concentração das proteínas eluídas. Um pequeno pico foi observado durante a eluição. Análises das três frações do pico por SDS-PAGE (canaletas 2, 3 e 4, Figura 22 B) mostraram que havia grande quantidade de Cpn60, NahE e alguns contaminantes. Apesar da quantidade de Cpn60 parecer maior que a de NahE durante eluição, devido seu baixo coeficiente de extinção molar (5.960) quando comparado com o coeficiente de NahE (38.637), a proteína Cpn60 pouco contribui para a absorvância a 280 nm. Dessa forma, os valores de absorvância no pico de eluição dependem em maior grau da quantidade de NahE.

As frações das canaletas 3 e 4 que continham tanto NahE quanto a chaperona Cpn60, foram dialisadas separadamente contra tampão fosfato de sódio 20 mM pH 7,4, NaCl 500 mM, para remoção do imidazol. Após a diálise, verificou-se um precipitado na fração correspondente a canaleta 4, que continha menor concentração da chaperona Cpn60. A amostra foi centrifugada e as frações solúvel e insolúvel analisadas através de SDS-PAGE (Figura 22 C). Notou-se uma grande precipitação da proteína NahE presente na fração insolúvel, enquanto a chaperona Cpn60 manteve-se na fração solúvel.

Em purificações realizadas no nosso laboratório com outras proteínas coexpressas com chaperonas, é possível observar que as chaperonas não interagem com a coluna, sendo eluídas juntamente com as demais proteínas de *E. coli*. Dessa forma, esse resultado nos leva a acreditar que a chaperona se mantém

ligada à NahE numa tentativa de auxiliar seu enovelamento, e está sendo co-purificada com a mesma.

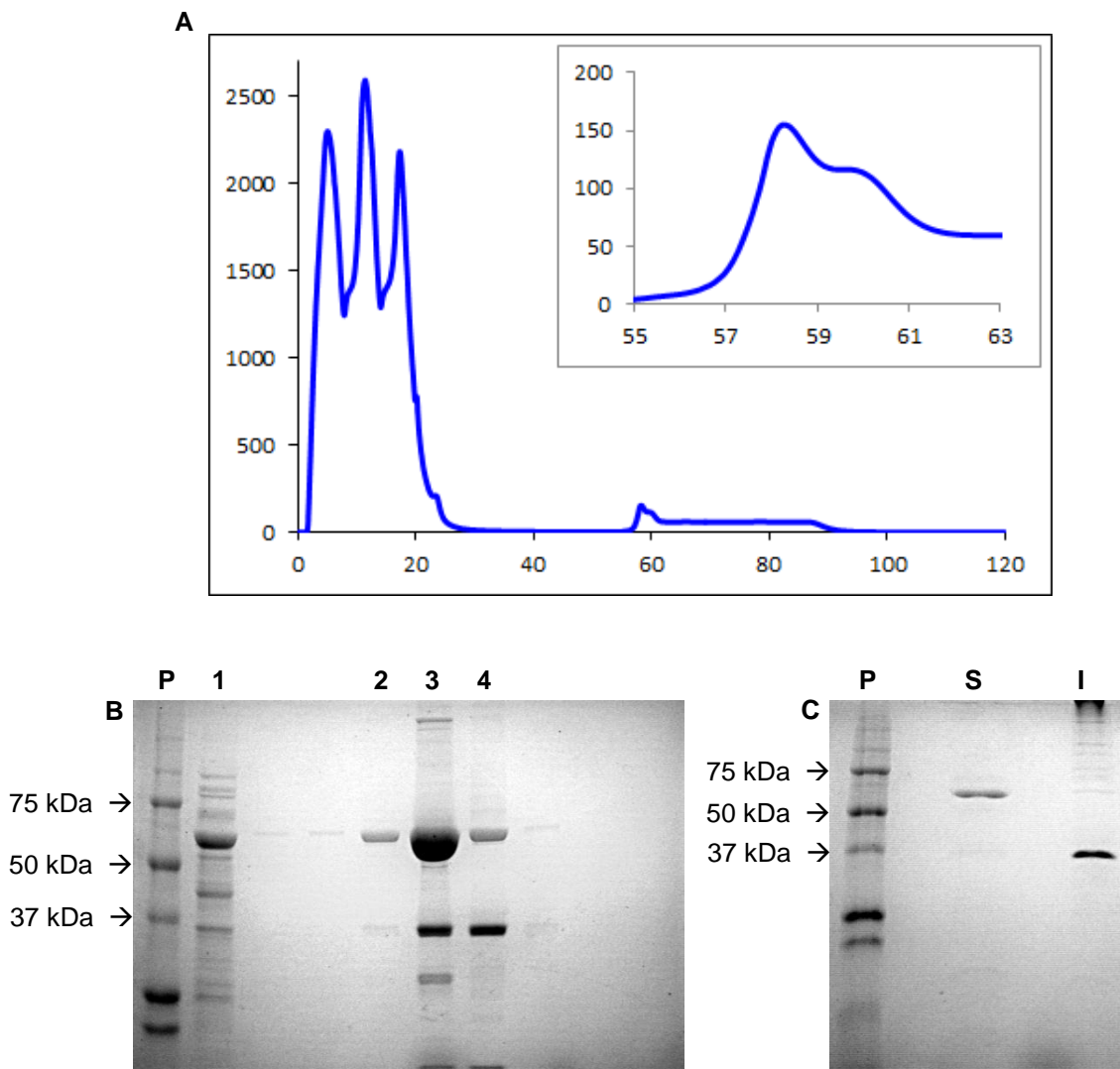


Figura 22: Purificação da proteína NahE por cromatografia de afinidade ao níquel.

A - Cromatograma obtido durante a purificação. A curva em azul representa a absorvância a 280 nm medida em mAU em função do volume de eluição da coluna em mL. Na caixa em detalhe o pico obtido durante eluição com imidazol.

B - SDS-PAGE contendo algumas frações da purificação. **P** - padrão de proteínas Bio-Rad; **1** - fração que não interagiu com a coluna entre os volumes de eluição 0 e 30 mL; **2, 3 e 4** - frações correspondentes ao pico obtido durante a eluição com imidazol entre os volumes de eluição 57 e 61 mL.

C - SDS-PAGE contendo as frações solúvel (**S**) e insolúvel (**I**) após diálise e centrifugação das frações do pico (**Pico**); **P** - padrão de proteínas Bio-Rad.

5.7 Caracterização do estado de agregação de NahE e Cpn60

Para verificar o estado de agregação, a fração de NahE contendo a chaperona Cpn60 (fração 10 do cromatograma, canaleta 4 do gel, Figura 22) foi analisada através de ensaio de DLS. Pelos dados exibidos na Tabela 10, pode-se observar que 68,2% da massa total das partículas em suspensão da amostra é formada por moléculas com raio médio de 4,5 nm e massa molecular aproximada de 112 kDa, o que corresponderia a um oligômero formado por uma molécula de chaperona Cpn60 e Cpn10 (aproximadamente 57 kDa e 10 kDa) e uma molécula da proteína NahE (38,6 kDa). Outros 31,3% da massa apresenta um raio médio de 8,5 nm e uma massa estimada de 498 kDa, sugerindo um oligômero ainda maior entre essas proteínas.

Tabela 10: Análise da proteína NahE por espalhamento de luz dinâmico.

Raio (nm)	% de Polidispersividade	Massa (kDa)	% de Massa
4,5	13,1	112	68,2
8,5	21,8	498	31,3
76,5	14,0	85.937	0,2
3396,0	0	615.073.000	0,3

Uma hipótese para o fato observado é de que, na tentativa de auxiliar num correto enovelamento, as chaperonas Cpn60 e Cpn10 interagiriam fortemente com a proteína NahE enovelada incorretamente. Ao realizar a etapa de cromatografia de afinidade, a cauda de histidina de NahE interagiria com a coluna, parte das chaperonas eram lavadas, mas outra parte permaneceria interagindo com NahE. Quando as proteínas fossem eluídas, ocorreria a formação de um precipitado de NahE nas frações que contivessem poucas chaperonas. Por estar num enovelamento incorreto, NahE na ausência dessas chaperonas, voltaria a ser insolúvel e precipitaria, enquanto que na fração com chaperonas, NahE manteria-se complexada com essas proteínas e seria carregada junto com elas para a fração

solúvel. Após esse resultado, o gene *nahE* foi clonado em vetor pMCSG19C utilizando o sistema LIC.

5.8 Clonagem e expressão do gene *nahE* utilizando sistema LIC

Como exibido no item 5.1, o gene *nahE* foi amplificado por PCR utilizando uma T_m de 45°C. O produto de amplificação foi purificado e tratado com a enzima T4 DNA polimerase para formar as caudas do sistema LIC.

Células de *E. coli* DH5 α contendo os vetores pMCSG19C e pMCSG7 foram crescidas e tiveram seus plasmídeos extraídos e purificados. A digestão do vetor pMCSG19C (aproximadamente 7500 pb) com a enzima de restrição Sspl ocorreu de forma eficiente segundo recomendações do fabricante, utilizando como teste 500 μ g de vetor e enzima na concentração de 5 U/ μ g de DNA, incubados a 37°C por 16 h (Figura 23). Ocorreu o aparecimento de duas bandas de tamanhos aproximados de 4000 pb e 3500 pb, correspondentes à clivagem inespecífica da enzima Sspl em outro local do vetor. Essa clivagem inespecífica, chamada de atividade estrela, pode ser evitada, ou ao menos diminuída, utilizando quantidades menores de enzima, diminuindo o tempo de reação ou a temperatura de incubação. Após a padronização da digestão do plasmídeo, 3 μ g do vetor foram digeridos utilizando 5 U/ μ g de Sspl. A banda correspondente ao vetor linearizado, na altura de aproximadamente 7000 a 8000 pb foi excisada do gel, purificada e tratada com a enzima T4 DNA polimerase.

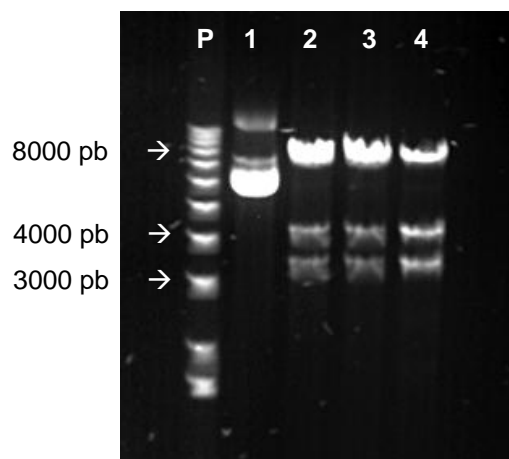


Figura 23: Digestão do vetor pMCSG19C utilizando a enzima SspI.

P - 1 kb plus DNA Ladder; **1** - vetor fechado; **2 a 4** - Digestão utilizando 5, 10 e 15 U/μg de SspI.

O produto da mistura de inserto e vetor tratados com T4 DNA polimerase foram inseridos em *E. coli* BL21, sendo obtidos vários transformantes. A maioria dos clones testados por PCR de colônia apresentaram-se como positivos para o gene *nahE*. Um clone foi selecionado para o teste de expressão. O processo desde a amplificação do gene, digestão e purificação do vetor, os respectivos tratamentos utilizando T4 DNA polimerase, transformação das células e identificação dos clones positivos, prontos para testes de expressão, levou cerca de 4 dias, mostrando que o processo de clonagem utilizando o sistema LIC é realmente rápido.

Proteínas expressas no vetor pMCSG19C apresentam uma cauda de fusão amino-terminal composta pela MBP, um sítio de clivagem reconhecido pela protease TVMV, uma cauda de poli-histidina, um sítio de clivagem reconhecido pela TEV, e, por fim, a sequência da proteína recombinante. A função da MBP é auxiliar na solubilidade da proteína recombinante, carregando a mesma para a fração solúvel. O próprio vetor apresenta o gene da protease TVMV, sendo que sua expressão é controlada pelo promotor P_{L-tetO} [65], diferente do que controla a expressão da proteína recombinante. Quando incorporado a organismos que não produzem a proteína repressora Tet, como a cepa BL21(DE3) de *E. coli*, o vetor pMCSG19C expressa a protease TVMV constitutivamente. Uma vez que a proteína alvo é expressa e em fração solúvel, a cauda de fusão MBP é clivada *in vivo* pela TVMV protease, liberando a proteína recombinante ainda com uma cauda de histidina amino-terminal. Essa cauda auxiliará na purificação por cromatografia de afinidade

ao níquel, e por fim, será clivada pela protease TEV. Utilizando esse sistema, restarão três resíduos de aminoácidos na região amino-terminal da proteína, sendo eles uma serina, uma asparagina e o terceiro dependerá do iniciador desenhado (como o códon deste obrigatoriamente inicia-se com uma guanina, o terceiro resíduo poderá ser uma valina, alanina, aspartato ou glicina, sendo que para NahE, foi optado por uma alanina).

NahE expressa nesse sistema e fusionada a cauda de histidina possui uma massa teórica de 39,33 kDa. Quando clivada pela protease TEV, sua massa teórica vai para 36,909 kDa. Como a MBP é clivada da proteína recombinante *in vivo*, ela também deverá aparecer no gel de poliacrilamida-SDS, com uma massa de 42,351 kDa.

A expressão em *E. coli* BL21 foi induzida pela adição de IPTG em concentração final de 1 mM em meio seletivo LB incubado a 37°C por 3 h. Amostras foram coletadas a cada hora de expressão (Figura 24). Após o término da expressão, foi feita a lise em tampão fosfato de sódio 20 mM pH 7,4, NaCl 500 mM. Observa-se, após a indução, o aparecimento das bandas de NahE e MBP, com pesos moleculares aproximados do valor esperado. Análises das frações solúvel e insolúvel indicaram que à 37°C, tanto NahE quanto MBP ficaram presentes na fração insolúvel.

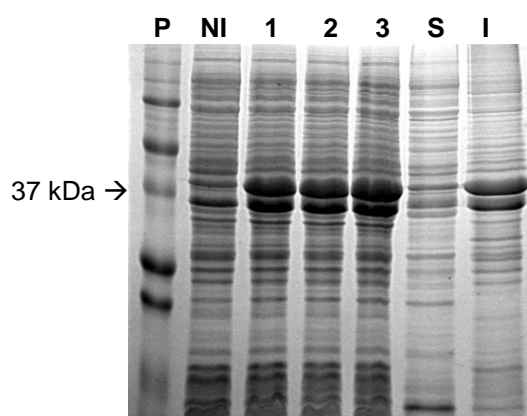


Figura 24: Expressão da proteína NahE no vetor pMCSG19C em BL21 a 37°C, IPTG 1 mM.

P - padrão de proteínas Bio-Rad; **NI** - expressão não induzida; **1 a 3** - tempos de expressão de 1 a 3 h; **S** - fração solúvel; **I** - fração insolúvel.

Uma nova expressão, para controle negativo, foi realizada utilizando *E. coli* BL21 sem nenhum vetor de expressão. Foi adicionado IPTG em concentração final de 1 mM e a cultura incubada a 37°C por 3 h. Foram analisadas em um mesmo gel as amostras coletadas antes e após a adição do IPTG na cultura de *E. coli* BL21 sem vetor (NI e TF, Figura 25), juntamente com a fração solúvel e insolúvel da expressão de NahE a 37°C no vetor pMCSG19C (S e I, Figura 25).

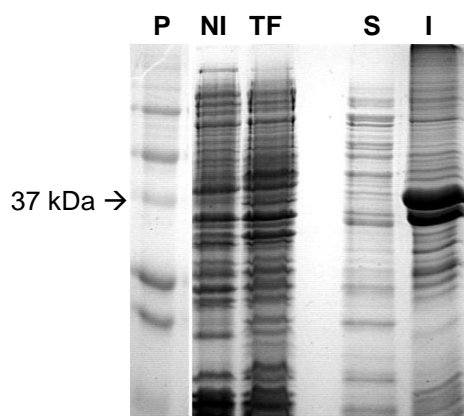


Figura 25: Controle negativo em BL21 induzida a 37°C, IPTG 1 mM.

P - padrão de proteínas Bio-Rad; **NI** - expressão não induzida em BL21 sem vetor; **TF** - tempo final de indução em BL21 sem vetor; **S** - fração solúvel da expressão de NahE; **I** - fração insolúvel da expressão de NahE.

Como pode-se observar, há uma banda relativamente forte na mesma altura da banda esperada para NahE, tanto no tempo antes da indução, como no induzido. Essa banda é de uma proteína nativa de *E. coli*, presente também em outras expressões de proteínas recombinantes estudadas pelo grupo BioEst. A presença dessa banda na fração solúvel pode levar a conclusão errônea de que há uma pequena porção solúvel de NahE.

Uma nova expressão foi realizada a 30°C utilizando o mesmo clone de NahE no vetor pMCSG19C. A expressão nessa temperatura foi induzida com IPTG nas concentrações finais de 1, 0,5 e 0,25 mM por 5 h. Foram coletadas frações no tempo final de expressão. A lise foi realizada em tampão Tris-HCl 50 mM pH 7,4, sacarose 1%, tween-20 1% e glicerol 1%. Amostras das frações solúveis e insolúveis foram analisadas por SDS-PAGE (Figura 26). Como pode ser observado pelo gel, a 30°C, grande parte da MBP ficou em fração solúvel após a lise, entretanto, NahE manteve-

se insolúvel. O resultado obtido foi bastante semelhante utilizando as três concentrações do agente indutor.

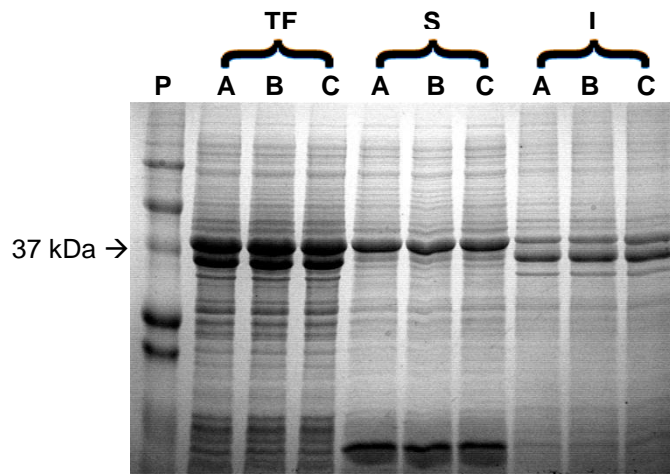


Figura 26: Expressão da proteína NahE no vetor pMCSG19C em BL21 a 30°C, IPTG a 1, 0,5 e 0,25 mM.

P - padrão de proteínas Bio-Rad; **TF** - tempo final da expressão; **A**, **B** e **C** - indução com respectivamente 1, 0,5 e 0,25 mM de IPTG; **S** - fração solúvel; **I** - fração insolúvel.

Também foi realizada uma expressão do mesmo clone a 22°C por 16 h utilizando IPTG na concentração final de 0,5 mM. A lise ocorreu em tampão Tris-HCl 50 mM pH 7,4, sacarose 1%, tween-20 1% e glicerol 1%. Como pode ser observado na Figura 27, a 22°C NahE permaneceu expressa em fração insolúvel.

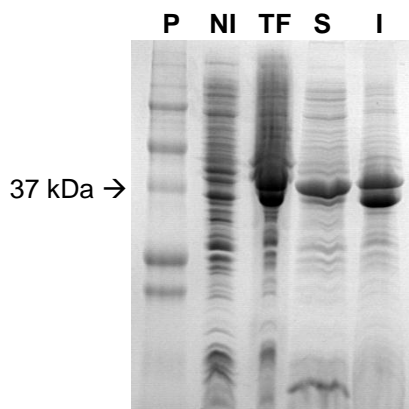


Figura 27: Expressão da proteína NahE no vetor pMCSG19C em BL21 a 22°C, IPTG a 0,5 mM.

P - padrão de proteínas Bio-Rad; **NI** - expressão não induzida; **TF** - tempo final de expressão; **S** - fração solúvel; **I** - fração insolúvel.

Como nova tentativa de obter NahE solúvel, o produto da mistura de inserto e vetor tratados com T4 DNA polimerase foram inseridos em *E. coli* ArticExpress, sendo obtidos vários transformantes. Todos os clones testados por PCR de colônia apresentaram-se como positivos para o gene *nahE*. Quatro clones foram selecionados para o teste de expressão.

A expressão dos clones em ArticExpress foi induzida com IPTG na concentração final de 0,5 mM por 24 h à 12°C. O clone de *nahE* no vetor pMCSG19C inserido em *E. coli* BL21 também foi expresso com IPTG 0,5 mM a 12°C por 24 h. A lise ocorreu em tampão Tris-HCl 50 mM pH 7,4, sacarose 1%, tween-20 1% e glicerol 1%. As amostras coletadas antes da indução, após o término da expressão, e das frações solúveis e insolúveis foram analisadas por SDS-PAGE (Figura 28).

Pode ser observado nas amostras do tempo final de indução as bandas de expressão de NahE e MBP, e nos clones em ArticExpress, também é observada a banda da chaperona Cpn60. À 12°C há uma grande solubilidade de Cpn60 e MBP. A maior parte de NahE manteve-se em fração insolúvel, entretanto ocorreu o aparecimento de uma pequena banda de NahE na fração solúvel nos clones A1, A2 e A4 de ArticExpress. A fração solúvel desses clones foi unida e utilizada numa etapa de cromatografia de afinidade.

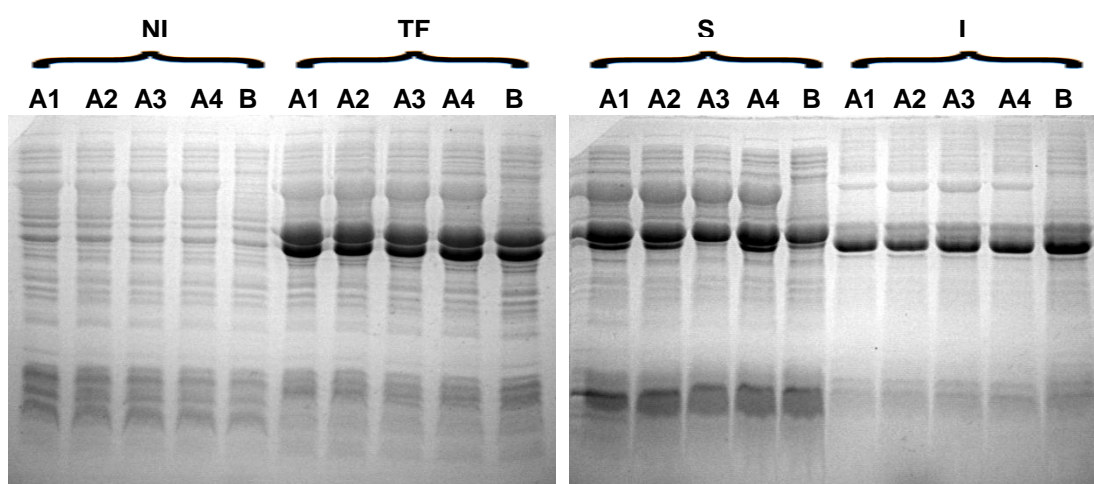


Figura 28: Expressão da proteína NahE no vetor pMCSG19C em ArticExpress e BL21 a 12°C, IPTG a 0,5 mM.

NI - expressão não induzida; **TF** - tempo final de expressão; **S** - fração solúvel; **I** - fração insolúvel; **A1** a **A4** - colônias de ArticExpress; **B** - colônia de BL21.

5.9 Purificação de NahE expressa em vetor com sistema LIC

As amostras da fração solúvel da expressão de NahE no vetor pMCSG19C em ArticExpress foi purificada em cromatografia de afinidade ao níquel. A coluna foi previamente equilibrada com tampão fosfato de sódio 20 mM pH 7,4, NaCl 500 mM, imidazol 30 mM. A amostra foi injetada e após a curva de absorbância a 280 nm retornar a valores próximos de zero (Figura 29 A), foi utilizado o step de imidazol na concentração de 500 mM para eluir as proteínas que interagiram com a coluna. Frações ao longo de todo processo foram coletadas e algumas analisadas por SDS-PAGE (Figura 29 B).

Como pode ser observado no gel, a MBP e parte da chaperona Cpn60 não interagiram com a coluna sendo coletadas nas primeiras frações da purificação (canaleta 1, Figura 29 B). Durante a eluição da coluna foi obtido um pico de proteína que foi coletado em 3 frações. A primeira fração do pico (canaleta 2) continha basicamente Cpn60, a segunda fração continha uma mistura de Cpn60 e NahE (canaleta 3), e a terceira fração continha somente NahE (canaleta 4). As duas frações contendo NahE foram dialisadas contra tampão fosfato de sódio 20 mM, NaCl 500 mM para retirada do imidazol. Ocorreu a formação de um precipitado na fração que continha apenas NahE. A fração com uma mistura de Cpn60 e NahE manteve-se transúcida e sem precipitados. A amostra em que ocorreu o precipitado foi centrifugada e analisada por SDS-PAGE (dado não exibido), o resultado obtido foi semelhante ao da Figura 22 C, em que NahE precipitou na ausência de Cpn60.

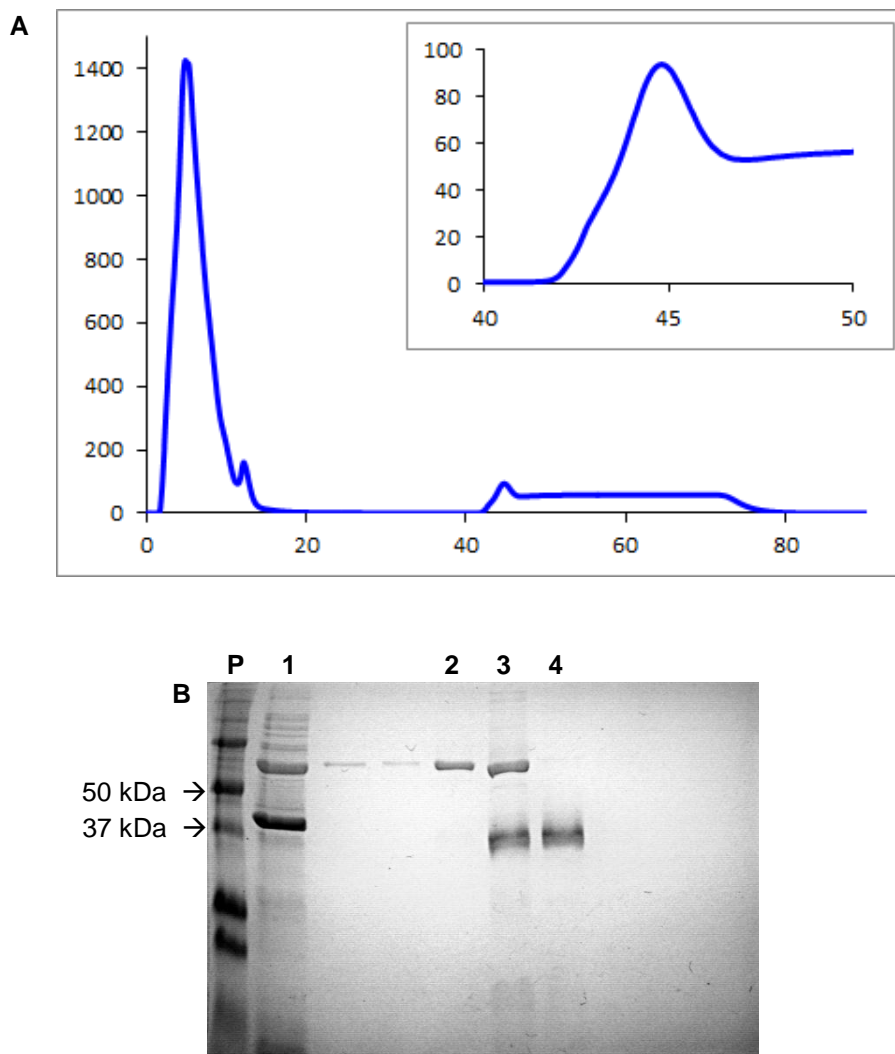


Figura 29: Purificação da proteína NahE por cromatografia de afinidade ao níquel.

A - Cromatograma obtido durante a purificação. A curva em azul representa a absorvância a 280 nm medida em mAU em função do volume de eluição da coluna em mL. Na caixa em detalhe o pico obtido durante eluição com imidazol.

B - SDS-PAGE contendo algumas frações da purificação. **P** - padrão de proteínas Bio-Rad; **1** - fração que não interagiu com a coluna; **2, 3 e 4** - frações correspondentes ao pico obtido durante a eluição com imidazol.

5.10 Reenovelamento de NahE

Devido a insolubilidade de NahE, optou-se pelo procedimento de expressão insolúvel em vetor pMCSG7, tratamento da fração insolúvel, purificação sob condições desnaturantes e posterior reenovelamento. O vetor do sistema LIC foi escolhido para expressão de NahE pela facilidade em obter clones positivos rapidamente, quando comparados com o sistema do vetor pET28a-TEV. A digestão do vetor pMCSG7 (aproximadamente 5300 pb) com a enzima de restrição SspI ocorreu segundo o descrito para o vetor pMCSG19C no item 5.8. A banda correspondente ao vetor linearizado foi excisada do gel, purificada e tratada com a enzima T4 DNA polimerase. O produto da mistura de inserto e vetor tratados com T4 DNA polimerase foram inseridos em *E. coli* BL21, sendo obtidos vários transformantes positivos. Novamente o processo de clonagem e obtenção de clones positivos utilizando o sistema LIC foi bastante rápido.

Proteínas expressas no vetor pMCSG7 apresentam uma cauda amino-terminal de poli-histidina e um sítio de clivagem reconhecido pela TEV. Ao utilizar a protease TEV, restarão três resíduos de aminoácidos na região amino-terminal da proteína, sendo eles uma serina, uma asparagina e o terceiro que dependerá do iniciador desenhado, neste caso, uma alanina. NahE expressa nesse sistema e fusionada a cauda de histidina possui uma massa teórica de 39,37 kDa. Quando clivada pela protease TEV, sua massa teórica vai para 36,909 kDa.

A expressão em *E. coli* BL21 foi induzida pela adição de IPTG em concentração final de 1 mM em 300 mL de meio seletivo LB incubado a 37°C por 3 h. Amostras foram coletadas antes da indução e no tempo final de expressão. Após o término da expressão, foi feita a lise em tampão fosfato de sódio 20 mM pH 7,4, NaCl 500 mM. Observa-se, após a indução, o aparecimento da banda de NahE com peso molecular aproximado do valor esperado (TF, Figura 30 A). Como esperado, NahE foi expressa em fração insolúvel. O tratamento da fração insolúvel ocorreu conforme descrito no item 4.20. Amostras foram coletadas ao longo do tratamento e analisadas por SDS-PAGE (T1 a T3 Figura 30 A). Após o tratamento da fração insolúvel, verificou-se a solubilização de NahE em uréia na concentração de 5 M (Figura 30 B). Procedeu-se então purificação por cromatografia de afinidade com tampões contendo uréia 5 M (Figura 31 A).

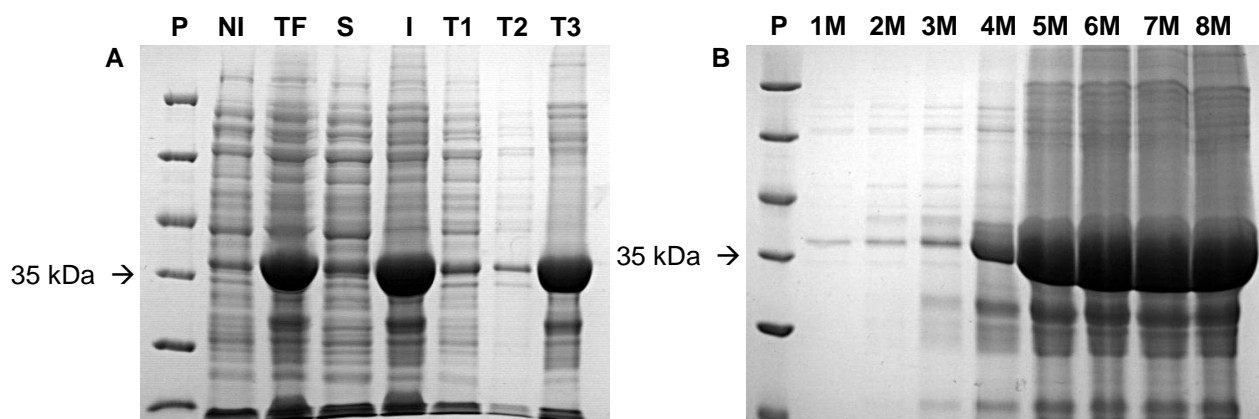


Figura 30: Expressão de NahE em pMCSG7 em BL21 a 37°C e tratamento da fração insolúvel.

P - padrão de proteínas Fermentas; **NI** - expressão não induzida; **TF** - tempo final de expressão; **S** - fração solúvel; **I** - fração insolúvel; **T1** e **T2** - fração solúvel durante tratamento do precipitado; **T3** - fração final do tratamento do pellet, utilizada para a etapa de purificação.

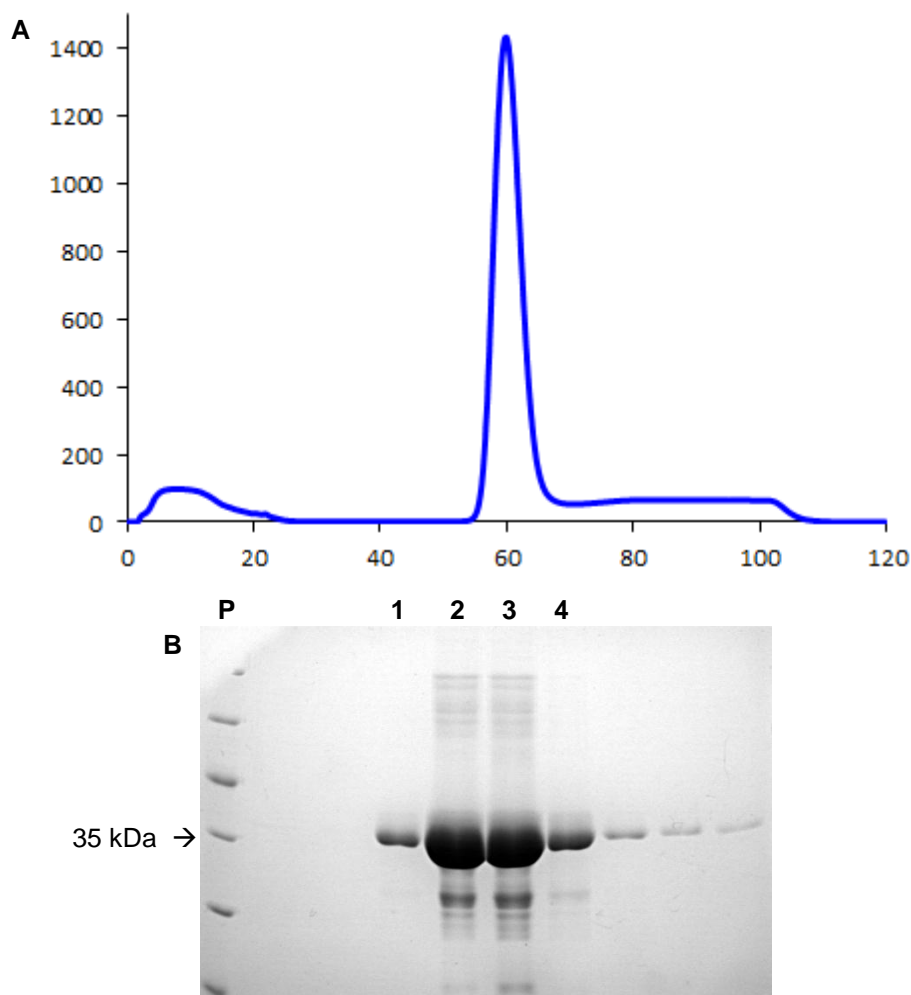


Figura 31: Purificação da proteína NahE por cromatografia de afinidade em condições desnaturantes.

A - Cromatograma obtido durante a purificação. A curva em azul representa a absorvância a 280 nm medida em mAU em função do volume de eluição da coluna em mL.

B - SDS-PAGE contendo algumas frações da purificação. **P** - padrão de proteínas Fermentas; **1 a 4** - frações correspondentes ao pico obtido durante a eluição com imidazol.

Após a purificação, as frações contendo NahE foram unidas. Parte da amostra foi diluída em tampão fosfato de sódio 20 mM pH 7,4, NaCl 100 mM até a concentração de uréia atingir 2 M. Outra parte da amostra foi diluída na mesma proporção utilizando o mesmo tampão porém contendo L-arginina na concentração final de 500 mM. As amostras, após incubação por 16 h a 4°C, foram dialisadas contra tampão fosfato de sódio 20 mM pH 7,4, NaCl 100 mM, para remoção da uréia e arginina. Houve formação de precipitado na amostra onde o reenovelamento ocorreu sem arginina. Amostras coletadas antes e após a diálise foram analisadas por SDS-PAGE. Como pode ser observado na canaleta 3 Figura 32, ocorreu precipitação de NahE no reenovelamento sem arginina, restando pouca ou nenhuma NahE na fração solúvel (canaleta 2). No reenovelamento com arginina, NahE manteve-se solúvel (canaleta 1).

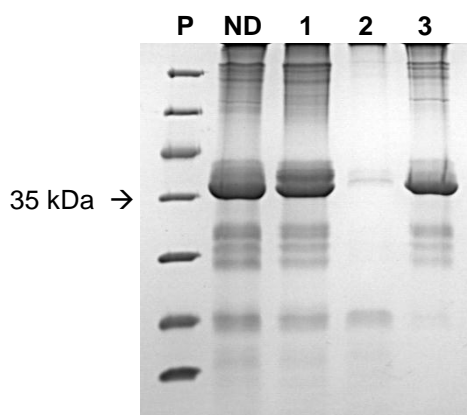


Figura 32: Reenovelamento da proteína NahE.

P - padrão de proteínas Fermentas; **ND** - fração não dialisada; **1** - fração solúvel após reenovelamento contendo arginina; **2** - fração solúvel após reenovelamento na ausência de arginina; **3** - fração insolúvel após reenovelamento na ausência de arginina.

Parte da fração solúvel do reenovelamento contendo arginina foi submetida a cromatografia de afinidade em coluna de níquel-sepharose. Como pode ser observado pelo cromatograma dessa purificação (Figura 33), toda a amostra protéica foi coletada nas primeiras frações. Não houve nenhum pico durante a eluição da coluna, indicando que nenhuma proteína interagiu com a resina. Era esperado que, por NahE ter a cauda de histidina amino-terminal, ficasse retida na coluna e fosse eluída durante o gradiente com imidazol. A subida da curva azul durante a eluição deve-se à absorvância do imidazol a 280 nm.

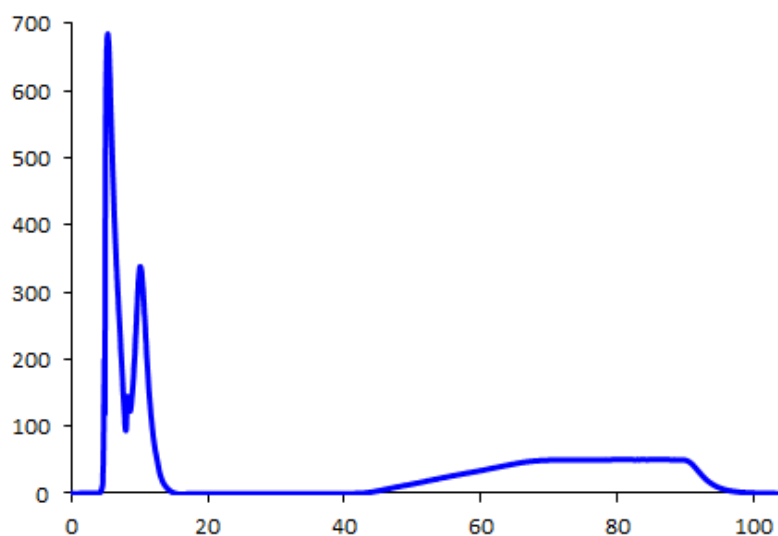


Figura 33: Cromatograma obtido durante a purificação por cromatografia de afinidade ao níquel da proteína NahE renovelada.

A curva em azul representa a absorvância a 280 nm medida em mAU em função do volume de eluição da coluna em mL.

Outra parte da fração solúvel do reenovelamento de NahE foi purificada por cromatografia de exclusão molecular. Como observado no cromatograma (Figura 34), houve um pico eluído próximo de 8 mL equivalente ao volume morto da coluna, no qual são eluídas proteínas totalmente excluídas dos poros.

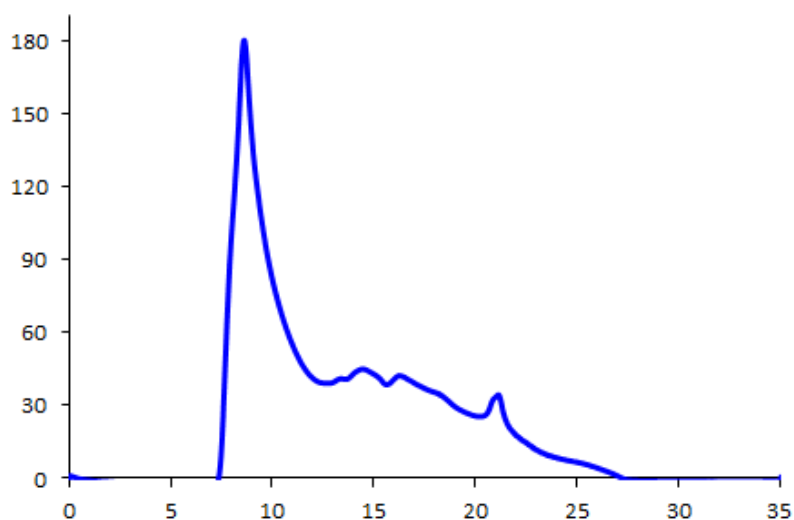


Figura 34: Cromatograma obtido durante a purificação por cromatografia de exclusão molecular da proteína NahE renovelada.

A curva em azul representa a absorvância a 280 nm medida em mAU em função do volume de eluição da coluna em mL.

Esse resultado indica um agregado protéico que atravessa a coluna rapidamente sendo eluído antes de proteínas de massas moleculares menores. Isso justificaria a falta de interação com a coluna de afinidade. Se NahE, após o processo de renovelamento, se encontra enovelada incorretamente formando grandes agregados solúveis, é possível que as caudas de histidinas estejam escondidas nesse aglomerado. Algumas das caudas, ocasionalmente expostas no aglomerado, poderiam ser ineficientes em retê-lo na coluna de afinidade.

Foi possível sequenciar cerca de 40% do gene de NahE no vetor pMCSG7 do clone utilizado para a expressão. Não há mutações nas bases cobertas pelo sequenciamento (Figura 35).

```

nahE      ATGTTGAATAAAGTTATTAAAAACCACGCGTCTTACCGCTGAAGATATCAACGGTGCCTGGACTATAATGCCACACCGTC 80
nahE-pMCSG7 ATGTTGAATAAAGTTATTAAAAACCACGCGTCTTACCGCTGAAGATATCAACGGTGCCTGGACTATAATGCCACACCGTC 80
*****

nahE      GACGCCTGATGCTTCTGATTGGCGCAGCACTAACACTGTGGACTTAGACGAGACTGCCCGCATAGTTGAAGAGCTGATTG 160
nahE-pMCSG7 GACGCCTGATGCTTCTGATTGGCGCAGCACTAACACTGTGGACTTAGACGAGACTGCCCGCATAGTTGAAGAGCTGATTG 160
*****

nahE      CAGCTGGTGTCAACGGTATTTTAAGTATGGGTACTTTTGGTGAGTGCGCCACGTTGACCTGGGAGGAGAAACGTGATTAT 240
nahE-pMCSG7 CAGCTGGTGTCAACGGTATTTTAAGTATGGGTACTTTTGGTGAGTGCGCCACGTTGACCTGGGAGGAGAAACGTGATTAT 240
*****

nahE      GTTTCGACGGTTGTTCGAGACCATTTCGTGGTCTGTGCCTTATTTCTGTGGCAGCAGCGCCCTGAATACCCGAGAAGTCAT 320
nahE-pMCSG7 GTTTCGACGGTTGTTCGAGACCATTTCGTGGTCTGTGCCTTATTTCTGTGGCAGCAGCGCCCTGAATACCCGAGAAGTCAT 320
*****

nahE      CCGCCAGACCCGAGAGCTTATCGATATTGGCGCTAACGGCACCATGCTAGGCGTGCCGATGTGGGTGAAGATGGACCTGC 400
nahE-pMCSG7 CCGCCAGACCCGAGAGCTTATCGATATTGGCGCTAACGGCACCATGCTAGGCGTGCCGATGTGGGTGAAGATGGACCTGC 400
*****

nahE      CCACAGCGGTCCAGTTCATTCGTGATGTTGCAGGCGCGGTACCGGAGGCTGCCATTGCGATTACGCCAACCCCGAAGCA 480
nahE-pMCSG7 CC----- 402
**

nahE      TTCAAATTCGACTTCCCTCGCCATTTTGGGCAGAGATGTCCAAAATTCCTCAGTAGTGACTGCCAAGTATCTAGGCAT 560
nahE-pMCSG7 -----

nahE      CGGAATGCTTGACTTGGACCTGAAATTTGGCGCCTAACATCCGCTTCCTCCACACGAGGACGACTATTACGCGGCCGCAC 640
nahE-pMCSG7 -----

nahE      GCATCAATCCCGAGCGCATAACCGGTTCTGGTCAAGCGGGCCATGTGCGGCCCGGCTACCGCTATCATGTTGCGTGAT 720
nahE-pMCSG7 -----

nahE      GAAGTGGAGCGGGCCAAGAGTACCGGTGACTGGATCAAGGCCAAAGCCATCTCCGATGATATGCGTGCAGCCGATTTCGAC 800
nahE-pMCSG7 -----

nahE      ATTGTTTCCGCGTGGCGACTTTTCGGAGTTCTCGAAGTATAACATCGGGCTTGAAAAGGCACGGATGGACGCGGCTGGTT 880
nahE-pMCSG7 -----

nahE      GGCTCAAGGCTGGTCCCTGCCGCTCCTCCCTACAATCTTGTTCAGAAAGATTACCTCGTGGTGCACAGAAATCAGGCAAG 960
nahE-pMCSG7 -----

nahE      GCGTGGGCGCGCTGCACGCTAAATACAGTAAATAA 996
nahE-pMCSG7 -----

```

Figura 35: Sequenciamento de *nahE* em vetor pMCSG7.

5.11 Caracterização da estrutura secundária de NahE reenovelada

Para analisar as estruturas secundárias por Dicroísmo Circular, a amostra protéica deve estar numa faixa correta de concentração de acordo com o caminho óptico da cubeta utilizada. Uma vez que a técnica consiste em uma medição de absorvância, em última instância, há um detector de fótons. Se a absorvância da amostra + tampão + cubeta é muito alta, não chegará luz suficiente ao detector e, por tanto, um espectro significativo não será registrado. Da mesma forma, se a absorvância da amostra + tampão + cubeta é muito baixa, então a absorção diferencial de luz circularmente polarizada à direita e à esquerda estará abaixo do limiar de detecção do sensor. Através do aparelho Jasco, utilizado para os experimentos de CD, é possível observar a alta tensão (HT) que controla a intensidade luminosa sobre a amostra. A voltagem HT é diretamente proporcional à absorvância. Como a absorvância sempre aumenta à medida que o comprimento de onda diminui de 260 a 178 nm, é normal a voltagem HT aumentar a fim de que possam chegar fótons suficientes ao detector para um registro confiável. Entretanto, se essa tensão fica acima de 700 V, o detector está saturado e a amplitude do espectro irá oscilar descontroladamente e dados de má qualidade serão gravados [39].

No caso da enzima NahE reenovelada, mesmo diluindo diversas vezes a amostra, a voltagem HT do aparelho atingiu níveis limites de operação logo após o início da curva do espectro a 260 nm. Dessa forma não foi possível caracterizar as estruturas secundárias em NahE após o processo de reenovelamento.

5.12 Expressão e purificação de NahK em vetor pET28a-TEV

A proteína NahK expressa no vetor pET28a-TEV apresenta a cauda de histidina e sua massa teórica foi estimada em 30,7 kDa. Utilizando a protease TEV para clivagem da cauda, a massa de NahK vai para 28,91 kDa. Nessa construção utilizando o sítio da enzima NheI no iniciador direto, após clivagem por TEV, irão

restar no amino-terminal da proteína os resíduos de aminoácidos glicina, histidina, metionina, alanina e serina.

O primeiro teste de expressão de NahK foi feito em *E. coli* BL21 em 500 mL de meio LB seletivo a 30°C por 5 horas sob agitação de 200 rpm. A indução foi realizada com IPTG na concentração final de 0,5 mM. A lise foi realizada em tampão fosfato de sódio 20 mM pH 7,4, NaCl 500 mM. As amostras coletadas antes da indução, após o término, e da fração solúvel e insolúvel foram analisadas por SDS-PAGE (Figura 36).

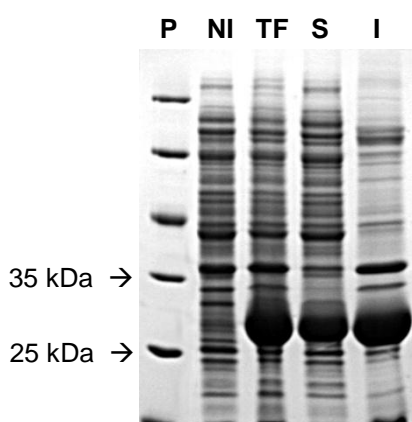


Figura 36: Expressão da proteína NahK no vetor pET28a-TEV em BL21 a 30°C, IPTG 0,5 mM.

P₁ - padrão de proteínas Fermentas; **NI** - expressão não induzida; **TF** - tempo final de expressão; **S** - fração solúvel; **I** - fração insolúvel.

Verificou-se o aparecimento de uma banda de aproximadamente 30 kDa após a indução da expressão. Apesar da presença de uma banda maior na fração insolúvel, há uma banda considerável na fração solúvel. A amostra foi então submetida à cromatografia de afinidade em coluna HisTrap HP no sistema ÄKTAprime plus.

A coluna foi previamente equilibrada com tampão fosfato de sódio 20 mM pH 7,4, NaCl 500 mM, imidazol 30 mM. A amostra foi injetada e após a curva de absorvância a 280 nm retornar a valores próximos de zero (Figura 37 A), foi utilizado um gradiente de imidazol de 30 mM a 500 mM para eluir as proteínas que interagiram com a coluna. Frações ao longo de todo processo foram coletadas e algumas analisadas por SDS-PAGE (Figura 37 B).

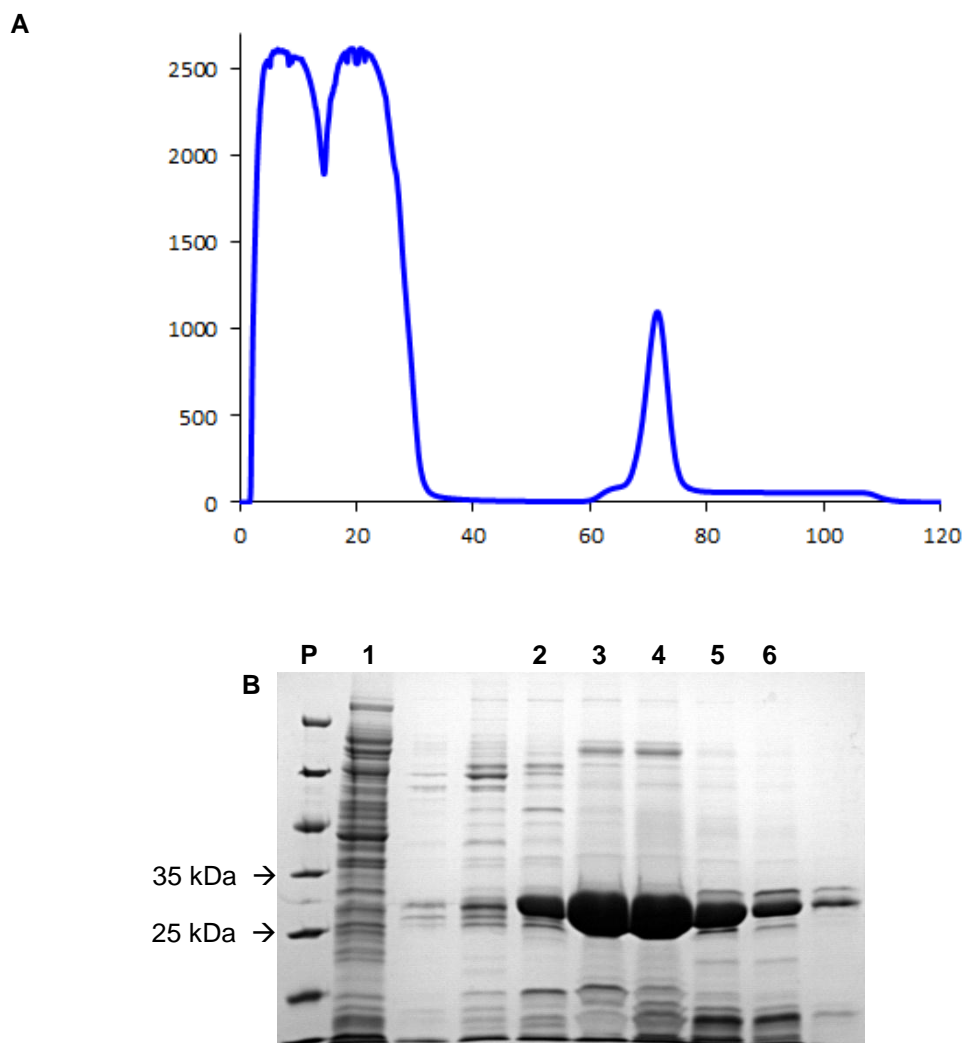


Figura 37: Purificação da proteína NahK por cromatografia de afinidade ao níquel.

A - Cromatograma obtido durante a purificação. A curva em azul representa a absorvância a 280 nm medida em mAU em função do volume de eluição da coluna em mL.

B – SDS-PAGE contendo algumas frações da purificação. **P** – padrão de proteínas Fermentas; **1** – fração que não interagiu com a coluna; **2 a 6** – frações correspondentes ao pico obtido durante a eluição com imidazol.

As frações correspondentes ao pico de eluição (canaletas 2 a 6, Figura 37 B) foram reunidas e dialisadas contra tampão fosfato de sódio 20 mM pH 7,4, NaCl 100 mM, para retirada do imidazol e excesso de sal. Devido a remoção do sal, ocorreu a formação de precipitado, que foi centrifugado e analisado por SDS-PAGE. Uma pequena parte da proteína NahK tornou-se insolúvel, apesar disso, grande parte da proteína manteve-se em fração solúvel (dado não exibido). A fração solúvel da amostra dialisada foi submetida à cromatografia de troca iônica.

O pI teórico de NahK fusionada com a cauda de histidina é de 5,65. A proteína possui uma carga efetiva negativa em pH 7,4, já que seus resíduos carregados estão desprotonados. Dessa maneira, a proteína é capaz de fazer pontes salinas com a resina da coluna de troca iônica que está carregada positivamente. Passando pela coluna um tampão em pH 7,4 mas com concentrações cada vez maiores de NaCl, ocorre a eluição das proteínas ligadas na coluna, já que há competição pelas cargas positivas entre os íons cloreto e as proteínas, ambos carregados negativamente.

O cromatograma obtido na purificação em coluna de troca iônica utilizando o tampão fosfato de sódio 20 mM e NaCl 100 mM, e amostras coletadas ao longo da purificação analisadas por SDS-PAGE estão representadas na Figura 38. Como pode ser observado pela curva de absorbância a 280 nm no cromatograma, grande parte das proteínas não interagiram com a coluna, enquanto que uma pequena porção foi eluída durante o gradiente de NaCl. As frações analisadas por SDS-PAGE mostram a presença em grande quantidade de NahK tanto nas frações que não interagiram com a coluna, quanto na fração correspondente ao pico durante a eluição, e que os mesmos contaminantes estavam presentes em todas as frações (Figura 38 B).

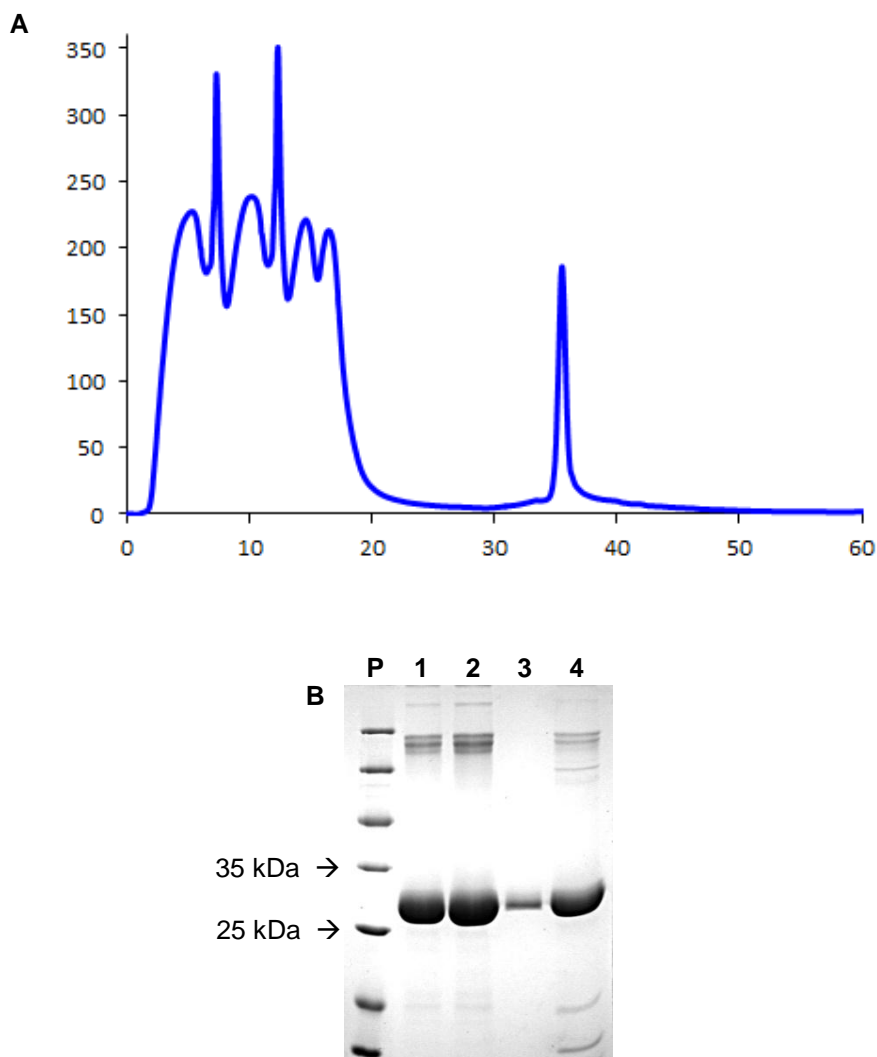


Figura 38: Purificação da proteína NahK por cromatografia de troca iônica.

A - Cromatograma obtido durante a purificação. A curva em azul representa a absorvância a 280 nm medida em mAU em função do volume de eluição da coluna em mL.

B - SDS-PAGE contendo algumas frações da purificação. **P** - padrão de proteínas Fermentas; **1** e **2** - frações correspondentes ao volume de eluição de 0 a 20 mL que não interagiram com a coluna; **3** - fração obtida durante a lavagem da coluna correspondente ao volume de eluição de 20 a 30 mL; **4** - fração correspondente ao pico obtido durante a eluição entre 30 e 40 mL.

Todas as frações contendo NahK provenientes da troca iônica foram unidas, dialisadas contra tampão Tris-HCl 50 mM pH 7,4, NaCl 50 mM, e submetidas a uma nova cromatografia de troca iônica em coluna RESOURCE Q equilibrada com o mesmo tampão da diálise. Devido a menor concentração de NaCl, esperava-se que NahK ficasse totalmente retida na coluna. Amostras coletadas durante a purificação foram analisadas por SDS-PAGE (Figura 39 B). Mesmo com a diminuição de NaCl na amostra, observa-se que uma grande parte da amostra não interage com a

coluna, e que uma porção menor é eluída durante o gradiente crescente de NaCl que se inicia nos 30 ml do volume de eluição (Figura 39 A). Através do gel verificase a presença de NahK em grande quantidade tanto nas frações que não interagiram com a coluna, quanto nas frações correspondentes ao pico durante a eluição. Visto que essa etapa de cromatografia de troca iônica não estava gerando resultados satisfatórios, ela foi removida das estratégias para purificação de NahK.

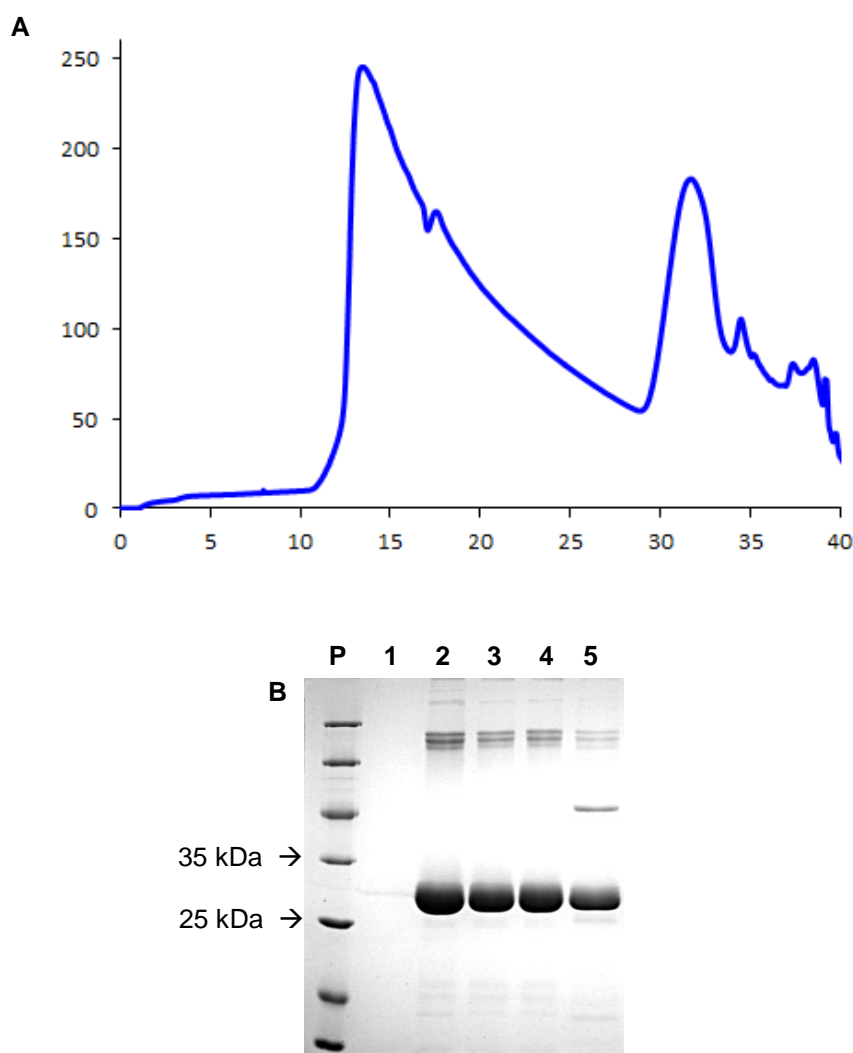


Figura 39: Purificação da proteína NahK por cromatografia de troca iônica.

A - Cromatograma obtido durante a purificação. A curva em azul representa a absorvância a 280 nm medida em mAU em função do volume de eluição da coluna em mL.

B – SDS-PAGE contendo algumas frações da purificação. **P** - padrão de proteínas Fermentas; **1 a 3** - frações que não interagiram com a coluna correspondentes ao volume de eluição de 0 a 30 mL; **4 e 5** - frações provenientes da eluição correspondentes ao volume de eluição de 30 a 40 mL.

As frações contendo NahK foram novamente unidas e submetidas a uma nova purificação por cromatografia de afinidade com o intuito de concentrar a amostra, que estava diluída após as tentativas de purificação por troca iônica. O cromatograma dessa etapa está representado na Figura 40. Posteriormente, as frações do pico de eluição foram unidas, dialisadas contra tampão Tris-HCl 50 mM pH 7,4, NaCl 50 mM e dosadas em espectrofotômetro, atingindo uma concentração de aproximadamente 3,45 mg/mL (9 mL totais, correspondendo a aproximadamente 31 mg de proteína). Parte da amostra foi utilizada nos testes de clivagem pela protease TEV.

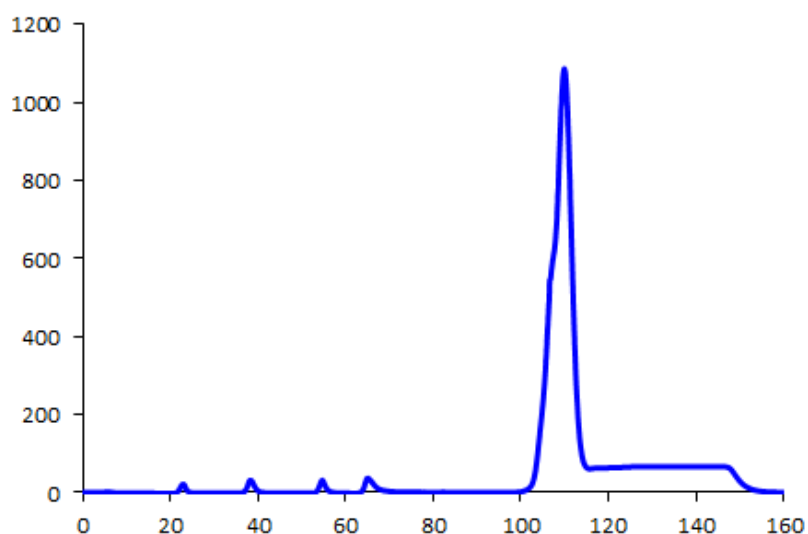


Figura 40: Cromatografia de afinidade ao níquel de NahK para concentração da amostra.

A curva em azul representa a absorbância a 280 nm medida em mAU em função do volume de eluição da coluna em mL.

5.13 Clivagem de NahK pela protease TEV

Alíquotas contendo 0,1 mg de NahK foram submetidas a clivagem com 0,005 mg (5% m/m) e 0,02 mg (20% m/m) de TEV, por 3, 6 e 16 h a 25°C. Após o tempo de clivagem, as amostras foram incubadas em banho maria a 100°C e misturadas com matriz para serem analisadas por MALDI-TOF.

A massa teórica de NahK fusionada à cauda de histidina é de 30,725 kDa, e de 28,919 kDa clivada pela TEV. Pelas análises dos perfis de massas obtidos (Figura 41), a clivagem por 3 e 6 h utilizando 5% e 20% de TEV apresentaram resultados semelhantes, apresentando picos de massa de aproximadamente 28,890 kDa. A clivagem utilizando 5% e 20% por 16 h apresentou picos de massas menores que o esperado, de 28,770 kDa e 25,829 kDa. A concentração de 5% de TEV por 3 h foi selecionada para a proteólise em maior escala.

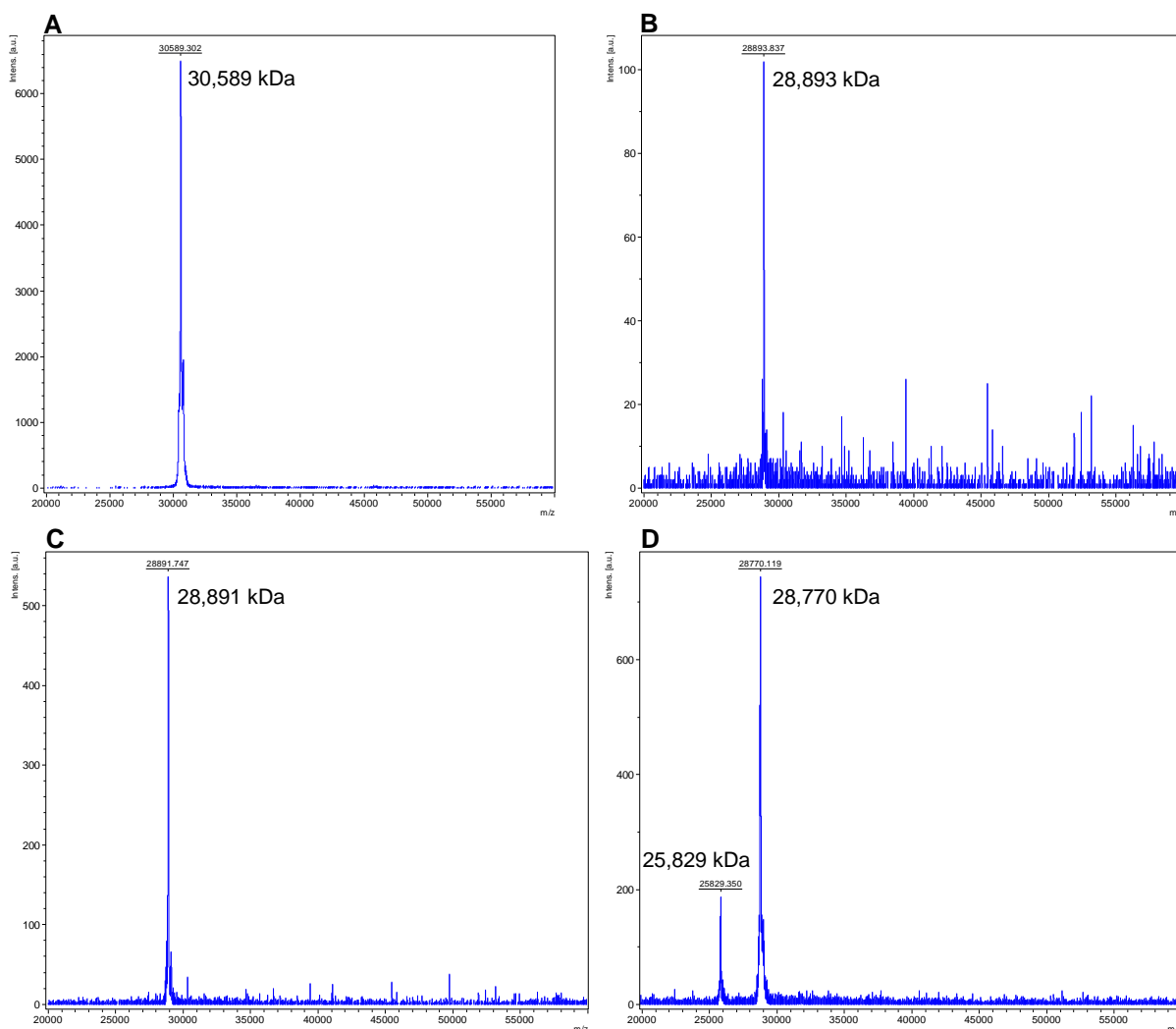


Figura 41: Perfis de massas da clivagem de NahK utilizando 5% de TEV.

A - NahK fusionada com a cauda de histidina; **B**, **C** e **D** - Clivagem utilizando 5% de TEV por 3, 6 e 16 h respectivamente.

5.14 Purificação de NahK sem cauda de histidina

Com o objetivo de separar a protease TEV e NahK recombinante, agora sem cauda de histidina, uma segunda etapa de purificação foi realizada por cromatografia de afinidade ao níquel. Uma vez que a protease TEV possui a cauda de histidina, ela ficará retida na coluna juntamente com algumas moléculas de NahK eventualmente não clivadas, assim como os contaminantes que interagiram com a coluna durante a primeira etapa de cromatografia de afinidade.

Uma quantidade maior de NahK, aproximadamente 15 mg, foi clivada com 5% m/m de TEV a 25°C por 3 h. Uma pequena quantidade da amostra não clivada e clivada de NahK, assim como uma amostra da protease TEV foi analisada por MALDI-TOF (Figura 42).

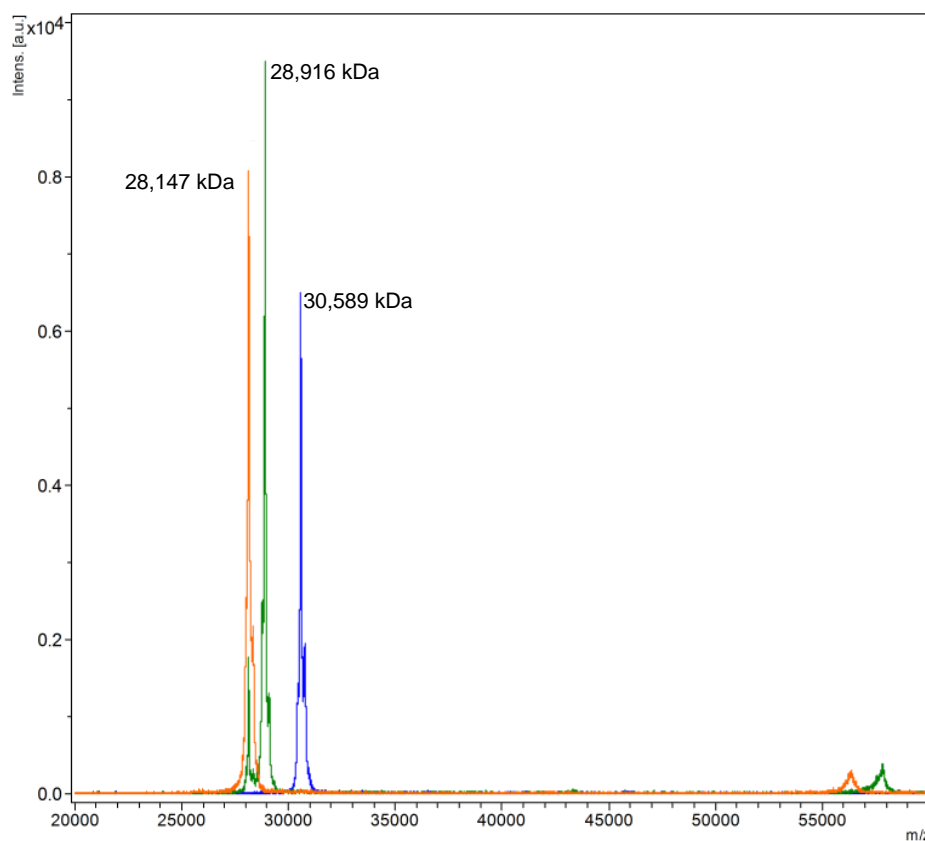


Figura 42: Perfil de massas de NahK não clivada (curva azul), clivada (curva verde), e protease TEV (curva vermelha).

Pelo perfil de massas é possível observar o pico correspondente a massa de NahK não clivada (curva azul, massa teórica de 30,725 kDa, massa observada de 30,589 kDa) e a massa da protease TEV (curva laranja, massa teórica de 28,137 kDa, massa observada de 28,147 kDa). Após a clivagem é possível verificar a diminuição de massa de NahK (pico maior, curva verde, massa teórica de 28,919 kDa, massa observada de 28,916 kDa) e visualizar também a presença da protease TEV (pico menor, curva verde).

A amostra contendo a proteína clivada foi injetada em coluna previamente equilibrada com tampão Tris-HCl 50 mM pH 7,4, NaCl 50 mM, imidazol 30 mM. Pelo cromatograma pode ser observado que a maior parte da fração protéica não interagiu com a coluna (Figura 43 A). Pequenos picos foram observados durante a eluição. Algumas frações foram analisadas por SDS-PAGE (Figura 43 B).

Como esperado, após a clivagem, a proteína NahK recombinante não mais interagiu com a coluna de afinidade (canaletas 1 a 3). Os picos obtidos durante a eluição da coluna (correspondentes às bandas nas canaletas 4 a 7) são, provavelmente, de NahK não clivada, protease TEV e contaminantes.

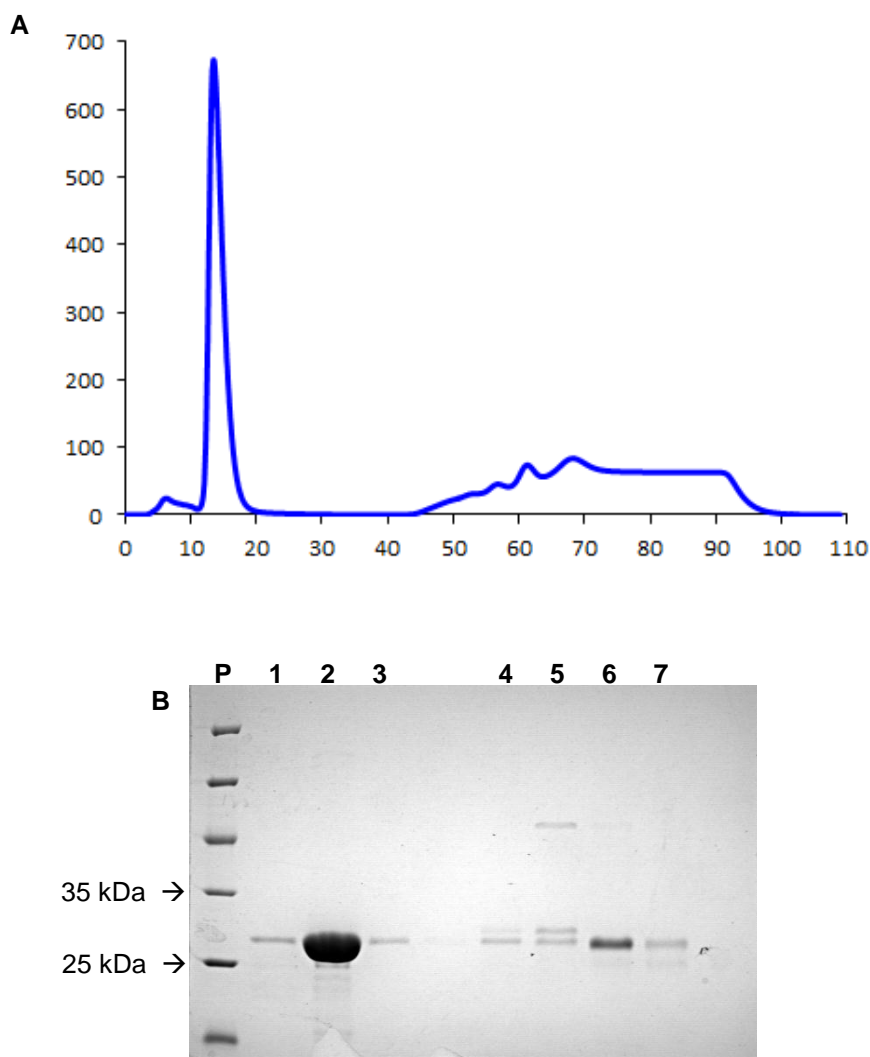


Figura 43: Purificação da proteína NahK clivada por cromatografia de afinidade.

A - Cromatograma obtido durante a purificação. A curva em azul representa a absorvância a 280 nm medida em mAU em função do volume de eluição da coluna em mL.

B - SDS-PAGE contendo algumas frações da purificação. **P** - padrão de proteínas Fermentas; **1 a 3** - frações que não interagiram com a coluna correspondentes aos volumes de eluição de 0 a 20 mL; **4 a 7** - frações provenientes da eluição com imidazol correspondentes ao volume de eluição de 45 a 75 mL.

A fração contendo NahK clivada foi submetida por fim a uma cromatografia de exclusão molecular, onde as proteínas são separadas por massa. Para isso 500 μ L de NahK clivada foram injetados em coluna Superdex 200 10/300 GL (GE Healthcare) previamente equilibrada com tampão Tris-HCl 50 mM pH 7,4, NaCl 50 mM. O perfil cromatográfico e o gel de poliacrilamida-SDS das frações coletadas durante a purificação estão representados na Figura 44. Através do volume de

retenção do pico foi possível estimar que NahK se comporta em solução como uma partícula de massa de 30 kDa, indicando que NahK é monomérica

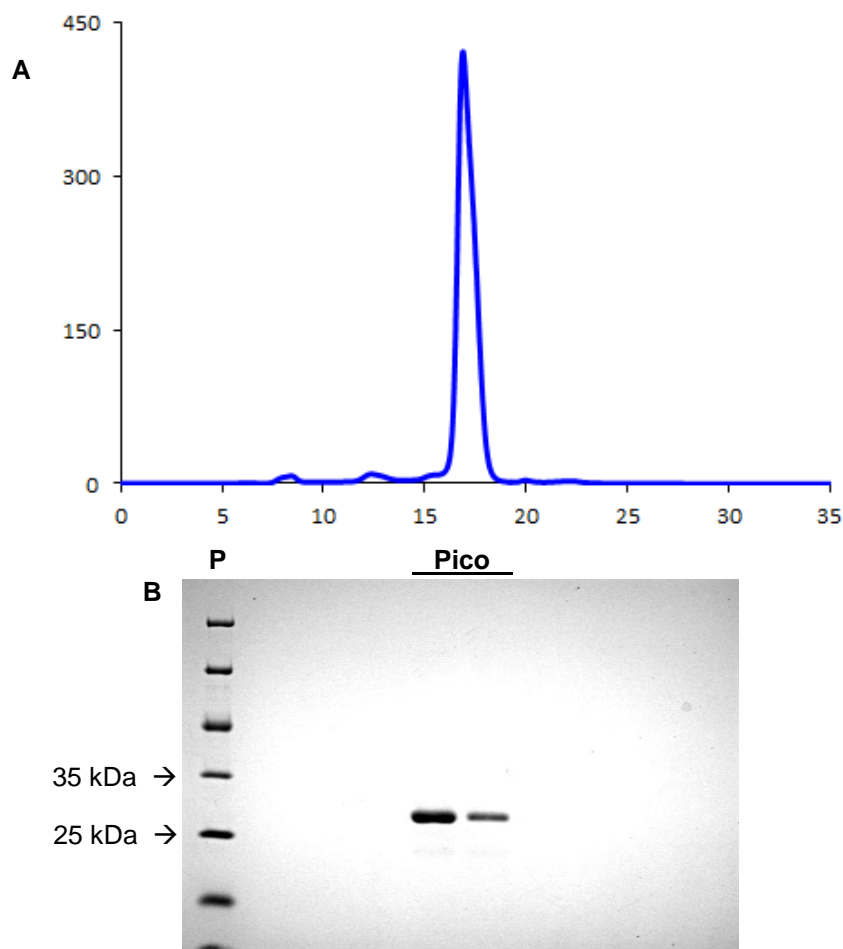


Figura 44: Purificação da proteína NahK clivada por cromatografia de exclusão molecular.

A - Cromatograma obtido durante a purificação. A curva em azul representa a absorvância a 280 nm medida em mAU em função do volume de eluição da coluna em mL.

B - SDS-PAGE contendo algumas frações da purificação. **P** - padrão de proteínas Fermentas; **Pico** - frações correspondentes ao pico principal obtido durante a purificação.

Com todas as etapas de expressão, purificação e clivagem pela TEV estabelecidas para NahK a proteína foi expressa em 2 L de meio seletivo LB líquido a 30°C por 5 horas sob agitação de 200 rpm, induzida por IPTG em concentração de 0,5 mM. A lise foi feita em 40 mL de tampão fosfato de sódio 20 mM pH 7,4, NaCl 500 mM. A fração solúvel foi purificada por cromatografia de afinidade ao níquel. Um pico de proteína foi obtido durante a eluição da coluna e suas frações foram unidas e dialisadas contra tampão Tris-HCl 50 mM pH 7,4, NaCl 100 mM para remover

imidazol e diminuir quantidade de NaCl. Após diálise, a amostra foi centrifugada e a fração solúvel dosada em espectrofotômetro, sendo sua concentração estimada em 12,8 mg/mL. A quantidade total obtida de NahK foi de aproximadamente 120 mg, sendo que 40 mg de proteína foram utilizados na clivagem por protease TEV na proporção de 5%, a 25°C por 3 horas. A amostra, após o tempo de clivagem, foi submetida a um segundo passo de cromatografia de afinidade. A fração correspondente à NahK clivada e purificada foi dialisada contra tampão Tris-HCl 50 mM pH 7,4, NaCl 50 mM para remoção do imidazol e diminuir a concentração de NaCl. Posteriormente, a amostra de NahK clivada, assim como a amostra não clivada, foram concentradas e purificadas através de cromatografia de exclusão molecular. As frações purificadas das proteínas NahK clivada e não clivada foram concentradas até chegar a uma concentração de 11 a 12 mg/mL, estando assim prontas para os ensaios de DLS, CD e testes de cristalização.

5.15 Caracterização do estado de agregação de NahK

Os dados obtidos no ensaio de DLS da proteína NahK estão representados na Tabela 11. Pode se observar que 99,4% da massa é formada por moléculas com raio médio de 2,4 nm e massa molecular de aproximadamente 25 kDa. A massa teórica de NahK clivada é de 28,919 kDa, portanto, pode-se inferir que a proteína apresenta-se como monômero em solução. Este dado confirma o resultado obtido através da cromatografia de exclusão molecular, que indicou, pelo volume de retenção do pico de NahK, que a proteína comportava-se como uma partícula de aproximadamente 30 kDa.

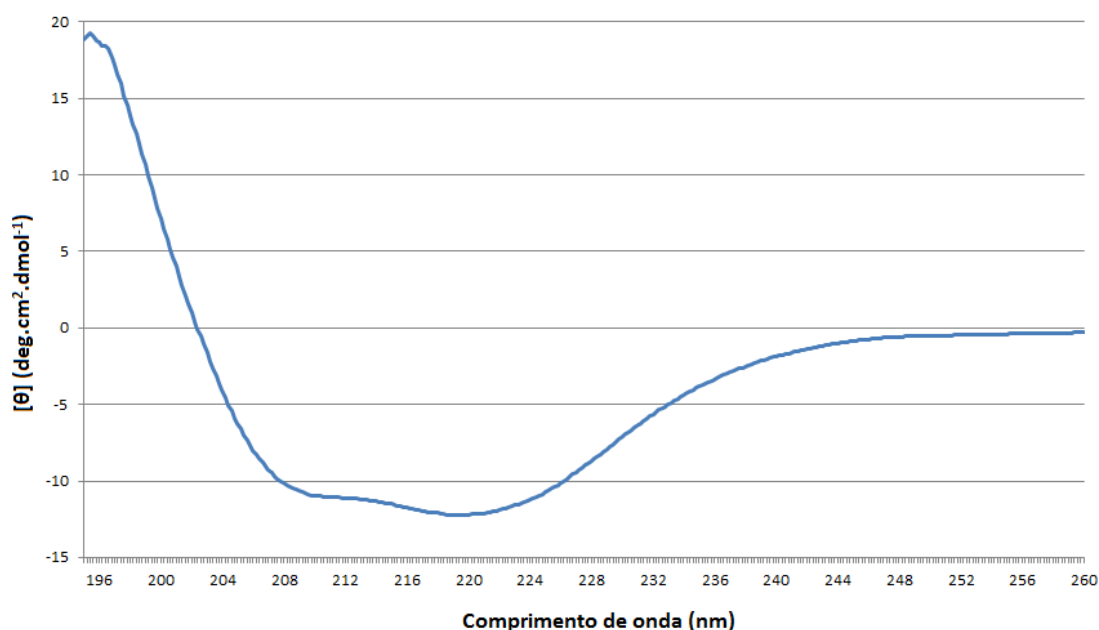
A polidispersividade de 99,4% da massa das moléculas foi inferior a 1%, indicando que NahK encontra-se monodispersa, ou seja, a maior parte das proteínas NahK na solução são partículas com o mesmo tamanho. Esse resultado sugere que a amostra está adequada para os ensaios de cristalização.

Tabela 11: Análise da proteína NahK por espalhamento dinâmico de luz.

Raio (nm)	% de Polidispersividade	Massa (kDa)	% de Massa
2,4	< 1	25	99,4
22,1	< 1	4.705	0,1
190,1	12,5	723.896	< 0,1

5.16 Caracterização da estrutura secundária de NahK

A estrutura secundária de NahK foi caracterizada através de Dicroísmo Circular. O espectro obtido para a proteína em Tris-HCl 5 mM pH 7,4, NaCl 5 mM a 20°C está representado na Figura 45. O gráfico foi gerado com a elipicidade molar ($\text{deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$) em função do comprimento de onda (nm). Através desse perfil é possível concluir que NahK encontra-se enovelada com estruturas em hélices- α e fitas- β , devido à presença das bandas negativas em 222, 218 e 208 nm e positivas próximas de 196 nm.

**Figura 45:** Espectro de Dicroísmo Circular da proteína NahK em Tris-HCl pH 7,4 5mM, NaCl 5mM a 20°C.

Através da análise do espectro obtido durante a desnaturação da proteína de 20 a 90°C, em intervalos de 5°C, verifica-se a perda de estrutura secundária à medida que a temperatura é elevada (Figura 46). Ocorre uma mudança perceptível no padrão do espectro entre 60 e 70°C, indicando que a desnaturação em maior grau de NahK ocorre nessa faixa de temperatura.

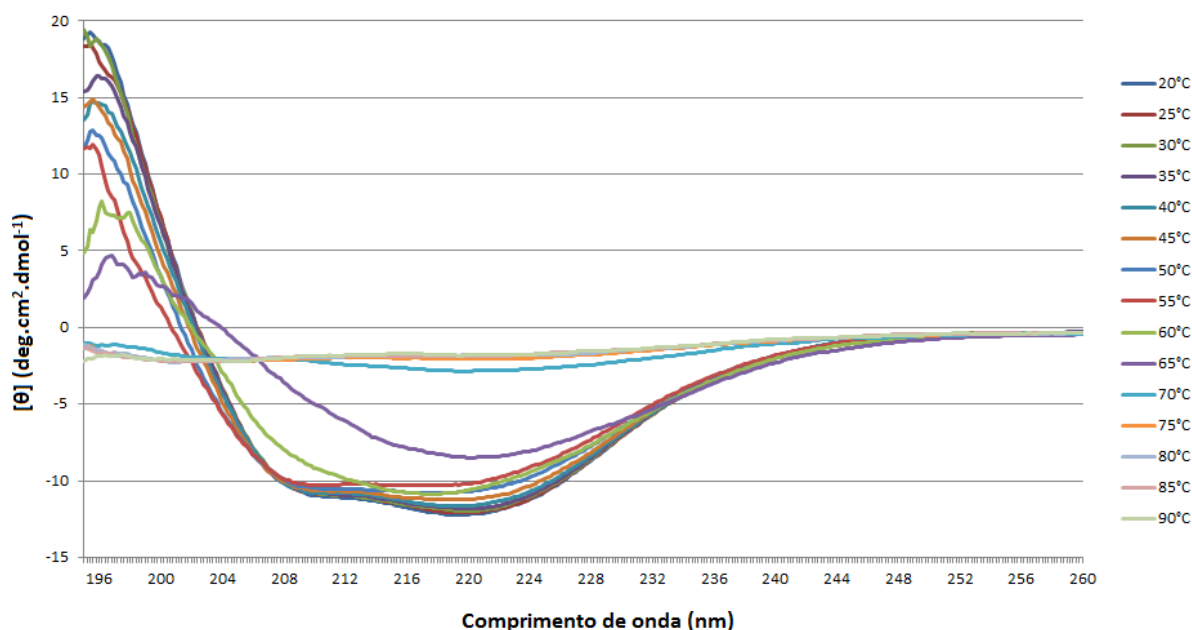


Figura 46: Espectro de Dicroísmo Circular da proteína NahK em Tris-HCl pH 7,4 5mM, NaCl 5mM em diferentes temperaturas.

A deconvolução dos dados brutos obtidos do CD foi feita com o programa CDNN 2.1. Com os dados da deconvolução, foi gerado um gráfico de barras exibindo a porcentagem das estruturas secundárias em função da temperatura (Figura 47). Pode-se observar a perda de estruturas em α -hélice e o ganho de estruturas aleatórias e em folhas- β com o aumento da temperatura. O ganho de estruturas aleatórias juntamente com a perda de α -hélices é evidenciado no intervalo entre as temperaturas de 60 e 70°C.

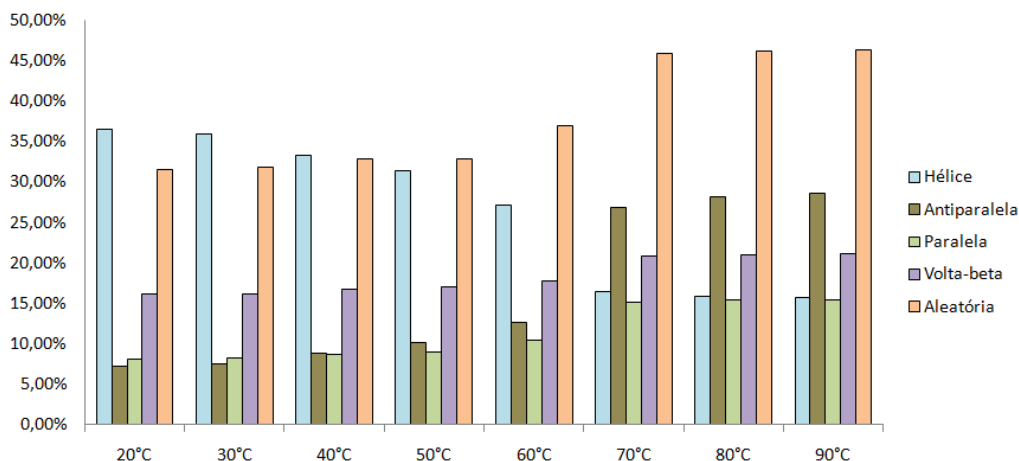


Figura 47: Deconvolução dos dados do espectro da Figura 46 feito pelo programa CDNN 2.1.

A porcentagem de estruturas hélices- α a 20°C foi de aproximadamente 37%, de fitas- β de aproximadamente 32% e de estruturas randômicas de aproximadamente 32%. Esse resultado é consistente com a porcentagem de resíduos de aminoácidos presentes nas estruturas secundárias da proteína HpcG de *E. coli* [55] que possui 41% de identidade com NahK.

5.17 Ensaios de cristalização de NahK

A maioria das 96 condições de cristalização testadas utilizando uma concentração de 12 mg/mL de NahK gerou um precipitado amorfo constituído de proteína insolúvel e provavelmente desenovelada. Entretanto, em algumas condições, surgiram nucleações com a formação de microcristais; em outras, pequenos monocristais foram observados (Figura 48). Melhores resultados foram obtidos nas gotas que continham a proporção de 2 μ L de proteína para cada μ L da condição de cristalização. Em alguns casos houve nucleação e crescimento de pequenos cristais tanto em amostras de NahK clivada, quanto com NahK não clivada, evidenciando que, em algumas condições, a cauda de histidina não foi um empecilho para a cristalização. Alguns cristais cresceram em dois dias, enquanto outros levaram cerca de 3 a 4 semanas.

As condições com formação de microcristais ou pequenos monocristais foram amplificadas variando-se a concentração de precipitante e o pH do tampão. Por

exemplo, a condição nº 22 do kit Crystall Screen 1 é acetato de sódio trihidratado 200 mM, Tris-HCl 100 mM pH 8,5, 30% PEG 4000. Dessa maneira, foram realizados ensaios de cristalização utilizando acetato de sódio trihidratado 200 mM, Tris-HCl 100 mM pH 6,5 a pH 10,5, e PEG 4000 nas concentrações de 18 a 34%. Com esse método, a partir de uma condição de cristalização, são geradas outras 24. Para a condição utilizada como exemplo, monocristais maiores e menos ramificados foram obtidos em pH 8,5 a 10,5 e concentração de 26 a 30% de PEG 4000.

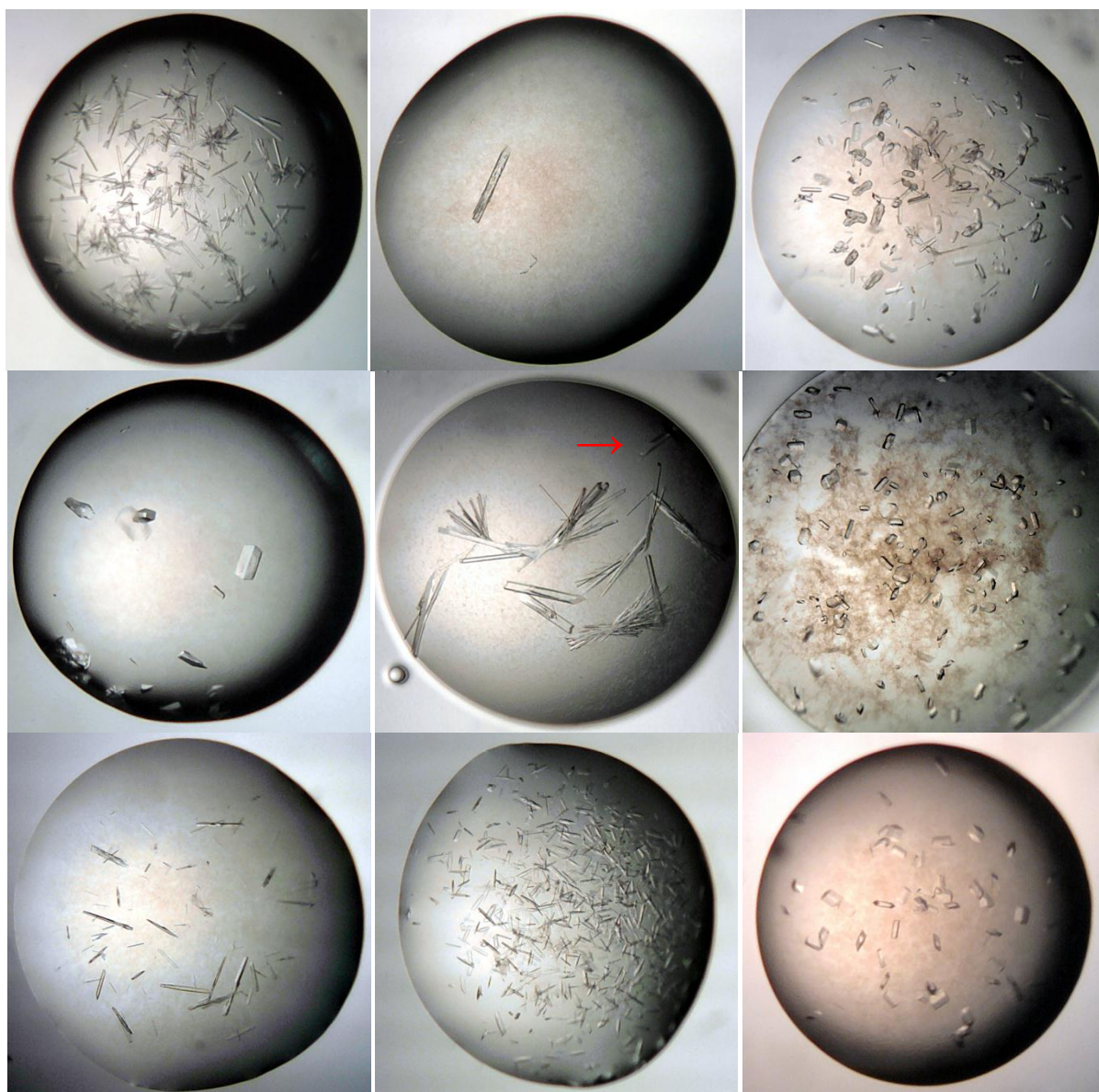


Figura 48: Cristais de NaH₂PO₄ em condições promissoras de cristalização.

Alguns cristais foram submetidos à difração de raios-X na linha de luz MX1 do Laboratório Nacional de Luz Sincrotron. A maioria dos cristais obtidos dos ensaios é de proteína, exibindo padrões de difração com centenas de reflexões. Alguns difrataram muito bem, gerando reflexões intensas em resolução superior a 2.0 angstroms. O padrão de difração de um cristal de NahK obtido da condição de cristalização acetato de sódio 200 mM, Tris-HCl 100 mM pH 10,5, 30% polietilenoglicol 4000 (seta vermelha Figura 48) está representado na Figura 49.

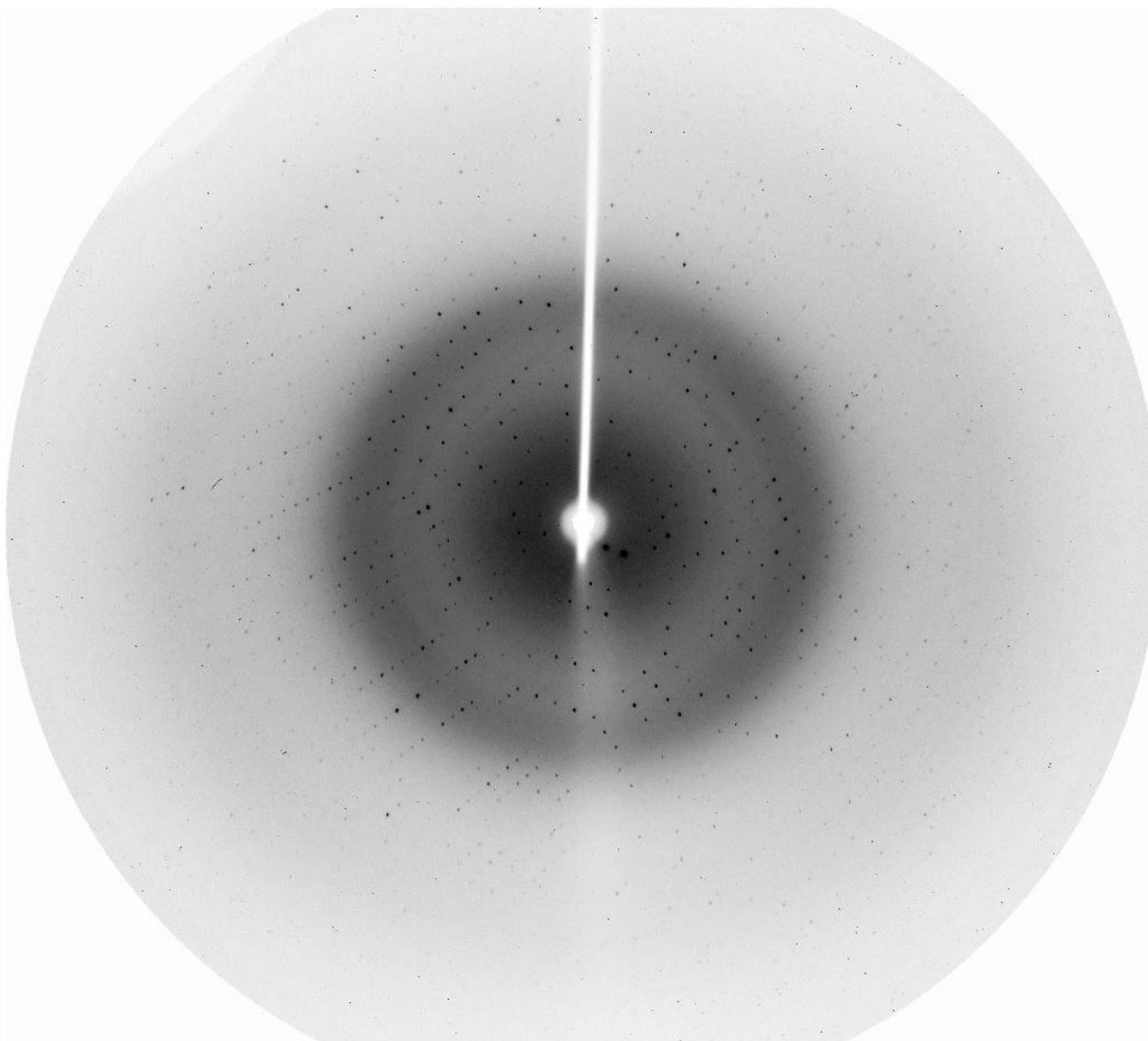


Figura 49: Padrão de difração de cristal de NahK.

Podem-se notar centenas de reflexões de intensidades diferentes do centro à borda da imagem (reflexões correspondentes a dados de baixa e de alta resolução, respectivamente).

Para esse cristal, foi coletado um conjunto de 477 imagens de 30° a 507°. Cada imagem foi obtida expondo o cristal ao feixe de raios-X por 45 segundos, rotacionado de 1° durante esse tempo. O detector ficou a uma distância de 85 mm do cristal. As imagens foram analisadas pelo programa iMosflm, sendo o cristal identificado como pertencente ao grupo espacial monoclinico $P2_1$. Outros dados obtidos para esse conjunto de imagens estão representados na Tabela 12.

Tabela 12: Dados coletados do cristal de NahK. Dados de alta resolução estão exibidos entre parênteses.

Número de imagens	477
Oscilação (°)	1,0
Grupo espacial	$P2_1$
Parâmetros de célula unitária (Å, °)	a=77,49; b=43,75; c=79,49; β =118,21
Resolução (Å)	19,97 - 1,72 (1,82 – 1,72)
Nº total de reflexões	481.814 (68.068)
Nº de reflexões únicas	50.040 (7.284)
Redundância	9,6 (9,3)
Rmerge (%)	7,1 (48,5)
Completeza (%)	99,9 (100,0)
$\langle I \rangle / \langle \sigma(I) \rangle$	24,2 (4,9)
Mosaicidade (°)	0,51

Os dados de difração foram obtidos no último trimestre de 2010, bem próximo da data sugerida para término desta dissertação. Assim, os passos seguintes de determinação estrutural com base nos dados coletados serão desenvolvidos nos próximos meses. Os dados serão usados em procedimentos de Substituição Molecular utilizando a proteína HpcG de *E. coli* [55] como modelo de busca (código PDB 2EB4). Estas proteínas compartilham 41% de identidade.

6. DISCUSSÃO

O petróleo é um dos recursos naturais do qual a sociedade é bastante dependente. A exploração e utilização de seus derivados acaba por contaminar o ambiente com grandes quantidades de hidrocarbonetos aromáticos policíclicos. Diante da toxicidade desses compostos torna-se evidente a necessidade de remediação das áreas contaminadas. Diversos microrganismos adquiriram a capacidade de utilizar esses poluentes como fonte de energia. A biorremediação, como excelente alternativa, utiliza desses microrganismos ou suas enzimas para degradar os contaminantes à moléculas inertes como CO₂ e H₂O.

A bactéria *Pseudomonas putida* G7 se destaca por possuir um plasmídeo que codifica enzimas capazes de degradar completamente o naftaleno, que é um dos compostos tóxicos presentes nos derivados de petróleo. É importante o estudo aprofundado das enzimas participantes dessa via de degradação a fim de utilizá-las futuramente como biorremediadores.

A fim de entender melhor os processos biológicos envolvidos nas interações proteína-proteína e proteína-ligantes para o reconhecimento de substratos e atividades enzimáticas, a cristalografia de proteínas tem-se mostrado de extrema importância. Através dessa técnica, é permitindo identificar detalhes importantes em nível atômico das estruturas das proteínas e associá-los com suas funções específicas.

Para que as enzimas sejam caracterizadas estruturalmente, o uso de proteínas recombinantes torna-se essencial já que são necessárias grandes quantidades de proteína para realizar os ensaios de cristalização e para o crescimento de cristais de qualidade para difração e coleta de dados. As vantagens de se usar um sistema de expressão heterólogo incluem um grande rendimento na obtenção da proteína, facilidade em purificá-la por meio de caudas de afinidade e possibilidade de incluir modificações na cadeia polipeptídica ou expressão de domínios isolados.

Diante da problemática das contaminações por HAPs, a possibilidade do uso da biorremediação para controle desses contaminantes e a necessidade de caracterização estrutural das enzimas da via de degradação do naftaleno, as enzimas NahE e NahK foram selecionadas como alvo para estudos estruturais neste trabalho.

As predições *in silico* sugeriram a ausência de domínios de hélices transmembrana para as duas proteínas, sendo assim, os iniciadores para amplificação dos genes foram desenhados de forma a amplificá-los em sua totalidade para que a expressão das proteínas recombinantes fossem mais semelhante possível de suas sequências de resíduos nativas.

Os genes *nahE* e *nahK* foram amplificados utilizando temperaturas de anelamento próximas das T_m previamente estabelecidas durante o desenho dos iniciadores. Embora os sequenciamentos de alguns clones (tanto para *nahE*, quanto para *nahK*) permanecessem inconclusivos em algumas regiões dos genes, e em alguns clones houvessem mutações, foi possível confirmar a janela de leitura e que as construções trabalhadas correspondem às sequências gênicas alvo.

Inicialmente, NahE recombinante foi expressa na fração insolúvel a 37°C em *E. coli* BL21(DE3) utilizando o vetor pET32a-TEV. Diversos testes de expressão em temperaturas mais baixas foram realizados na tentativa de obter NahE expressa em fração solúvel, sem sucesso. Normalmente, a expressão de proteínas solúveis corresponde o primeiro gargalo em todo o processo de obtenção de proteínas recombinantes para estudos estruturais.

Devido à insolubilidade de NahE em BL21(DE3), foram realizados testes de expressão em *E. coli* ArticExpress(DE3) utilizando o vetor pET32a-TEV. Também foram realizados testes utilizando o vetor pMCSG19C expresso tanto em *E. coli* BL21(DE3) quanto em ArticExpress(DE3). Apesar da obtenção de pequenas quantidades de NahE em fração solúvel utilizando ArticExpress, verificou-se que, com o uso dessa cepa tanto com a expressão no vetor pET28a-TEV, quanto com a expressão no vetor pMCSG19C, ocorreu uma falsa solubilidade de NahE ao formar complexos de chaperoninas com a proteína recombinante.

O processo de reenovelamento possibilitou a obtenção de NahE em fração solúvel e sua purificação parcial, entretanto, foi possível identificar por cromatografia de exclusão molecular que NahE reenovelada permanecia solúvel na forma de um grande aglomerado, impossibilitando assim sua utilização para estudos estruturais.

Apesar dos resultados de solubilidade indesejáveis para NahE, foi possível confirmar a grande eficiência do sistema de Clonagem Independente de Ligação,

sendo que o processo, desde a amplificação do gene até a obtenção de clones expressando a proteína, foi excepcionalmente rápido.

O gene *nahK* foi clonado em pET28a-TEV e a proteína recombinante foi expressa parcialmente solúvel em *E. coli* BL21(DE3) a 30°C. Mesmo sendo observada grande quantidade de proteína recombinante presente na fração insolúvel, o rendimento de proteína solúvel após primeira etapa de cromatografia de afinidade atingiu aproximadamente 60 mg/L de expressão.

A proteína NahK apresenta um *pI* teórico de 5,18. Os tampões utilizados tanto para as etapas de lise das células de expressão quanto para as etapas de purificação utilizam pH 7,5. NahK em solução no pH dos tampões utilizados apresenta uma carga líquida negativa, possibilitando o uso de troca aniônica como estratégia para purificação. Apesar disso, as tentativas de purificação de NahK utilizando esse tipo de cromatografia não se mostraram efetivas, sendo adotada apenas duas cromatografia de afinidade para purificação.

Após a primeira etapa de purificação por afinidade e clivagem com protease TEV, a proteína recombinante NahK permaneceu com 5 resíduos de aminoácidos (GHMAS) além de sua sequência nativa. O peso molecular da enzima foi determinado por espectrometria de massa e o valor registrado foi bastante próximo do teórico. Um elevado nível de pureza foi obtido após duas etapas de cromatografia de afinidade.

Algumas características estruturais de NahK recombinante foram determinadas por Espalhamento Dinâmico de Luz e Dicroísmo Circular. Os ensaios de DLS indicaram a monodispersividade das amostras de NahK e uma estrutura monomérica em solução.

Através dos ensaios de CD foi possível verificar a presença de estruturas secundárias em NahK recombinante. A curva de desnaturação apresentou estabilidade até aproximadamente 60°C, sendo que a 65°C ocorre uma mudança abrupta na forma da curva, indicando grandes mudanças estruturais. Através da deconvolução dos dados foi possível confirmar a presença de estruturas secundárias na proteína, que apresenta hélices- α , fitas- β e regiões aleatórias. A deconvolução dos dados da curva de desnaturação evidencia uma perda acentuada

de estruturas em hélices- α bem como o aumento de estruturas randômicas e em fitas- β à medida que ocorria o aumento de temperatura.

Os resultados dos géis de SDS-PAGE após as etapas de purificação, juntamente com os ensaios de DLS da proteína NahK evidenciaram que a amostra estava adequada aos testes de cristalização, apresentando elevado grau de pureza e monodispersividade.

Os testes de cristalização foram realizados utilizando 96 condições precipitantes. Os resultados desse ensaio foram bastante promissores sendo identificados diversas condições onde houveram nucleações e crescimento de pequenos cristais. A partir dessas condições promissoras, outras condições de cristalização foram testadas, sendo que houve um rápido crescimento de monocristais na solução de acetato de sódio 200 mM, Tris-HCl 100 mM pH 10,5, 30% polietilenoglicol 4000.

A difração de vários cristais provindos dessa condição foi testada e obtiveram-se padrões de difração característicos de cristais de proteína. O cristal que apresentou maior poder de difração, com reflexões intensas em alta resolução, foi utilizado para a coleta de um conjunto de dados. A partir desse conjunto será possível, através de Substituição Molecular, a resolução da estrutura tridimensional de NahK recombinante.

7. CONCLUSÃO

- A proteína recombinante NahE apresentou-se insolúvel quando expressa em *E. coli* BL21(DE) em diversas temperaturas.
- Quando expressa em *E. coli* ArticExpress(DE3), NahE recombinante formou complexos com chaperoninas, que mascararam sua insolubilidade. Não foi possível obter NahE solúvel livre de chaperoninas.
- O processo de reenovelamento não se mostrou efetivo para a obtenção de NahE recombinante em forma solúvel e monodispersa.
- A proteína recombinante NahK apresentou-se parcialmente solúvel quando expressa em *E. coli* BL21(DE3) a 30°C.
- A massa molecular da proteína recombinante NahK, após clivagem da cauda de afinidade, determinada por espectrometria de massa MALDI-TOF, foi bastante próxima do valor teórico esperado.
- Em solução, NahK recombinante apresentou-se monodispersa e monomérica.
- Foi possível obter grandes quantidades de NahK recombinante com elevado grau de pureza utilizando duas etapas de cromatografia de afinidade.
- A proporção de estruturas secundárias de NahK recombinante é de aproximadamente 37% de hélices- α , 32% de fitas- β e 32% de estruturas randômicas. Esses valores foram próximos dos valores observados nas estruturas secundárias da proteína HpcG de *E. coli* que possui 41% de identidade com NahK.
- A condição de cristalização contendo acetato de sódio 200 mM, Tris-HCl 100 mM pH 10,5, 30% polietilenoglicol 4000 permite o crescimento de monocristais de NahK que difratam em resoluções maiores que 2 angstroms.

8. PERSPECTIVAS

- Concluir a determinação estrutural de NahK recombinante por meio da técnica de Substituição Molecular;
- Estabelecer colaborações internacionais com grupos capazes de produzir os substratos específicos de NahK para teste de atividade enzimático;
- Realizar testes de atividade de NahK com moléculas análogas ao substrato;
- Resolver a estrutura tridimensional de NahK complexada com ligantes análogos ao substrato;
- Clonar e coexpressar os genes *nahK* e *nahL* a fim de se obter cristais para a resolução da estrutura tridimensional desse complexo;
- Medir a interação de NahK e análogos do substrato por Calorimetria de Titulação Isotérmica;
- Utilizar os produtos da reação enzimática de NahE para auxiliar em seu reenovelamento e estabilidade da estrutura;
- Clonar o gene *nahE* em vetores de expressão pMCSG9 e pMCSG10, que utiliza caudas de fusão amino-terminal de Proteína Ligadora de Maltose e Glutathione S-Transferase;
- Utilizar outras cepas bacterianas para expressão de NahE recombinante.

9. REFERÊNCIAS BIBLIOGRÁFICAS

1. SOLOMONS, T.W.G. and C.B. FRYHLE, *Química orgânica*. 7ª ed 2001: ed. Rio de Janeiro.
2. VAN HAMME, J.D., A. SINGH, and O.P. WARD, *Recent advances in petroleum microbiology*. *Microbiol Mol Biol Rev*, 2003. **67**(4): p. 503-49.
3. ANTOS, G.J. and A.M. AITANI, *Catalytic Naphtha Reforming*. 2th ed, 2004, New York: CRC Press.
4. YEUNG, P.Y., R.L. JOHNSON, and J.G. XU, *Biodegradation of petroleum hydrocarbons in soil as affected by heating and forced aeration*. *Journal of Environmental Quality*, 1997. **26**(6): p. 1511-1516.
5. NETTO, A.D.P., et al., *Avaliação da contaminação humana por hidrocarbonetos policíclicos aromáticos (HPAs) e seus derivados nitrados (nHPAs): uma revisão metodológica*. *Química Nova*, 2000. **23**(6): p. 765-773.
6. PÉREZ, S., et al., *Occurrence of polycyclic aromatic hydrocarbons in sewage sludge and their contribution to its toxicity in the toxalert 100 bioassay*. *Chemosphere*, 2001. **45**(6-7): p. 705-12.
7. BRÄNDLI, R., et al., *Occurrence and relevance of organic pollutants in compost, digestate and organic residue*. EPF Lausanne - ENAC - ISTE - CECOTOX / Agroscope FAL Reckenholz, 2004.
8. JACQUES, R.J.S., F.M. BENTO, and F.A.O. CAMARGO, *Biodegradação de hidrocarbonetos aromáticos policíclicos*. *Ciência e Natura*, 2007. **29**(1): p. 7-24.
9. JOHNSEN, A.R., L.Y. WICK, and H. HARMS, *Principles of microbial PAH-degradation in soil*. *Environ Pollut*, 2005. **133**(1): p. 71-84.
10. BAMFORTH, S.M. and I. SINGLETON, *Bioremediation of polycyclic aromatic hydrocarbons: current knowledge and future directions*. *Journal of Chemical Technology and Biotechnology*, 2005. **80**: p. 723–736.
11. TIBURTIUS, E.R.L., P. PERALTA-ZAMORA, and E.S. LEAL, *Contaminação de águas por BTXs e processos utilizados na remediação de sítios contaminados*. *Química Nova*, 2004. **27**(3): p. 441-446.

12. KRZEMINSKI, J., et al., *Synthesis of Fjord region diol epoxides as potential ultimate carcinogens of dibenzo[a,l]pyrene*. Chem Res Toxicol, 1994. **7**(2): p. 125-9.
13. STANSBURY, K.H., J.W. FLESHER, and R.C. GUPTA, *Mechanism of aralkyl-DNA adduct formation from benzo[a]pyrene in vivo*. Chem Res Toxicol, 1994. **7**(2): p. 254-9.
14. BAIRD, W.M., L.A. HOOVEN, and B. MAHADEVAN, *Carcinogenic polycyclic aromatic hydrocarbon-DNA adducts and mechanism of action*. Environ Mol Mutagen, 2005. **45**(2-3): p. 106-14.
15. RUBIN, H., *Synergistic mechanisms in carcinogenesis by polycyclic aromatic hydrocarbons and by tobacco smoke: a bio-historical perspective with updates*. Carcinogenesis, 2001. **22**(12): p. 1903-30.
16. CERNIGLIA, C.E., *Fungal metabolism of polycyclic aromatic hydrocarbons: past, present and future applications in bioremediation*. J Ind Microbiol Biotechnol, 1997. **19**(5-6): p. 324-33.
17. TROQUET, J., C. LARROCHE, and C.-G. DUSSAP, *Evidence for the occurrence of an oxygen limitation during soil bioremediation by solid-state fermentation*. Biochemical Engineering Journal, 2003. **13**: p. 103-112.
18. EL FANTROUSSI, S. and S.N. AGATHOS, *Is bioaugmentation a feasible strategy for pollutant removal and site remediation?* Curr Opin Microbiol, 2005. **8**(3): p. 268-75.
19. PAUL, D., et al., *Assessing microbial diversity for bioremediation and environmental restoration*. Trends Biotechnol, 2005. **23**(3): p. 135-42.
20. ALCADE, M., et al., *Environmental biocatalysis: from remediation with enzymes to novel green processes*. TRENDS in Biotechnology, 2006. **24**(6): p. 281-287.
21. SUTHERLAND, T.D., et al., *Enzymatic bioremediation: from enzyme discovery to applications*. Clin Exp Pharmacol Physiol, 2004. **31**(11): p. 817-21.
22. DUA, M., et al., *Biotechnology and bioremediation: successes and limitations*. Appl Microbiol Biotechnol, 2002. **59**(2-3): p. 143-52.

23. HARITASH, A.K. and C.P. KAUSHIK, *Biodegradation aspects of polycyclic aromatic hydrocarbons (PAHs): a review*. J Hazard Mater, 2009. **169**(1-3): p. 1-15.
24. VALDMAN, E., et al., *On-line detection of low naphthalene concentrations with a bioluminescent sensor*. Process Biochemistry, 2004. **39**: p. 1217–1222.
25. STOHS, S.J., S. OHIA, and D. BAGCHI, *Naphthalene toxicity and antioxidant nutrients*. Toxicology, 2002. **180**(1): p. 97-105.
26. SOTA, M., et al., *Genomic and functional analysis of the IncP-9 naphthalene-catabolic plasmid NAH7 and its transposon Tn4655 suggests catabolic gene spread by a tyrosine recombinase*. J Bacteriol, 2006. **188**(11): p. 4057-67.
27. HABE, H. and T. OMORI, *Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria*. Biosci Biotechnol Biochem, 2003. **67**(2): p. 225-43.
28. RHODES, G., *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models*. 2^a ed2000: Academic Press.
29. BERMAN, H.M., et al., *The Protein Data Bank*. Nucleic Acids Res, 2000. **28**(1): p. 235-42.
30. BROWN, E.N., et al., *Determining Rieske cluster reduction potentials*. J Biol Inorg Chem, 2008. **13**(8): p. 1301-13.
31. CARREDANO, E., et al., *Substrate binding site of naphthalene 1,2-dioxygenase: functional implications of indole binding*. J Mol Biol, 2000. **296**(2): p. 701-12.
32. THOMPSON, L.C., et al., *2-Hydroxychromene-2-carboxylic acid isomerase: a kappa class glutathione transferase from Pseudomonas putida*. Biochemistry, 2007. **46**(23): p. 6710-22.
33. KITA, A., et al., *An archetypical extradiol-cleaving catecholic dioxygenase: the crystal structure of catechol 2,3-dioxygenase (metapyrocatechase) from Pseudomonas putida mt-2*. Structure, 1999. **7**(1): p. 25-34.

34. SUBRAMANYA, H.S., et al., *Enzymatic ketonization of 2-hydroxymuconate: specificity and mechanism investigated by the crystal structures of two isomerases*. *Biochemistry*, 1996. **35**(3): p. 792-802.
35. BANACHOWICZ, E., *Light scattering studies of proteins under compression*. *Biochim Biophys Acta*, 2006. **1764**(3): p. 405-13.
36. DEV, S. and A. SUROLIA, *Dynamic light scattering study of peanut agglutinin: size, shape and urea denaturation*. *J Biosci*, 2006. **31**(5): p. 551-6.
37. FERRÉ-D'AMARÉ, A.R. and S.K. BURLEY, *Use of dynamic light scattering to assess crystallizability of macromolecules and macromolecular assemblies*. *Structure*, 1994. **2**(5): p. 357-9.
38. RANJBAR, B. and P. GILL, *Circular dichroism techniques: biomolecular and nanostructural analyses- a review*. *Chem Biol Drug Des*, 2009. **74**(2): p. 101-20.
39. GREENFIELD, N.J., *Using circular dichroism spectra to estimate protein secondary structure*. *Nat Protoc*, 2006. **1**(6): p. 2876-90.
40. SREERAMA, N. and R.W. WOODY, *Computation and analysis of protein circular dichroism spectra*. *Methods Enzymol*, 2004. **383**: p. 318-51.
41. WILKINS, M.R., et al., *Protein identification and analysis tools in the ExPASy server*. *Methods Mol Biol*, 1999. **112**: p. 531-52.
42. GASTEIGER, E., et al., *ExPASy: The proteomics server for in-depth protein knowledge and analysis*. *Nucleic Acids Res*, 2003. **31**(13): p. 3784-8.
43. SCHEER, M., et al., *BRENDA, the enzyme information system in 2011*. *Nucleic Acids Res*, 2011. **39**(Database issue): p. D670-6.
44. EATON, R.W. and P.J. CHAPMAN, *Bacterial metabolism of naphthalene: construction and use of recombinant bacteria to study ring cleavage of 1,2-dihydroxynaphthalene and subsequent reactions*. *J Bacteriol*, 1992. **174**(23): p. 7542-54.
45. FINN, R.D., et al., *The Pfam protein families database*. *Nucleic Acids Res*, 2010. **38**(Database issue): p. D211-22.

46. COPLEY, R.R. and P. BORK, *Homology among (betaalpha)(8) barrels: implications for the evolution of metabolic pathways*. J Mol Biol, 2000. **303**(4): p. 627-41.
47. ALTSCHUL, S.F., et al., *Basic local alignment search tool*. J Mol Biol, 1990. **215**(3): p. 403-10.
48. SONNHAMMER, E.L., G. VON HEIJNE, and A. KROGH, *A hidden Markov model for predicting transmembrane helices in protein sequences*. Proc Int Conf Intell Syst Mol Biol, 1998. **6**: p. 175-82.
49. SLABINSKI, L., et al., *XtalPred: a web server for prediction of protein crystallizability*. Bioinformatics, 2007. **23**(24): p. 3403-5.
50. YANG, Z.R., et al., *RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins* Bioinformatics 2005. **21**: p. 3369-3376.
51. HARAYAMA, S., P.R. LEHRBACH, and K.N. TIMMIS, *Transposon mutagenesis analysis of meta-cleavage pathway operon genes of the TOL plasmid of Pseudomonas putida mt-2*. J Bacteriol, 1984. **160**(1): p. 251-5.
52. HARAYAMA, S., et al., *Physically associated enzymes produce and metabolize 2-hydroxy-2,4-dienoate, a chemically unstable intermediate formed in catechol metabolism via meta cleavage in Pseudomonas putida*. J Bacteriol, 1989. **171**(11): p. 6251-8.
53. STANLEY, T.M., et al., *Expression and stereochemical and isotope effect studies of active 4-oxalocrotonate decarboxylase*. Biochemistry, 2000. **39**(12): p. 3514.
54. JOHNSON, W.H., Jr., et al., *4-Oxalocrotonate tautomerase, its homologue YwhB, and active vinylpyruvate hydratase: synthesis and evaluation of 2-fluoro substrate analogues*. Biochemistry, 2004. **43**(32): p. 10490-501.
55. IZUMI, A., et al., *Structure and mechanism of HpcG, a hydratase in the homoprotocatechuate degradation pathway of Escherichia coli*. J Mol Biol, 2007. **370**(5): p. 899-911.
56. TSUDA, M. and T. LINO, *Naphthalene degrading genes on plasmid NAH7 are on a defective transposon*. Mol Gen Genet, 1990. **223**(1): p. 33-9.

57. SAMBROOK, J., E.F. FRITSCH, and T. MANIATIS, *Molecular Cloning: A Laboratory Manual*. 1989, New York: Cold Spring Harbor Laboratory.
58. Invitrogen. *Manual do SYBR® Safe DNA Gel Stain*. 17 de junho de 2007 [cited 2011 Janeiro]; Available from: http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Nucleic-Acid-Purification-and-Analysis/Nucleic-Acid-Gel-Electrophoresis/nucleic_acid_gel_electrophoresis/Nucleic-Acid-Stains/SybrSafe-A-better-DNA-gel-stain.html.
59. TAYLOR, R.G., D.C. WALKER, and R.R. MCINNES, *E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing*. *Nucleic Acids Res*, 1993. **21**(7): p. 1677-8.
60. LOENEN, W.A., *Tracking EcoKI and DNA fifty years on: a golden story full of surprises*. *Nucleic Acids Res*, 2003. **31**(24): p. 7059-69.
61. VIEIRA, J. and J. MESSING, *The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers*. *Gene*, 1982. **19**(3): p. 259-68.
62. MARTIN, J. and F.U. HARTL, *The effect of macromolecular crowding on chaperonin-mediated protein folding*. *Proc Natl Acad Sci U S A*, 1997. **94**(4): p. 1107-12.
63. CARNEIRO, F.R., et al., *Spectroscopic characterization of the tumor antigen NY-REN-21 and identification of heterodimer formation with SCAND1*. *Biochem Biophys Res Commun*, 2006. **343**(1): p. 260-8.
64. STOLS, L., et al., *A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site*. *Protein Expr Purif*, 2002. **25**(1): p. 8-15.
65. DONNELLY, M.I., et al., *An expression vector tailored for large-scale, high-throughput purification of recombinant proteins*. *Protein Expr Purif*, 2006. **47**(2): p. 446-54.
66. ASLANIDIS, C. and P.J. DE JONG, *Ligation-independent cloning of PCR products (LIC-PCR)*. *Nucleic Acids Res*, 1990. **18**(20): p. 6069-74.

67. STOCKI, S.A., R.L. NONAY, and L.J. REHA-KRANTZ, *Dynamics of bacteriophage T4 DNA polymerase function: identification of amino acid residues that affect switching between polymerase and 3' --> 5' exonuclease activities*. J Mol Biol, 1995. **254**(1): p. 15-28.
68. ASLANIDIS, C., P.J. DE JONG, and G. SCHMITZ, *Minimal length requirement of the single-stranded tails for ligation-independent cloning (LIC) of PCR products*. PCR Methods Appl, 1994. **4**(3): p. 172-7.
69. SANGER, F., S. NICKLEN, and A.R. COULSON, *DNA sequencing with chain-terminating inhibitors*. Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
70. FOX, J.D. and D.S. WAUGH, *Maltose-binding protein as a solubility enhancer*. Methods Mol Biol, 2003. **205**: p. 99-117.
71. *Purification of Expressed Proteins from Inclusion Bodies*, 2006, Cold Spring Harbor Protocols. doi:10.1101/pdb.prot4089.
72. CHOW, M.K., et al., *REFOLD: an analytical database of protein refolding methods*. Protein Expr Purif, 2006. **46**(1): p. 166-71.
73. HASPEL, J., et al., *System for cleavable Fc fusion proteins using tobacco etch virus (TEV) protease*. Biotechniques, 2001. **30**(1): p. 60-1, 64-6.
74. BATTYE, T.G., et al., *iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM*. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 271-81.