

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

CARACTERIZAÇÃO ESTRUTURAL DA PROTEÍNA Sm21.7
RECOMBINANTE DE *Schistosoma mansoni*

MARIANA AMALIA FIGUEIREDO COSTA

MARIANA AMALIA FIGUEIREDO COSTA

TESE DE DOUTORADO

CARACTERIZAÇÃO ESTRUTURAL DA PROTEÍNA Sm21.7
RECOMBINANTE DE *Schistosoma mansoni*

Tese apresentada ao colegiado do curso de pós-graduação em Bioquímica e Imunologia do Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais (UFMG), como requisito parcial para obtenção do grau de Doutor em Bioquímica e Imunologia.

Orientador: Prof. Dr. Ronaldo Alves Pinto Nagem

Belo Horizonte

2014

"A mente que se abre a uma nova ideia
nunca mais volta ao seu tamanho original."
Albert Einstein

Dedico a realização desse sonho a meus pais, Antônio e Marília, ao meu esposo João Paulo e a meus irmãos Veber e Vinícius pelo apoio incondicional.

Agradecimentos

Ao meu orientador, prof. Dr. Ronaldo Nagem, por todos os ensinamentos transmitidos, pela paciência, pela ótima convivência, pela confiança em mim depositada e por abrir todas as portas para o meu futuro;

Ao prof. Dr. Afredo Goes, por ser um grande colaborador desse trabalho, pela confiança depositada em nosso grupo, pelas nossas longas e animadas conversas nos momentos dos cafezinhos;

Aos amigos do BIOEST e aos agregados: Juliana, Débora, Samuel, Rubens, Simara, Cíntia, Sandro, Fernandes, Marina, Marininha, Mariana, Lorena, Brisa, Luís, Andrea, Deia e Bruninha;

Aos amigos do Laboratório de Imunologia Celular e Molecular pela amizade e troca de conhecimentos;

Aos amigos do Laboratório de Enzimologia e Físico-química de Proteínas pelos intermináveis empréstimos de reagentes e pela amizade e carinho de todos;

Ao Jamil pela nossa amizade pelo conhecimento transmitido e pelos auxílios em todos os momentos;

À Jacque pela nossa amizade e pelos conselhos maravilhosos;

Aos professores do Departamento de Bioquímica e Imunologia que tanto acrescentaram à minha formação;

Ao LNLS pela utilização de toda a estrutura e pelo apoio de todos os funcionários e estudantes;

Ao professor Tiago Brandão (Departamento de Química) pela confiança e orientação;

Ao LVTA pelo apoio nos experimentos de espectrometria de massas;

À Utah State University, em especial aos professores Dr. Alvan Hengge e Dr. Sean Johnson pela oportunidade de trabalhar em seus laboratórios;

Aos funcionários do Departamento de Bioquímica e Imunologia pelo apoio;

Ao CNPq e CAPES pelos financiamentos;

Aos meus irmãos, Veber e Vinícius, pelo incentivo e pelo carinho;

Ao meu esposo, João Paulo, pelo apoio, paciência, compreensão, carinho e por ser uma pessoa tão especial;

Aos meus pais, Antônio e Marília, os grandes responsáveis por essa vitória, pelo apoio incondicional, pelo carinho e por acreditarem em meus sonhos.

Resumo

A esquistossomose é uma doença inflamatória crônica que representa um grave problema de saúde pública nos países tropicais e subtropicais. A droga de escolha para o tratamento, Praziquantel, é eficaz contra vermes adultos, mas não contra formas imaturas e, além disso, não previne a reinfeção. Um importante candidato para constituir uma vacina antiesquistossomose é a proteína Sm21.7 de *Schistosoma mansoni*, que foi capaz de reduzir a carga parasitária em modelo murino de imunização. Sm21.7 é constituída por um domínio de mão EF N-terminal e um domínio de cadeia leve de dineína C-terminal. No presente trabalho, o gene codificador de Sm21.7 foi clonado e expresso em *Escherichia coli* e a proteína recombinante (rSm21.7) foi purificada. Análises de Dicroísmo Circular demonstraram que a proteína encontrava-se parcialmente enovelada, com aproximadamente 37% de α -hélice e 15% de fitas- β . rSm21.7 comportou-se como um dímero em análises de *cross-linking* com glutaraldeído. Cristais de rSm21.7 foram obtidos utilizando PEG monometileter 2.000 e PEG 8.000 como precipitantes em diferentes condições. Imagens de difração de raios X de um cristal nativo (2,05Å de resolução) e um cristal derivado de NaI (1,95Å de resolução) foram coletados na linha W01B-MX2 do Laboratório Nacional de Luz Síncrotron (LNLS). Ambos os cristais pertenciam ao sistema cristalino hexagonal com parâmetros de célula unitária semelhantes $a=b=108,5$; $c=55,8$ Å. As fases foram obtidas por SIRAS e usadas para gerar o primeiro mapa de densidade eletrônica no qual um modelo parcial foi construído. Esse modelo passou por vários ciclos de refinamento para obtenção de uma estrutura final. O modelo possuía apenas 96 resíduos da parte C-terminal (de Gln89 a Asn184 – domínio de cadeia leve de dineína) e a clivagem da proteína na gota de cristalização foi confirmada por SDS-PAGE de cristais dissolvidos. O domínio de cadeia leve de dineína da proteína que foi cristalizado possui 2 α -hélices e quatro fitas- β que formam uma folha- β antiparalela. A determinação da estrutura do domínio C-terminal de rSm21.7 auxiliou na localização de epítomos responsáveis pela resposta protetora conferida pela proteína.

Abstract

Schistosomiasis is an inflammatory chronic disease that represents a major health problem in tropical and subtropical countries. The drug of choice for the treatment, Praziquantel, is effective in killing adult worms, but fails to kill immature forms and prevent reinfection. One important candidate to compose an anti-schistosomiasis vaccine is the protein Sm21.7 from *Schistosoma mansoni*, that is capable to reduce the worm burden in a murine immunization model. Sm21.7 has an N-terminal EF hand domain and a C-terminal dynein light chain domain. In the present work, the Sm21.7 gene was cloned and the recombinant protein (rSm21.7) expressed in *Escherichia coli* and purified to homogeneity. Circular Dichroism analysis showed that the protein was partly folded with 37% of α -helices and 15% of β -strands. rSm21.7 showed to behave as a dimer in a cross-linking with glutaraldehyde experiment. Crystals of rSm21.7 suitable for X-ray diffraction studies were obtained using PEG monomethyl ether 2,000 and PEG 8,000 as precipitants in different conditions. X-ray diffraction images of a native crystal (at 2.05 Å resolution) and a NaI derivative (at 1.95 Å resolution) were collected at W01B-MX2 beamline at the Laboratório Nacional de Luz Síncrotron (LNLS). Both crystals belong to the hexagonal system with similar unit cell parameters $a=b=108.5$, $c=55.8$ Å. SIRAS-derived phases were used to generate the first electron density map where a partial model of rSm21.7 was automatically constructed. The model consists of 96 C-terminal amino acid residues of the full-length protein (from Gln89 to Asn184 – dynein light chain domain) and the cleavage of the protein in the crystallization drop was confirmed by SDS-PAGE using sample from dissolved crystals. The Dynein Light Chain domain which was crystallized has 2 α -helices and 4 β -strands that form an antiparallel β -sheet. The structure of the rSm21.7 C-terminal domain allowed the localization of epitopes responsible for the protective response elicited by the protein.

Lista de Figuras

Figura 1: Distribuição mundial das espécies causadoras da esquistossomose	4
Figura 2: Prevalência da infecção por <i>S. mansoni</i> no Brasil	5
Figura 3: Ciclo de vida dos parasitas do gênero <i>Schistosoma</i>	8
Figura 4: Estrutura do praziquantel (PZQ)	12
Figura 5: Sequência de aminoácidos da proteína Sm21.7 e localização dos domínios que a constituem.	16
Figura 6: Alinhamento de proteínas de trematódeos contendo motivos de mão EF e de cadeia leve de dineína (DLC)	17
Figura 7: Estrutura da proteína FhCaBP4 de <i>F. hepatica</i> obtida por modelagem molecular.	18
Figura 8: Alinhamento do motivo de mão EF canônico com o motivo de mão EF da proteína Sm21.7.....	19
Figura 9: Motivo de mão EF ligado ao cálcio (PDB 1XO5) (Gentry et al., 2005).	20
Figura 10: Estruturas do motivo de mão EF e do domínio de mão EF.	20
Figura 11: Formação de dímeros por proteínas que possuem domínios de mão EF.	21
Figura 12: Peptídeo <i>Swallow</i> ligado ao dímero de cadeia leve de dineína PDB (3E2B) 24	
Figura 13: Estrutura tridimensional da cadeia leve de dineína – DLC8 (PDB 1PWK) . 25	
Figura 14: Estrutura dimérica da cadeia leve de dineína (PDB-4DS1).....	26
Figura 15: Mapa circular do vetor de clonagem pGEM®-T	31
Figura 16: Mapa circular do vetor de expressão pET-28a(+).	35
Figura 17: Mapa linear do vetor modificado pET28a(TEV)	35
Figura 18: Representação esquemática das técnicas de gota suspensa e gota depositada para Cristalização de Macromoléculas	45
Figura 19: Amplificação do gene codificador da proteína Sm21.7	49
Figura 20: PCR de colônia das bactérias transformadas com Sm21.7-pGEM-T	50
Figura 21: PCR de colônia das bactérias transformadas com Sm21.7-pET28a(TEV)... 51	
Figura 22: Expressão da proteína rSm21.7 em bactérias <i>E. coli</i> BL21(DE3).....	52
Figura 23: Purificação da proteína rSm21.7 por cromatografia de afinidade em coluna de níquel..	53
Figura 24: Purificação da proteína rSm21.7 por cromatografia de exclusão molecular. 55	
Figura 25: Espectrometria de massas da proteína rSm21.7 purificada.....	56
Figura 26: Espectro de Dicroísmo Circular entre 190 e 260 nm resultante da análise da proteína rSm21.7 a 20°C.....	57
Figura 27: Desnaturação térmica da proteína rSm21.	58
Figura 28: Estimativa da porcentagem de estrutura secundária da proteína rSm21.7 entre 20 e 90°C	59
Figura 29: Espalhamento Dinâmico da Luz da proteína rSm21.7.....	60
Figura 30: Análise de <i>cross-linking</i> da proteína rSm21.7	61
Figura 31: Cristais da proteína rSm21.7 obtidos no <i>screening</i> inicial pelo método da gota suspensa.	64

Figura 32: Cristais da proteína rSm21.7 obtidos após a otimização das condições iniciais	65
Figura 33: Cristais da proteína rSm21.7 obtidos pelo método da gota depositada	66
Figura 34: Padrão de difração de raios X dos cristais nativo e derivado de rSm21.7	68
Figura 35: Mapa de densidade eletrônica da região da proteína rSm21.7 demonstrando o resíduo Gln89	70
Figura 36: Canal formado pela disposição das proteínas nativas rSm21.7 no cristal	71
Figura 37: SDS-PAGE dos cristais da proteína rSm21.7	72
Figura 38: Espectrometria de massas dos cristais de rSm21.7 dissolvidos	73
Figura 39: Degradação da proteína rSm21.7 visualizada por SDS-PAGE.....	76
Figura 40: Análise da proteólise limitada da proteína rSm21.7 por tripsina.....	77
Figura 41: Geis de eletroforese bi e uni-dimensionais de extratos de parasitas <i>S. mansoni</i>	79
Figura 42: Estrutura tridimensional do domínio de DLC da proteína rSm21.7	82
Figura 43: Sobreposição das estruturas do DLC da proteína rSm21.7 (magenta) e de Dyn2 de <i>S. cerevisiae</i> (PDB – 4HT6) (ciano)	83
Figura 44: Gráfico RMSD para todos os aminoácidos em relação à estrutura secundária	84
Figura 45: Estrutura tridimensional do homodímero de rSm21.7	87
Figura 46: Fitas- β da interface do homodímero de rSm21.7	88
Figura 47: Interações entre os monômeros de rSm21.7	88
Figura 48: Análise da superfície eletrostática da interface do dímero de rSm21.7	90
Figura 49: Análise de superfície eletrostática das proteínas rSm21.7 e Dyn2 (PDB 4HT6)	92
Figura 50: Sequência da proteína Sm21.7 contendo os peptídeos sintetizados na membrana e avaliados pela reatividade	93
Figura 51: Peptídeos reativos localizados na estrutura do domínio de cadeia leve de dineína da proteína rSm21.7	94

Lista de Tabelas

Tabela 1: Iniciadores direto e reverso utilizados para amplificação do gene codificador da proteína Sm21.7.....	28
Tabela 2: Espalhamento Dinâmico da Luz da proteína rSm21.7	61
Tabela 3: Estatísticas de coleta de dados e processamento dos cristais nativo e derivado de rSm21.7.....	69
Tabela 4: Dados de DLS ao longo do tempo da proteína rSm21.7 com e sem PMSF...	80
Tabela 5: Estatísticas de refinamento da estrutura tridimensional da proteína rSm21.7	81
Tabela 6: Estatísticas de coleta de dados e processamento do cristal nativo coletado no NSLS.	86
Tabela 7: Estatísticas de refinamento da estrutura tridimensional da proteína rSm21.7.	87
Tabela 8: Análise dos elementos de estrutura secundária da proteína rSm21.7.....	89

Lista de abreviaturas e símbolos

°C – grau Celsius

[θ]MRW – elipcidade molar

μg – micrograma (10^{-6} gramas)

μL – microlitro (10^{-6} litros)

μM – micromolar (10^{-6} Molar)

% – porcentagem

^{45}Ca – isótopo do cálcio

2D – bidimensional

1D – unidimensional

Å – ångström

ATP – Adenosina Trifosfato

BSA – *Bovine Serum Albumin* – Albumina Sérica Bovina

CD – *Circular Dichroism* – Dicroísmo Circular

cDNA – *complementary DNA* – DNA complementar

C-terminal – carboxi-terminal

Da – Daltons

DAB – 3,3'-*Diaminobenzidine* – 3, 3'-Diaminobenzidina

DALY – *Disability-Adjusted Life Year*

DLC – *Dynein Light Chain* – Cadeia Leve de Dineína

DLS – *Dynamic Light Scattering* – Espalhamento Dinâmico da Luz

DNA – *Deoxyribonucleic Acid* – Ácido Desoxirribonucleico

DTT – ditioneitol

EDTA – *Ethylenediamine Tetraacetic Acid* – Ácido Etilenodiamino Tetra-acético

g – força centrífuga

g – gramas

hid – hidrodinâmico

HIV – *Human Immunodeficiency Virus* – Vírus da Imunodeficiência Humana

IDH – Índice de Desenvolvimento Humano

IPTG – *Isopropyl β -D-1-Thiogalactopyranoside* – Isopropil β -D-1-Tiogalactopiranosídeo

K – Kelvin

kDa – quiloDaltons (10^3 Daltons)

L – Litro
 LB – Luria-Bertani
 LNLS – Laboratório Nacional de Luz Síncrotron
 M – Molar
 mdeg – millidegrees
 mg – miligramas (10^{-3} gramas)
 mL – mililitro (10^{-3} litros)
 MM – Massa Molecular
 mm – milímetros (10^{-3} metros)
 mM – milimolar (10^{-3} Molar)
 MWCO – *Molecular Weight Cut Off*
 ng – nanogramas (10^{-9} gramas)
 nm – nanômetros (10^{-9} metros)
 nNOS – óxido nítrico sintase neuronal
 NSLS – *National Synchrotron Light Source*
 N-terminal – amino-terminal
 OMS – Organização Mundial de Saúde
 OX – oxamniquine
 pb – pares de bases
 PBS – *Phosphate Buffered Saline* – Tampão Fosfato Salino
 PCR – *Polymerase Chain Reaction* – Reação em Cadeia da Polimerase
 PDB – Protein Data Bank
 PEG – polietilenoglicol
 pH – potencial hidrogeniônico
 Pd – polidispersividade
 PMSF – *Phenylmethylsulfonyl Fluoride* – Fluoreto de Fenilmetilsulfonil
 psi – *pound force per square inch*
 p/v – peso por volume
 PZQ – praziquantel
 RMN – Ressonância Magnética Nuclear
 RMSD – *root-mean-square deviation*
 RNA – *Ribonucleic Acid* – Ácido Ribonucleico
 rP22 – proteína P22 recombinante de *Schistosoma mansoni*
 rpm – rotações por minuto

rSm21.7 – proteína Sm21.7 recombinante de *Schistosoma mansoni*

RT-PCR – *Reverse Transcription Polymerase Chain Reaction* – Transcrição Reversa seguida de Reação em Cadeia da Polimerase

SDS – *Sodium Dodecyl Sulfate* – Dodecil Sulfato de Sódio

SDS-PAGE – *Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis* – Eletroforese em Gel de Poliacrilamida contendo SDS

SIRAS – *Single Isomorphous Replacement with Anomalous Scattering* – Substituição Isomórfica Simples com Espalhamento Anômalo

SWAP – *Soluble Worm Antigenic Preparation* – Preparação Antigênica Solúvel do Parasita

TAE – Tris, Ácido acético, EDTA

TEMED – N, N, N', N'-tetrametiletilenodiamina

TEV – *Tobacco Etch Virus*

TFA – *Trifluoroacetic Acid* – Ácido Trifluoroacético

UV – Ultra violeta

V – Volts

v/v – volume por volume

V_M – Coeficiente de Matthews

Abreviações (uma e três letras) dos resíduos de aminoácidos

A	Ala	Alanina
C	Cys	Cisteína
D	Asp	Aspartato
E	Glu	Glutamato
F	Phe	Fenilalanina
G	Gly	Glicina
H	His	Histidina
I	Ile	Isoleucina
K	Lys	Lisina
L	Leu	Leucina
M	Met	Metionina
N	Asn	Asparagina
P	Pro	Prolina

Q	Gln	Glutamina
R	Arg	Arginina
S	Ser	Serina
T	Thr	Treonina
V	Val	Valina
W	Trp	Triptofano
Y	Tyr	Tirosina

Índice

1 – Introdução e justificativa	1
2 – Revisão da literatura	3
2.1 – A parasitose causada por <i>Schistosoma</i> e sua distribuição mundial.....	3
2.2 – Ciclo de vida dos parasitas do gênero <i>Schistosoma</i>	5
2.3 – Patologia da esquistossomose	9
2.4 – Formas de controle do <i>Schistosoma</i> e tratamento da esquistossomose	10
2.5 – Desenvolvimento de vacinas contra a esquistossomose	14
2.6 – Domínios de mão EF e de cadeia leve de dineína (DLC) - os domínios constituintes da proteína Sm21.7	19
3 – Objetivos.....	27
3.1 – Objetivo geral	27
3.2 – Objetivos específicos	27
4 – Materiais e métodos.....	28
4.1 – Subclonagem do gene codificador da proteína Sm21.7 no vetor pET28a(TEV)	28
4.1.1 – Amplificação do gene codificador da proteína Sm21.7.....	28
4.1.2 – Eletroforese em gel de agarose	29
4.1.3 – Purificação do produto de PCR e bandas do gel.....	29
4.1.4 – Clonagem do gene codificador da proteína Sm21.7 no vetor pGEM®-T	30
4.1.5 – Extração de DNA plasmidial bacteriano	33
4.1.6 – Preparação de bactérias eletrocompetentes.....	34
4.1.7 – Subclonagem do gene codificador da proteína Sm21.7 no vetor de expressão pET28a(TEV).....	34
4.2 – Expressão da proteína Sm21.7 recombinante (rSm21.7) em sistema procariótico e lise bacteriana.....	37
4.2.1 – Eletroforese em gel de poliacrilamida contendo SDS (SDS-PAGE)	38
4.3 – Purificação da proteína rSm21.7 por cromatografias de afinidade e exclusão molecular	39
4.4 – Análises biofísicas e bioquímicas da proteína rSm21.7 purificada	40
4.4.1 – Análise da estrutura secundária da proteína rSm21.7 por Dicroísmo Circular (CD).....	40
4.4.2 – Avaliação do estado oligomérico da proteína rSm21.7 através de Espalhamento Dinâmico da Luz (DLS)	41
4.5 – Ensaio de cross-linking da proteína rSm21.7	42

4.6 – Proteólise limitada da proteína rSm21.7	42
4.7 – Ensaio de <i>Western Blot</i> com anticorpo anti-histidinas	43
4.8 – Determinação da estrutura da proteína rSm21.7 por cristalografia de raios X..	44
4.8.1 – Testes de cristalização da proteína rSm21.7	44
4.8.2 – Coleta de dados de difração de raios X.....	46
4.8.3 – Processamento dos dados de difração de raios X	47
4.8.4 – Determinação das fases e refinamento da estrutura tridimensional da proteína rSm21.7	47
5 – Resultados e discussão.....	49
5.1 – Subclonagem do gene codificador da proteína Sm21.7 no vetor de expressão pET28a(TEV)	49
5.2 – Expressão, lise e purificação da proteína rSm21.7	51
5.3 – Análise da estrutura secundária da proteína rSm21.7	56
5.4 – Análise do estado de oligomerização da proteína rSm21.7	59
5.4.1 – Estado oligomérico de rSm21.7 – DLS	59
5.4.2 – Estado oligomérico de rSm21.7 – <i>Cross-linking</i>	61
5.5 – Cristalização da proteína rSm21.7 e coleta de dados de difração de raios X	63
5.6 – A proteína rSm21.7 possui um domínio suscetível à ação de proteases.....	75
5.7 – Determinação da estrutura tridimensional da proteína rSm21.7.....	80
5.8 – Localização de epítomos reativos na estrutura do domínio de cadeia leve de dineína da proteína rSm21.7	93
6 - Conclusões	96
7 – Perspectivas	97
8 – Referências Bibliográficas.....	98
9 – Apêndice	114

1 – Introdução e justificativa

As proteínas são as macromoléculas mais versáteis nos sistemas vivos e atuam em funções cruciais em essencialmente todos os processos biológicos. Funcionam como catalisadores, transportam e armazenam outras moléculas, transmitem impulsos nervosos, fornecem apoio mecânico e proteção imunitária, geram movimento e controlam movimento de diferenciação (BERG; TYMOCZKO; STRYER, 2012).

Essa notável gama de funções surge da existência de milhares de proteínas, cada uma enovelada em uma estrutura tridimensional distinta, que permite que ela interaja com uma ou mais de um variado conjunto de moléculas (BERG; TYMOCZKO; STRYER, 2012).

Assim, a determinação da estrutura tridimensional de proteínas é um passo importante no entendimento de suas funções, assim como na caracterização de mecanismos de ação. Atualmente, o conhecimento da estrutura e função de proteínas proporciona não apenas a caracterização das mesmas e identificação de funções biológicas, mas também auxilia no desenvolvimento de produtos biotecnológicos.

Nesse contexto de determinação da estrutura de proteínas, a cristalografia de raios X se destaca como a principal técnica utilizada. As primeiras estruturas cristalográficas de proteínas determinadas foram as da mioglobina (KENDREW; PARRISH, 1957) e da hemoglobina (GREEN; INGRAM; PERUTZ, 1954), que culminaram no Prêmio Nobel de Química em 1.962. Atualmente o PDB (Protein Data Bank) possui mais de 80 mil estruturas de proteínas resolvidas por cristalografia (<http://www.rcsb.org/pdb/statistics/holdings.do>).

O campo de desenvolvimento de vacinas também pode beneficiar-se da disponibilidade da estrutura tridimensional de proteínas. O termo vacinologia estrutural está sendo usado recentemente para definir casos em que a estrutura tridimensional dos antígenos vacinais fornece informações que podem promover o desenho/desenvolvimento de novos e melhorados antígenos (DORMITZER; GRANDI; RAPPUOLI, 2012). Essa técnica está sendo utilizada para aumentar o conhecimento sobre candidatos a vacinas que foram previamente identificados, identificar novos candidatos e melhorar a imunogenicidade, estabilidade, segurança e facilidade de produção de antígenos contra muitos patógenos (DORMITZER; GRANDI; RAPPUOLI, 2012).

Alguns estudos com as bactérias *Neisseria meningitidis* sorogrupo B (MASIGNANI *et al.*, 2003) e *Streptococcus agalactiae* (NUCCITELLI *et al.*, 2011) e o vírus sincicial respiratório (SWANSON *et al.*, 2011) obtiveram resultados promissores utilizando a vacinologia estrutural. Nesses trabalhos a disponibilidade da estrutura tridimensional proporcionou a melhoria das características bioquímicas dos antígenos vacinais e guiaram para a combinação de epítomos que aumentaram a resposta imune no hospedeiro (DORMITZER; GRANDI; RAPPUOLI, 2012). Em um futuro próximo, a tecnologia poderá ser utilizada para o melhor entendimento da natureza dos epítomos antigênicos que são reconhecidos pelo sistema imune (DORMITZER; GRANDI; RAPPUOLI, 2012).

No contexto de vacinologia estrutural, o desenvolvimento de uma vacina contra a esquistossomose, uma doença parasitária causada por espécies de *Schistosoma* que ameaça 700 milhões de pessoas no mundo, pode ter um enorme impacto na redução da morbidade e mortalidade decorrentes da infecção.

A proteína Sm21.7 de *Schistosoma mansoni* é capaz de desenvolver uma resposta imune protetora contra a infecção desafio quando utilizada como uma vacina (AHMED *et al.*, 2001, 2006; REZENDE *et al.*, 2011). A determinação de sua estrutura tridimensional, com a consequente localização de epítomos reativos, poderia guiar para o desenvolvimento de uma vacina com propriedades imunológicas melhoradas.

Assim, no presente trabalho é apresentada a caracterização estrutural parcial da proteína antigênica Sm21.7 recombinante de *S. mansoni*.

2 – Revisão da literatura

2.1 – A parasitose causada por *Schistosoma* e sua distribuição mundial

A esquistossomose é uma doença inflamatória crônica, causada por trematódeos do gênero *Schistosoma*, que representa um importante problema na saúde pública mundial (Revisado por GRYSEELS *et al.*, 2006). A Organização Mundial de Saúde (OMS) considera que a esquistossomose é a segunda doença parasitária em importância socioeconômica no mundo, sendo a malária a primeira (Revisado por ELBAZ; ESMAT, 2013; THÉTIOT-LAURENT *et al.*, 2013; VAN DER WERF *et al.*, 2003).

Essa doença tropical negligenciada é endêmica em 74 países, sendo todos classificados como países em desenvolvimento ou subdesenvolvidos (Figura 1). Estimativas indicam que aproximadamente 240 milhões de pessoas estão atualmente infectadas pelo parasita e que mais de 700 milhões vivem em áreas que oferecem risco de infecção (<http://www.who.int/schistosomiasis/en/>). Adicionalmente, mais de 200.000 mortes ocorrem anualmente em decorrência da doença (THÉTIOT-LAURENT *et al.*, 2013).

Um indicador da importância da esquistossomose na saúde pública é o DALY (*Disability Adjusted Life Years*), uma medida desenvolvida pela OMS para quantificar o impacto global de uma doença sobre a população. Esta medida avalia os anos de vida perdidos por morte prematura e os anos de vida que o indivíduo esteve incapacitado em decorrência da doença (ANAND; HANSON, 1997). Assim, um DALY pode ser considerado como um ano de vida saudável perdida. Atualmente, as perdas anuais, medidas em DALY, devido à esquistossomose são de aproximadamente 4,5 milhões de anos de vida, demonstrando o quanto essa doença pode ser debilitante (Revisado por STEINMANN *et al.*, 2006).

As cinco espécies de *Schistosoma* capazes de infectar humanos tendem a ocorrer em regiões geográficas específicas. A infecção pela espécie *Schistosoma haematobium*, responsável pela esquistossomose urinária, ocorre em países da África. As outras quatro espécies causam a esquistossomose intestinal e/ou hepática. Infecções por *Schistosoma mansoni* ocorrem na África, América do Sul, Oeste do Mediterrâneo e Caribe; por *Schistosoma intercalatum* no oeste da África e por *Schistosoma japonicum* e *Schistosoma mekongi* na Ásia, mais precisamente na China, Indonésia e Filipinas

(Figura 1) (Revisado por GRYSEELS *et al.*, 2006; THÉTIOT-LAURENT *et al.*, 2013; U OLVEDA, 2013).

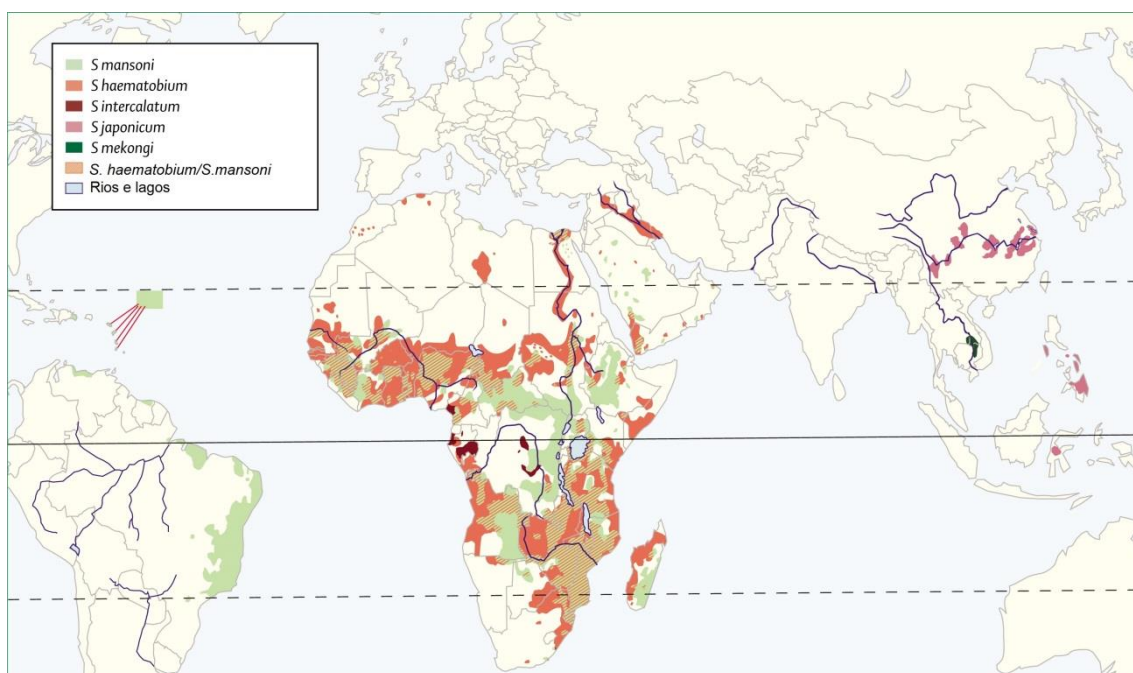


Figura 1: Distribuição mundial das espécies causadoras da esquistossomose. As espécies causadoras da doença ocorrem em regiões geográficas específicas, como demonstradas no mapa. A espécie *S. haematobium* ocorre em países da África, *S. mansoni* ocorre na África, América do Sul, Oeste do Mediterrâneo e Caribe, *S. intercalatum* ocorre no oeste da África e *S. japonicum* e *S. mekongi* ocorrem na Ásia (China, Indonésia e Filipinas). Em alguns lugares da África ocorre a sobreposição das espécies *S. mansoni* e *S. haematobium* (Adaptado de GRYSEELS *et al.*, 2006).

No Brasil, a infecção por *S. mansoni* ocorre em uma vasta área endêmica, que se estende do Maranhão ao estado de Minas Gerais. Entretanto, existem focos isolados da doença em outras localidades (COURA; AMARAL, 2004). A prevalência estimada da esquistossomose no Brasil foi de aproximadamente 5,3% no período entre 2005 e 2009, variando entre zero e 70% nas diferentes municipalidades estudadas (Figura 2) (SCHOLTE *et al.*, 2014). Essa porcentagem indicava aproximadamente 10 milhões de pessoas infectadas.

Nesse estudo, as maiores prevalências foram encontradas nas regiões Norte e Nordeste do Brasil, provavelmente devido à existência de locais apropriados para a proliferação do hospedeiro intermediário. Já as menores prevalências foram estimadas no Centro-Oeste, Sul e parte do Sudeste (exceto o estado de Minas Gerais), sendo estas duas últimas as regiões que possuem maiores IDH (Índice de Desenvolvimento

Humano) no Brasil, evidenciando a relação entre a prevalência da esquistossomose e condições econômicas precárias (SCHOLTE *et al.*, 2014).

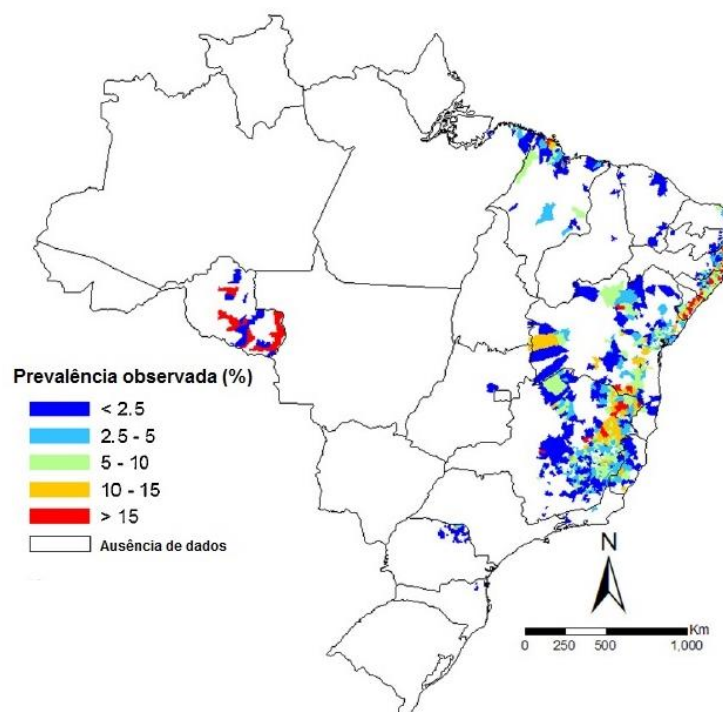


Figura 2: Prevalência da infecção por *S. mansoni* no Brasil. Estimativas entre os anos de 2005 e 2009 indicaram que as áreas de maior prevalência da esquistossomose encontravam-se nas regiões Norte, Nordeste e no Estado de Minas Gerais. Já as de menor prevalência estavam nas regiões Sul e Sudeste do país, como demonstrado no mapa do território brasileiro (Adaptado de SCHOLTE *et al.*, 2014).

2.2 – Ciclo de vida dos parasitas do gênero *Schistosoma*

A transmissão da esquistossomose depende de fatores ambientais e comportamentais. Geralmente essa doença ocorre em localidades nas quais o acesso ao suprimento de água limpa é limitado e que dependem de atividades relacionadas ao contato com a água (DOENHOFF; PICA-MATTOCCIA, 2006). Essa relação entre transmissão da esquistossomose e contato com água ocorre devido ao parasita passar por vários estágios de desenvolvimento que ocorrem nesse ambiente, incluindo um período de desenvolvimento em um caramujo aquático específico, que é o hospedeiro intermediário (KING, CHARLES H., 2009).

O ciclo de vida desses parasitas (Figura 3) necessita de dois hospedeiros: um caramujo aquático no qual o parasita se reproduz assexuadamente e um hospedeiro definitivo mamífero no qual ocorre a reprodução sexuada (DOENHOFF; PICAMATTOCCIA, 2006).

Os ovos do parasita são depositados na água juntamente com as excretas de indivíduos infectados e, então, uma larva microscópica denominada miracídio é liberada de cada ovo (Revisado por THÉTIOT-LAURENT *et al.*, 2013). Tais ovos podem ficar viáveis por 7 dias até atingirem um local contendo água para liberação do miracídio (Revisado por GRYSEELS *et al.*, 2006). Para sobreviver, essa pequena larva precisa encontrar e penetrar em um caramujo aquático, que atua como hospedeiro intermediário. Cada espécie de *Schistosoma* infecta um molusco específico. Assim, os caramujos do gênero *Bulinus* são hospedeiros intermediários para *S. haematobium* e *S. intercalatum*; os do gênero *Biomphalaria* para *S. mansoni*; os do gênero *Oncomelania* para *S. japonicum* e os do gênero *Neotricula* para *S. mekongi* (Revisado por THÉTIOT-LAURENT *et al.*, 2013). O miracídio utiliza-se de movimentos mecânicos e atividade proteolítica para penetrar no hospedeiro intermediário (WARREN, 1978).

Uma vez dentro do caramujo, o miracídio transforma-se em um esporocisto e multiplica-se. Os novos esporocistos migram para as glândulas digestivas do molusco, onde se transformam, através de reprodução assexuada, em novas larvas de cauda bifurcada denominadas cercárias, que deixam o caramujo e permanecem no ambiente aquático. A reprodução assexuada permite um aumento no número de parasitas. Assim, a partir da quarta semana após a infestação do caramujo com o miracídio, cerca de 1.500 cercárias são liberadas todos os dias (Revisado por THÉTIOT-LAURENT *et al.*, 2013).

A infecção de humanos ocorre durante o contato com a água, quando ocorre a penetração das cercárias na pele, sendo essa ação auxiliada por secreções glandulares do parasita (WARREN, 1978). Geralmente a infecção ocorre em atividades cotidianas como higiene, recreação, pescaria e irrigação. Durante a penetração na pele, as cercárias perdem sua cauda bifurcada e a forma do parasita resultante, denominada esquistossômulo, entra nos capilares e vasos linfáticos (Revisado por THÉTIOT-LAURENT *et al.*, 2013).

Uma vez nos capilares sanguíneos, os esquistossômulos são passivamente levados pelo fluxo sanguíneo até atingir o coração e, então, os pulmões (Revisado por EL RIDI; TALLIMA, 2013). Dos pulmões, os esquistossômulos se dirigem para o

sistema porta hepático onde se alimentam e se desenvolvem, transformando-se em machos e fêmeas adultos. Esses, posteriormente, migram acasalados para os plexos venosos vesicais e pélvicos (Revisado por NEVES, 2005).

Os vermes adultos são brancos ou acinzentados e possuem um corpo cilíndrico com comprimento variando entre 7 e 20 mm. Possuem duas ventosas terminais, um tegumento complexo, um trato digestivo cego e órgãos reprodutivos. Diferente de outros trematódeos, *Schistosoma* possui sexos separados e o corpo do macho contém uma fenda, o canal ginecóforo, no qual ele abriga a fêmea, que é mais longa e fina (Revisado por GRYSEELS *et al.*, 2006).

Posteriormente, as fêmeas começam a depositar ovos, que podem ficar presos na mucosa intestinal ou serem arrastados para o fígado. Os ovos que conseguirem chegar à luz intestinal vão para o exterior junto com as fezes (Revisado por NEVES, 2005). Ovos no plexo venoso pélvico nas infecções por *S. haematobium* migram para o sistema urinário, passam pela parede da bexiga e são excretados na urina (WARREN, 1978). Alcançando a água, os ovos liberam o miracídio, fechando o ciclo de vida do parasita.

A fêmea pode liberar de 200 a 2.000 ovos por dia a partir do quadragésimo dia após a infecção. O parasita vive de 5 a 10 anos e é capaz de evadir a resposta imune por vários mecanismos (BOROS, 1989). O potencial de reprodução teórico de um par de vermes é de 600 bilhões de parasitas (GRYSEELS *et al.*, 2006).

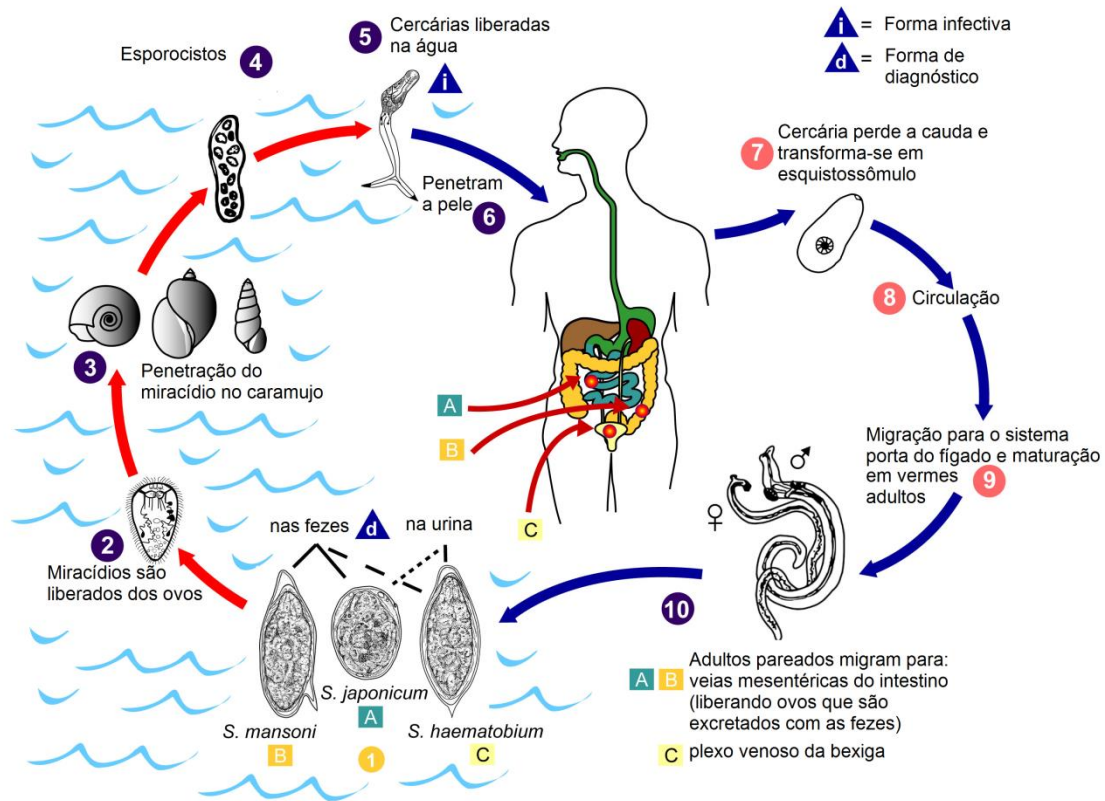


Figura 3: Ciclo de vida dos parasitas do gênero *Schistosoma*. O ciclo de vida desse parasita inclui dois hospedeiros, um caramujo aquático e um mamífero, que é o hospedeiro definitivo. (1) Os ovos do parasita excretados juntamente com as fezes ou urina do hospedeiro mamífero alcançam um ambiente aquático e, então, (2) os miracídeos são liberados desses ovos. (3) O miracídio penetra em um caramujo, que é específico para cada espécie de *Schistosoma* e (4) transforma-se em esporocisto. Este, por sua vez, sofre reprodução assexuada, (5) originando as cercárias, que são liberadas na água. (6) A cercária penetra na pele humana durante o contato com a água e, então, (7) perde sua cauda bifurcada transformando-se em esquistossômulo, que (8) entra na circulação sanguínea. Esses últimos (9) são levados para o sistema porta do fígado onde maturam em vermes adultos. (10) Os vermes adultos pareados migram para as veias mesentéricas do intestino ou para o plexo venoso da bexiga (*S. haematobium*) de onde a fêmea libera centenas de ovos por dia. Muitos desses ovos são excretados com as fezes ou urina (Adaptado de <http://www.cdc.gov/parasites/schistosomiasis/biology.html>).

2.3 – Patologia da esquistossomose

Uma das características mais marcantes do *Schistosoma* é sua longevidade, que atinge, em média, entre 5 e 10 anos. Entretanto, existem relatos de parasitas sobrevivendo por 30 anos no hospedeiro definitivo. Durante esse tempo, a fêmea produz centenas de ovos por dia, mas muitos desses ovos não são excretados com as fezes ou urina e ficam retidos nos tecidos, como os do intestino, fígado e do sistema urinário (*S. haematobium*). Como são intensamente imunogênicos, os ovos induzem reações inflamatórias, os granulomas, que são responsáveis pela maior parte da patogênese e morbidade da doença (DOENHOFF; PICA-MATTOCCIA, 2006).

A penetração percutânea da cercária provoca uma erupção temporária que, às vezes, persiste por dias como uma lesão papulopruriginosa, especialmente após infecções primárias como as que ocorrem em turistas e migrantes (BOTTIEAU *et al.*, 2006; GRYSEELS *et al.*, 2006).

A esquistossomose aguda (Febre de Katayama) é uma reação de hipersensibilidade sistêmica contra o esquistossômulo que migra, ocorrendo em poucas semanas após a infecção primária (GAZZINELLI *et al.*, 1985). Essa doença se inicia subitamente com febre, calafrios, fraqueza, perda de peso, dores de cabeça, náusea, vômito, diarreia, tosse, hepatoesplenomegalia e eosinofilia (LAMBERTUCCI, 1993).

As principais lesões em infecções crônicas estabelecidas são devidas aos ovos que são mantidos nos tecidos durante a migração perivesicular ou peri-intestinal ou após a embolização no fígado, baço, pulmões ou sistema cerebroespinhal. Os ovos secretam enzimas proteolíticas que provocam reações eosinofílicas inflamatórias e granulomatosas, que são progressivamente substituídas por depósitos fibróticos (CHEEVER; HOFFMANN; WYNN, 2000). Com o tempo, o número de granulomas pode aumentar muito e prejudicar o funcionamento adequando do órgão (Revisado por THÉTIOT-LAURENT *et al.*, 2013).

Os sinais crônicos da doença podem ser gastrointestinais, hepáticos, neurológicos ou geniturinários. A migração dos ovos pela parede do intestino provoca inflamação granulomatosa da mucosa, pseudopolipose, microulcerações e sangramento superficial. Os sintomas mais comuns são dor abdominal crônica ou intermitente e desconforto, perda de apetite e diarreia com ou sem sangue (Revisado por GRYSEELS *et al.*, 2006; U OLVEDA, 2013).

A esquistossomose inflamatória hepática é uma reação aos ovos capturados nos espaços perisinusoidais periportais do fígado. Características típicas incluem aumento do lobo esquerdo do fígado e esplenomegalia nodular. A deposição massiva de colágeno nos espaços periportais leva à progressiva oclusão das veias portais, hipertensão portal, esplenomegalia, circulação venosa colateral e varizes gastrointestinais. O sangramento das varizes gastrointestinais é a complicação mais séria, comumente fatal, da esquistossomose fibrótica hepática (Revisado por GRYSEELS *et al.*, 2006).

Os ovos de *S. haematobium* provocam inflamação granulomatosa, ulceração e pseudopolipose das paredes vesiculares e dos ureteres (CHEEVER; HOFFMANN; WYNN, 2000). Lesões crônicas podem evoluir para fibrose ou calcificação da bexiga e ureteres inferiores, resultando em hidroureter e hidronefrose. A compressão crônica pode, eventualmente, levar ao dano do parênquima e falência renal (Revisado por GRYSEELS *et al.*, 2006). Além disso, a relação entre esquistossomose urinária e câncer de bexiga é bem documentada em várias localidades da África (AMONKAR; MURALI; KRISHNAMURTHY, 2001; GELFAND, 1967; MOSTAFA; SHEWEITA; O'CONNOR, 1999).

2.4 – Formas de controle do *Schistosoma* e tratamento da esquistossomose

As intervenções para o controle da esquistossomose incluem a administração de drogas em massa e/ou quimioterapia de casos individuais, bem como melhoria das condições sanitárias, intervenções ambientais para controle do hospedeiro intermediário e educação sanitária para evitar o contato com águas possivelmente contaminadas (ELBAZ; ESMAT, 2013).

As mudanças comportamentais da população, melhorias das condições sanitárias, fornecimento seguro de água, programas educacionais e a procura por cuidados com a saúde são ações que podem reduzir a transmissão da esquistossomose, mas tais mudanças são difíceis de serem implementadas devido às precárias condições de vida onde a doença ocorre (GRYSEELS *et al.*, 2006).

Já o controle do caramujo com moluscidas é caro e logisticamente complexo. São necessários mão de obra e materiais em abundância para a aplicação eficiente, que pode ser um processo demorado e que necessita de vigilância epidemiológica e

malacológica adicionais. Além disso, a toxicidade dos moluscicidas para outros organismos aquáticos, incluindo peixes, levanta preocupações ecológicas (Revisado por GRYSEELS *et al.*, 2006). Outro fator que prejudica esse tipo de controle é o fato de que a população de caramujos pode ser reduzida drasticamente, mas raramente eliminada, então o retratamento regular é necessário (SOUZA, 1995). Formas alternativas para o controle de caramujos podem ser os métodos físicos ou a utilização de competidores biológicos, mas tais métodos são difíceis para serem colocados em prática (COWIE, 2001; Revisado por GRYSEELS *et al.*, 2006).

Sendo assim, atualmente o controle da esquistossomose é baseado principalmente no tratamento de pessoas infectadas com as drogas recomendadas.

O tratamento da esquistossomose foi revolucionado na década de 1.970 pela introdução de duas drogas relativamente seguras e que podem ser administradas oralmente: oxamniquine (OX) e praziquantel (PZQ). Apesar de ambas estarem em uso atualmente, OX é uma droga efetiva apenas contra *S. mansoni* e seu custo é elevado, sendo seu uso restrito ao Brasil (DOENHOFF; PICA-MATTOCCIA, 2006). Além disso, essa droga pode provocar efeitos colaterais mais pronunciados que o PZQ, como sonolência e ataques epiléticos (GRYSEELS *et al.*, 2006).

O PZQ é a droga de escolha para o tratamento da esquistossomose, pois é eficaz contra todas as espécies de *Schistosoma* que infectam humanos (DAVIS; BILES; ULRICH, 1979; GÖNNERT; ANDREWS, 1977; ISHIZAKI; KAMO; BOEHME, 1979; KATZ; ROCHA; CHAVES, 1979; NASH *et al.*, 1982; WEBBE; JAMES, 1977). Devido a isso seu uso se torna especialmente importante em áreas onde mais de uma espécie está presente, como áreas onde infecções por *S. mansoni* e *S. haematobium* coexistem (ELBAZ; ESMAT, 2013).

Os efeitos colaterais do PZQ são suaves e incluem náusea, vômitos, urticária e edema. Entretanto, em infecções pesadas, pode ocorrer cólica aguda com diarreia sanguinolenta pouco tempo após o tratamento (STELMA *et al.*, 1995).

Com relação à sua estrutura, o PZQ possui um centro assimétrico (Figura 4) e as preparações comerciais da droga são constituídas por uma mistura racêmica de iguais quantidades de *levo* e *dextro* isômeros. Contudo, apenas a forma *levo* tem atividade contra o *Schistosoma*, não sendo demonstrado efeito terapêutico para a forma *dextro* (TANAKA *et al.*, 1989).

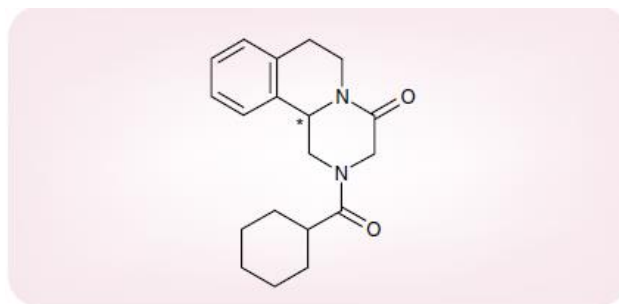


Figura 4: Estrutura do praziquantel (PZQ). O PZQ é a droga de escolha para o tratamento da esquistossomose. Em sua estrutura, essa droga possui um centro assimétrico, que está marcado com um “*” na presente imagem (DOENHOFF; PICAMATTOCCIA, 2006).

Diferentemente de outras drogas antiesquistossomose, PZQ não é mutagênico em um grande número de espécies, como bactérias, leveduras, insetos e mamíferos e sua toxicidade em ratos, camundongos, coelhos e cães é muito baixa. Além disso, PZQ não causa problemas relacionados à reprodução ou teratogenicidade (FROHBURG, 1984), e a administração em crianças e mulheres grávidas é segura e permitida (WHO, 2002).

Estudos indicam que o tratamento com PZQ sempre alcança 60% de cura ou mais. Entretanto, após décadas de tratamento em massa com o PZQ, observou-se que a morbidade relacionada à esquistossomose foi significativamente reduzida, mas que a reinfecção não pode ser controlada dessa forma. Assim, para sustentar os ganhos obtidos com o uso de PZQ, a droga precisa ser administrada periodicamente por um período de tempo indefinido. Porém, o uso contínuo e extensivo da droga pode levar ao desenvolvimento de parasitas resistentes (U OLVEDA, 2013).

De fato, testes laboratoriais já demonstraram que, quando o parasita é submetido a doses subótimas da droga, pode ocorrer o desenvolvimento de resistência ao PZQ (FALLON; DOENHOFF, 1994). Outros estudos estimaram que a dose efetiva de isolados de *S. mansoni* resistentes foi cerca de três vezes maior que para o parasita controle (CIOLI *et al.*, 2004).

Outras evidências do desenvolvimento de parasitas resistentes foram encontradas em áreas endêmicas. Uma delas veio do Senegal, onde, na década de 1.990, a transmissão alcançou proporções epidêmicas no norte do país. Quando PZQ foi utilizado na tentativa de controle, a cura foi de apenas 18%, indicando que os pacientes poderiam estar infectados com parasitas resistentes à droga (STELMA *et al.*, 1995).

Outro exemplo de resistência ocorreu no Egito, onde, no período de 1.997 a

1.999, muitos indivíduos tratados com PZQ não foram completamente curados. Desses indivíduos infectados e não curados, foram isolados parasitas menos suscetíveis ao PZQ (ISMAIL et al., 1996, 1999).

Além do possível desenvolvimento de resistência dos parasitas ao PZQ, outro problema associado a essa droga é sua falta de eficácia contra formas jovens do parasita. Assim, em modelos experimentais de infecção, o *Schistosoma* mostra uma distribuição bifásica de sensibilidade à droga. A suscetibilidade ao PZQ é reduzida a baixos níveis no início da infecção e é apenas gradualmente readquirida (BOTROS et al., 2005; SABAH et al., 1986; XIAO; CATTO; WEBSTER, 1985).

A resistência das formas jovens é, atualmente, o problema mais sério no uso clínico do PZQ (THÉTIOT-LAURENT et al., 2013) e a consequência prática desse fenômeno é que, em áreas de transmissão muito ativa, muitas pessoas provavelmente possuem formas imaturas no momento do tratamento, resultando em baixas taxas de cura nesses casos (ELBAZ; ESMAT, 2013).

Além disso, apesar do tratamento em massa com o PZQ, ainda existe um considerável espalhamento da doença, principalmente no Oeste da África. Adicionalmente aos fatores já citados com relação à droga, a construção de barragens e o desenvolvimento de esquemas de irrigação contribuem para esse espalhamento da esquistossomose, pois constituem novos focos de infecção (CAPRON et al., 2001).

Pelo exposto, sabe-se que os programas de controle baseados somente na quimioterapia têm um efeito apenas temporário na transmissão. Em áreas endêmicas, uma vez que o tratamento em massa termine, a prevalência da doença pode retornar aos níveis anteriores em 18 a 24 meses (Revisado por GRAY et al., 2010). Assim, o controle baseado apenas no tratamento com PZQ precisa ser cuidadosamente avaliado em termos de efetividade e sustentabilidade em longo prazo (GRAY et al., 2010).

Tendo em vista as limitações associadas ao controle da esquistossomose, muitos consideram que a melhor estratégia para tal é a imunização com uma vacina combinada ao tratamento com drogas (BERGQUIST, N. ROBERT; LEONARDO; MITCHELL, 2005). Uma vacina antiesquistossomose iria contribuir muito para o decréscimo da morbidade associada a essa doença através de respostas imunes protetoras que levam a uma redução da carga parasitária e da produção de ovos (Revisado por SIDDIQUI; SIDDIQUI; GANLEY-LEAL, 2011). Na opinião de especialistas, uma vacina capaz de conferir uma proteção de 50% já seria eficaz em reduzir consideravelmente a morbidade e mortalidade decorrentes da infecção por *Schistosoma* (BERGQUIST, N. ROBERT;

LEONARDO; MITCHELL, 2005; BERGQUIST, ROBERT; UTZINGER; MCMANUS, 2008).

2.5 – Desenvolvimento de vacinas contra a esquistossomose

Há alguns anos, indícios demonstraram que o desenvolvimento de uma vacina contra a esquistossomose é um objetivo possível de ser alcançado. O primeiro indício foi que a imunização de camundongos com uma preparação contendo cercárias irradiadas resultou em proteção contra infecções desafio com redução da carga parasitária (LI HSÜ; HSÜ; BURMEISTER, 1981; MINARD *et al.*, 1978). Já o segundo demonstrou que populações humanas em áreas endêmicas que são frequentemente expostas à infecção desenvolvem algum grau de proteção naturalmente (HAGAN, 1992; HAGAN *et al.*, 1991).

Nesse contexto de desenvolvimento de uma vacina, mais de 100 antígenos de *Schistosoma* já foram identificados, dos quais mais de um quarto mostraram algum nível de proteção em modelo murino de esquistossomose. Acredita-se que proteínas expostas que realizam tais funções seriam alvos efetivos para o desenvolvimento de uma vacina (BERGQUIST, ROBERT; UTZINGER; MCMANUS, 2008; BETHONY *et al.*, 2011; MCMANUS; LOUKAS, 2008; SIDDIQUI; SIDDIQUI; GANLEY-LEAL, 2011).

Um forte candidato para composição de uma vacina antiesquistossomose é a proteína Sm21.7 de *S. mansoni*. Sua identificação ocorreu por *screening* de uma biblioteca de expressão do esporocisto utilizando soro de coelhos imunizados com cercárias irradiadas (FRANCIS; BICKLE, 1992).

Análises de *Northern Blot* demonstraram que o RNA (ácido ribonucleico) mensageiro para Sm21.7 está presente no esporocisto, no esquistossômulo e no verme adulto (FRANCIS; BICKLE, 1992). Já análises de RT-PCR (Transcrição Reversa seguida de Reação em Cadeia da Polimerase) demonstraram que a presença de transcritos ocorre em ovos, cercárias e no verme adulto (FITZSIMMONS, COLIN M. *et al.*, 2007). A proteína antigênica, por sua vez, foi encontrada em preparações de esquistossômulo e verme adulto (CURWEN *et al.*, 2004; FRANCIS; BICKLE, 1992).

Em estudos de proteção, a proteína Sm21.7 foi testada como antígeno em vacinas recombinante (AHMED *et al.*, 2001) e de DNA (ácido desoxirribonucleico)

(AHMED *et al.*, 2006). A vacinação de camundongos *Swiss Albino* e *C57BL/6* com a proteína recombinante conferiu 40 e 71%, respectivamente, de redução no número de vermes adultos após a infecção desafio. Exame patológico do fígado dos camundongos revelou uma redução no número, tamanho e celularidade do granuloma (AHMED *et al.*, 2001).

A proteção induzida por Sm21.7 também foi avaliada por vacina de DNA e atingiu níveis significativos (41 e 53%) em camundongos *Swiss Albino* e *C57BL/6*, respectivamente. Houve também um decréscimo no número, tamanho e celularidade dos granulomas nos animais vacinados quando comparados aos controles. Ainda foi observada a redução na viabilidade e fecundidade dos parasitas e também na eclosão de ovos, além de uma redução de 62% e 67% no número de ovos no fígado e intestino, respectivamente (AHMED *et al.*, 2006).

Recentemente, um grupo colaborador identificou a proteína P22 de *S. mansoni*, uma proteína da fração PIII de preparações do verme adulto. Para obtenção dessa fração PIII, o extrato solúvel do verme adulto (SWAP) foi submetido à cromatografia de troca iônica, originando as frações PI a PVI (HIRSCH; GOES, 1996). De todas essas frações, PIII apresentou os resultados mais promissores em ensaios de imunização, sendo capaz de produzir efeitos protetores e imunomodulatórios no modelo de esquistossomose murina (HIRSCH *et al.*, 1997). Um cDNA (DNA complementar) codificador da proteína P22 foi isolado de uma biblioteca de verme adulto utilizando soro anti-PIII e mostrou 100% de identidade a Sm21.7 (REZENDE *et al.*, 2011). Diante disso, observa-se que P22 e Sm21.7 são diferentes denominações para a mesma proteína.

Camundongos imunizados com a proteína recombinante P22 (rP22) exibiram um decréscimo de 51% e 22,5% na carga parasitária e ovos no fígado, respectivamente. Além disso, vacinação com rP22 produziu uma redução de 60% no tamanho do granuloma do fígado e 71% da fibrose em testes murinos, sugerindo que rP22 é um forte candidato para compor uma vacina contra o *S. mansoni* (REZENDE *et al.*, 2011).

Análises de imunolocalização demonstraram que Sm21.7 localiza-se no tegumento do parasita (AHMED *et al.*, 2001; REZENDE *et al.*, 2011). Esse tegumento representa a interface entre o microrganismo e o hospedeiro e é por meio dele que o parasita realiza algumas funções vitais que permitem sua sobrevivência. Diante disto, o estudo desta estrutura celular, seus componentes e como ela interage com o sistema imune pode fornecer informações importantes sobre o controle da doença,

especialmente aquelas relacionadas à procura por novas drogas e desenvolvimento de uma vacina.

Devido às proteínas P22 e Sm21.7 apresentarem 100% de identidade, a denominação Sm21.7 será utilizada no presente trabalho para facilitar o acompanhamento.

A sequência de nucleotídeos de Sm21.7 codifica uma proteína de 184 resíduos de aminoácidos com massa molecular de aproximadamente 21,7 kDa (FRANCIS; BICKLE, 1992). Uma busca no BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>), demonstrou que Sm21.7 pertence a uma família de proteínas contendo um motivo de mão EF N-terminal (Arg17-Gly67) e um domínio de cadeia leve de dineína (DLC) C-terminal (Lys81-Thr181) (FRANCIS; BICKLE, 1992; REZENDE *et al.*, 2011) (Figura 5).

```

MDSPMEKFIQTYLTLLRDGDETVETSKLSESCRKEKLDMKQVNEWIALFDVDKDQKITFE 60
EFCRGLGLKQNMRIERNHIKTVQSGREQSLPEGVSIIASTMPKPKQVEVTQLFKDIYNE 120
VKKDPDMNKVVVTKFKSELERRYGRVWQVNAVTHSYWASFSEPFQSIQFYDNKIIILAWR 180
TPSN 184

```

Figura 5: Sequência de aminoácidos da proteína Sm21.7 e localização dos domínios que a constituem. A sequência da proteína Sm21.7 apresenta um motivo de mão EF, destacado em cinza claro; um domínio de cadeia leve de dineína, destacado em cinza escuro e um domínio provisório de calmodulina, sublinhado em vermelho. Os números ao final de cada linha indicam a posição dos resíduos de aminoácidos (Adaptado de REZENDE *et al.*, 2011).

Proteínas contendo essa organização de domínios (domínio de mão EF N-terminal e domínio de cadeia leve de dineína C-terminal) foram encontradas exclusivamente no tegumento de trematódeos, constituindo assim, alvos interessantes para o desenvolvimento de fármacos e vacinas.

Exemplos de tais proteínas, além de Sm21.7, são: Sm22.6 (JEFFS *et al.*, 1991; STEIN; DAVID, 1986), Sm20.8 (HOFFMANN, 1997; MOHAMED *et al.*, 1998) e Sm21.6 (LOPES *et al.*, 2009) de *S. mansoni*. Em *S. japonicum* foram identificadas Sj22.6 (SANTIAGO *et al.*, 1998) e Sj21.7 (HAFALLA *et al.*, 1999). Já em *S. haematobium* foi identificada a proteína Sh22.6 (FITZSIMMONS, C. M. *et al.*, 2004).

Em outros trematódeos foram identificadas: FgCaBP1, FgCaBP2, FgCaBP3 e FgCaBP4 (SUBPIATTANA; GRAMS; VICHASRI-GRAMS, 2012; VICHASRI-

GRAMS *et al.*, 2006) de *Fasciola gigantica* e FhCaBP3 (BANFORD *et al.*, 2013) e FhCaBP4 (ORR *et al.*, 2012) de *Fasciola hepatica*.

No verme *Opisthorchis viverrini* também foi identificada uma proteína com esses domínios (OvCaBP) (SENAWONG *et al.*, 2012), assim como em *Clonorchis sinensis*, no qual as seguintes proteínas foram identificadas: CsTegu21.6 (KIM, YU-JUNG *et al.*, 2012), CsTP21.1 (CHEN *et al.*, 2011), CsTP22.3 (ZHOU, ZHENWEN *et al.*, 2008), CsTP20.8 (ZHOU, ZHENWEN *et al.*, 2007) e CsTP31.8 (HUANG *et al.*, 2007). A figura 6 apresenta um alinhamento de algumas dessas proteínas, destacando os motivos de mão EF e de cadeia leve de dineína.

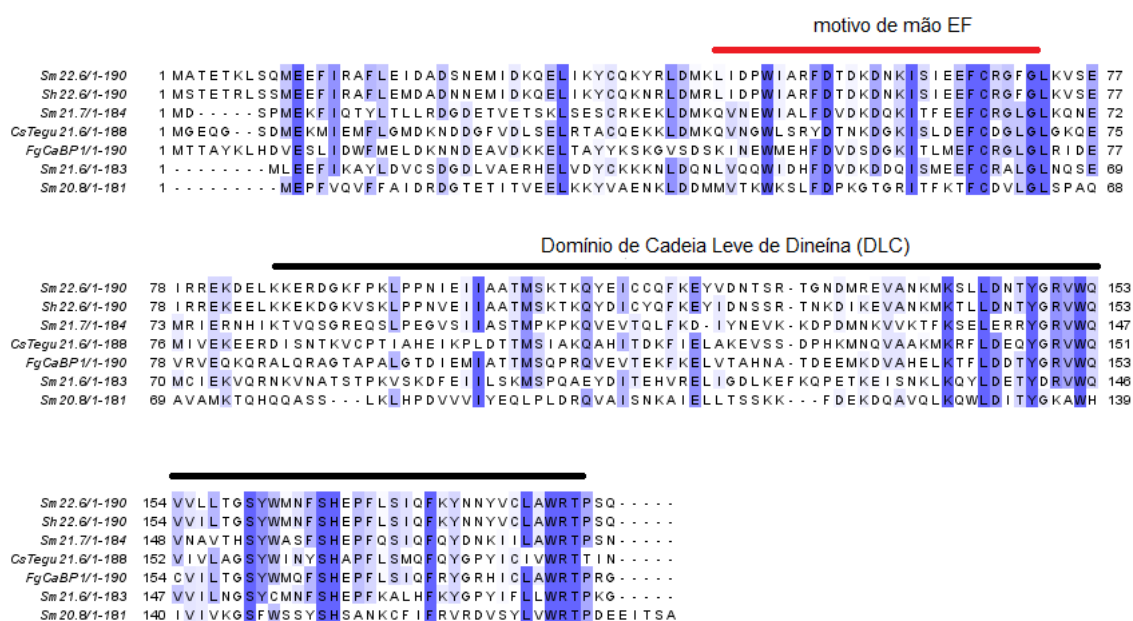


Figura 6: Alinhamento de proteínas de trematódeos contendo motivos de mão EF e de cadeia leve de dineína (DLC). Algumas proteínas identificadas em trematódeos possuem em sua sequência um ou dois domínios de mão EF na parte N-terminal e um DLC na parte C-terminal. Na imagem, o motivo de mão EF e o DLC da proteína Sm21.7 estão destacados em vermelho e preto, respectivamente. No alinhamento, os tons de azul detacam a conservação dos aminoácidos, sendo maior a conservação quanto mais escura a tonalidade do azul. As proteínas alinhadas foram: Sm22.6 de *S. mansoni* (resíduos 1 a 190), Sh22.6 *S. haematobium* (resíduos 1 a 190), Sm21.7 *S. mansoni* (resíduos 1 a 184), CsTegu21.6 de *C. sinensis* (resíduos 1 a 188), FgCaBP de *F. gigantica* (resíduos 1 a 190), Sm21.6 de *S. mansoni* (resíduos 1 a 183) e Sm20.8 de *S. mansoni* (resíduos 1 a 181). A imagem foi gerada no Jalview (WATERHOUSE *et al.*, 2009).

Análises de modelagem molecular demonstraram que a proteína FhCaBP4 de *F. hepatica* apresenta os domínios de mão EF e de cadeia leve de dineína bem enovelados e interligados por um segmento flexível (ORR *et al.*, 2012) (Figura 7).

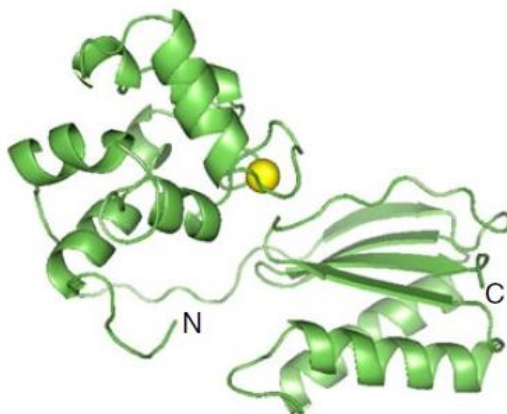


Figura 7: Estrutura da proteína FhCaBP4 de *F. hepatica* obtida por modelagem molecular. A estrutura dessa proteína é constituída por um domínio de mão EF na parte N-terminal (região superior esquerda) e por um domínio de cadeia leve de dineína na porção C-terminal (região da direita), que são ligados por um segmento flexível. O modelo foi gerado pelo Phyre2 server por homologia e o modelo final foi solvatado com minimização de energia pelo YASARA. A esfera amarela representa um íon cálcio ligado ao domínio de mão EF (ORR *et al.*, 2012).

Em proteínas contendo motivo de mão EF, um *loop* constituído de 12 resíduos é responsável pela ligação ao íon de cálcio. Esse loop possui 12 resíduos com o seguinte padrão: X*Y*Z*-Y*-X**-Z, no qual X, Y, Z, -Y, -X, -Z são os resíduos responsáveis pela ligação ao cálcio e “*” representam os outros resíduos (KRETSINGER; NOCKOLDS, 1973) (Figura 8). Apesar da existência do motivo de mão EF na proteína Sm21.7, os experimentos de ligação ao ^{45}Ca tanto com a proteína Sm21.7 recombinante quanto com a proteína derivada do parasita não mostraram evidências de ligação a esse íon, sugerindo a presença de um domínio não funcional (FRANCIS; BICKLE, 1992). Esse fato pode ser explicado pela substituição de uma glicina na posição 6 no *loop* de ligação ao cálcio por uma glutamina (Figura 8). Quase todas as estruturas de mão EF possuem uma glicina nessa posição e a glutamina presente em Sm21.7 poderia desestabilizar a conformação e a geometria específicas para ligação ao ^{45}Ca (FRANCIS; BICKLE, 1992). Dentre as proteínas de trematódeos citadas anteriormente com domínio de mão EF e domínio de cadeia leve de dineína, além de Sm21.7, apenas Sm20.8 não liga-se ao íon cálcio (MOHAMED *et al.*, 1998).

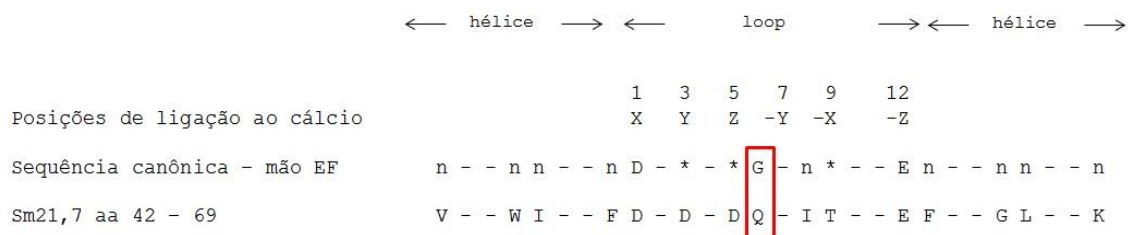


Figura 8: Alinhamento do motivo de mão EF canônico com o motivo de mão EF da proteína Sm21.7. Nesse alinhamento, observa-se na posição 6 do *loop* a presença de uma glicina na sequência canônica, enquanto em Sm21.7 observa-se uma glutamina (destacadas em vermelho), que poderia ser responsável por Sm21.7 não ligar-se a ^{45}Ca . No motivo canônico, o *loop* contém 6 posições responsáveis pela ligação do íon cálcio: X, Y, Z, -Y, -X, -Z. Nas sequências apresentadas, “*” representa um oxigênio na cadeia lateral (D,E,N,Q,S,T). “-” representa qualquer resíduo. “n” representa um resíduo apolar (Adaptado de FRANCIS; BICKLE, 1992).

2.6 – Domínios de mão EF e de cadeia leve de dineína (DLC) - os domínios constituintes da proteína Sm21.7

Analisando individualmente cada um dos domínios constituintes da proteína Sm21.7, pode-se descrever algumas características das proteínas que possuem domínios de mão EF ou de cadeia leve de dineína.

Proteínas contendo o motivo de mão EF que se ligam ao cálcio regulam diversos processos celulares, pois são sensíveis e capazes de responder a alterações nas concentrações desse íon (GOPALAKRISHNA; ANDERSON, 1985).

O motivo de mão EF foi descoberto por R.H. Kretsinger na análise da estrutura cristalográfica da parvalbumina (KRETSINGER; NOCKOLDS, 1973). Esse motivo é constituído por duas α -hélices quase perpendiculares (chamadas hélices E e F na parvalbumina) separadas por um *loop* de 12 resíduos (KRETSINGER; NOCKOLDS, 1973; STRYNADKA; JAMES, 1989) (Figuras 9 e 10). Como citado anteriormente, no *loop*, os resíduos 1, 3, 5, 7, 9 e 12 (representados por X, Y, Z, -Y, -X e -Z, respectivamente, nas figuras 8 e 9), interagem com os íons cálcio através de suas cadeias laterais (Revisado por MICHIELS *et al.*, 2002). Esses resíduos são: X e Y, ácido aspártico ou asparagina; Z, ácido aspártico, asparagina ou serina; -Y, uma cadeia lateral contendo um grupo carbonil; -X, na grande maioria das estruturas é uma molécula de água e -Z, um ácido glutâmico ou aspártico (Figura 9) (FRANCIS; BICKLE, 1992; LEWIT-BENTLEY; RÉTY, 2000).

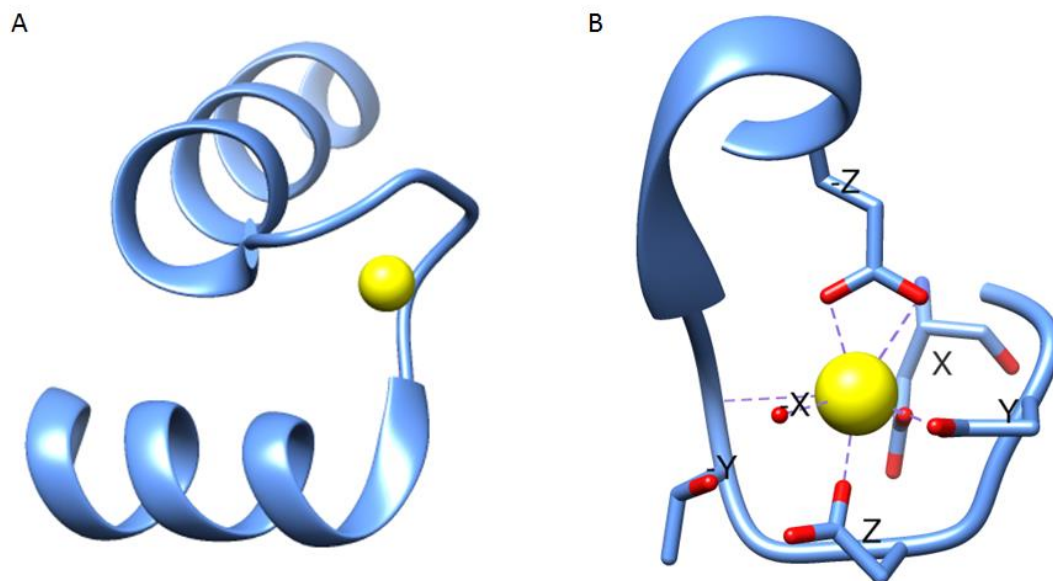


Figura 9: Motivo de mão EF ligado ao cálcio (PDB 1XO5) (GENTRY et al., 2005). (A) Motivo de mão EF com as α -hélices E e F quase perpendiculares e o *loop* de ligação ao cálcio entre as mesmas. (B) Destaque do *loop* de ligação ao cálcio evidenciando os resíduos cujas cadeias laterais interagem com o íon. A esfera amarela representa o íon cálcio. A imagem foi gerada pelo programa UCSF Chimera (PETTERSEN et al., 2004).

Proteínas com mão EF geralmente são constituídas por múltiplos domínios, cada um contendo um par de motivos de mão EF (Figura 10). Algumas superfamílias de proteínas contendo domínios de mão EF são naturalmente encontradas como dímeros de domínios idênticos ou domínios intimamente relacionados (Figura 11), introduzindo outra dimensão de complexidade na arquitetura molecular (YAP, K. L. *et al.*, 1999).

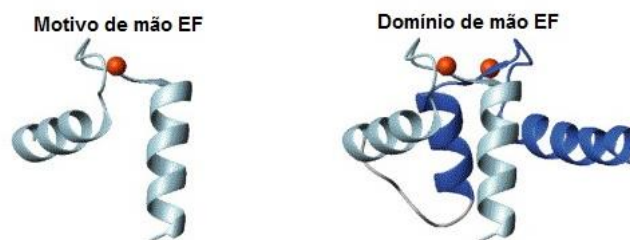


Figura 10: Estruturas do motivo de mão EF e do domínio de mão EF. A estrutura de um motivo de mão EF isolado possui a organização de hélice-*loop*-hélice. Já o domínio de mão EF, possui dois motivos de mão EF interligados. A esfera vermelha representa o íon cálcio (Adaptado de BHATTACHARYA; BUNICK; CHAZIN, 2004).

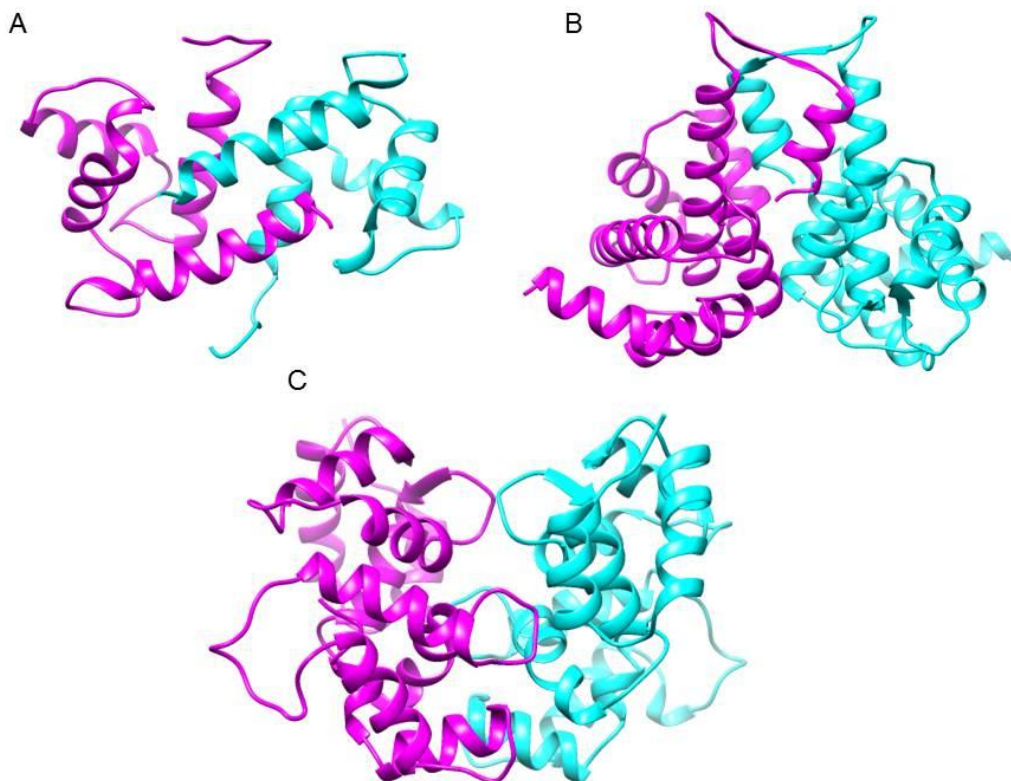


Figura 11: Formação de dímeros por proteínas que possuem domínios de mão EF. As proteínas contendo esses domínios podem ser encontradas como dímeros, como por exemplo: (A) Proteína S100 (PDB 1QLK – (DROHAT et al., 1998)) constituída por um dímero, sendo cada monômero formado por 2 motivos de mão EF, (B) Calpaína (PDB 1DVI – (BLANCHARD et al., 1997)), na qual cada monômero contém 5 motivos de mão EF e (C) Neurocalcina que também possui 5 motivos em cada monômero (PDB 1BJF – (VIJAY-KUMAR; KUMAR, 1999)). Cada monômero é destacado em ciano ou magenta na figura (Adaptado de YAP, K. L. et al., 1999). A imagem foi gerada pelo programa UCSF Chimera (PETTERSEN et al., 2004).

O motivo de mão EF é encontrado em mais de 50% das proteínas que se ligam ao cálcio (Revisado por ZHOU, YUBIN *et al.*, 2006) e a família das proteínas que possui esse motivo inclui 66 subfamílias (LEWIT-BENTLEY; RÉTY, 2000; ZHOU, YUBIN *et al.*, 2006), dentre as quais se destacam troponina C, cadeia leve de miosina, parvalbumina (COLLINS, 1976), calmodulina (CHIN; MEANS, 2000; HOEFLICH; IKURA, 2002) e S100 (KAWASAKI; NAKAYAMA; KRETSINGER, 1998; LEWIT-BENTLEY; RÉTY, 2000; ZHOU, YUBIN *et al.*, 2006).

A mais bem caracterizada das proteínas com motivo de mão EF é a calmodulina, uma proteína de 17 kDa que possui duas cabeças globulares ligadas por uma α -hélice flexível, sendo que cada cabeça globular contém dois motivos de mão EF (HOEFLICH;

IKURA, 2002; SPYRACOPOULOS *et al.*, 2001; ZHANG; TANAKA; IKURA, 1995). A flexibilidade do *loop* interdomínios confere à calmodulina a habilidade de reconhecer e ligar-se a mais de 350 proteínas alvo (TJANDRA *et al.*, 2008; YAP, KYOKOL. *et al.*, 2000), dentre as quais estão as serina/treonina cinases (MAL *et al.*, 2002).

A ligação ao cálcio na calmodulina induz uma alteração conformacional que resulta na exposição de resíduos hidrofóbicos (GOPALAKRISHNA; ANDERSON, 1985; MAKI *et al.*, 2002), sendo estes importantes na interação com algumas proteínas alvo e também com algumas drogas (MATSUSHIMA *et al.*, 2000; ZHANG; YUAN, 1998), caracterizando a calmodulina como uma proteína chamada sensora (CHIN; MEANS, 2000).

Além das proteínas sensoras, existe um grupo de proteínas contendo mão EF que não sofrem alterações de conformação significativas com a ligação do cálcio. Tais proteínas funcionam como tampões ou transportadores de cálcio. Um exemplo importante desse grupo de proteínas é a parvalbumina (CHIN; MEANS, 2000).

A grande versatilidade do motivo de mão EF é claramente refletida no crescente número de estruturas tridimensionais resolvidas, revelando uma grande variedade de conformações, organização do domínio e resposta estrutural à ligação do cálcio (GRABAREK, 2006).

O outro fragmento da proteína Sm21.7 é constituído por um domínio de cadeia leve de dineína. As moléculas de dineína pertencem a duas classes: as dineínas de axonema, responsáveis pelo deslizamento dos microtúbulos uns contra os outros para gerar movimentos flagelar e ciliar e as dineínas citoplasmáticas, envolvidas no transporte pelos microtúbulos (DOWNING; SUI, 2007; GENNERICH; VALE, 2009; VALE, 2003).

A dineína citoplasmática é uma molécula motora implicada em uma variedade de funções, incluindo movimento retrógrado de organelas, migração nuclear, alinhamento do fuso mitótico e transporte axonal (HIROKAWA, 1998; KARKI; HOLZBAUR, 1999; KING, S. M., 2000; STUCHELL-BRERETON *et al.*, 2011; VALLEE; SHEETZ, 1996). O complexo que constitui a dineína possui moléculas de cadeia pesada (530 kDa cada), de cadeias intermediárias (74 kDa), de cadeias leves intermediárias (52 a 61 kDa) e de cadeias leves (8 a 22 kDa) (FAN, JING-SONG *et al.*, 2001; HUGHES *et al.*, 1995; PASCHAL *et al.*, 1992), algumas existindo como homodímeros no complexo (Revisado por STUCHELL-BRERETON *et al.*, 2011).

Cada subunidade da cadeia pesada contém um domínio de ligação aos microtúbulos e seis domínios de ATPases associadas a diversas atividades celulares, que convertem a energia da hidrólise de ATP (adenosina trifosfato) em movimento da dineína pelos microtúbulos. A região N-terminal das cadeias pesadas interage com os componentes das cadeias intermediárias e cadeias leves intermediárias. As cadeias intermediárias, por sua vez, através de sua região N-terminal, interagem com as cadeias leves (Revisado por STUCHELL-BRERETON *et al.*, 2011).

A cadeia leve da dineína citoplasmática (DLC8) foi originalmente identificada como uma cadeia da dineína em *Chlamydomonas* (MITCHELL; ROSENBAUM, 1986; PFISTER; FAY; WITMAN, 1982; PIPERNO; LUCK, 1979). O DLC8 contém 89 resíduos e é altamente conservado ao longo da evolução (90% de identidade de *Caenorhabditis elegans* a humanos) (ROMES; TRIPATHY; SLEP, 2012).

A contribuição das cadeias leves para o funcionamento da dineína permanece uma questão em aberto. Entretanto, sabe-se que elas são necessárias para a realização de funções como transporte intraflagelar retrógrado (PAZOUR; WILKERSON; WITMAN, 1998), migração nuclear, localização do complexo motor (BECKWITH *et al.*, 1998), e também para a correta orientação do fuso mitótico (DUNSCH *et al.*, 2012).

Alguns dados da literatura sugerem que as cadeias leves estabilizam a estrutura da dineína, entretanto, o mecanismo dessa atuação não é bem determinado. Estudos com *Saccharomyces cerevisiae* sugerem que a estabilização pode ser consequência da interação entre cadeias leves e moléculas de dinactina (STUCHELL-BRERETON *et al.*, 2011), que é uma proteína multimérica que se liga à molécula de dineína, sendo necessária para suas atividades (SCHROER, 2004).

Apesar das cadeias leves de dineína, juntamente com as cadeias pesadas, intermediárias e leves intermediárias, atuarem como um complexo, uma significativa proporção de moléculas identificadas como DLC não está associada com o complexo. Esse fato levanta a hipótese de que essas moléculas exercem funções não relacionadas às da dineína (Revisado por BARBAR, 2008).

Assim, alguns dados sugerem que as cadeias leves de dineína podem atuar como reguladores de mais de 100 proteínas diversas (Revisado por BARBAR, 2008; ZAYTSEVA *et al.*, 2014).

Já foi demonstrada a associação de cadeias leves de dineína com uma grande variedade de polipeptídeos incluindo a óxido nítrico sintase neuronal (nNOS) (FAN, J. S. *et al.*, 1998; JAFFREY; SNYDER, 1996), miosina V (BENASHSKI *et al.*, 1997), o

fator proapoptótico *Bim*, fatores transcricionais *Swallow* (SCHNORRER; BOHMANN; NUSSLEIN-VOLHARD, 2000) e *Trps1* (KAISER *et al.*, 2003), proteínas virais (RAUX; FLAMAND; BLONDEL, 2000) e proteína ligadora de p53 (LO *et al.*, 2005). Também já foi reportada a interação de cadeias leves de dineína com um peptídeo que atua no processo de duplicação dos centríolos durante a divisão celular (SLEVIN *et al.*, 2014) e com o receptor de orexina, um peptídeo neurotransmissor (DUGUAY *et al.*, 2011).

Estudos de RMN (Ressonância Magnética Nuclear) e cristalografia revelaram que, quando ligados a um dímero de cadeias leves de dineína, os polipeptídeos *Bim*, nNOS e *Swallow* ocupam uma fenda semelhante na interface dimérica da cadeia leve (Figura 12) (Revisado por BARBAR, 2008).

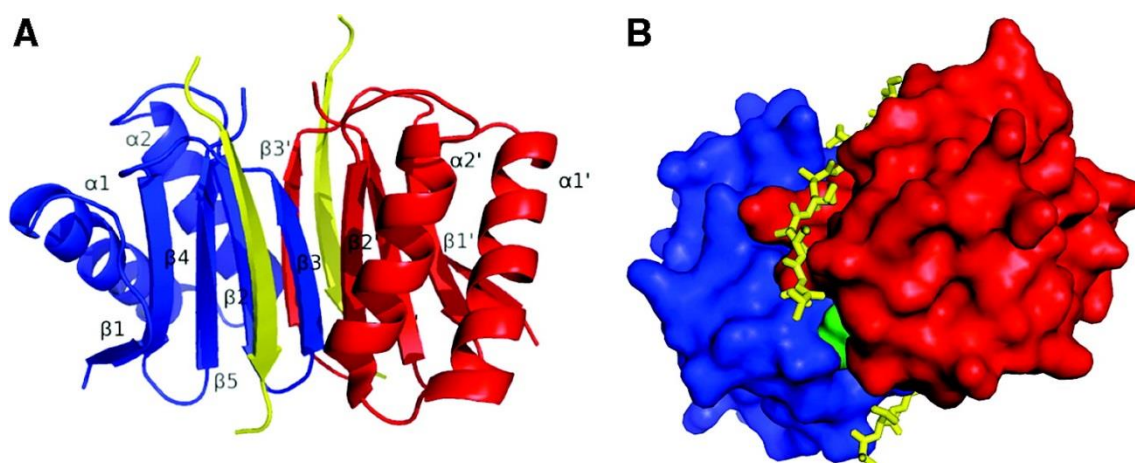


Figura 12: Peptídeo *Swallow* ligado ao dímero de cadeia leve de dineína PDB (3E2B). (A) Representação do dímero do domínio de cadeia leve de dineína, sendo que um monômero está representado em azul e o outro em vermelho. O peptídeo está representado em amarelo. (B) Representação de superfície demonstrando o peptídeo *Swallow* (em amarelo) na interface do dímero (BARBAR, 2008).

Além dessas associações com outras proteínas e peptídeos foi demonstrado que a cadeia leve de dineína (Dyn2p) de *Yarrowia lipolytica* atua na biogênese e maturação dos peroxissomos. Foi observado que a ausência de Dyn2p da levedura *Y. lipolytica* resulta em danos na função do peroxissomo e na biogênese dos mesmos. Dyn2p também atua no endereçamento de proteínas para a matriz peroxissomal (CHANG *et al.*, 2013). Adicionalmente, foi demonstrado que leveduras *S. cerevisiae* com o gene para a cadeia leve de dineína DYN2 deletado foram incapazes de utilizar ácido oleico como fonte de carbono, o metabolismo do qual exige peroxissomos funcionais (SMITH

et al., 2006). Esse papel da cadeia leve de dineína adiciona-se à lista de funções não relacionadas à motilidade que podem ser atribuídas a componentes individuais do complexo dineína (CHANG *et al.*, 2013).

Como citado anteriormente, DLC8 é capaz de formar dímeros e a estrutura de cada monômero contém 4 ou 5 fitas- β antiparalelas e duas α -hélices (Figura 13) (FAN, JING-SONG *et al.*, 2001). Cadeias leves de dineína dimerizam-se através de fitas- β que se estendem pelos dois monômeros da molécula (Figura 14) (FAN, JING-SONG *et al.*, 2001; LIANG *et al.*, 1999; ROMES; TRIPATHY; SLEP, 2012).

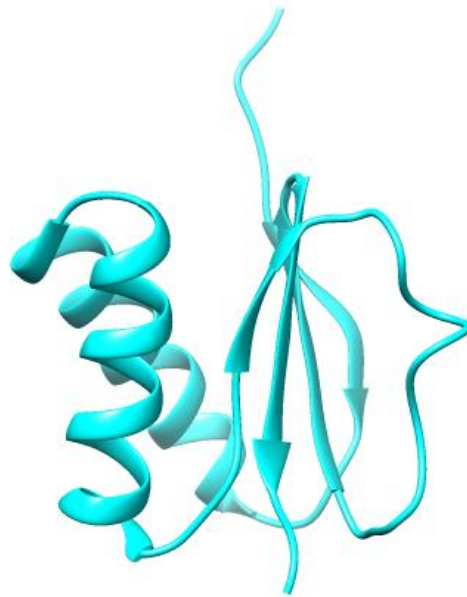


Figura 13: Estrutura tridimensional da cadeia leve de dineína – DLC8 (PDB 1PWK). A estrutura deste polipeptídeo possui 4 fitas- β formando uma folha- β antiparalela e duas α -hélices (Adaptado de WANG *et al.*, 2003). A imagem foi gerada pelo programa UCSF Chimera (PETTERSEN *et al.*, 2004).

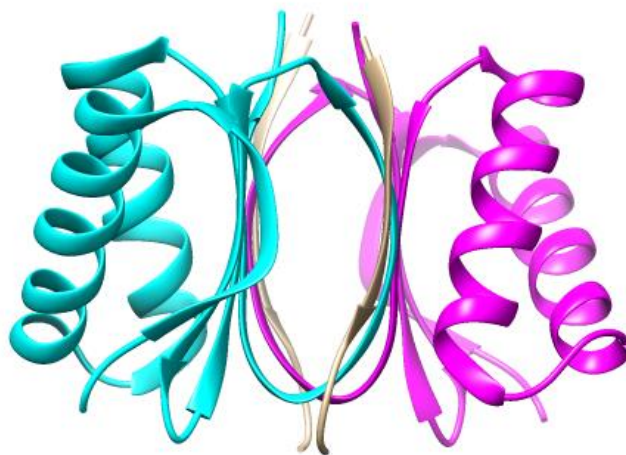


Figura 14: Estrutura dimérica da cadeia leve de dineína (PDB-4DS1). O domínio de cadeia leve de dineína pode formar dímero através da região da molécula monomérica que possui uma folha- β antiparalela. Cada um dos monômeros está destacado em ciano e magenta (Adaptado de ROMES; TRIPATHY; SLEP, 2012). A imagem foi gerada pelo programa UCSF Chimera (PETTERSEN et al., 2004).

Tendo em vista as informações aqui expostas e a justificativa descrita evidenciando a vacinologia estrutural como técnica promissora no desenvolvimento de novas vacinas, o presente trabalho tem como objetivo principal caracterizar estruturalmente a proteína recombinante Sm21.7 através de análises bioquímicas e biofísicas.

3 – Objetivos

3.1 – Objetivo geral

Caracterizar a estrutura molecular da forma recombinante da proteína Sm21.7 de *S. mansoni*.

3.2 – Objetivos específicos

- 1- Subclonar o gene codificador da proteína Sm21.7 no vetor de expressão pET28a(TEV);
- 2- Expressar a proteína recombinante em sistema procariótico;
- 3- Purificar a proteína recombinante por cromatografias de afinidade e exclusão molecular;
- 4- Realizar ensaios de análise de estrutura secundária através da técnica de Dicroísmo Circular (CD);
- 5- Analisar o estado de oligomerização da proteína recombinante através de Espalhamento Dinâmico da Luz (DLS);
- 6- Realizar ensaios de cristalização e coleta de dados de difração de raios X;
- 7- Determinar a estrutura tridimensional da proteína.

4 – Materiais e métodos

4.1 – Subclonagem do gene codificador da proteína Sm21.7 no vetor pET28a(TEV)

4.1.1 – Amplificação do gene codificador da proteína Sm21.7

O gene codificador da proteína Sm21.7 (Gene ID 8347700) foi amplificado através da técnica de PCR (Reação em Cadeia da Polimerase) utilizando-se um plasmídeo recombinante previamente construído, pETDEST42-rP22 (REZENDE *et al.*, 2011), como DNA molde. Os iniciadores usados na reação de amplificação (Tabela 1) foram desenhados com base na sequência de nucleotídeos depositada em banco de dados e, para permitir a clonagem no vetor de expressão, foram adicionados sítios de restrição para as enzimas *NdeI* e *XhoI* aos iniciadores direto e reverso, respectivamente.

Tabela 1: Iniciadores direto e reverso utilizados para amplificação do gene codificador da proteína Sm21.7. Os iniciadores direto e reverso utilizados para amplificação possuem sítios de restrição para as enzimas *NdeI* e *XhoI*, respectivamente, que estão sublinhados.

Iniciador	Sequência	T _m
Direto:	5´ <u>CATATGG</u> ATAGTCCAATGGAAAAATTTATTC 3´	50°C
Reverso:	5´ <u>CTCGAGT</u> CTAGTTACTTGGTGTACGCCAAG 3´	56°C

Sendo assim, para amplificação do gene de interesse, uma reação com volume final de 50 µL foi preparada. Em um microtubo foram adicionados 5 µL de tampão 10x (600 mM Tris-SO₄ pH 8,9, 180 mM Sulfato de amônio), 1 µL de uma mistura dos 4 desoxinucleotídeos (concentração final de 0,2 mM de cada desoxinucleotídeo), 2 µL de MgSO₄ (concentração final de 2 mM), 0,2 µL de *PlatinumR Taq High Fidelity* (1 unidade de enzima) (Invitrogen), 1 µL de cada iniciador (concentração final de 0,2 µM), 3,0 µL do DNA molde e água livre de DNase e RNase para completar o volume.

A reação foi executada em um termociclador MJ Mini Personal Thermal Cycler (Bio-Rad) utilizando-se os seguintes parâmetros:

- Desnaturação a 94°C por 2 minutos;
- 25 ciclos de desnaturação a 94°C por 30 segundos, anelamento a 55°C por 30 segundos e extensão a 68°C por 1 minuto;
- Extensão final a 68°C por 5 minutos.

Ao término dos ciclos de amplificação, 5 µL do produto da PCR foram analisados por eletroforese em gel de agarose para verificar a amplificação da sequência codificadora da proteína Sm21.7. Posteriormente, o restante do produto de PCR foi purificado utilizando-se o Kit *Illustra GFX PCR DNA and Gel Band Purification* (GE Healthcare) seguindo-se as instruções do fabricante. A concentração do produto de PCR purificado, em ng/µL, foi estimada através da absorbância a 260 nm medida em um equipamento nanodrop ND1000 (NanoDrop).

4.1.2 – Eletroforese em gel de agarose

Para visualização do DNA proveniente dos procedimentos da clonagem, foi utilizada a eletroforese em gel de agarose 1% corado com *SYBR® safe* (Life Technologies). Em tal processo, 0,3 g de agarose foram adicionados a 30 mL de TAE 0,5x (20 mM Tris, 0,0571 % ácido acético, 50 mM EDTA (Ácido Etilenodiamino Tetra-acético) pH 8,0) e a mistura foi aquecida até a completa dissolução da agarose. Posteriormente, quando a temperatura atingiu aproximadamente 60°C, 1 µL de *SYBR® safe* foi adicionado e todo o volume (30 mL) foi despejado em um suporte adequado para a solidificação do gel. As amostras foram, então, misturadas em tampão de amostra (50% (v/v) glicerol, 0,25% (p/v) azul de bromofenol e 0,25% (p/v) xileno cianol) e carregadas nas canaletas. Em seguida, o gel foi submetido a uma corrente de 80 V por 30 minutos e as bandas foram visualizadas por iluminação UV.

4.1.3 – Purificação do produto de PCR e bandas do gel

O kit *Illustra GFX PCR DNA and Gel Band Purification* (GE Healthcare) foi utilizado para purificação de bandas provenientes de géis de agarose ou de produtos de PCR. Nesse procedimento de purificação, as bandas de interesse visualizadas no gel de agarose foram excisadas e 10 µL de *Capture buffer 3* foram adicionados para cada 10 mg de gel. Então, esse material foi aquecido à 60°C até a completa dissolução dos fragmentos do gel. Já para purificação de produtos de PCR, 500 µL de *Capture Buffer 3*

foram adicionados à até 100 µL de produto de PCR e o material foi homogeneizado. Posteriormente, todo o volume dos procedimentos citados acima foram carregados em uma coluna fornecida pelo fabricante (GFX MicroSpin) seguida por uma incubação de 1 minuto à temperatura ambiente. Então, a coluna foi centrifugada a 16.000 *g* por 1 minuto e o *flow through* foi descartado. Em seguida, 500 µL de *Washing buffer type 1* foram adicionados à coluna e a mesma foi centrifugada novamente nas mesmas condições anteriores. A coluna foi transferida para um novo tubo para a eluição do DNA, que ocorreu através da adição de aproximadamente 30 µL de água livre de DNase e RNase seguida pela incubação à temperatura ambiente e pela centrifugação à 16.000 *g* por 1 minuto. O *Capture buffer 3* e o *Washing buffer 1* são fornecidos no kit.

4.1.4 – Clonagem do gene codificador da proteína Sm21.7 no vetor pGEM®-T

O sistema de clonagem pGEM®-T (Promega) (Figura 15) é amplamente utilizado para clonagem de produtos de PCR. Tal sistema foi construído através da clivagem do vetor pGEM®-5Zf(+) com a enzima *EcoRV* seguida da adição de uma timidina 3' a ambas as extremidades. Essa timidina aumenta a eficiência da ligação do produto de PCR no plasmídeo pela prevenção da recircularização do vetor e pelo fornecimento de uma extremidade coesiva para os produtos de PCR, que geralmente possuem uma adenosina nas extremidades 3' (<http://www.promega.com>).

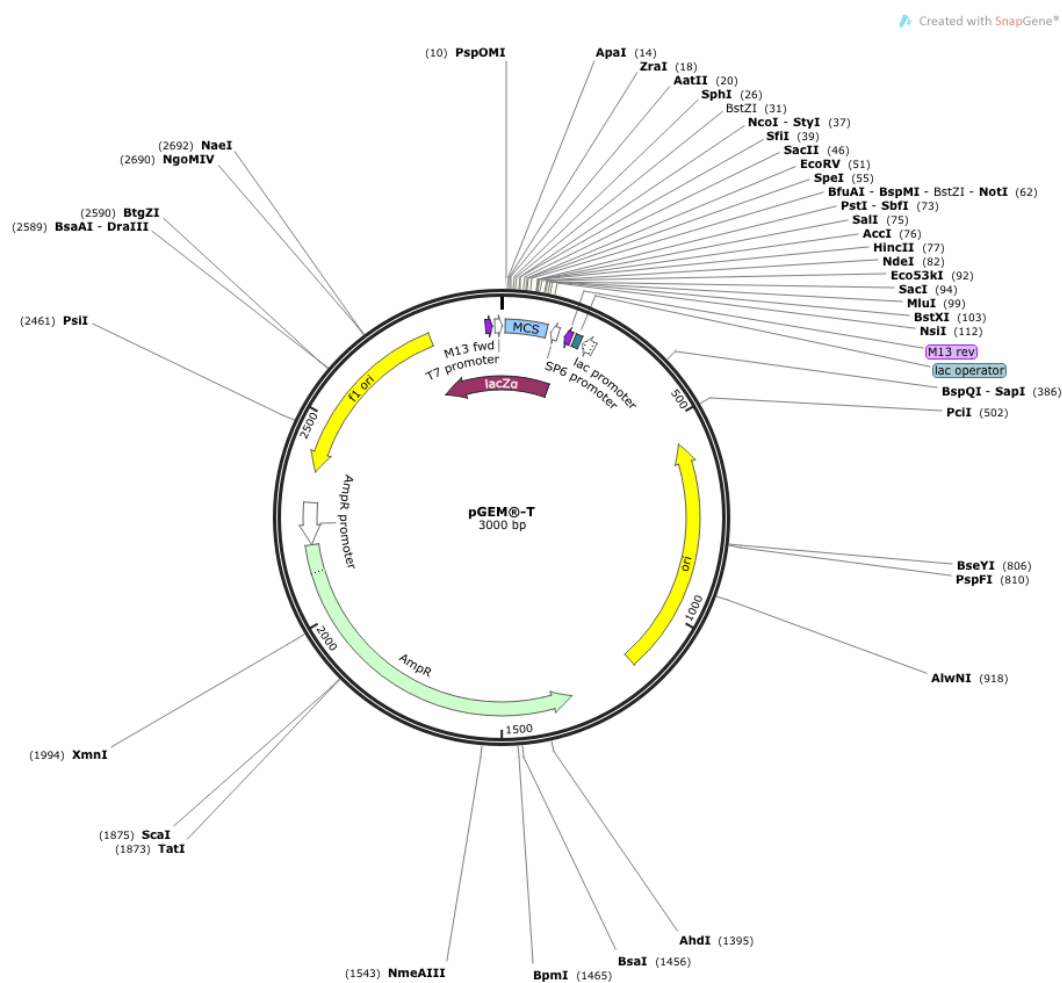


Figura 15: Mapa circular do vetor de clonagem pGEM®-T. A imagem do vetor de 3.000 pb utilizado para clonagem de produtos de PCR mostra que o mesmo possui resistência a ampicilina e um sítio múltiplo de clonagem. Também apresenta a origem f1, que permite a produção de vetor de fita simples, a origem de replicação em bactérias (ori), o operador lac e o promotor T7.

Para clonagem no sistema pGEM®-T, inicialmente, a concentração de inserto necessária para a reação de ligação foi calculada segundo a relação:

$$\text{concentração do inserto} = \frac{\text{concentração do vetor} \times \text{tamanho do inserto}}{\text{tamanho do vetor}} \times 3,$$

na qual:

concentração do vetor = 50 ng

tamanho do inserto = 555 pb

tamanho do vetor = 3.000 pb.

Posteriormente, uma reação de 10 µL para ligação do inserto ao vetor foi preparada. A concentração de inserto determinada pela relação citada acima foi de 28 ng, aos quais foram misturados 1 µL de vetor (50 ng), 5 µL de Tampão 2x, 1 µL de T4 DNA ligase (3 unidades de enzima) e água livre de DNase e RNase para completar o volume final. A reação foi incubada por 16 horas a 4°C.

Imediatamente após o término da reação de ligação, o plasmídeo recombinante Sm21.7-pGEM-T foi transformado em bactérias eletrocompetentes *Escherichia coli* DH5α. Essa bactéria foi desenvolvida por Hanahan e possui uma mutação *endA1* que auxilia na redução da degradação por nucleases e *recA1* que reduz a recombinação homóloga tornando o inserto mais estável (HANAHAHAN, 1983).

Para a transformação do plasmídeo em *E. coli* DH5α, 50 µL de bactérias eletrocompetentes foram misturados a 3 µL do plasmídeo recombinante em uma cubeta adequada. A mistura foi submetida a um choque em eletroporador Micropulser (Bio-Rad) para entrada do plasmídeo nas bactérias. Imediatamente após o choque, foram adicionados às bactérias 300 µL de meio Luria-Bertani (LB) (1% (p/v) peptona, 0,5% (p/v) extrato de levedura, 0,5% (p/v) NaCl) sem antibiótico e as mesmas foram incubadas a 37°C e 250 rpm por uma hora. Posteriormente, 100 µL e 200 µL dessas bactérias foram plaqueadas em meio LB-ágar (1% (p/v) peptona, 0,5% (p/v) extrato de levedura, 0,5% (p/v) NaCl, 1,5% (p/v) ágar bacteriológico) contendo ampicilina a uma concentração final de 100 µg/mL. As placas foram incubadas a 37°C por 16 horas.

As colônias isoladas crescidas na placa foram analisadas por PCR de colônia para confirmação da transformação das bactérias com o plasmídeo recombinante. Nessa reação de PCR de 10 µL, o DNA molde foi obtido através da raspagem da colônia com uma ponteira estéril, que em seguida foi levemente inserida por alguns segundos no tubo contendo os reagentes da PCR. Tais reagentes consistiam em 1 µL de tampão 10x (100 mM Tris-HCl pH 8,8, 500 mM KCl, 0,8% (v/v) Nonidet P40), 1 µL de uma mistura dos 4 desoxinucleotídeos (concentração final de 0,2 mM de cada desoxinucleotídeo), 0,6 µL de de MgCl₂ (concentração final de 1,5 mM), 0,05 µL de Taq DNA polimerase (0,25 unidades de enzima) (Fermentas), 1 µL de cada iniciador citado no item 4.1.1 (concentração final de 0,2 µM) e água livre de DNase e RNase para completar o volume.

A reação foi executada em um termociclador (MJ Mini Personal Thermal Cycler – Bio-Rad) utilizando-se os seguintes parâmetros:

- Desnaturação a 95°C por 2 minutos;
- 25 ciclos de desnaturação a 95°C por 30 segundos, anelamento a 55°C por 30 segundos e extensão a 72°C por 1 minuto;
- Extensão final a 72°C por 5 minutos.

Posteriormente, o produto de PCR foi analisado em gel de agarose 1%, como descrito no item 4.1.2.

As colônias que se apresentaram positivas na PCR foram inoculadas em meio LB contendo ampicilina a uma concentração final de 100 µg/mL e incubadas por 16 horas a 37°C e 200 rpm. Em seguida o DNA plasmidial dessas bactérias foi extraído utilizando-se o kit *Axyprep Plasmid Miniprep* (Axygen) seguindo-se as instruções do fabricante. Após a extração, a concentração, em ng/µL, foi estimada em equipamento nanodrop ND1000 (Nanodrop).

4.1.5 – Extração de DNA plasmidial bacteriano

O DNA plasmidial das bactérias foi extraído utilizando-se o kit *Axyprep Plasmid Miniprep* (Axygen). Inicialmente, as bactérias foram inoculadas em meio LB e incubadas a 37°C por 16 horas. Posteriormente, 4 mL dessa cultura foram centrifugados a 12.000 *g* por 1 minuto. O sobrenadante foi descartado e o *pellet* contendo as bactérias foi ressuspensionado em 250 µL do *Buffer S1*. Em seguida, 250 µL do *Buffer S2* foram adicionados e os tubos foram invertidos de 4 a 6 vezes para homogeneização. Imediatamente após a última etapa citada, 350 µL do *Buffer S3* foram adicionados e os tubos invertidos de 6 a 8 vezes. Esse material foi, então, centrifugado a 12.000 *g* por 10 minutos. O sobrenadante foi carregado na coluna fornecida pelo fabricante e essa foi posteriormente centrifugada a 12.000 *g* por 1 minuto seguindo-se com o descarte do *flow through*. Um volume de 500 µL do *Buffer W1* foi adicionado à coluna, que foi novamente centrifugada e o *flow through* descartado. Para remoção do sal, 700 µL do *Buffer W2* foram adicionados e a coluna centrifugada. Essa lavagem com o *Buffer W2* foi repetida e, em seguida, a coluna foi novamente centrifugada a 12.000 *g* por 1 minuto para retirar qualquer resíduo da solução. O DNA foi, então, eluído em 30 µL de água livre de DNase e RNase através de centrifugação a 12.000 *g* por 1 minuto após a incubação da coluna por 1 minuto à temperatura ambiente. Os reagentes *Buffer S1*, *S2*, *S3*, *W1* e *W2* são fornecidos no kit.

4.1.6 – Preparação de bactérias eletrocompetentes

Para preparar bactérias eletrocompetentes, 5 mL de meio LB sem antibióticos foram inoculados com uma colônia bacteriana seguindo-se a incubação por 16 horas a 37°C e 200 rpm. Em seguida, essa cultura foi adicionada a 500 mL de meio LB sem antibiótico e as bactérias foram incubadas a 37°C com agitação de 200 rpm até atingir uma absorbância a 600 nm entre 0,5 e 0,7. As células foram, então, resfriadas em gelo por 20 minutos e, em todos os passos subsequentes, os materiais foram mantidos em gelo. As bactérias foram, então, sedimentadas por centrifugação a 4.000 *g* por 15 minutos a 4°C. O sobrenadante foi cuidadosamente descartado e o *pellet* ressuspendido em 500 mL de glicerol 10% (v/v) gelado. Novamente o material foi centrifugado nas mesmas condições anteriores e o sobrenadante descartado. O *pellet* foi ressuspendido em 250 mL de glicerol 10% gelado e novamente centrifugado e, em seguida, o sobrenadante foi descartado. O *pellet* foi ressuspendido em 20 mL de glicerol 10% gelado, centrifugado e o sobrenadante descartado. Finalmente, as bactérias foram ressuspendidas em aproximadamente 3 mL de glicerol 10% gelado, aliquotadas e imediatamente estocadas a -80°C.

4.1.7 – Subclonagem do gene codificador da proteína Sm21.7 no vetor de expressão pET28a(TEV)

O sistema pET de expressão de proteínas heterólogas em *E. coli* vem sendo amplamente utilizado nos últimos anos. Estes plasmídeos possuem o promotor T7, que é acessado pela T7 RNA polimerase, que irá gerar altas concentrações de RNA mensageiro do gene que estiver sob seu controle. Além disso, alguns plasmídeos pET possuem uma sequência do operador lac depois do promotor T7, o que torna a expressão mais controlada (<http://www.csun.edu>).

O vetor pET28a(+) (Figura 16) possui a sequência codificadora de uma cauda de histidinas N-terminal, o que facilita a purificação da proteína recombinante por cromatografia de afinidade. O vetor também possui um sítio de clivagem para a protease trombina para posterior remoção da cauda de histidinas (<http://www.csun.edu>).

Esse vetor foi modificado no LNLS (Laboratório Nacional de Luz Síncrotron) substituindo-se o sítio de clivagem para trombina pelo sítio de clivagem para a protease TEV (*Tobacco Etch Virus*) (Figura 17) (CARNEIRO *et al.*, 2006). Esse vetor modificado foi utilizado para subclonagem do gene codificador da proteína Sm21.7.

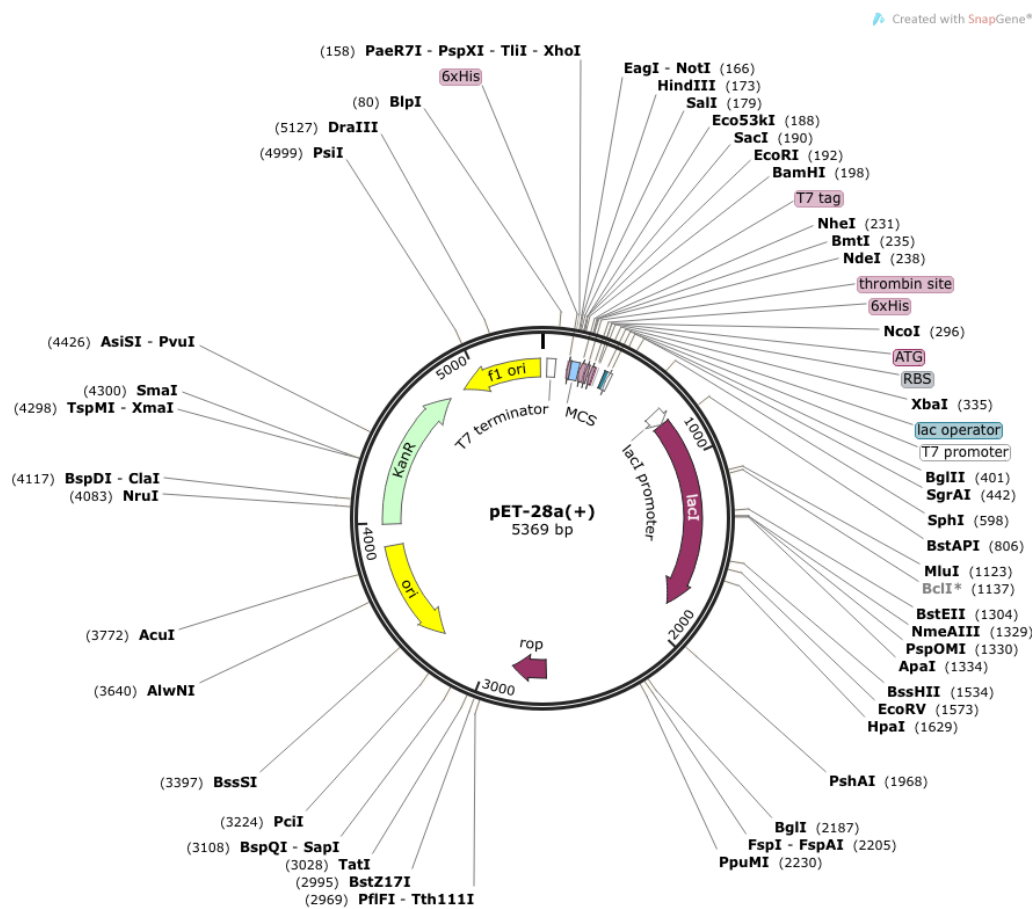


Figura 16: Mapa circular do vetor de expressão pET-28a(+). A imagem demonstra que o vetor possui 5.369 pb, resistência à canamicina e um sítio múltiplo de clonagem. Também apresenta a origem f1, que permite a produção de vetor de fita simples, a origem de replicação em bactérias (ori), o operador lac e o promotor T7.

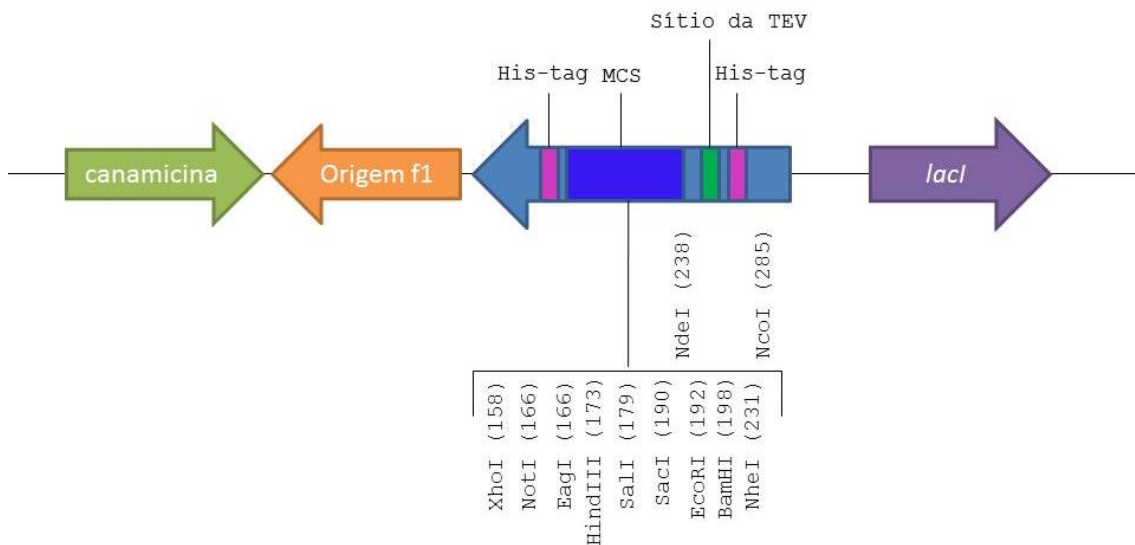


Figura 17: Mapa linear do vetor modificado pET28a(TEV). No vetor modificado, produzido no LNLS, o sítio de clivagem para a trombina foi substituído pelo sítio de clivagem para a protease TEV.

Para subclonagem no vetor de expressão pET28a(TEV) (CARNEIRO *et al.*, 2006), o plasmídeo recombinante Sm21.7-pGEM-T e o vetor pET28a(TEV) foram digeridos com as enzimas de restrição *NdeI* (Thermo Scientific) e *XhoI* (Thermo Scientific). A escolha das enzimas foi feita de forma que o menor número de aminoácidos fosse adicionado à sequência nativa e baseando-se na análise feita pelo software *Nebcutter* (<http://tools.neb.com/NEBcutter2/>) para verificar a presença de sítios de clivagem na própria sequência a ser clonada.

A reação de digestão do vetor Sm21.7-pGEM-T, com um volume final de 40 µL, foi preparada adicionando-se 4 µL de Tampão O 10x (50 mM Tris-HCl pH 7,5, 10 mM MgCl₂, 100 mM NaCl, 0,1 mg/mL BSA (Albumina Sérica Bovina) (Fermentas)), 2 µL da enzima *NdeI* (20 unidades de enzima) e 4 µL da enzima *XhoI* (40 unidades de enzima) a 30 µL do vetor (4,2 µg). Já a digestão do pET28a(TEV) foi preparada da mesma forma, mas com 1 µL de *NdeI* (10 unidades de enzima) e 2 µL de *XhoI* (20 unidades de enzima) adicionados a 30 µL do plasmídeo (2,2 µg). Os volumes finais foram completados com água livre de DNase e RNase. As reações ocorreram pela incubação por 16 horas a 37°C.

Em seguida, os produtos das reações de digestão foram analisados em gel de agarose como citado em 4.1.2 e tanto a banda correspondente ao inserto liberado do plasmídeo recombinante quanto o vetor pET28a(TEV) linearizado foram excisados do gel e submetidas ao procedimento de purificação do DNA utilizando-se o kit *Illustra GFX PCR DNA and Gel Band Purification* (GE Healthcare) (item 4.1.3). O DNA recuperado do gel foi quantificado em um equipamento nanodrop ND1000 (Nanodrop).

A próxima etapa da clonagem consistiu na ligação do inserto liberado pelas enzimas de restrição ao vetor pET28a(TEV) linearizado com as mesmas enzimas. Inicialmente, a concentração de inserto a ser utilizada foi calculada segundo a relação citada em 4.1.4, na qual o tamanho do vetor pET-28a(TEV) foi de 5.357 pb e os outros parâmetros os mesmos.

Esse cálculo resultou em 15,5 ng do inserto (1 µL), que foi utilizado para ligação a 50 ng do vetor (5,7 µL). Essas quantidades foram utilizadas na reação de ligação, à qual também foram adicionados 4 µL do tampão 5x concentrado (250 mM Tris-HCl pH 7,6, 50 mM MgCl₂, 5 mM ATP, 5 mM DTT (Ditiotreitol), 25% (p/v) Polietilenoglicol (PEG) 8.000)), 1 µL da enzima T4 DNA ligase (Invitrogen) (1 unidade de enzima) e água livre de DNase e RNase para completar o volume de 20 µL. A reação, foi, então, incubada a 16°C por 16 horas.

Imediatamente após esse tempo de incubação, o plasmídeo recombinante Sm21.7-pET28a(TEV) foi transformado em bactérias eletrocompetentes *E. coli* BL21(DE3) segundo protocolo descrito em 4.1.4. Após a transformação, 100 µL das bactérias transformadas foram plaqueadas em meio LB-ágar contendo canamicina a uma concentração final de 50 µg/mL. As placas foram incubadas por 16 horas a 37°C e as colônias positivas foram selecionadas por PCR de colônia (item 4.1.4). O DNA plasmidial das colônias foi extraído como descrito em 4.1.5 e sequenciado utilizando-se iniciadores T7 direto e reverso.

A bactéria *E. coli* BL21(DE3) é adequada para a expressão de proteínas a partir de construções que utilizam o promotor T7. A denominação DE3 indica que a bactéria carrega uma cópia cromossômica do gene T7 RNA polimerase sob controle do promotor lacUV5, que pode ser induzido pela adição de IPTG (Isopropil β-D-1-tiogalactopiranosídeo). Essa é a bactéria mais utilizada para expressão de proteínas, pois é deficiente para as proteases Lon e OmpT.

4.2 – Expressão da proteína Sm21.7 recombinante (rSm21.7) em sistema procariótico e lise bacteriana

Uma vez que o gene codificador da proteína Sm21.7 foi clonado em um vetor de expressão e o plasmídeo recombinante transformado em *E. coli* BL21(DE3), uma colônia foi selecionada e utilizada para a expressão da proteína recombinante. Inicialmente, 20 mL de meio LB contendo canamicina a uma concentração final de 50 µg/mL foram inoculados com uma colônia e esse material foi incubado a 37°C e 200 rpm por 16 horas. Em seguida, essa cultura foi utilizada para inocular 1L de meio LB também contendo canamicina e esse inóculo foi mantido a 37°C com agitação até a absorbância a 600 nm atingir um valor entre 0,4 e 0,8. Ao atingir essa fase do crescimento, a expressão da proteína Sm21.7 recombinante foi induzida pela adição de 0,5 mM de IPTG (concentração final) seguida da incubação a 37°C por 4 horas com agitação.

Para verificar a expressão da proteína rSm21.7, foram retiradas alíquotas da cultura de bactérias antes da indução com IPTG e ao fim da expressão. Essas alíquotas foram centrifugadas a 17.000 *g* por um minuto, o sobrenadante foi descartado e o *pellet* foi ressuspensionado em 50 µL de tampão de amostra (125 mM Tris-HCl pH 6,8, 2% (p/v)

SDS (Dodecil Sulfato de sódio), 5% (v/v) β -mercaptoetanol, 0,08% (p/v) azul de bromofenol) para análise em SDS-PAGE (Eletroforese em gel de poliacrilamida contendo SDS) (ver item 4.2.1). Após coloração e descoloração do gel, a expressão da proteína recombinante foi verificada pela observação de uma banda com a massa molecular esperada apenas no extrato celular ao final da expressão. A massa molecular teórica da proteína rSm21.7 foi obtida através de análise da sequência de aminoácidos no programa EXPASY Proteomics Server ("<http://www.expasy.org>"; (ARTIMO *et al.*, 2012).

Ao fim das 4 horas de expressão, a cultura de 1L foi centrifugada a 10.000 *g* por 5 minutos, o sobrenadante foi descartado e o *pellet* foi ressuscitado em 15 mL de tampão de lise (50 mM Tris-HCl pH 7,5, 1% (v/v) glicerol, 1% (v/v) Tween 20, 1% (p/v) sacarose e 1mM PMSF (Fluoreto de Fenilmetilsulfonil)). A lise das bactérias foi realizada através de três ciclos de homogeneização a 15.000 psi em um equipamento Emulsiflex-C3 (Avestin) e, imediatamente após, as frações solúvel e insolúvel foram separadas por centrifugação durante 1 hora a 10.000 *g* e 4°C. A presença da proteína recombinante nas frações solúvel ou insolúvel também foi verificada através de SDS-PAGE como descrito no item 4.2.1.

4.2.1 – Eletroforese em gel de poliacrilamida contendo SDS (SDS-PAGE)

A eletroforese em gel de poliacrilamida foi realizada segundo Laemmli, 1970 (LAEMMLI, 1970) em um sistema Mini-Protean Tetra Cell (Bio-Rad).

Foram preparados géis de separação com 15% de poliacrilamida e de empilhamento com 4%. Primeiramente, para preparar o gel de separação, foram misturados: 2,5 mL de uma solução contendo 30% (p/v) de acrilamida/0,8% (p/v) de bis-acrilamida, 1,25 mL de Tris-HCl 1,5 M pH 8,8, 50 μ L de SDS 10% (p/v), 50 μ L de persulfato de amônio 10% (p/v), 5 μ L de TEMED (N, N, N', N'-tetrametiletilenodiamina) e água para um volume final de 5 mL. Essa mistura foi aplicada em aparato adequado e mantida a temperatura ambiente por alguns minutos até a polimerização.

Posteriormente, o gel de empilhamento foi preparado contendo 130 μ L de 30% de acrilamida/0,8% de bis-acrilamida, 250 μ L Tris-HCl 0,5 M pH 6,8, 10 μ L SDS 10%, 10 μ L de persulfato de amônio 10%, 1 μ L de TEMED e água para um volume de 1 mL.

Essa mistura foi aplicada sobre o gel de separação até a polimerização na presença dos espaçadores que formam as canaletas.

As corridas dos géis foram realizadas em tampão constituído de glicina 192 mM, SDS 0,1% (p/v) e Tris 25 mM pH 8,3 em uma voltagem constante de 200 V, após aquecimento das amostras por 5 minutos a 100°C.

Para visualização das bandas, o gel foi incubado por 2 a 16 horas em solução de 15% (v/v) etanol, 7% (v/v) ácido acético e 0,1% (p/v) *Coomassie Brilliant Blue R250*. A seguir, uma solução de 10% (v/v) etanol e 5% (v/v) ácido acético foi utilizada para a remoção do excesso de corante, deixando visíveis as bandas de proteínas.

4.3 – Purificação da proteína rSm21.7 por cromatografias de afinidade e exclusão molecular

A clonagem no vetor pET28a(TEV) incorpora, anteriormente à sequência codificadora da proteína de interesse, uma cauda de histidinas N-terminal seguida por um sítio de clivagem para a protease TEV (CARNEIRO *et al.*, 2006). Essa construção permite a purificação da proteína recombinante por cromatografia de afinidade em coluna de níquel.

Sendo assim, após a lise celular e centrifugação do lisado, o sobrenadante, que constitui a fração solúvel, acrescido de 30 mM de imidazol, foi aplicado em uma coluna contendo níquel (His-Trap HP – 5 mL – GE Healthcare) acoplada a um sistema AKTA prime (GE Healthcare). Essa coluna foi previamente equilibrada com 30 mL (6 volumes de coluna) de tampão de ligação (20 mM Fosfato de sódio pH 7,5, 500 mM NaCl, 30 mM Imidazol, 1 mM PMSF). Após a ligação das proteínas com afinidade ao níquel, a coluna foi lavada com 20 mL do tampão de ligação para remoção de moléculas inespecíficas. Finalmente, as proteínas ligadas ao níquel foram eluídas pela passagem de um gradiente de 20 mL de tampão de eluição (20 mM Fosfato de sódio pH 7,5, 500 mM NaCl, 500 mM Imidazol, 1mM PMSF). Para essa proposta, o tampão de eluição foi automaticamente misturado ao tampão de ligação dentro do sistema AKTA prime (GE Healthcare) durante a corrida cromatográfica.

Para avaliação da pureza e integridade das amostras eluídas, as mesmas foram aplicadas em um gel de poliacrilamida 15%, que foi posteriormente corado com *Coomassie Brilliant Blue R250* (item 4.2.1).

A amostra da proteína rSm21.7 eluída da cromatografia de afinidade foi concentrada para um volume de 5 mL (Vivaspin 20 MWCO 10 kDa - GE Healthcare) para realização de uma cromatografia de exclusão molecular, que foi a última etapa de purificação. Essa etapa foi realizada em uma coluna HiLoad 16/60 Superdex 200 prep grade (GE Healthcare) acoplada a um sistema AKTA prime (GE Healthcare). Os 5 mL da amostra foram aplicados na coluna previamente equilibrada com 2 volumes de coluna do tampão 50 mM Fosfato de Sódio pH 7,5, 50 mM NaCl, 1 mM PMSF e as proteínas foram eluídas no mesmo tampão. As frações eluídas nessa etapa de purificação também foram analisadas em SDS-PAGE.

Após os dois passos de purificação, a proteína também foi analisada por espectrometria de massas utilizando-se a matriz Ácido sinapínico no equipamento MALD-TOF Autoflex III (Bruker Daltonics).

4.4 – Análises biofísicas e bioquímicas da proteína rSm21.7 purificada

4.4.1 – Análise da estrutura secundária da proteína rSm21.7 por Dicroísmo Circular (CD)

O fenômeno do CD consiste na absorção diferencial da luz circularmente polarizada por uma molécula assimétrica. O espectro de moléculas em α -hélice, caracteristicamente, possui dois picos negativos a 222 e 208 nm. Já para estruturas em folha- β , o espectro consiste em um pico negativo próximo a 215 nm e um positivo próximo a 198 nm (SREERAMA; WOODY, 2004).

Para análise da estrutura secundária através de CD, a proteína rSm21.7 purificada e concentrada para 17 mg/mL foi diluída 50 vezes em água milliQ, ficando a uma concentração de aproximadamente 0,34 mg/mL. O tampão em que a proteína se encontrava (50 mM Fosfato de sódio pH 7,5, 50 mM NaCl, 1 mM PMSF) também foi diluído 50 vezes, sendo utilizado como referência.

Antes das medidas, a amostra foi centrifugada a 17.000 g por 10 minutos para sedimentação de qualquer material em suspensão. A análise foi realizada em uma cubeta de quartzo de 0,5 mm de caminho óptico em um espectropolarímetro Jasco J-815 (Jasco) em combinação com um controlador de temperatura PFD 425S/15 (Jasco). A temperatura foi variada entre 20 e 90°C e as medidas foram feitas de 10 em 10°C.

Foram utilizados absorvância entre 190 e 260 nm, *datapich* de 1 nm e cada leitura consistiu de 5 acumulações.

Os dados de CD são inicialmente obtidos em termos de elipcidade, medida em milidegrees (mdeg). Entretanto, para proteínas, os dados são geralmente expressos em elipcidade molar média, que pode ser determinada pelo cálculo:

$$[\theta] \text{ MRW} = \theta / (10 \times C_r \times I)$$

No qual:

θ = elipcidade obtida pelo espectropolarímetro em mdeg;

$C_r = (1.000 \times n \times cg) / M_r$

n = número de resíduos da proteína

cg = concentração da proteína em g/mL

M_r = massa molecular da proteína

I = caminho óptico

Após a obtenção dos dados em elipcidade molar média, os mesmos foram plotados em um gráfico de comprimento de onda por elipcidade molar média em cada uma das temperaturas analisadas para verificar o perfil e a presença dos picos característicos das estruturas secundárias.

A partir dos dados de elipcidade, pode-se estimar a porcentagem de estrutura secundária constituinte da proteína. Essa estimativa foi feita através do programa *CDNN* (BÖHM; MUHR; JAENICKE, 1992) para cada uma das temperaturas analisadas, o que permitiu avaliar a desnaturação térmica da proteína.

4.4.2 – Avaliação do estado oligomérico da proteína rSm21.7 através de Espalhamento Dinâmico da Luz (DLS)

O DLS é uma técnica que permite estimar o raio hidrodinâmico de uma partícula em solução (incluindo proteínas), além de fornecer informações sobre a polidispersividade da amostra. Nessa técnica são medidas as flutuações na intensidade da luz espalhada que se deve ao movimento Browniano das partículas (JACHIMSKA; WASILEWSKA; ADAMCZYK, 2008).

O ensaio de DLS foi realizado para determinação do estado oligomérico da proteína rSm21.7 em solução utilizando-se um equipamento Zetasizer nano ZS90 (Malvern). Antes do ensaio, a amostra foi centrifugada por 10 minutos a 17.000 *g* para retirada de qualquer material suspenso. O DLS foi realizado a 25°C em uma cubeta de vidro contendo um volume de 1,5 mL da proteína a uma concentração de 3,5 mg/mL. Após a leitura, os resultados foram analisados no programa Malvern Zetasizer.

Outra análise relacionada realizada foi o DLS ao longo do tempo, no qual a cubeta contendo a amostra foi mantida no equipamento durante 250 leituras de 12 acumulações cada com intervalos de 3 minutos entre as leituras. Posteriormente, em dias alternados durante 30 dias, foram feitas avaliações através de 10 leituras de 12 acumulações cada com intervalos de 3 minutos entre as leituras. Cada um desses conjuntos de leituras foi feito em dias determinados para acompanhar a agregação da amostra. As amostras analisadas foram mantidas a 4°C nos períodos entre as análises.

4.5 – Ensaio de *cross-linking* da proteína rSm21.7

O estado oligomérico da proteína rSm21.7 também foi avaliado por *cross-linking* com glutaraldeído. Esse ensaio foi realizado segundo Santos e colaboradores (SANTOS *et al.*, 2012) com pequenas modificações.

Basicamente, amostras de proteína purificada foram diluídas para uma concentração de 5 µM em tampão 20 mM Hepes pH 7,5 em um volume total de 100 µL. Posteriormente foram adicionadas as concentrações de 0,025%, 0,05%, 0,1%, 0,2% e 0,4% de glutaraldeído e as amostras foram incubadas em gelo por 5 ou 30 minutos. A reação foi finalizada pela adição de 10 µL de 1 M Tris-HCl pH 8,0 e as amostras analisadas por SDS-PAGE como descrito em 4.2.1.

4.6 – Proteólise limitada da proteína rSm21.7

Já foi demonstrado que a técnica de proteólise limitada pode fornecer importantes informações sobre a estrutura das proteínas, assim complementando os resultados e informações que podem ser obtidos utilizando-se técnicas biofísicas e espectroscópicas (FONTANA *et al.*, 2004).

A proteólise limitada deriva-se do fato de que um segmento específico flexível de uma proteína enovelada pode estar suficientemente exposto para ligar-se ao sítio ativo da protease. Assim, a proteólise limitada ocorre preferencialmente em *loops* que apresentam certa flexibilidade conformacional, enquanto a parte da proteína com estrutura compacta permanece resistente à ação das proteases (FONTANA *et al.*, 2004).

No ensaio de proteólise limitada com rSm21.7 foram testadas as proteases pepsina, tripsina, proteinase K, quimotripsina e subtilisina. Foram incubados 50 µg da proteína rSm21.7 purificada com 1 µg de cada uma das proteases em um volume final de 17 µL que foi completado com água. Para cada protease foram avaliados os tempos 0, 5 minutos, 10 minutos, 30 minutos, 1 hora, 2 horas, 3 horas e 4 horas a uma temperatura de 25°C. Ao final de cada um dos tempos, a cada uma das reações foi adicionado 3 µL de TFA (Ácido Trifluoroacético) 3% para cessar a ação da protease.

As amostras submetidas à proteólise limitada foram, então, analisadas por SDS-PAGE (item 4.2.1) no qual 10 µL de cada uma das amostras testadas foram aplicados no gel. O restante da amostra (10 µL) foi separado em um gel e transferido para uma membrana de nitrocelulose (Bio-Rad) para realização de um *Western Blot* utilizando-se anticorpo anti-histidinas.

4.7 – Ensaio de *Western Blot* com anticorpo anti-histidinas

Para realização de ensaios de *Western Blot*, as amostras foram separadas em gel de poliacrilamida 15% e, posteriormente, transferidas para uma membrana de nitrocelulose, como citado anteriormente. A transferência ocorreu durante 1 hora a 100 V, em tampão 25 mM Tris, 187 mM glicina e 20% (v/v) metanol. Em seguida, a membrana foi corada com solução PonceauS (0,1% (p/v) em 1% (v/v) ácido acético) para verificar a transferência e, imediatamente após, foi lavada várias vezes com água para retirada da coloração. A membrana foi, então, incubada por 16 horas em solução de bloqueio (PBS (Tampão Fosfato Salino)-Caseína 3% (p/v)) e posteriormente lavada por 3 vezes (20 minutos cada) com solução de lavagem (PBS-Tween-20 0.05%). Em seguida foi adicionado à membrana o anticorpo primário anti-histidinas (anti-his Santa Cruz) diluído 1:100 em solução de incubação (PBS-Caseína 0,25% (p/v)). Esta primeira etapa de incubação ocorreu durante 1 hora. Após 3 lavagens de 10 minutos, houve a incubação com o anticorpo secundário de coelho conjugado com peroxidase diluído

1:5.000 em solução de incubação. Então, a membrana foi lavada 3 vezes (5 minutos cada) com solução de lavagem e 2 vezes (5 minutos cada) com PBS 1x. Para revelação, 10 mg de DAB (diamino benzidina) foram dissolvidos em 10 mL de PBS (Solução A) e 5 mg de cloronaftol foram diluídos em 1,66 mL de metanol e 8,32 mL de PBS (Solução B). Essas duas soluções foram misturadas e a elas foram acrescentados 8,32 µL de peróxido de hidrogênio. A membrana foi incubada nessa solução até o aparecimento das bandas e a reação de revelação foi cessada pela adição de PBS 1x à membrana.

4.8 – Determinação da estrutura da proteína rSm21.7 por cristalografia de raios X

4.8.1 – Testes de cristalização da proteína rSm21.7

Após a cromatografia de exclusão molecular, a proteína rSm21.7 foi submetida a testes de cristalização pelas técnicas da gota suspensa e da gota depositada (Figura 18), sendo essa última realizada em um equipamento HoneyBee 936 (DigiLab Global) no LNLS. Em ambas as técnicas, a cristalização ocorre por difusão de vapor, na qual um poço contendo a solução de cristalização e uma gota formada pela mesma solução e pela proteína encontram-se em um sistema vedado. Ocorre, então, através da difusão de vapor, uma redução do volume da gota que perde água, possibilitando a supersaturação da concentração da proteína e consequente formação dos cristais (LI; CHANG, 2009).

A diferença entre as duas técnicas citadas acima está no fato de que, na gota suspensa, a gota é depositada de cabeça para baixo em uma lâmina, como mostrado na figura 18A. Já na gota depositada a gota é depositada sobre uma cavidade (figura 18B). Essa última é bastante utilizada em condições de baixa tensão superficial da gota.

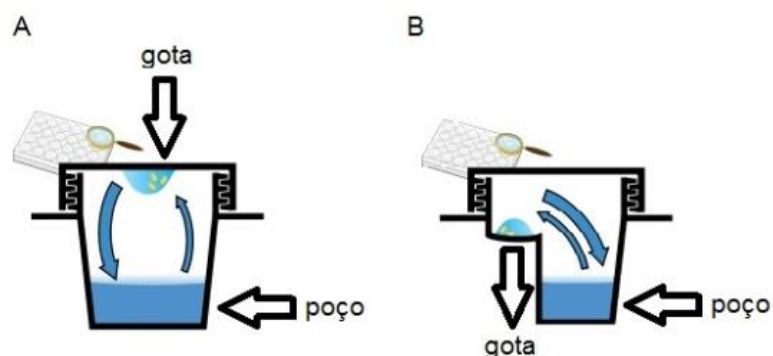


Figura 18: Representação esquemática das técnicas de gota suspensa e gota depositada para Cristalização de Macromoléculas. Nos dois métodos uma gota contendo a solução de cristalização e a proteína encontra-se em um sistema vedado que contém a solução de cristalização. (A) Gota suspensa: a gota fica de cabeça para baixo em uma lamínula. (B) Gota depositada: a gota é colocada sobre uma pequena cavidade. Em ambos os métodos a cristalização acontece por difusão de vapor.

Em um *screening* inicial, foram avaliadas 96 condições de cristalização constituintes dos kits comerciais *Crystal Screen* e *Crystal Screen 2* (Hampton Research) em placas de 24 poços (Hampton Research) através da técnica da gota suspensa. Essas soluções foram preparadas no laboratório de acordo com as formulações do kit. Cada condição de cristalização do kit é constituída de um tampão em pH específico, um sal e um precipitante, mas em algumas soluções o tampão e/ou sal estão ausentes.

Nesses testes, 1 mL da solução de cristalização foi adicionado a cada poço e nas lamínulas foram adicionadas gotas com 1 μL da proteína e 1 μL da solução do poço. Já nos testes de gota depositada, foram utilizados 1 μL de proteína e 1 μL da solução de poço em uma placa de 96 poços. Cada poço foi preenchido com 80 μL da solução de cristalização. Foram avaliadas quase 600 condições por essa técnica. A concentração da proteína utilizada foi aproximadamente 20 mg/mL, tanto nos testes da gota suspensa quanto nos testes da gota depositada.

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

Além do pH e concentrações de sal e precipitante, outra variação foi inserida nos testes de otimização. Foram testados os seguintes volumes e proporções de proteína

para solução de cristalização: 2 µL proteína:2 µL solução de cristalização, 2 µL proteína:1 µL solução de cristalização e 1 µL proteína:2 µL solução de cristalização.

Os testes de cristalização com as condições otimizadas também foram realizados pela técnica da gota suspensa. Todas as placas de cristalização foram incubadas a 18°C e a presença de cristais foi observada através de estereoscópio com aumento de até 100 vezes.

Alguns testes de cristalização foram realizados no NSLS (National Synchrotron Light Source) – Brookhaven National Laboratory.

4.8.2 – Coleta de dados de difração de raios X

As coletas de dados de difração de raios X foram realizadas na linha W01B-MX2 (GUIMARÃES *et al.*, 2009) no LNLS e na linha X25 no NSLS. Foi coletado um conjunto de dados de um cristal nativo e um conjunto de dados de um cristal derivado de iodo no LNLS. Conjuntos de dados nativos também foram coletados no NSLS.

Para realização da coleta de dados do cristal nativo no LNLS, o mesmo foi transferido por alguns segundos com o auxílio de um *loop* de *nylon* para uma solução crioprotetora constituída por 10% etilenoglicol em solução de cristalização. O cristal foi, então, congelado em um fluxo de nitrogênio a 100K.

Durante a coleta de dados do cristal nativo, o mesmo foi exposto por 16 segundos para coleta de cada imagem de difração. Ao todo foram coletadas 260 imagens com 0,5 graus por imagem, totalizando 130 graus de rotação.

Para obtenção do cristal derivado, foi utilizada a estratégia *quick cryosoaking* (DAUTER; DAUTER; RAJASHANKAR, 2000; NAGEM; DAUTER; POLIKARPOV, 2001; NAGEM; POLIKARPOV; DAUTER, 2003), na qual um cristal nativo foi transferido por 30 segundos para a mesma solução crioprotetora citada anteriormente acrescida de iodeto de sódio para uma concentração final de 0,4M. O cristal foi, então, congelado em nitrogênio e exposto por 6 segundos para coleta de cada uma das imagens. Ao todo foram coletadas 300 imagens com 0,5 graus por imagem, totalizando 150 graus de rotação. Em ambos os casos as imagens foram gravadas em um detector Rayonix MarMosaic 225 CCD.

O cristal coletado no NSLS foi exposto por 0,5 segundos para coleta de cada uma das 720 imagens de difração. Nesse caso, as imagens foram coletadas a cada 0,5 graus de rotação em um detector Pilatus 6M.

4.8.3 – Processamento dos dados de difração de raios X

O processamento de dados na Cristalografia de Macromoléculas é a conversão de um conjunto de dados de difração em uma lista de reflexões de Bragg com intensidades medidas. Na cristalografia moderna, os dados de difração consistem de um conjunto de imagens coletadas em um detector 2D, onde cada imagem é coletada em uma orientação específica do cristal.

Nessa etapa, os cristais são caracterizados, ou seja, seus parâmetros de rede e suas simetrias são determinados. O processamento dos dados foi realizado utilizando-se o programa HKL-2000 (OTWINOWSKI; MINOR, 1997).

4.8.4 – Determinação das fases e refinamento da estrutura tridimensional da proteína rSm21.7

O problema das fases foi resolvido utilizando o método Substituição Isomórfica Simples com Espalhamento Anômalo (SIRAS – *Single Isomorphous Replacement with Anomalous Scattering*) no software PHENIX (ADAMS *et al.*, 2010). A localização dos átomos pesados foi obtida através da contribuição do espalhamento anômalo presente no conjunto de dados derivado, mas as contribuições isomórficas e anômalas foram usadas para determinação das fases. As fases derivadas foram usadas para gerar o mapa de densidade eletrônica, permitindo a construção de um modelo preliminar da proteína.

A partir do modelo inicial gerado, o refinamento da estrutura foi realizado através de métodos automáticos pelo programa PHENIX (ADAMS *et al.*, 2010) e manuais pelo programa Coot (EMSLEY *et al.*, 2010).

Os resultados do refinamento foram acompanhados pela redução do fator R_{work} , definido pela equação:

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

na qual:

$|F_{obs}|$: amplitude medida do fator de estrutura

$|F_{calc}|$: amplitude calculada do fator de estrutura

Ao término dos refinamentos, a estrutura foi validada pelo programa ADIT, disponível em <http://validate.rcsb.org/>, sendo esse, o mesmo programa utilizado antes do depósito de uma estrutura no PDB.

Até o momento, o refinamento da estrutura cristalográfica obtida a partir dos dados coletados no NSLS não foi concluído e apenas um modelo preliminar será apresentado.

5 – Resultados e discussão

5.1 – Subclonagem do gene codificador da proteína Sm21.7 no vetor de expressão pET28a(TEV)

O gene codificador da proteína Sm21.7 foi amplificado por PCR utilizando-se como DNA molde um plasmídeo recombinante previamente construído pETDEST42-rP22 (REZENDE *et al.*, 2011). A amplificação desse gene foi verificada através de eletroforese em gel de agarose, no qual foi observado um fragmento de aproximadamente 500 pb, condizente com o peso molecular do gene (555 pb) (Figura 19).

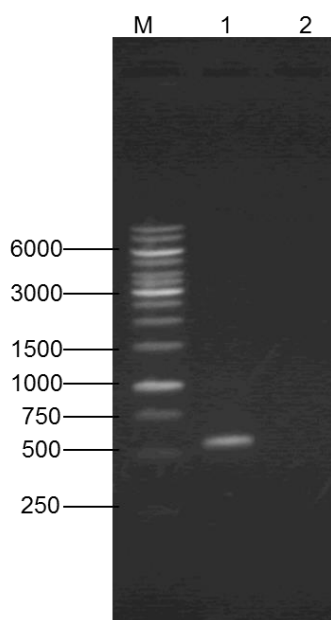


Figura 19: Amplificação do gene codificador da proteína Sm21.7. O gene codificador da proteína de interesse foi amplificado por PCR, que foi posteriormente analisada por eletroforese em gel de agarose corado com SYBR® safe (Invitrogen), no qual pode-se visualizar: M – marcador de peso molecular em pares de bases (*O'GeneRuler™ 1 kb DNA Ladder* – Fermentas), 1 – produto de PCR referente à amplificação do gene de interesse (Sm21.7), 2 – controle negativo da reação.

Após a purificação desse produto de PCR, o mesmo foi ligado ao vetor pGEM®-T e o plasmídeo recombinante resultante foi transformado em bactérias *E. coli* DH5α. As colônias positivas foram selecionadas por PCR de colônia utilizando-se os iniciadores direto e reverso citados em 4.1.1. A análise em gel de agarose desses

produtos de PCR confirmou que as três colônias testadas possuíam o gene de 555 pb ligado ao pGEM®-T (Figura 20) .

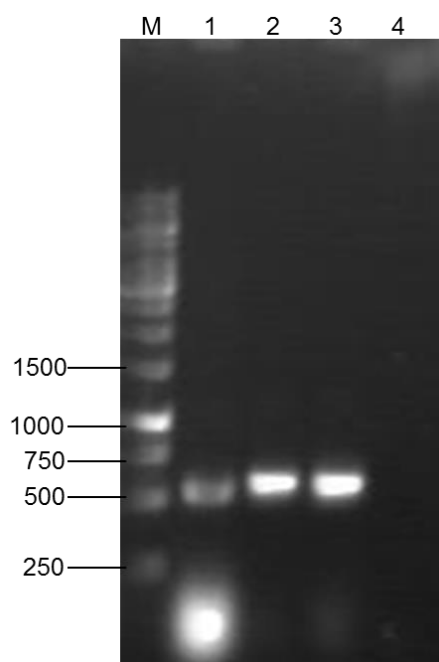


Figura 20: PCR de colônia das bactérias transformadas com Sm21.7-pGEM-T. Os produtos da PCR de colônia resultante da clonagem no vetor pGEM®-T foram analisados por eletroforese em gel de agarose corado com *SYBR® safe* (Invitrogen), no qual pode-se visualizar: M – marcador de massa molecular em pares de bases (*O'GeneRuler™ 1 kb DNA Ladder* – Fermentas), 1 a 3 – colônias 1 a 3, 4 – controle negativo da reação.

Com o objetivo de realizar a subclonagem no vetor de expressão pET28a(TEV), o plasmídeo Sm21.7-pGEM-T e o vetor pET28a(TEV) foram digeridos para liberação do inserto e linearização do vetor, respectivamente. Após a ligação do inserto ao vetor pET28a(TEV), o plasmídeo recombinante foi transformado em bactérias *E. coli* BL21(DE3) e as colônias positivas selecionadas por PCR de colônia. Apenas duas das colônias testadas, canaletas 3 e 7 da figura 21, não possuíam o plasmídeo Sm21.7-pET28a(TEV).

Após a subclonagem, algumas colônias foram selecionadas, seu DNA plasmidial extraído e sequenciado.

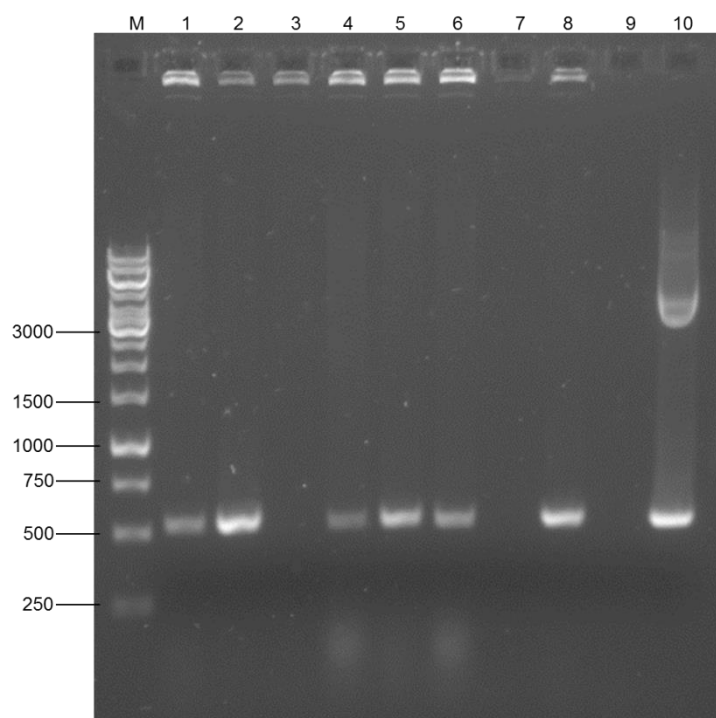


Figura 21: PCR de colônia das bactérias transformadas com Sm21.7-pET28a(TEV). Os produtos da PCR de colônia resultante da subclonagem no vetor pET28a(TEV) foram analisados por eletroforese em gel de agarose corado com SYBR®safe (Invitrogen), no qual pode-se visualizar: M – marcador de massa molecular em pares de bases (*O'GeneRuler™ 1 kb DNA Ladder* – Fermentas), 1 a 8 – colônias 1 a 8, 9 – controle negativo da reação, 10 – controle positivo da reação.

5.2 – Expressão, lise e purificação da proteína rSm21.7

De acordo com análises *in silico* da sequência de aminoácidos da proteína Sm21.7 (EXPASY Proteomics Server – <http://www.expasy.org/>), a massa molecular teórica da mesma é de 21.688,7 Da (184 resíduos de aminoácidos). Incluindo a cauda de histidinas e o sítio de clivagem da protease TEV, utilizados na construção escolhida, a massa molecular é de 23.688,8 Da (200 resíduos de aminoácidos – rSm21.7). Tendo em vista essa informação, a expressão da proteína rSm21.7 em bactérias *E. coli* BL21(DE3) foi analisada por SDS-PAGE, no qual foi visualizada uma banda de aproximadamente 25 kDa no extrato bacteriano obtido após 4 horas de indução (Figura 22). A ausência dessa banda no extrato celular antes da indução indica que uma proteína com a massa molecular esperada (incluindo a cauda de histidinas e o sítio de clivagem da protease TEV) foi expressa. As frações solúvel e insolúvel também foram analisadas por SDS-

PAGE, o qual demonstrou que a proteína rSm21.7 foi expressa na forma solúvel (Figura 22).

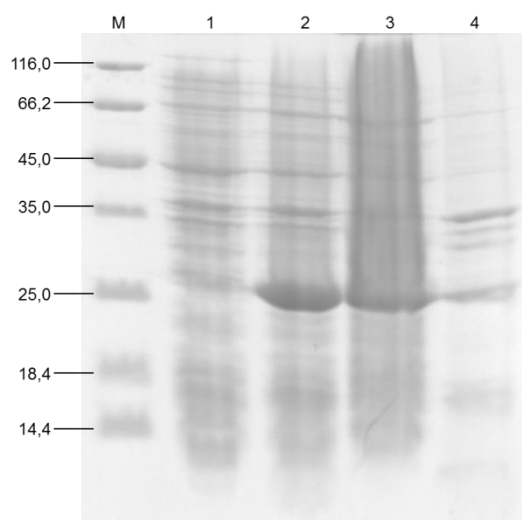


Figura 22: Expressão da proteína rSm21.7 em bactérias *E. coli* BL21(DE3). Após a expressão da proteína recombinante a 37°C e lise bacteriana, os extratos celulares foram analisados por SDS-PAGE. Esta imagem representa um SDS-PAGE corado com *Coomassie Brilliant Blue R250* dos extratos celulares obtidos durante a expressão da proteína rSm21.7 e após a lise. Pode-se visualizar no mesmo: M – marcador de massa molecular em kDa (*Unstained Protein Molecular Weight Marker* – Fermentas), 1 – extrato celular antes da indução com IPTG, 2 – extrato celular após 4 horas de indução com IPTG, 3 – fração solúvel, 4 – fração insolúvel.

Expressões em larga escala foram feitas com o intuito de purificar a proteína recombinante em grande quantidade. Considerando que rSm21.7, encontrava-se, pelo menos em parte, na fração solúvel e foi expressa com uma cauda de histidinas N-terminal, o primeiro passo de purificação foi uma cromatografia de afinidade em coluna de níquel. A figura 23A apresenta o cromatograma referente à purificação da proteína rSm21.7 por afinidade. Pode-se observar um pico de eluição da proteína recombinante com aproximadamente 80% do tampão de eluição (linha pontilhada no eixo secundário), o que representa 400 mM de imidazol.

Após a cromatografia de afinidade, as amostras eluídas foram analisadas por SDS-PAGE (figura 23B) e verificou-se a necessidade de um segundo passo de purificação e/ou polimento da amostra.

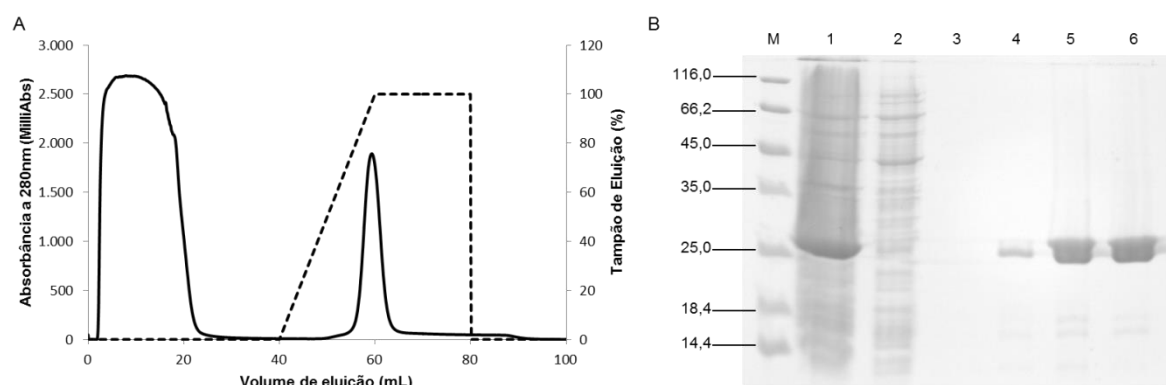


Figura 23: Purificação da proteína rSm21.7 por cromatografia de afinidade em coluna de níquel. Após a lise celular, a proteína recombinante foi purificada por cromatografia de afinidade em coluna de níquel e a pureza das amostras avaliada por SDS-PAGE. (A) – Cromatograma obtido na cromatografia de afinidade no qual a linha pontilhada representa a aplicação do tampão de eluição. (B) – SDS-PAGE corado com *Coomassie Brilliant Blue R250*, no qual pode-se visualizar: M – marcador de massa molecular em kDa (*Unstained Protein Molecular Weight Marker* – Fermentas), 1 – fração solúvel, 2 – frações referentes ao *flow through*, 3 – frações referentes à lavagem, 4 a 7 – frações correspondentes ao pico de eluição da proteína rSm21.7.

A segunda etapa de purificação consistiu em uma cromatografia de exclusão molecular, que separa proteínas de acordo com a massa das mesmas. A amostra proveniente da cromatografia de afinidade foi aplicada em uma coluna de exclusão molecular, da qual a proteína foi eluída como um pico principal em aproximadamente 94 mL (Figura 24A).

De acordo com a calibração da coluna de exclusão molecular utilizando-se proteínas com massa molecular determinada (Figura 24C), estimou-se que esse volume de eluição corresponde ao de uma partícula com massa de aproximadamente 31 kDa. Este valor não define precisamente o estado oligomérico que rSm21.7 teria em solução, mas sugere que a proteína poderia se encontrar tanto na forma monomérica quanto dimérica. Entretanto, esse tipo de análise não é a mais adequada e precisa para a determinação do estado oligomérico da proteína. Primeiramente, foi utilizada uma coluna cromatográfica preparativa, ao passo que a utilização de uma coluna analítica seria mais precisa na determinação da massa molecular. Em segundo lugar, a coluna utilizada abrange uma faixa de separação de massa molecular entre 10 e 600 kDa, o que poderia interferir no resultado, já que é uma faixa bastante ampla para esse tipo de análise e que a massa da proteína de interesse encontra-se quase no limite inferior de massa molecular. Adicionalmente, as concentrações de proteína utilizadas poderiam propiciar uma agregação da amostra. Finalmente, a calibração da coluna é feita com

proteínas globulares e o cálculo para proteínas não globulares pode sofrer alguns efeitos. Devido aos fatos citados, a cromatografia de exclusão molecular não se mostrou a forma mais adequada de avaliar a oligomerização dessa proteína e esse parâmetro será definido posteriormente por outras técnicas.

As frações obtidas nessa etapa de purificação foram analisadas em SDS-PAGE (Figura 24B) e apresentaram-se suficientemente puras para a realização dos testes posteriores.

Após a cromatografia de exclusão molecular, a proteína purificada foi analisada por espectrometria de massas e a massa derivada dessa análise foi de 23.564,4 Da (Figura 25). Uma explicação para a diferença entre a massa teórica (23.688,8 Da) e a massa obtida neste experimento é a ausência de uma calibração precisa no equipamento já que a mesma foi realizada para a análise de peptídeos até 8.000 Da e não de proteínas.

Vale destacar ainda que massas moleculares inferiores a 23 kDa também foram detectadas neste experimento o que poderia sugerir uma degradação protéica ou ainda contaminantes de massa molecular bem próximas à da proteína alvo. Com relação à presença de massas inferiores a 23 kDa, pode-se observar nos géis das amostras provenientes das purificações por afinidade e exclusão molecular, que a proteína apresenta o que pode ser um início de degradação ou a presença de formas truncadas da proteína, já que na região de 25 kDa (local esperado para a proteína de interesse) duas bandas mais proeminentes são vistas bem próximas uma da outra.

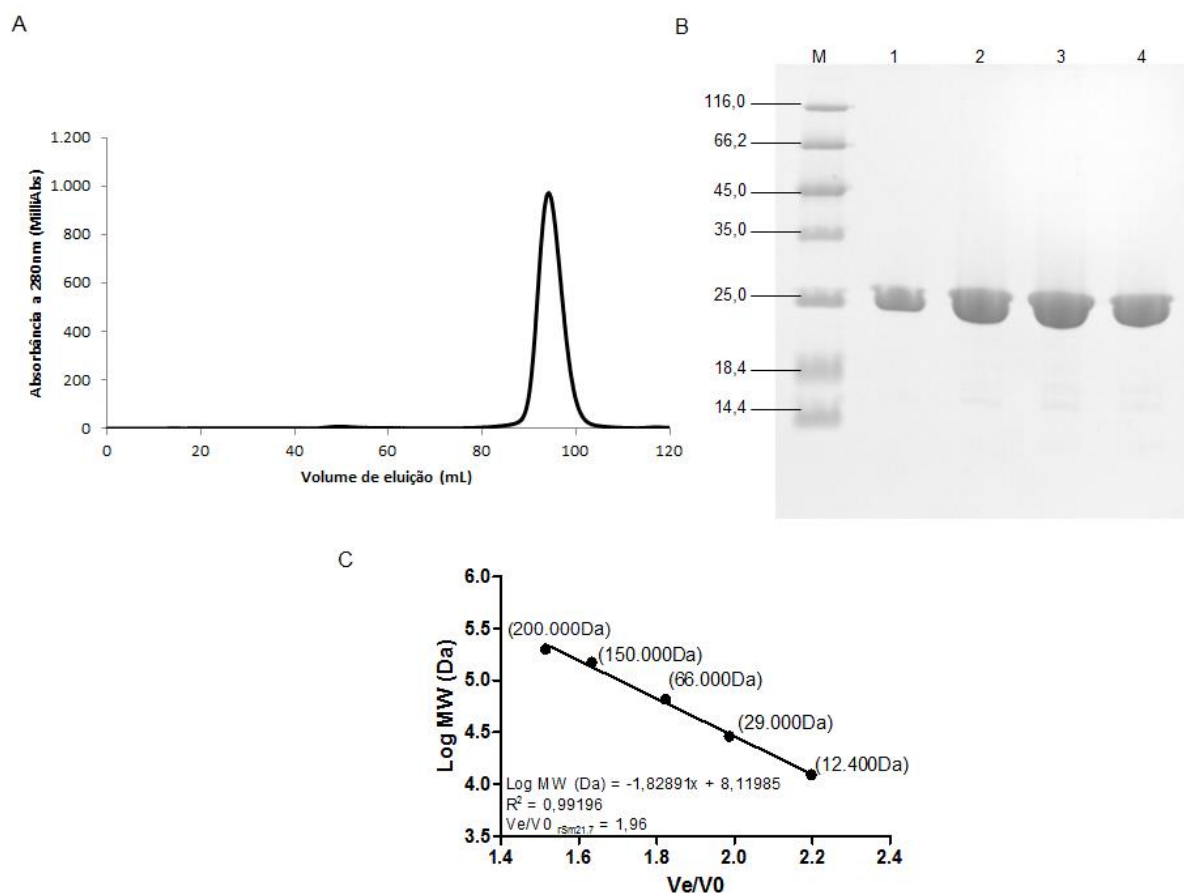


Figura 24: Purificação da proteína rSm21.7 por cromatografia de exclusão molecular. Após a cromatografia de afinidade, a proteína rSm21.7 foi submetida a uma cromatografia de exclusão molecular em coluna HiLoad 16/60 Superdex 200 prep grade (A) – Cromatograma obtido na cromatografia de exclusão molecular. (B) – SDS-PAGE corado com *Coomassie Brilliant Blue R250* no qual pode-se visualizar: M – marcador de massa molecular em kDa (*Unstained Protein Molecular Weight Marker* – Fermentas), 1 a 4 – frações eluídas no volume de corrida de aproximadamente 94 mL. (C) – Calibração da coluna HiLoad 16/60 Superdex 200 prep grade com as proteínas: citocromo C (12,4 kDa), anidrase carbônica (29 kDa), albumina bovina (66 kDa), álcool desidrogenase (150 kDa), amilase (200 kDa) (*Sigma Gel Filtration Molecular Weight Markers* – MWF200).

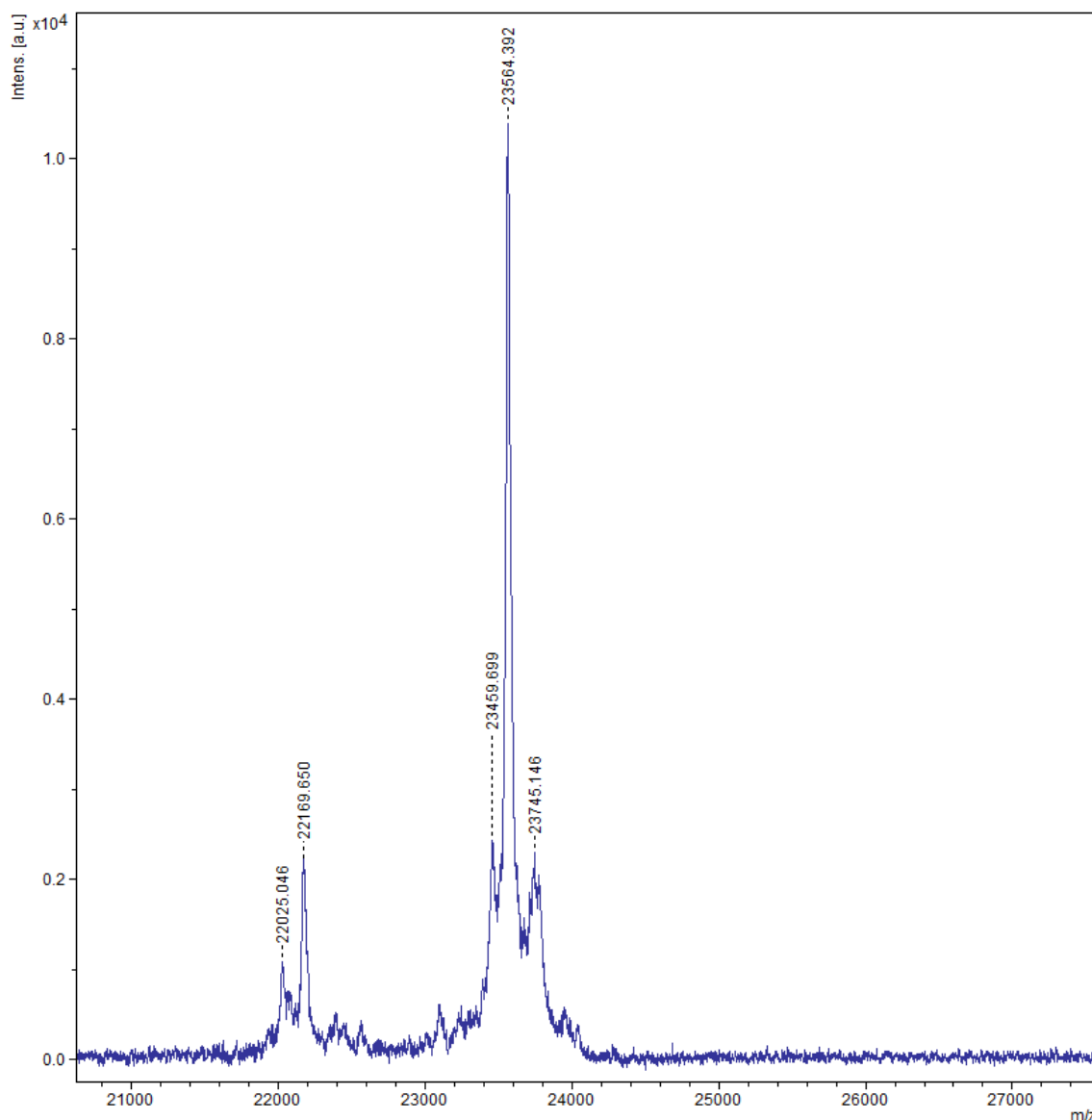


Figura 25: Espectrometria de massas da proteína rSm21.7 purificada. Após as cromatografias de afinidade e exclusão molecular, a proteína rSm21.7 foi analisada por espectrometria de massas.

5.3 – Análise da estrutura secundária da proteína rSm21.7

Com o objetivo de avaliar o enovelamento da proteína rSm21.7 foi realizado um ensaio de CD utilizando a amostra obtida logo após a última corrida cromatográfica (gel filtração). A análise do espectro de CD demonstrou que a proteína rSm21.7 possui, em tampão 50 mM Fosfato de sódio pH 7,5, 50mM NaCl e a uma temperatura de 20 °C, um espectro com dois picos negativos em aproximadamente 222 e 208 nm, que são característicos de estruturas em α -hélice e um pico positivo a 198 nm, demonstrando

também a existência de fitas- β na proteína (figura 26). Esses resultados podem ser observados através de um gráfico de elipcidade molar média plotado em função do comprimento de onda (figura 26).

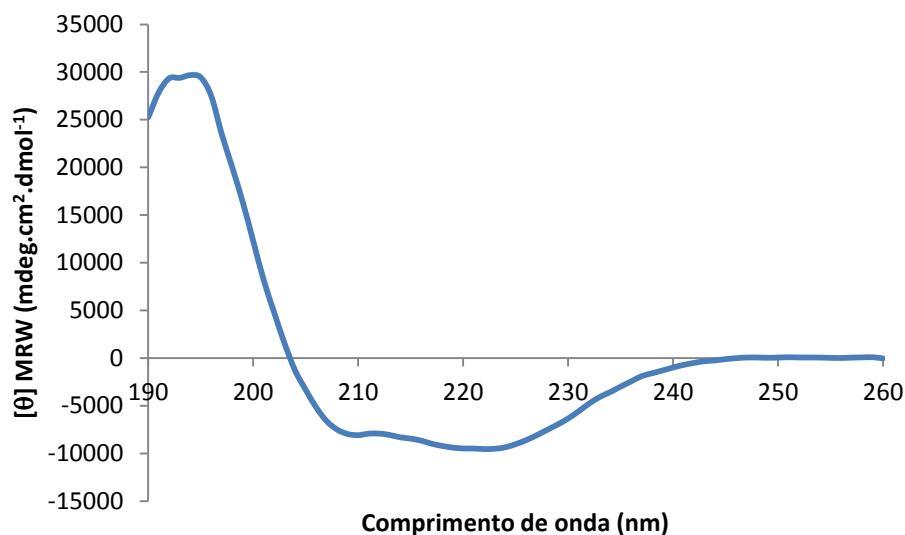


Figura 26: Espectro de Dicroísmo Circular entre 190 e 260 nm resultante da análise da proteína rSm21.7 a 20°C. Após a purificação da proteína rSm21.7, a mesma foi submetida a um ensaio de CD entre 190 e 260 nm. No gráfico, os valores calculados para elipcidade molar média foram plotados no eixo Y em função do comprimento de onda no eixo X.

Durante a desnaturação térmica da proteína, esse espectro característico de estrutura enovelada foi gradativamente sendo desfeito, indicando a perda dos elementos estruturais em α -hélice e fitas- β (Figura 27). Pode-se observar na figura 27 que esta perda de elementos estruturais ocorre principalmente a partir de 60°C.

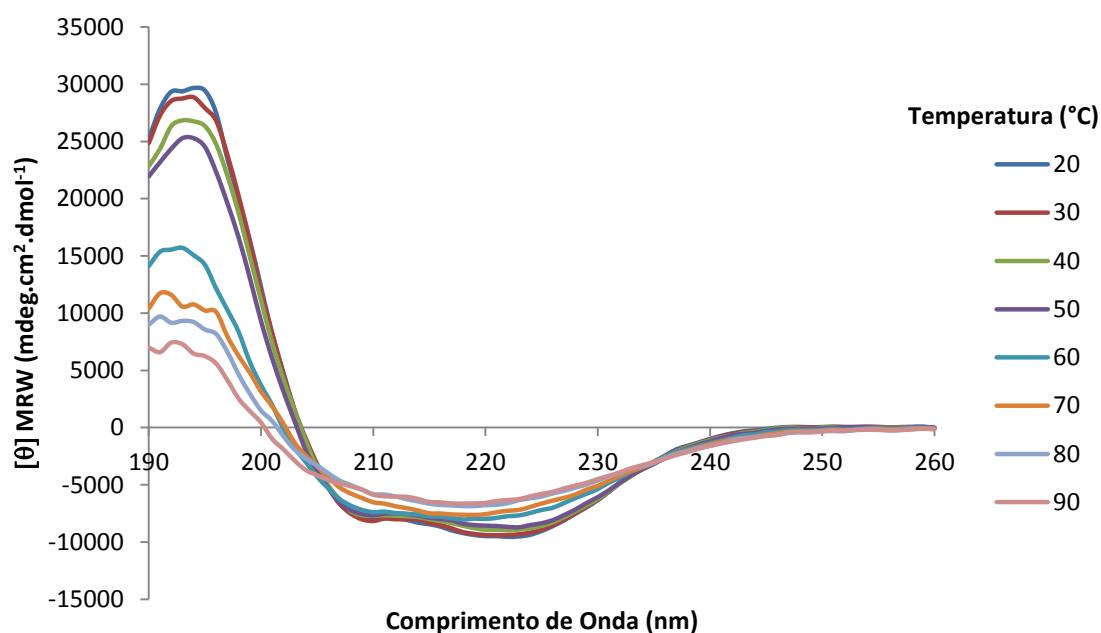


Figura 27: Desnaturação térmica da proteína rSm21.7. No experimento de CD a proteína rSm21.7 foi aquecida de 20 a 90°C para observação da perda de elementos estruturais em α -hélice e fitas- β .

A deconvolução dos dados de elipcidade molar através do programa *CDNN* permite que seja feita uma estimativa da porcentagem de elementos de estrutura secundária encontrados na proteína. Através dessa abordagem, a proteína rSm21.7 foi analisada em temperaturas variando entre 20 e 90°C (Figura 28).

A uma temperatura de 20°C, a proteína apresentou aproximadamente 37% de sua estrutura em α -hélice, 15% na forma de fitas- β e quase 50% de sua estrutura na forma de voltas- β e estrutura desordenada, denominada *random coil*. Através do aumento da temperatura, a porcentagem de α -hélice decresceu enquanto a porcentagem de estrutura desordenada aumentou em cerca de 10%. Esse fato sugere que o aumento da temperatura levou a uma desnaturação parcial da proteína, com perda de estrutura secundária helicoidal. Deve-se salientar ainda que, mesmo a 20°C, a proteína apresenta um conteúdo significativo de estruturas desordenadas (Figura 28).

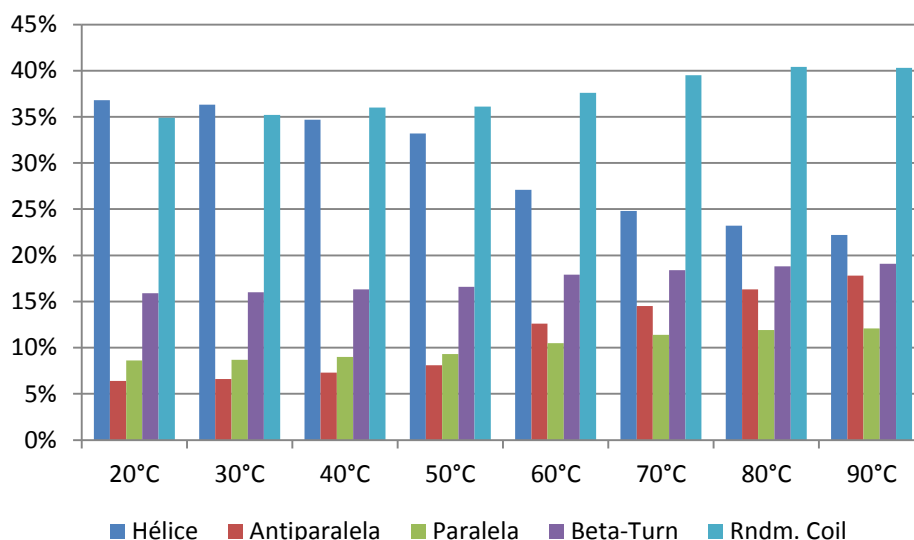


Figura 28: Estimativa da porcentagem de estrutura secundária da proteína rSm21.7 entre 20 e 90°C. Através da deconvolução dos dados de CD, a porcentagem de estrutura secundária da proteína rSm21.7 foi estimada. Os dados de elipicidade molar obtidos na desnaturação térmica foram utilizados no programa *CDNN* para fazer essa estimativa, que fornece as porcentagens de α -hélice, fitas- β paralelas e antiparalelas, voltas- β e estrutura randômica (*random coil*). Essas porcentagens foram plotadas no eixo Y em função das temperaturas no eixo X.

5.4 – Análise do estado de oligomerização da proteína rSm21.7

Os resultados da cromatografia de exclusão molecular da proteína rSm21.7 não possibilitaram um assinalamento conclusivo do seu estado oligomérico. Assim, foram realizados ensaios de DLS e de *cross-linking* para determinação do mesmo.

5.4.1 – Estado oligomérico de rSm21.7 – DLS

Nas análises de DLS observou-se que a 25°C e a uma concentração de aproximadamente 3,5 mg/mL, a amostra purificada da proteína rSm21.7 apresenta uma única população de moléculas com raio hidrodinâmico de aproximadamente 3,4 nm (Figura 29). Correlacionando-se este raio com sua massa molecular, o experimento sugere uma molécula com aproximadamente 60 kDa (Tabela 2).

Assim como nas análises por cromatografia de exclusão molecular esse resultado não foi preciso o suficiente para indicar o estado oligomérico da proteína quando em solução: dímero ou trímero?

Entretanto, a massa molecular estimada neste experimento depende, em última análise, da forma da molécula (coeficiente de difusão translacional) e, por isto, pode apresentar inconsistências no caso de proteínas não globulares. Adicionalmente, as concentrações de proteínas utilizadas podem influenciar nos resultado dos experimentos. Nessas condições, o DLS também não representa a melhor técnica para análise do estado de oligomerização da proteína rSm21.7, assim como a cromatografia de exclusão molecular. Outras análises sobre o estado de oligomerização de rSm21.7 serão apresentados posteriormente.

A análise de DLS demonstrou, porém, que a proteína rSm21.7 encontra-se monodispersa (13,5% de polidispersividade), o que representa um bom indicador da qualidade da amostra para posteriores testes de cristalização (Tabela 2). É bem conhecido no campo da Cristalografia de Macromoléculas que quanto maior for a polidispersividade das proteínas em solução, menor será a probabilidade de cristalização das mesmas (FERRÉ-D'AMARÉ; BURLEY, 1997; ZULAUF; D'ARCY, 1992).

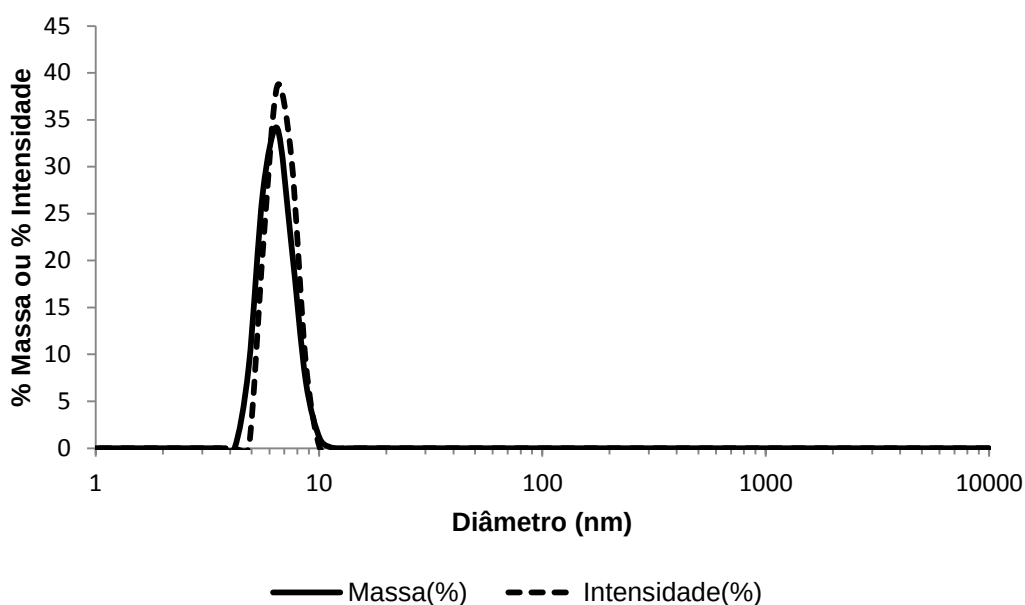


Figura 29: Espalhamento Dinâmico da Luz da proteína rSm21.7. A proteína recombinante foi purificada por cromatografias de afinidade e exclusão molecular e posteriormente o estado de oligomerização foi avaliado por DLS. O gráfico apresenta no eixo X o logaritmo do diâmetro e no eixo Y a porcentagem de massa ou de intensidade.

Tabela 2: Espalhamento Dinâmico da Luz da proteína rSm21.7. Após a purificação da proteína rSm21.7, o raio hidrodinâmico, a massa molecular e a polidispersividade da amostra foram determinados por DLS.

Raio hid. (nm)	Pd (%)	MM (kDa)	Intensidade (%)	Massa (%)
3,4	13,5	59,7	100	100

5.4.2 – Estado oligomérico de rSm21.7 – *Cross-linking*

Uma vez que não foi possível determinar precisamente o estado de oligomerização de rSm21.7 por cromatografia de exclusão molecular ou DLS, foi realizado um ensaio de *cross-linking* com glutaraldeído. As análises desse *cross-linking* em SDS-PAGE demonstraram que a proteína rSm21.7 pode existir como um dímero. Quando a proteína rSm21.7 foi incubada por 30 minutos com concentrações iguais ou superiores a 0,025% de glutaraldeído, foi possível visualizar uma banda com aproximadamente 45kDa, sugerindo a presença de um dímero, conforme destacado na figura 30. Já na incubação por 5 minutos, foi possível visualizar uma banda com aproximadamente 45kDa nas concentrações iguais ou superiores a 0,05% de glutaraldeído (resultados não mostrados).

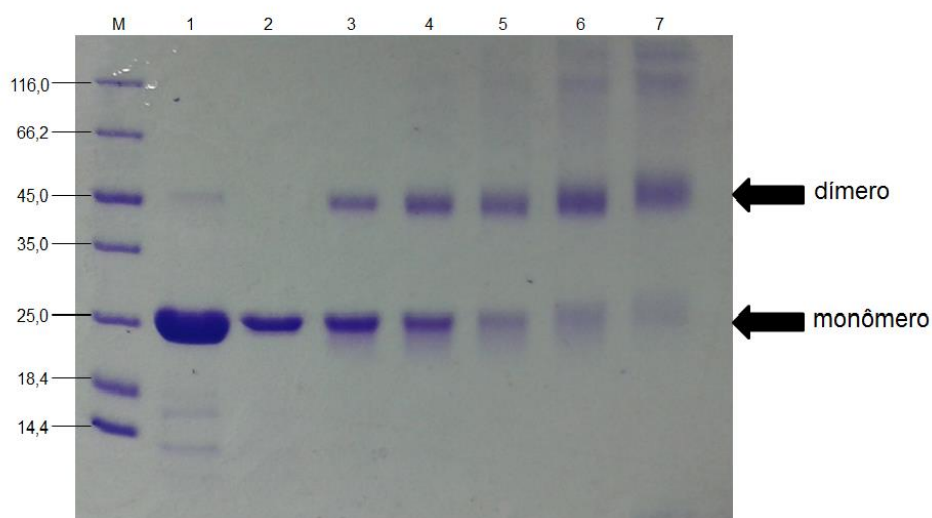


Figura 30: Análise de *cross-linking* da proteína rSm21.7. Amostras de proteína purificada foram incubadas com concentrações crescentes de glutaraldeído por 30 minutos e posteriormente analisadas por SDS-PAGE corado com *Coomassie Brilliant Blue R250*. M – marcador de massa molecular em kDa (*Unstained Protein Molecular Weight Marker* – Fermentas); 1 – rSm21.7 a 60 μ M; 2 – rSm21.7 a 5 μ M; 3 – 0,025% glutaraldeído; 4 – 0,05% glutaraldeído; 5 – 0,1% glutaraldeído; 6 – 0,2% glutaraldeído; 7 – 0,4% glutaraldeído.

Assim, tendo em vista as limitações apresentadas pela cromatografia de exclusão molecular e pelo DLS na determinação do estado oligomérico dessa molécula, pode-se sugerir que o *cross-linking* seja um teste adequado para esse objetivo. Essas análises com glutaraldeído indicam fortemente que a proteína comporta-se como um dímero.

Nesse contexto de oligomerização da proteína rSm21.7, já foi reportado que algumas das proteínas de trematódeos que apresentam um domínio de mão EF e um domínio de cadeia leve de dineína comportam-se como dímeros (BANFORD *et al.*, 2013; ORR *et al.*, 2012; SENAWONG *et al.*, 2012). Tal propriedade de dimerização foi demonstrada para as proteínas FhCaBP3 e FhCaBP4 de *F. hepatica* através de *cross-linking*, SDS-PAGE e gel filtração (BANFORD *et al.*, 2013; ORR *et al.*, 2012) e para FgCaBP1, por SDS-PAGE (VICHASRI-GRAMS *et al.*, 2006). Já para a proteína OvCaBP a dimerização foi demonstrada por *Western Blot*, que também identificou a presença de outras formas multiméricas, além do dímero (SENAWONG *et al.*, 2012).

Analisando individualmente ambos os domínios constituintes da proteína rSm21.7, verifica-se que tanto o domínio de mão EF quanto o domínio de cadeia leve de dineína podem formar dímeros (FAN, JING-SONG *et al.*, 2001; LIANG *et al.*, 1999; YAP, K. L. *et al.*, 1999).

Algumas proteínas contendo mão EF são naturalmente encontradas como dímeros e, a estrutura tridimensional de algumas delas já foram determinadas, como por exemplo, as proteínas S100 (Figura 11), calciclina, psoriasina e p11, bem como calpaína (Figura 11) e neurocalcina (Figura 11) (Revisado por YAP, K. L. *et al.*, 1999).

Tais proteínas podem formar dímeros de diferentes formas. As da subfamília S100 contêm dois domínios de mão EF formando homodímeros ou heterodímeros com outras proteínas S100 (HUNG; HSU; YU, 2013; POTTS *et al.*, 1995; REIS; BORTOT; CALIRI, 2014; TEIGELKAMP *et al.*, 1991).

Já a subunidade menor da calpaína, uma serino protease, contém 5 domínios de mão EF, sendo que o quinto domínio forma a interface do dímero (BLANCHARD *et al.*, 1997; LIN *et al.*, 1997; RAVULAPALLI *et al.*, 2005). Outro exemplo de dímero de mão EF é a neurocalcina. A interface dimérica nesse caso envolve apenas seis resíduos (VIJAY-KUMAR; KUMAR, 1999). Claramente, proteínas de diferentes subfamílias usam diferentes mecanismos de dimerização e essa diversidade pode ser responsável pela regulação de múltiplas proteínas alvo (YAP, K. L. *et al.*, 1999).

O domínio de cadeia leve de dineína também pode existir como monômero ou dímero (BENASHSKI *et al.*, 1997; LO *et al.*, 2005; ROMES; TRIPATHY; SLEP,

2012; WANG *et al.*, 2003) . O monômero consiste em uma molécula polarizada com uma face consistindo em um par de α -hélices (face voltada para o solvente) e a outra face hidrofóbica contendo 4 ou 5 fitas- β , que formam uma folha- β (Figura 13). Essa folha- β antiparalela constitui o domínio de dimerização e, uma das fitas- β sofre uma contribuição do monômero oposto, estabilizando o dímero (Figura 14) (ESPINDOLA *et al.*, 2000).

Esses dados da literatura suportam os resultados que demonstram a proteína rSm21.7 como um homodímero.

5.5 – Cristalização da proteína rSm21.7 e coleta de dados de difração de raios X

A Cristalografia de Macromoléculas é, atualmente, a principal técnica experimental para obtenção de estruturas tridimensionais de proteínas. Assim, com o objetivo de obter a estrutura da proteína rSm21.7, testes de cristalização foram realizados tanto pelo método da gota suspensa quanto pelo método da gota depositada.

Utilizando as soluções dos kits comerciais *Crystal Screen* e *Crystal screen 2* (Hampton Research) no *screening* inicial pela técnica da gota suspensa, duas condições propiciaram o crescimento de pequenos cristais de rSm21.7 em forma de placa (Figura 31). Tais condições foram: (i) 200 mM sulfato de amônio, 100 mM acetato de sódio pH 4,6, 30% (p/v) PEG monometil éter 2.000 e (ii) 200 mM sulfato de amônio, 100 mM cacodilato de sódio pH 6,5, 30% (p/v) PEG 8.000.

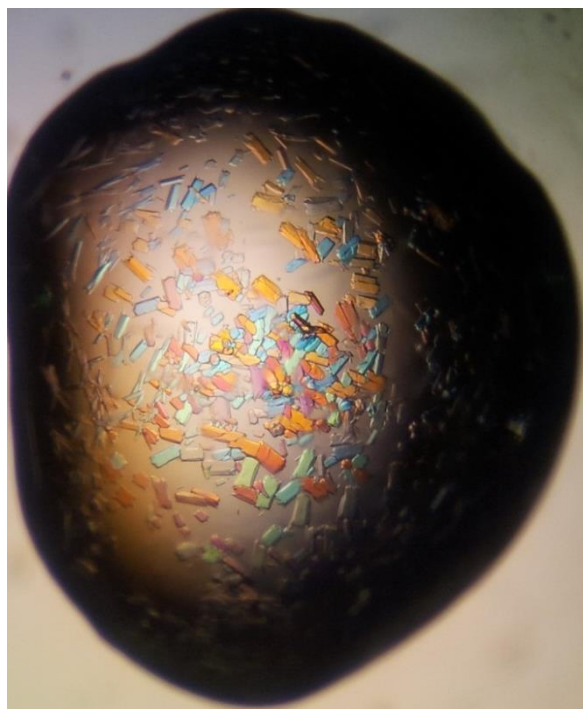


Figura 31: Cristais da proteína rSm21.7 obtidos no *screening* inicial pelo método da gota suspensa. A proteína recombinante purificada foi concentrada para 20 mg/mL para a realização dos testes de cristalização pelo método da gota suspensa. Assim, a proteína foi cristalizada em duas condições. Os cristais em forma de placas aqui demonstrados foram obtidos no *screening* inicial em uma condição contendo 200 mM sulfato de amônio, 100 mM cacodilato de sódio pH 6,5, 30% (p/v) PEG 8.000.

Objetivando a obtenção de cristais maiores, essas condições foram otimizadas com variações de pH, concentração de sal, de precipitante e volume/proporção das gotas, totalizando 288 novas condições testadas. As variações de pH para o tampão acetato de sódio foram feitas entre os pHs 4,0 e 5,5. Já para o tampão cacodilato de sódio os pHs variaram entre 5,5 e 7,4. As concentrações de PEG 8.000 e PEG monometil éter 2.000 variaram entre 20 e 40% (p/v) e de sulfato de amônio entre 100 e 300 mM.

Os cristais que resultaram nos melhores dados de difração (maior resolução) dentre aqueles testados foram obtidos após aproximadamente 30 dias de cristalização nas seguintes condições: (i) 100 mM sulfato de amônio, 25% (p/v) PEG monometil éter 2.000 e 100 mM acetato de sódio pH 5,5 (1 μ L de proteína e 2 μ da solução de poço) e (ii) 200 mM sulfato de amônio, 25% (p/v) PEG 8.000 e 100 mM cacodilato de sódio pH 5,5 (2 μ L de proteína e 2 μ L de solução de poço) (Figura 32). Em ambas as condições ocorreu a formação de um precipitado no fundo da gota a partir do qual os cristais em forma de placa cresceram

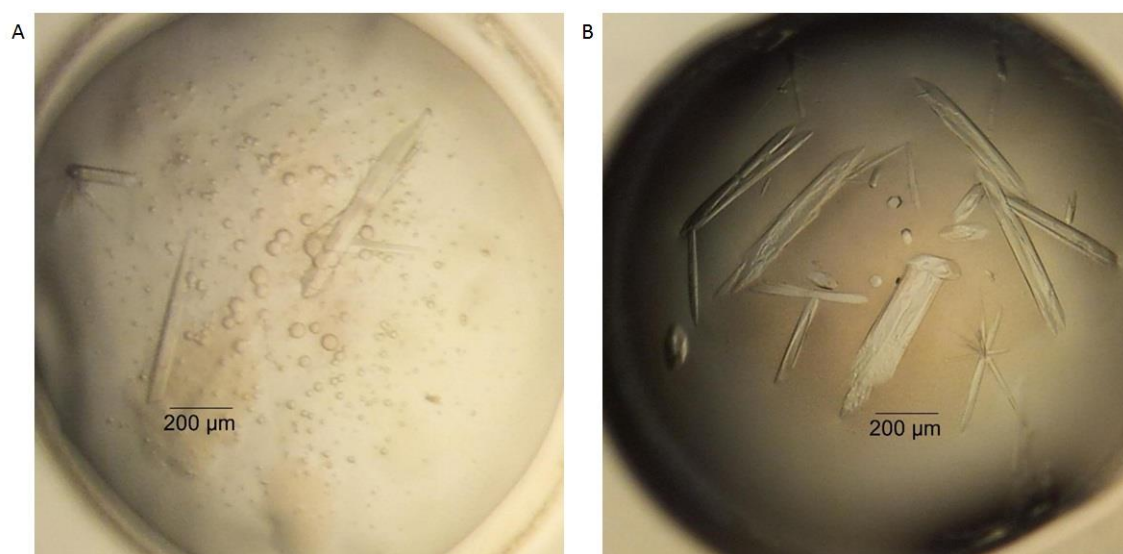


Figura 32: Cristais da proteína rSm21.7 obtidos após a otimização das condições iniciais. Após a obtenção de pequenos cristais no *screening* inicial, as condições de cristalização foram otimizadas e os cristais obtidos foram submetidos a experimentos de difração de raios X. Os cristais que alcançaram maior resolução foram obtidos nas seguintes condições: (A) 100 mM sulfato de amônio, 25% (p/v) PEG monometil éter 2.000 e 100 mM acetato de sódio pH 5,5 e (B) 200 mM sulfato de amônio, 25% (p/v) PEG 8.000 e 100 mM cacodilato de sódio pH 5,5. A concentração da proteína foi de aproximadamente 20 mg/mL.

Quando a técnica da gota depositada foi utilizada, também foi observada a formação de cristais (Figura 33). Contudo, estas condições não foram otimizadas por terem sido obtidas posteriormente à resolução da estrutura tridimensional da proteína.

Nas seguintes condições ocorreu o crescimento de cristais: (i) 200 mM formato de magnésio, 20% PEG 3.350; (ii) 200 mM cloreto de sódio, 100 mM bis-tris pH 5,5, 25% PEG 3.350; (iii) 200 mM acetato de amônio, 100 mM bis-tris pH 5,5, 25% PEG 3.350; (iv) 200 mM cloreto de magnésio, 100 mM hepes pH 7,0, 20% PEG 6.000; (v) 200 mM cloreto de cálcio, 100 mM tris pH 8,0, 20% PEG 6.000; (vi) 200 mM nitrato de sódio, 20% PEG 3.350; (vii) 200 mM nitrato de sódio, 100 mM bis-tris propano pH 6,5, 20% PEG 3.350; (viii) 200 mM formato de sódio, 100 mM bis-tris propano pH 6,5, 20% PEG 3.350; (ix) 200 mM acetato de sódio, 100 mM bis-tris propano pH 6,5, 20% PEG 3.350 (Figura 33) e (x) 200 mM brometo de sódio, 100 mM bis-tris propano pH 7,5, 20% PEG 3.350 (Figura 33).

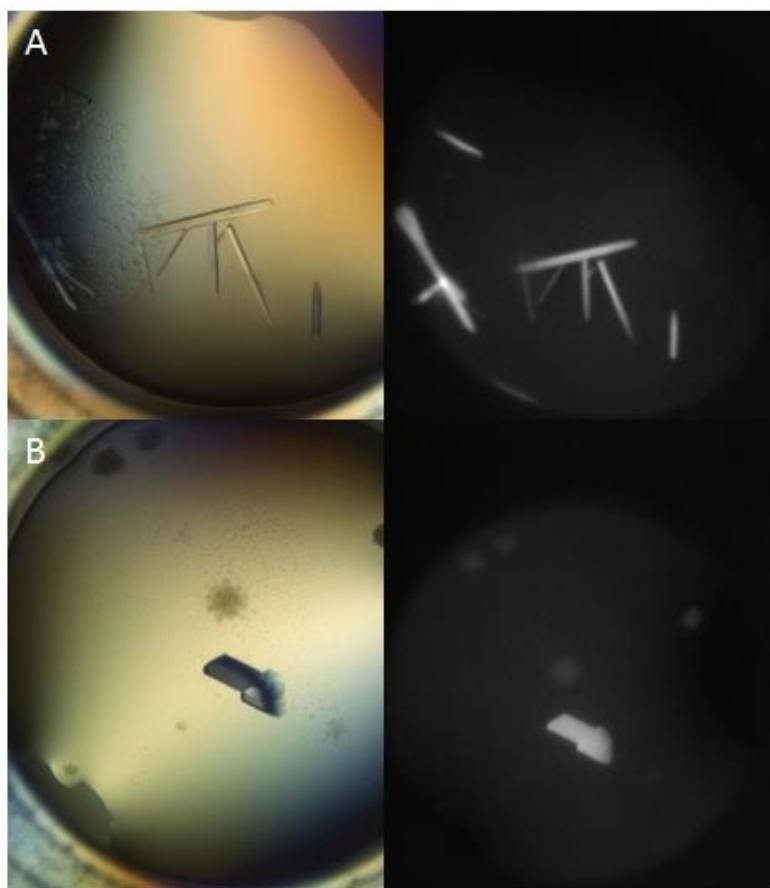


Figura 33: Cristais da proteína rSm21.7 obtidos pelo método da gota depositada. Um *screening* de aproximadamente 600 condições foi realizado no LNLS pelo método da gota depositada. Dessa forma, cristais foram obtidos em várias condições, duas das quais são: (A) 200 mM acetato de sódio, 100 mM bis-tris propano pH 6,5, 20% PEG 3.350 e (B) 200 mM brometo de sódio, 100 mM bis-tris propano pH 7,5, 20% PEG 3.350. Os ensaios foram realizados utilizando-se a proteína rSm21.7 a 20 mg/mL. As fotografias ao lado dos cristais foram obtidas com iluminação UV e o brilho dos cristais indica que os mesmos são constituídos de proteína.

Para a coleta de dados, um cristal nativo crescido na condição contendo 100 mM sulfato de amônio, 25% PEG monometil éter 2.000 e 100 mM acetato de sódio pH 5,5 (Figura 32) foi submetido a um experimento de difração de raios X. Este cristal apresentou um limite de difração até 2,05 Å de resolução (Figura 34). Os dados coletados foram, então, processados no programa HKL2000 (OTWINOWSKI; MINOR, 1997) no qual verificou-se que o cristal pertence ao sistema cristalino hexagonal (Tabela 3).

Após o processamento, o coeficiente de Mathews (MATTHEWS, 1968) foi estimado em $2,00 \text{ Å}^3 \text{Da}^{-1}$, com um conteúdo de solvente de 38,4%. Esse dado foi obtido assumindo a presença de uma molécula de proteína com 200 aminoácidos (cauda 6xHis,

sítio de clivagem para a protease TEV e a proteína completa) por unidade assimétrica do cristal.

Após a coleta de dados de difração e processamento dos mesmos, as tentativas de resolver a estrutura da proteína pelo método de Substituição Molecular não foram bem sucedidas. Em todas as abordagens testadas procuravam-se soluções que comportassem ambos os domínios na estrutura alvo. Assim, foram feitas diversas tentativas utilizando como molde o motivo de mão EF da Calmodulina de *Paramecium tetraurelia* (PDB - 1CLM) (RAO, S. T. *et al.*, 1993) e o um domínio semelhante ao da cadeia leve de dineína da proteína Hsv2p de *Kluyveromyces lactis* (PDB - 4EXV e 4AV9) (BASKARAN *et al.*, 2012). A primeira proteína apresenta uma identidade (similaridade) de 33% (52%) com o motivo de mão EF de rSm21.7. Já a segunda proteína apresenta 46% de similaridade (24% de identidade) com o domínio de cadeia leve de dineína de rSm21.7. Estas proteínas disponíveis no PDB são as que apresentaram a maior similaridade em sequência com a proteína alvo deste trabalho.

Como o problema das fases não foi resolvido pelo método de Substituição Molecular, o método de Átomos Pesados foi utilizado para faseamento e determinação da estrutura.

Para utilizar o método de Átomos Pesados, um cristal derivado foi obtido da seguinte forma. Um cristal nativo (crescido na condição 200 mM sulfato de amônio, 25% PEG 8.000 e 100 mM cacodilato de sódio pH 5,5) (Figura 32) foi transferido por cerca de 30 segundos para uma solução criogênica (constituída pela solução do poço com 10% etilenoglicol) contendo 0,4M de iodeto de sódio. Esse cristal foi, então, utilizado em um experimento de difração de raios X e apresentou um poder de difração até 1,95 Å de resolução (Figura 34) (Tabela 3).

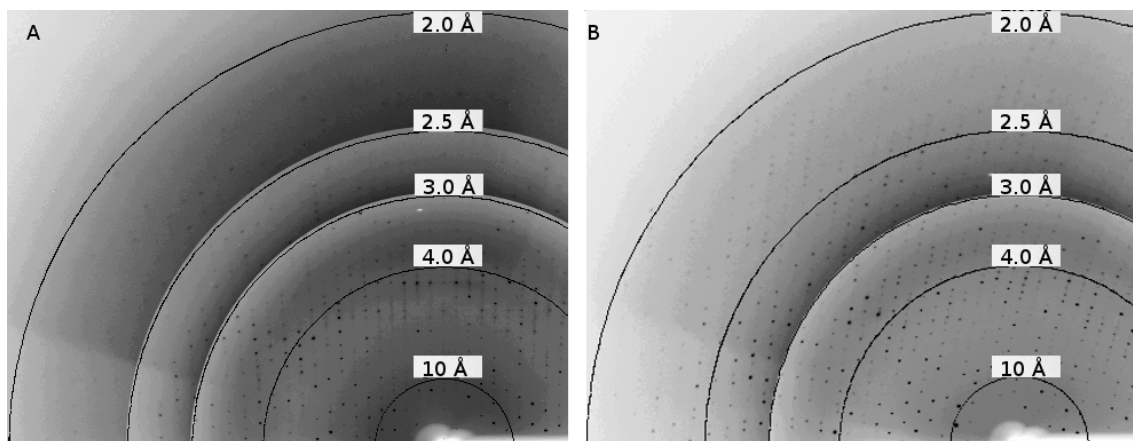


Figura 34: Padrão de difração de raios X dos cristais nativo e derivado de rSm21.7. Os cristais obtidos após a otimização das duas condições (método da gota suspensa) foram submetidos a experimentos de difração de raios X no LNLS. O cristal nativo (A) difratou até 2,05 Å enquanto o cristal derivado (B) difratou até 1,95 Å de resolução. As imagens foram coletadas em um detector Rayonix MarMosaic 225 CCD. Nessa figura, o brilho e contraste das imagens foram ajustados nas diferentes faixas de resolução para permitir uma melhor visualização dos pontos de difração.

O cristal derivado de iodo apresentou valores para os parâmetros de célula bem parecidos com os do cristal nativo sugerindo que ambos poderiam ser isomórficos. Além disto, uma diferença anômala de 6% nos pares de Friedel assegura que houve uma incorporação de ânions iodeto na estrutura cristalográfica do cristal derivado.

Utilizando o programa SHELXD (SHELDRICK, 2010), dois locais de incorporação de iodeto com ocupância bem significativa para uma derivatização rápida (1,00 e 0,95) foram encontrados na estrutura cristalográfica do derivado. Esses locais, juntamente com ambos os conjuntos de dados (nativo e derivado) foram utilizados para determinação das fases dos fatores de estrutura através do Método de SIRAS usando o programa PHENIX (ADAMS *et al.*, 2010).

Tabela 3: Estatísticas de coleta de dados e processamento dos cristais nativo e derivado de rSm21.7. Os dados entre parênteses indicam os valores da menor e maior resolução, respectivamente.

	nativo	derivado
Comprimento de onda (Å)	1,459	1,459
Temperatura (K)	100	100
Distância cristal-detector (mm)	120	105
Rotação por imagem (°)	0,5	0,5
Tempo de exposição por imagem (s)	16	6
Variação total de rotação (°)	130	150
Grupo espacial	P6 ₁ 22	P6 ₁ 22
Parâmetros de célula unitária (Å)	a=b=108,5, c=55,8	a=b=109,1, c=56,4
Resolução (Å)	50,00-2,05 (50,00-4,69/2,11-2,05)	50,00-1,95 (50,00-4,46/2,01-1,95)
Nº de observações	169.511	183.507
Nº. de reflexões únicas	22.353 (1.899/1.778)	27.084 (2.284/2.092)
Completeza (%)	96,7 (97,8/92,9)	98,9 (99,0/92,0)
<I/ σ(I)>	26,9 (47,5/2,7)	16,9 (52,4/2,1)
Multiplicidade	7,6 (8,3/4,9)	6,8 (9,4/3,1)
R _{merge} ^a (%)	6,8 (3,4/50,9)	10,7 (4,3/52,1)
Mosaicidade (°)	0,81	1,18
Coeficiente de Matthews (Å ³ /Da) ^b	4,20	4,30
Conteúdo de solvente (%) ^b	70,7	71,4
R _{anom} ^c (%)	-	5,98 ^e
R _{isom} ^d (%)	-	21,6 ^e

$$^a R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl) \times 100$$

^b Números calculados utilizando os 96 resíduos N-terminais por unidade assimétrica

$$^c R_{anom} = \sum_{hkl} |\langle I(hkl) \rangle - \langle I(\overline{hkl}) \rangle| / \sum_{hkl} \langle I(hkl) \rangle \times 100$$

$$^d R_{isom} = \sum_{hkl} |F_{deriv}(hkl) - F_{nat}(hkl)| / \sum_{hkl} F_{nat}(hkl) \times 100$$

^e O sinal anômalo foi observado até uma resolução de 3 Å

Após a determinação das fases dos fatores de estrutura, foi realizada uma etapa de modificação da densidade eletrônica seguida por uma construção automática do modelo contra o conjunto de dados nativo. Este procedimento resultou em um modelo parcial de rSm21.7 com um R_{work} (R_{free}) de 0,23 (0,25). O R_{work} é uma medida da concordância entre as amplitudes dos fatores de estrutura medidos experimentalmente e aquelas calculadas a partir da estrutura parcial que está sendo refinada.

Surpreendentemente, o modelo inicial continha apenas os 96 resíduos C-terminais constituintes do domínio de cadeia leve de dineína (de Gln89 até Asn184), possuindo uma massa molecular teórica de 11.276,8Da. Não foi observada nenhuma densidade eletrônica dos resíduos anteriores à Gln89 (Figura 35).

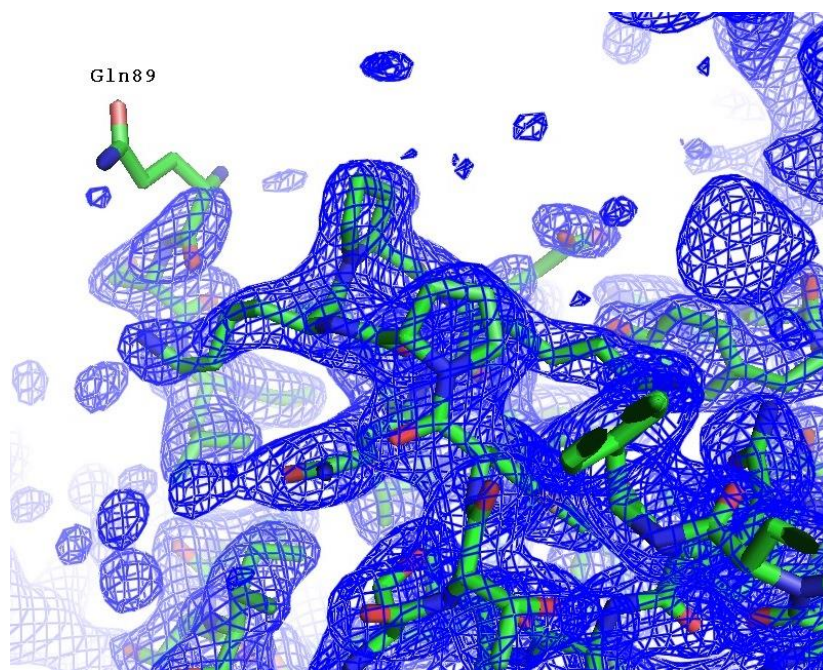


Figura 35: Mapa de densidade eletrônica da região da proteína rSm21.7 demonstrando o resíduo Gln89. Não foi observada densidade eletrônica dos resíduos anteriores ao Gln89. Os mapas ($2mF_{obs}-DF_{calc}$, Φ_{calc}) foram contornados a $1,0 \sigma$. A imagem foi gerada pelo programa Pymol (DELANO, W. L., 2007).

Após inspeção deste modelo inicial, verificou-se que todas as proteínas do cristal nativo apresentavam o resíduo Gln89 apontando para um canal de aproximadamente 80 Å de diâmetro como mostrado na figura 36.

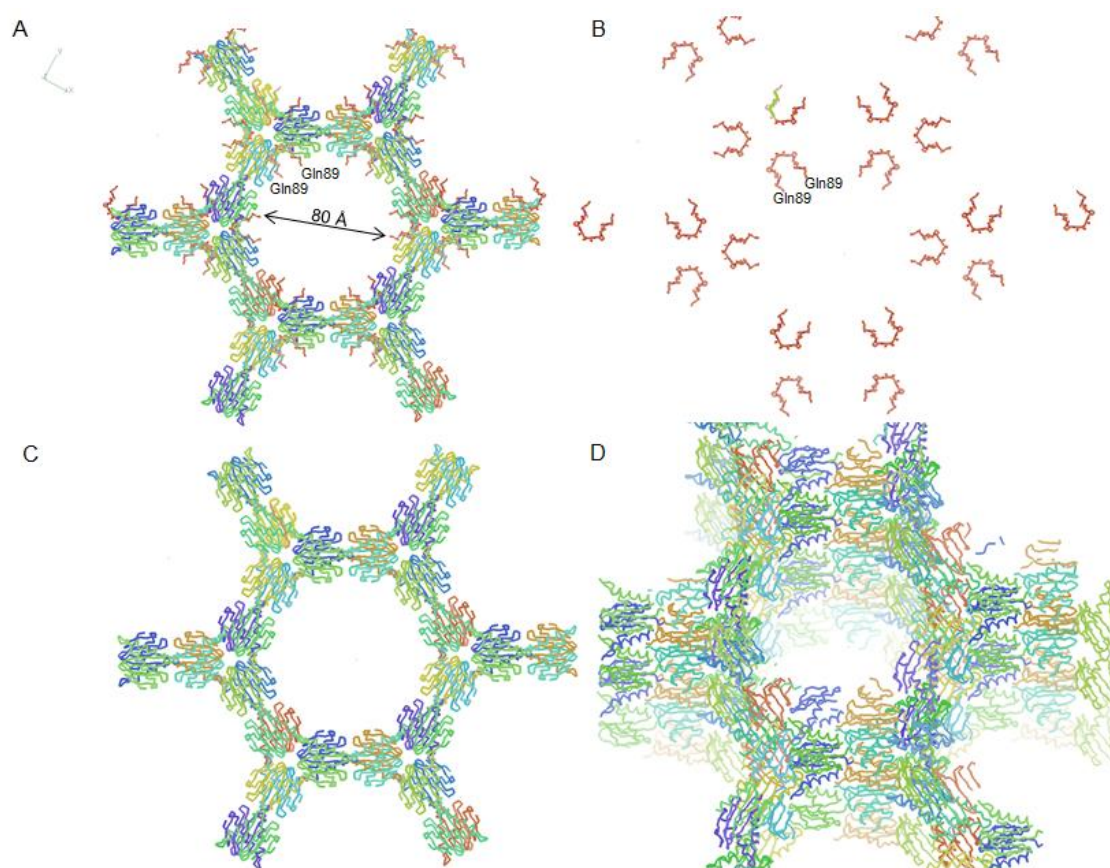


Figura 36: Canal formado pela disposição das proteínas nativas rSm21.7 no cristal. Nessas imagens as cadeias principais das moléculas de proteína são apresentadas. (A) Resíduos de Gln89 N-terminais apontados para o interior do canal de 80 Å. (B) Cadeia principal dos 5 resíduos N-terminais das moléculas da proteína. (C) Canal formado pelas proteínas. (D) Imagem C em uma orientação diferente. As imagens foram geradas pelo programa Coot (EMSLEY et al., 2010).

Diante desta observação, duas hipóteses foram levantadas. Na primeira, a proteína teria sofrido clivagem na própria gota de cristalização, uma vez que análises de SDS-PAGE e espectrometria de massas demonstraram que a mesma possuía aproximadamente 23 kDa logo após a purificação e no momento do preparo dos testes de cristalização. Já na segunda, a proteína rSm21.7 apresentaria uma região N-terminal parcialmente desenovelada que ocuparia os canais de 80 Å de diâmetro e que impossibilitaria uma definição de densidade eletrônica na região.

Com o objetivo de verificar se a proteína sofreu clivagem na gota de cristalização e apenas o domínio de cadeia leve de dineína foi cristalizado, um SDS-PAGE dos cristais foi realizado. Alguns cristais foram lavados na solução de cristalização e, posteriormente, dissolvidos nessa mesma condição para separação eletroforética e análise por espectrometria de massas (neste caso os cristais foram

lavados e dissolvidos em água). Observou-se que os cristais dissolvidos apresentavam uma massa molecular de aproximadamente 11 kDa (Figuras 37 e 38), indicando que a proteína sofreu uma clivagem após a preparação do experimento de cristalização.

A massa molecular teórica do fragmento que teve sua densidade eletrônica visualizada (Gln89 a Asn184) foi de 11.276,8 Da. Já a massa observada por espectrometria de massas foi de 11.609,3 Da, sugerindo a presença de outros três resíduos no fragmento cristalizado (Gly, Arg e Glu).

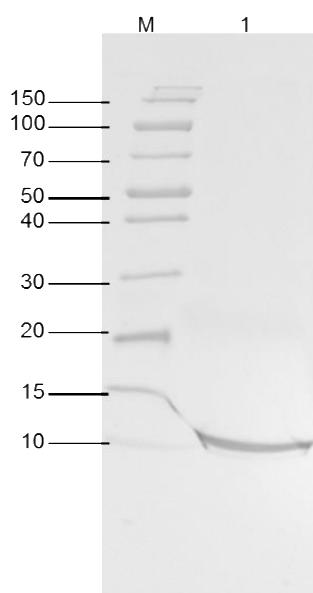


Figura 37: SDS-PAGE dos cristais da proteína rSm21.7. Alguns cristais da proteína rSm21.7 foram dissolvidos em solução de cristalização e posteriormente essa amostra foi separada eletroforéticamente. Pode-se visualizar no gel: M – Marcador de massa molecular em kDa (*PageRuler Broad Range Unstained Protein Ladder* – Thermo Scientific), 1 – cristais da proteína rSm21.7 dissolvidos.

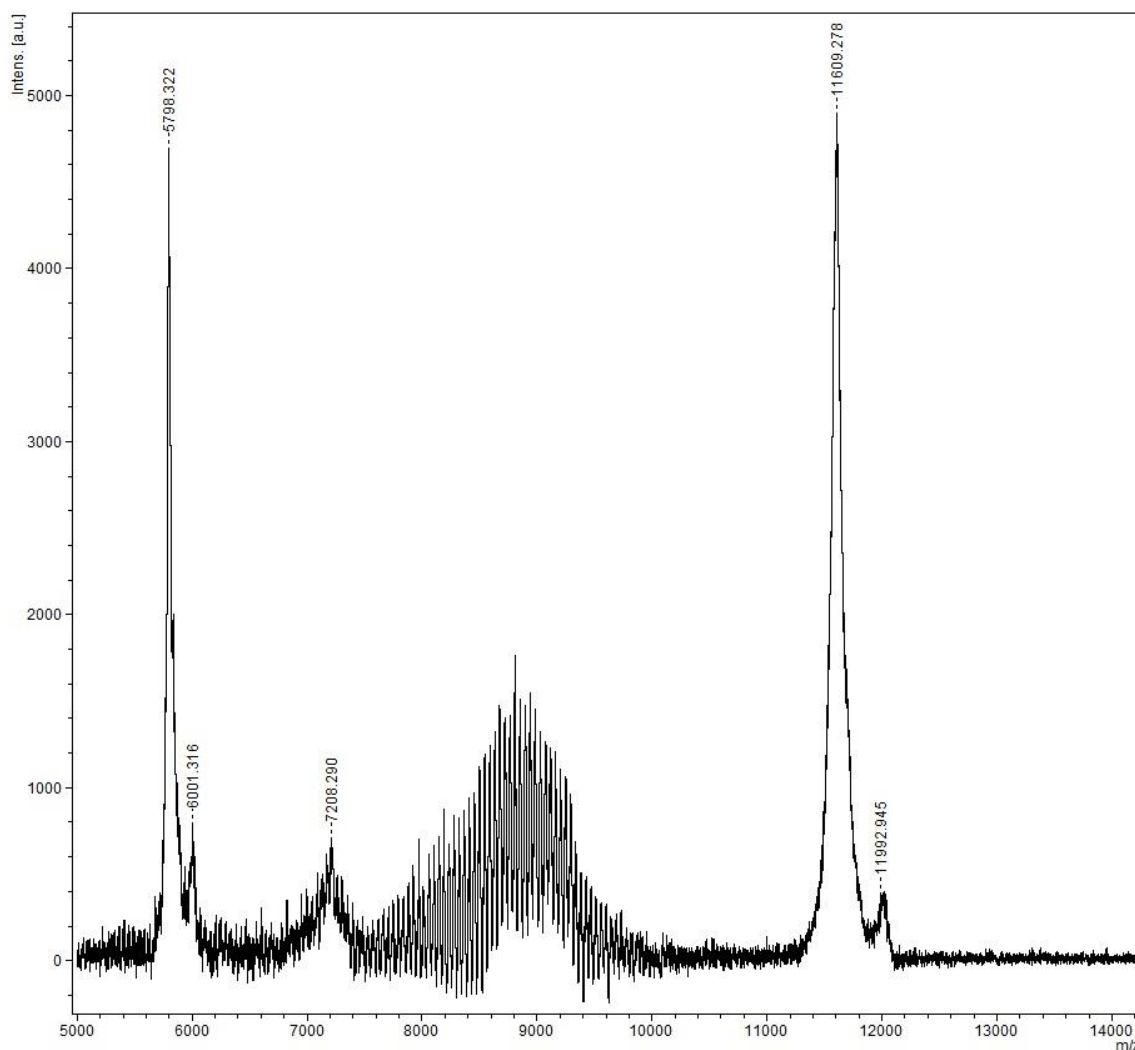


Figura 38: Espectrometria de massas dos cristais de rSm21.7 dissolvidos. Os cristais da proteína rSm21.7 foram dissolvidos e posteriormente analisados por espectrometria de massas.

Devido ao fato de a proteína estar clivada no cristal, o coeficiente de Matthews foi novamente calculado, considerando-se uma molécula de 11.276,8 Da no cristal. Assim o Coeficiente de Matthews (V_M) passou a ser $4.20 \text{ \AA}^3 \text{ Da}^{-1}$ com um conteúdo de solvente de 70.7%.

De fato essa não é uma situação incomum, já que outros exemplos da literatura demonstram a ocorrência da clivagem da proteína na gota de cristalização. Um polipeptídeo de *Streptomyces coelicolor* A3(2), denominado fator sigma, foi clonado, expresso, purificado e submetido a testes de cristalização. Entretanto os cristais apresentavam apenas a porção C-terminal da proteína com massa molecular menor que 11 kDa, contrastando com a massa de 25 kDa da proteína purificada (KIM, KEON

YOUNG *et al.*, 2014). O mesmo ocorreu com um membro da família das epsinas de levedura, que também se apresentou clivado na gota de cristalização (TANNER; PRAG, 2012).

De acordo com dados da literatura essa clivagem na gota de cristalização poderia ser resultado de uma ação autolítica ou ação de proteases contaminantes (por exemplo, do meio de cultura) (CHIOU *et al.*, 1999).

Diferentemente dos exemplos citados acima, alguns trabalhos nos quais a proteína foi clivada na gota de cristalização conseguiram demonstrar as causas da clivagem. A proteína CstF-77 de camundongos só foi cristalizada após sua clivagem na gota de cristalização por proteases provenientes de contaminação por fungos (BAI; AUPERIN; TONG, 2007), assim como a proteína CPSF-100 (Ydh1p) de leveduras (MANDEL *et al.*, 2006). Esses cristais foram reproduzidos adicionando-se proteases às gotas de cristalização (BAI; AUPERIN; TONG, 2007; MANDEL *et al.*, 2006).

Outros exemplos de cristalização que ocorreram após contaminação por fungos e consequente clivagem foram o da enzima 1-Deoxi-D-xilulose 5-fosfato sintase de *E. coli*, cujos resíduos 183 a 238 e 292 a 317 não apresentaram densidade eletrônica (XIANG *et al.*, 2007) e o de uma proteína de 51 kDa de *Bacillus sphaericus* (CHIOU *et al.*, 1999).

Uma técnica muito comum na Cristalografia de Macromoléculas é a proteólise *in situ*, que consiste exatamente na adição de proteases às gotas de cristalização para clivar regiões da proteína que podem estar inibindo a formação do cristal (WERNIMONT; EDWARDS, 2009)

Na verdade, muitos domínios ou fragmentos proteolíticos de proteínas foram cristalizados, quando a molécula inteira não foi ou dados de difração dos cristais não puderam ser coletados. O domínio catalítico da integrase de HIV (*Human Immunodeficiency Virus*) (DYDA *et al.*, 1994) e um subfragmento da miosina (RAYMENT *et al.*, 1993) são dois casos reportados na literatura.

Alguns exemplos de cristalização envolvendo proteólise *in situ* (adição de proteases às gotas de cristalização) são a topoisomerase (TANEJA *et al.*, 2006), o fragmento Vh de humanos (GAUR *et al.*, 2004), Ipa D, um componente do sistema de secreção de *Shigella flexneri* (JOHNSON *et al.*, 2006), a proteína semelhante à dihidrodipicolonato sintetase de humanos (BUNKER; LOOMES; BAKER, 2012), a proteína TraN de bactérias gram negativas (GOESSWEINER-MOHR *et al.*, 2012), a subunidade R2 da ribonuclease redutase de herpesvirus associado ao sarcoma de Kaposi

(GURMU *et al.*, 2010), neuroaminidase de *Streptococcus pneumoniae* (HSIAO *et al.*, 2009) e a subunidade β do hormônio luteinizante de bovinos (MCPHERSON; DAY; HARRIS, 2004).

No caso da proteína rSm21.7, assim como na cristalização da proteína de *Bacillus sphaericus* citada anteriormente, a clivagem pode ter influenciado na formação do cristal, uma vez que a proteína inteira precipitou na gota no momento da realização dos testes. É possível que o fragmento N-terminal tenha sido liberado na solução e então, os cristais cresceram apenas após a clivagem (CHIOU *et al.*, 1999).

Mesmo com um modelo apenas do domínio de cadeia leve de dineína da proteína rSm21.7, o mesmo foi submetido a ciclos de refinamento para obtenção de uma estrutura final desse domínio que será apresentada mais ao final da seção de Resultados.

5.6 – A proteína rSm21.7 possui um domínio suscetível à ação de proteases

Durante os testes com a proteína rSm21.7, observou-se que a mesma sofria um processo de degradação ao longo do tempo, e o polipeptídeo resultante possuía uma massa molecular menor que 14 kDa, como pode ser visualizado em SDS-PAGE (Figura 39).

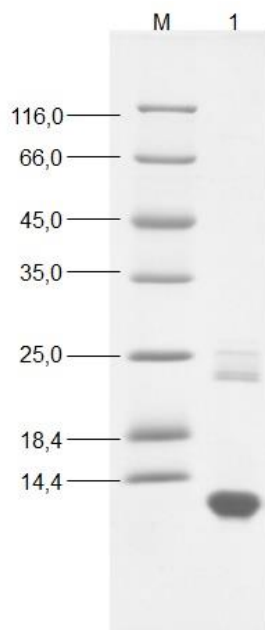


Figura 39: Degradação da proteína rSm21.7 visualizada por SDS-PAGE. Após a purificação da proteína rSm21.7, a mesma sofre degradação ao longo do tempo, como demonstrado em SDS-PAGE. Pode-se visualizar: M – marcador de massa molecular (*Unstained Protein Molecular Weight Marker* – Fermentas), 1 – proteína rSm21.7 purificada e mantida a 4°C por 20 dias.

Uma vez observada a degradação da rSm21.7, o inibidor de protease PMSF foi adicionado em todos os tampões de lise e purificação. Então, o comportamento da proteína foi avaliado. Observou-se que essa degradação ocorre mesmo na presença do PMSF. Entretanto essa degradação ocorre após um tempo mais longo quando comparado ao tempo necessário na ausência do inibidor. Todos os resultados apresentados até o item 5.5 foram obtidos com a proteína rSm21.7 purificada na presença de PMSF.

Tendo em vista que a proteína rSm21.7 perde uma parte de sua cadeia polipeptídica com o tempo e que apenas uma parte da proteína foi cristalizada, foram feitos ensaios de proteólise limitada para verificar a presença de domínios suscetíveis e resistentes à ação de proteases. Na proteólise limitada ocorre a utilização de uma temperatura que não é a ótima para a protease permitindo, assim, a conservação de domínios bem enovelados.

Nesse experimento, foram avaliadas as seguintes proteases: tripsina, quimotripsina, pepsina, proteinase K e subtilisina, além dos diferentes tempos de incubação. Para a maioria das proteases foi visualizada por SDS-PAGE uma banda resistente à ação das mesmas com massa molecular inferior a 14 kDa (Figura 40).

A figura 40 corresponde ao ensaio de proteólise limitada utilizando-se tripsina, que representa os melhores resultados obtidos. Pode-se observar que já no tempo inicial de incubação com a protease, que corresponde a apenas alguns segundos, houve a clivagem de um fragmento da proteína, resultando em outro fragmento com massa menor que 14 kDa (canaleta 2 da figura 40). Essa cadeia polipeptídica resultante foi resistente à ação da tripsina até o tempo de 4 horas de incubação (canaletas 3 a 9 da figura 40).

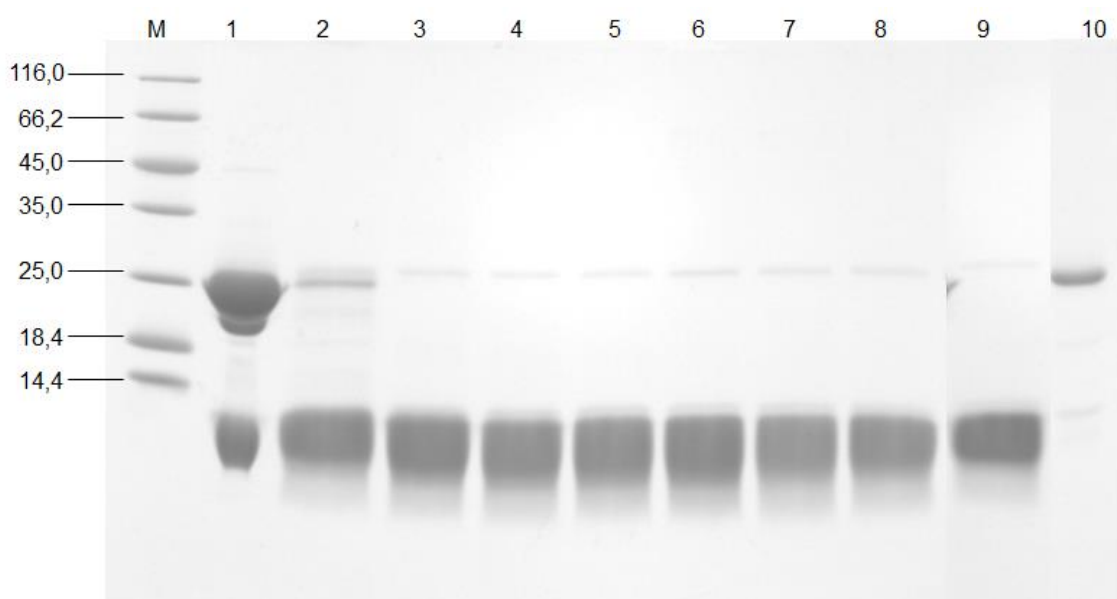


Figura 40: Análise da proteólise limitada da proteína rSm21.7 por tripsina
A proteína rSm21.7 purificada foi submetida a um ensaio de proteólise limitada. A tripsina foi uma das proteases utilizadas nesse ensaio, que foi avaliado por SDS-PAGE, no qual pode-se visualizar: M – marcador de massa molecular em kDa (*Unstained Protein Molecular Weight Marker* – Fermentas), 1 – proteína rSm21.7 não clivada pela tripsina, 2 – tempo 0, 3 – 5 minutos, 4 – 10 minutos, 5 – 30 minutos, 6 – 1 hora, 7 – 2 horas, 8 – 3 horas, 9 – 4 horas, 10 – controle (tripsina).

Uma vez que apenas a porção C-terminal da proteína rSm21.7 foi cristalizada e que a proteína possui uma cauda de histidinas N-terminal, foi realizado um ensaio de *Western Blot* com anticorpo anti-histidinas avaliando as mesmas amostras da proteólise limitada com tripsina. O objetivo desse teste foi verificar se a parte degradada pela ação da protease constituiu-se na parte N-terminal ou C-terminal.

Foi observado que as cadeias polipeptídicas resultantes da clivagem com tripsina não possuíam cauda de histidinas, demonstrando que a parte N-terminal da proteína rSm21.7 foi clivada (resultados não mostrados).

Geralmente, os locais nos quais as proteases atuam são segmentos polipeptídicos expostos flexíveis, como por exemplo, os que ligam dois domínios compactos, enquanto as regiões bem enoveladas permanecem resistentes à ação das mesmas (NOVOTNÝ; BRUCCOLERI, 1987).

Dados de modelagem molecular da proteína FhCaBP4 de *F. hepatica* demonstraram que a mesma possui dois domínios compactos e bem enovelados e um segmento que une os mesmos (Figura 7) (ORR *et al.*, 2012). Assim, caso esse tipo de segmento exista na rSm21.7, seria o local mais provável de ocorrência da proteólise limitada.

Dado o fato de que não foi visualizado no *Western Blot* nenhum fragmento de massa similar ao que seria o domínio N-terminal de mão EF com uma cauda de histidinas (12.080 Da = 23.689 Da (proteína completa) – 11.609 Da (domínio de cadeia leve de dineína)), é possível que esta porção tenha: (i) perdido apenas a cauda de histidinas ficando com uma massa semelhante ao domínio de cadeia leve de dineína impossibilitando distingui-las no gel ou, (ii) sido clivada em fragmentos menores que não podem ser visualizados no gel que foi utilizado.

Na literatura, foram encontrados alguns dados que sugerem a ocorrência de um processamento semelhante no próprio parasita *Schistosoma*. Análises de proteômica demonstraram um *spot* em gel 2D (Figura 41A) (CURWEN *et al.*, 2004) e uma banda em gel 1D (KNUDSEN *et al.*, 2005) (Figura 41B) com massa molecular menor que a esperada para a proteína Sm21.7 na comparação com o padrão de massa molecular apresentado nesses trabalhos. Não há nenhuma discussão nesses artigos com relação à massa obtida dos *spots*/bandas, mas as figuras dos géis evidenciam uma clivagem e/ou degradação dessa proteína nos extratos do parasita utilizados.

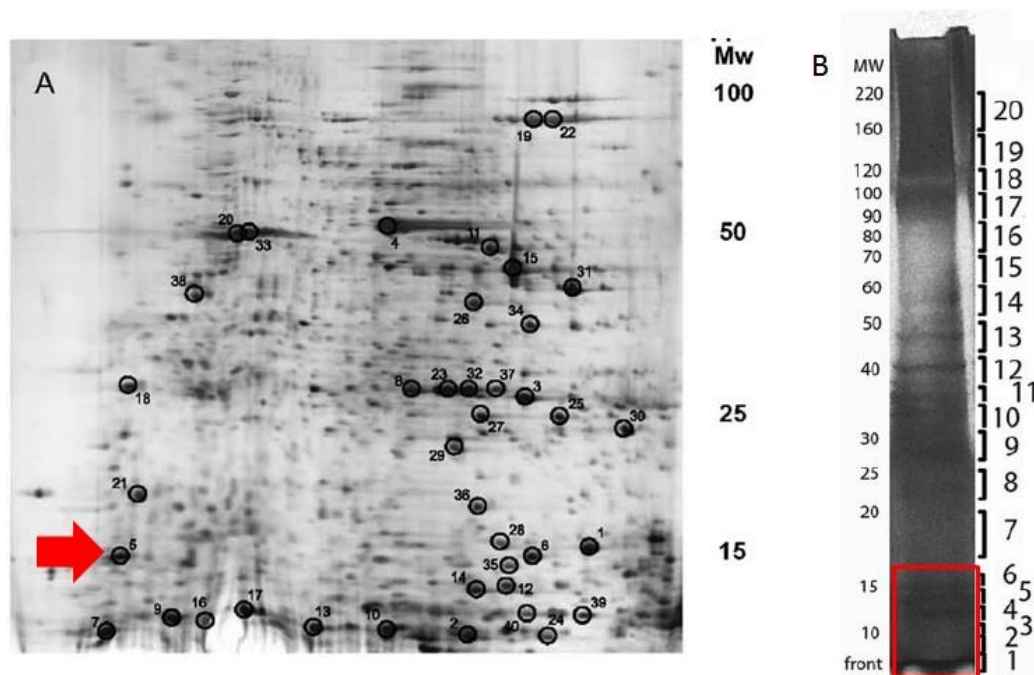


Figura 41: Géis de eletroforese bi e uni-dimensionais de extratos de parasitas *S. mansoni*. A separação eletroforética de extratos do parasita evidencia um processamento da proteína Sm21.7, uma vez que a massa da mesma apresenta um valor menor que o esperado. O spot correspondente à Sm21.7 está destacado pela seta vermelha em (A) com aproximadamente 15 kDa (CURWEN et al., 2004) e a região de onde foi isolada a proteína Sm21.7 está destacado em vermelho (B) com massas entre 10 e 15 kDa (KNUDSEN et al., 2005).

Novamente considerando que a proteína rSm21.7 sofre degradação com o tempo e que apenas o domínio de cadeia leve de dineína foi cristalizado, o raio hidrodinâmico e a polidispersividade da amostra foram analisados durante alguns dias após a purificação.

Foram realizados, então, ensaios de DLS ao longo do tempo com a proteína purificada na presença de PMSF e na ausência do mesmo. Observou-se que na presença de PMSF a amostra começa a ser clivada após aproximadamente 20 dias passados da purificação. Já na amostra que não possui PMSF observou-se que após 5 dias a proteína encontra-se clivada. Essa clivagem foi observada pela redução do raio hidrodinâmico e da massa molecular calculada no experimento de DLS (Tabela 4).

Nesses ensaios ainda foi observado que rSm21.7, já com seu raio hidrodinâmico reduzido, ainda encontra-se monodispersa, indicando que o fragmento de aproximadamente 2,4 nm de raio e 26,7 kDa de massa molecular é uma amostra de boa qualidade para cristalização. Este último resultado sugere ainda que, nestas condições

(clivada), a proteína se apresenta na forma de um dímero (2 x 11,6 kDa ~ 26 kDa) e que parece ser, então, o DLC o domínio responsável pela oligomerização de rSm21.7.

Tabela 4: Dados de DLS ao longo do tempo da proteína rSm21.7 com e sem PMSF. A amostra de proteína purificada foi analisada ao longo do tempo por DLS para acompanhar a degradação da mesma. Após 23 dias da purificação houve uma redução no raio hidrodinâmico na amostra com PMSF. Já na amostra sem PMSF essa redução já se inicia com 5 dias.

Amostra	Tempo	Diâmetro (nm)	Pd (%)	MM (kDa)	Intensidade (%)	Massa (%)
com PMSF	Imediatamente após	7,0 +/- 0,8	11,6	63,4 +/- 7,3	100	100
	1 dia	6,9 +/- 1,1	15,8	60,5 +/- 9,6	100	100
	13 dias	6,9 +/- 0,1	13,8	61,1 +/- 8,5	95	100
	23 dias	4,849	20,6	26,7	79,3	100
sem PMSF	Imediatamente após	6,6 +/- 0,7	10,5	55,5 +/- 5,8	100	100
	5 dias	5,8 +/- 0,3	5,8	40,1 +/- 2,3	57,1	100

5.7 – Determinação da estrutura tridimensional da proteína rSm21.7

Após a coleta dos dados de difração dos cristais nativo e derivado, as fases dos fatores de estrutura foram calculadas e um modelo inicial da proteína rSm21.7 pode ser construído. Esse modelo passou por ciclos manuais e automáticos de refinamento e a cada ciclo foi observada a redução do R_{work} com um valor final de 18,1% (R_{free} de 21,8). As estatísticas do refinamento podem ser visualizadas na tabela 5.

Tabela 5: Estatísticas de refinamento da estrutura tridimensional da proteína rSm21.7. Após os ciclos de refinamento a estrutura foi validada e as estatísticas finais estão mostradas.

R _{work} ^a (%)	18,1
R _{free} ^b (%)	21,8
N ^o de:	
átomos de proteína	798
moléculas de água	66
moléculas de etilenoglicol	8
íons SO ₄ ⁻²	5
R.m.s.d:	
Comprimento de ligação(Å)	0,003
Ângulo de ligação(°)	0,8
Gráfico de Ramachandran:	
Favoráveis(%)	96.81
Permitidos(%)	3,19

$$^a R_{work} = \sum_{hkl} |F_{obs}| - |F_{calc}| / \sum_{hkl} |F_{obs}|$$

^b R_{free} é o valor de R_{work} calculado para 5% do conjunto de dados não inclusos no refinamento

Após a construção do modelo, observou-se que apenas o domínio de cadeia leve de dineína foi cristalizado. Esse domínio possui 2 α -hélices e 4 fitas- β , que formam uma folha- β antiparalela. A folha- β possui em um de seus lados as α -hélices e do outro lado um *loop* conectando as fitas- β 2 e 3. A extremidade N-terminal possui um resíduo glutamato e a C-terminal possui uma asparagina (Figura 42).

Após todos os ciclos de refinamento, a estrutura foi validada no programa ADIT e demonstrou que o modelo possui os parâmetros adequados para posterior depósito no PDB.

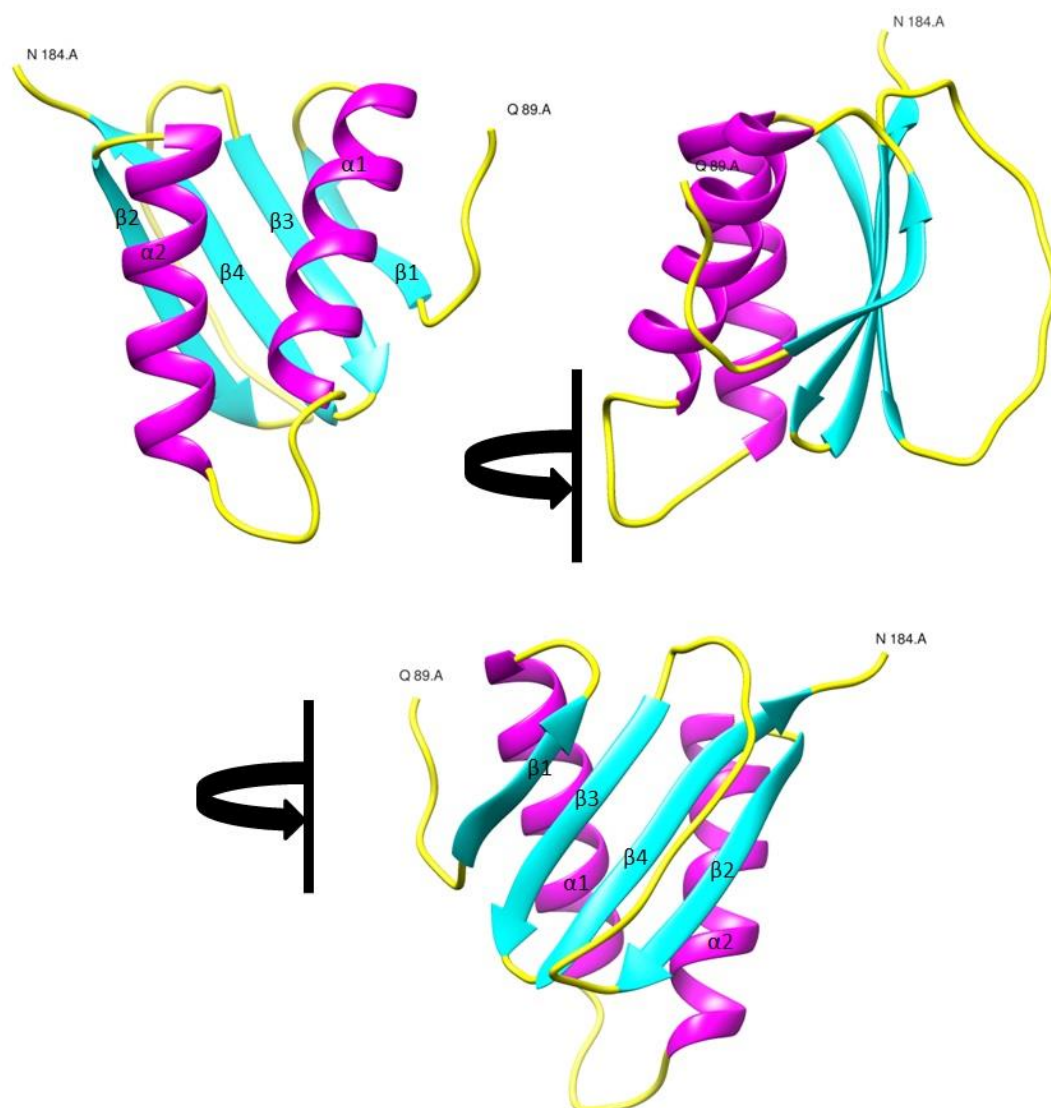


Figura 42: Estrutura tridimensional do domínio de DLC da proteína rSm21.7. Cristais da proteína rSm21.7 foram utilizados em experimentos de difração de raios X e, posteriormente, a estrutura do domínio de cadeia leve de dineína da proteína foi determinada. A imagem foi gerada pelo programa UCSF Chimera (PETTERSEN et al., 2004).

Durante as tentativas de faseamento por Substituição Molecular, uma procura através de *Blast* indicou que, dentre as proteínas depositadas no PDB, aquela com maior identidade de sequência (24%) em relação ao domínio de cadeia leve de dineína de Sm21.7 foi a proteína Hsv2p de *Kluyveromyces lactis* (PDB - 4EXV e 4AV9) (BASKARAN et al., 2012; KRICK et al., 2012).

Entretanto, uma análise no DALI server (http://ekhidna.biocenter.helsinki.fi/dali_server- (HOLM; ROSENSTROM, 2010)) após

a construção do modelo indicou um alinhamento estrutural com 19% de identidade entre o domínio de cadeia leve de dineína de rSm21.7 e uma DLC citoplasmática (Dyn2) de *Saccharomyces cerevisiae* (PDB 4HT6 - (RAO, L. *et al.*, 2013)). Essa foi a maior similaridade estrutural dentre as proteínas depositadas no PDB.

De fato, assim como o domínio de cadeia leve de dineína de rSm21.7, Dyn2 também possui 2 α -hélices e quatro fitas- β que formam uma folha- β antiparalela (Figura 43A) (RAO, L. *et al.*, 2013). O RMSD calculado a partir da sobreposição de 84 carbonos alfa das duas estruturas é de apenas 1,4 Å, indicando que, de uma forma geral, o enovelamento é o mesmo entre estas estruturas, como pode ser observado na figura 43C.

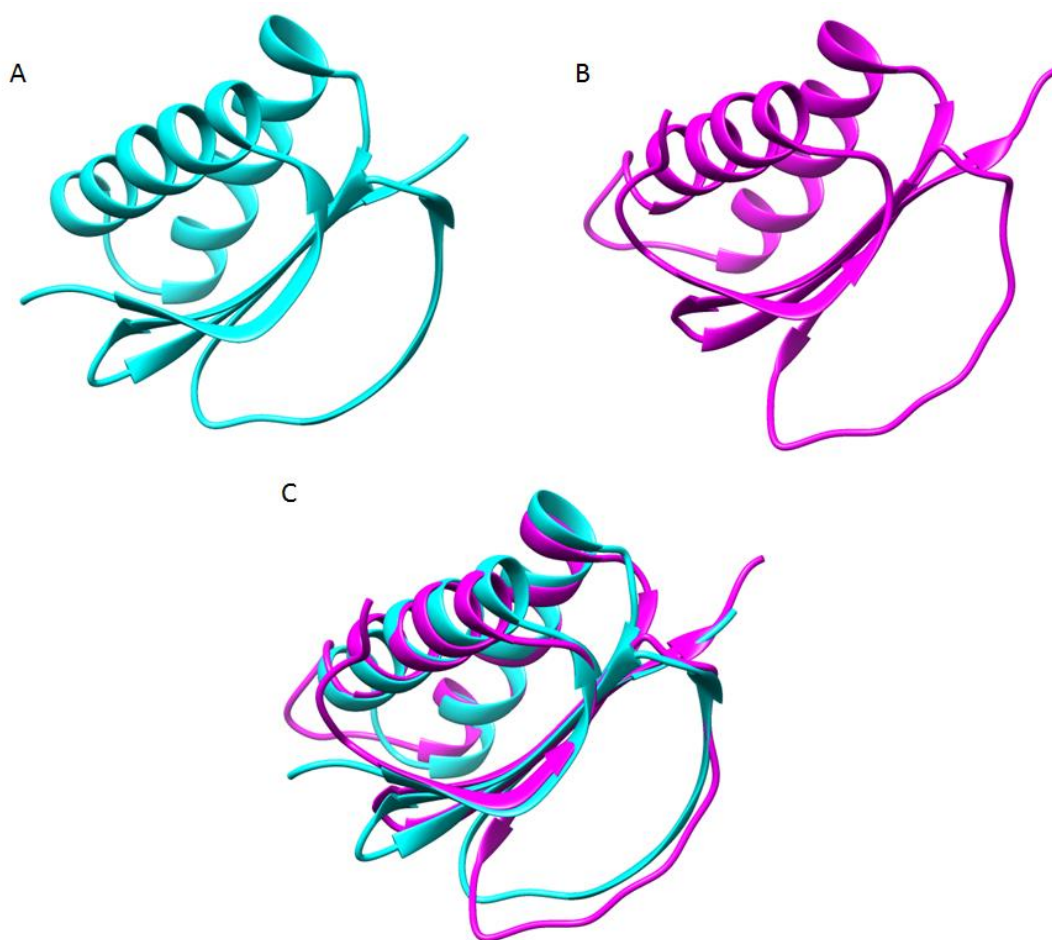


Figura 43: Sobreposição das estruturas do DLC da proteína rSm21.7 (magenta) e de Dyn2 de *S. cerevisiae* (PDB – 4HT6) (ciano). (A) Estrutura de Dyn 2. (B) Estrutura do domínio de cadeia leve de dineína da proteína rSm21.7. (C) Sobreposição das estruturas A e B. As estruturas possuem enovelamento semelhante, como indicado pelo RMSD de 1,4 Å, sendo ambas compostas por 2 α -hélices e fitas- β , que formam uma folha- β antiparalela. A imagem foi gerada pelo programa UCSF Chimera (PETTERSEN *et al.*, 2004).

Apesar da similaridade observada entre estes dois domínios é possível destacar algumas regiões onde existe uma maior diferença estrutural entre os domínios, com valores de RMSD elevados. Como mostrado na figura 44, as regiões com maiores valores de RMSD correspondem às regiões de *loop*.

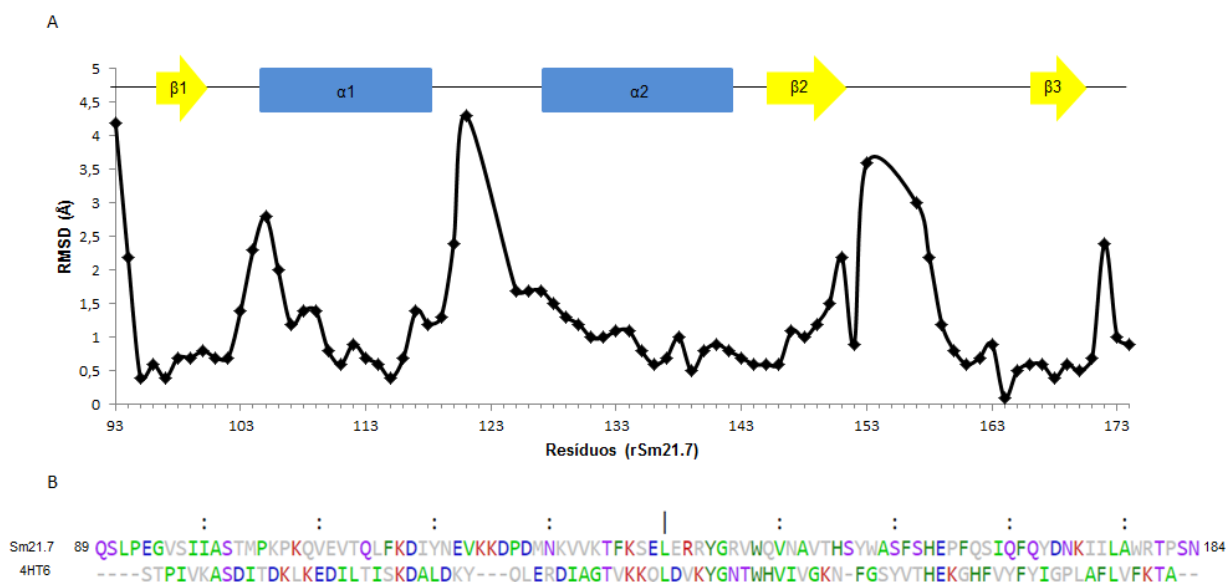


Figura 44: Gráfico RMSD para todos os aminoácidos em relação à estrutura secundária. Nesse gráfico fica evidente que os valores de RMSD mais elevados correspondem às regiões principalmente de *loop*. (A) Nessa imagem as estruturas em α -hélice destacam-se como barras azuis. Já as fitas- β destacam-se com setas amarelas. No gráfico observa-se os valores de RMSD dos resíduos do domínio de cadeia leve de dineína de rSm21.7 quando alinhados com a sequência de Dyn2 de *S. cerevisiae* (PDB 4HT6). (B) Alinhamento das sequências dos domínios de cadeia leve de dineína de rSm21.7 e Dyn2.

Uma análise de superfície da estrutura da proteína rSm21.7 realizada pelo programa PISA ((KRISSINEL; HENRICK, 2007) <http://www.ebi.ac.uk/pdbe/pisa/>), sugere a existência de um dímero cristalográfico, com uma área de contato de aproximadamente $1.056 \text{ \AA}^2$. De fato, apesar da observação de um monômero de rSm21.7 por unidade assimétrica, pode-se gerar este dímero pela aplicação de uma simetria cristalográfica (-y+1, -x+1, -z+5/6).

Assim, tendo em vista: (i) que é possível a visualização de um dímero na estrutura cristalográfica, (ii) que foi visualizada através de DLS uma estrutura dimérica em solução mesmo após a clivagem da parte N-terminal da proteína rSm21.7 e (iii) que domínios de cadeia leve de dineína podem ser encontrados na forma dimérica

(BANFORD *et al.*, 2013; FAN, JING-SONG *et al.*, 2001; LIANG *et al.*, 1999), algumas novas tentativas de cristalização da proteína foram realizadas no NSLS a fim de se obter uma estrutura cristalográfica que apresentasse um homodímero de rSm21.7 na unidade assimétrica com maior resolução que os dados iniciais.

Esses ensaios forneceram condições de cristalização a partir de uma solução inicial contendo PEG 3.350 e formiato de potássio. Após coleta de dados de difração e processamento dos dados (HKL-2000 - (OTWINOWSKI; MINOR, 1997)), observou-se que os cristais pertenciam a outros sistemas cristalinos: ortorrômbico ($P2_12_12_1$) (Tabela 7) e trigonal ($P1$) (resultados não mostrados). As estatísticas de coleta e processamento de um dos conjuntos de dados são apresentadas na tabela 6.

Após o processamento, o coeficiente de Mathews (MATTHEWS, 1968) foi estimado em $2,4 \text{ \AA}^3\text{Da}^{-1}$, com um conteúdo de solvente de 48,2%. Esse dado foi obtido assumindo a presença de duas moléculas de proteína com 11.277 Da cada por unidade assimétrica do cristal.

Tabela 6: Estatísticas de coleta de dados e processamento do cristal nativo coletado no NSLS. Os dados entre parênteses indicam os valores da menor e maior resolução, respectivamente.

Comprimento de onda (Å)	1,100
Temperatura (K)	100
Distância cristal-detector (mm)	200
Rotação por imagem (°)	0,5
Tempo de exposição por imagem (s)	0,5
Variação total de rotação (°)	360
Grupo espacial	P2 ₁ 2 ₁ 2 ₁
Parâmetros de célula unitária (Å)	a=56,01, b=59,28, c=64,54
Resolução (Å)	50,00-1,5 (50,00-4,07/1,53-1,5)
Nº de observações	398.859
Nº. de reflexões únicas	35.067 (1.916/1.732)
Completeza (%)	100 (99,9/99,5)
$\langle I/\sigma(I) \rangle$	50,2 (72,0/2,9)
Multiplicidade	11,4 (11,7/7,4)
R _{merge} ^a (%)	6,7 (5,3/47,5)
Mosaicidade (°)	0,72
Coeficiente de Matthews (Å ³ /Da) ^b	2,4
Conteúdo de solvente (%) ^b	48,2

$$^a R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)} \times 100$$

^b Valores calculados assumindo a presença de 192 (2 x 96) resíduos de aminoácidos por unidade assimétrica

Assim, após determinação dessa estrutura por Substituição Molecular, foi observado que o domínio de cadeia leve de dineína de rSm21.7 é capaz de formar um homodímero através da face que contém a folha-β (figura 45). Observa-se que neste dímero, o *loop* que conecta as fitas β2 e β3 agora adota a conformação de uma fita-β, assim como o DLC apresentado na figura 14. Para uma melhor visualização, na figura 46, estas duas novas fitas-β no dímero são apresentadas em cores destacadas. Mesmo com o refinamento desta estrutura incompleto, as estatísticas são apresentadas na tabela 7:

Tabela 7: Estatísticas de refinamento da estrutura tridimensional da proteína rSm21.7. O refinamento desse conjunto de dados ainda não foi concluído.

R_{work}^a (%)	18,2
R_{free} (%)	21,9
Nº de:	
átomos de proteína	1596
moléculas de água	244
R.m.s.d:	
Comprimento de ligação(Å)	0,04
Ângulo de ligação(º)	3,2
Gráfico de Ramachandran:	
Favoráveis(%)	90,7
Permitidos(%)	8,8
Não permitidos(%)	0,6

$$^a R_{work} = \sum_{hkl} |F_{obs}| - |F_{calc}| / \sum_{hkl} |F_{obs}|$$

Como já descrito, domínios de cadeia leve de dineína podem ser encontrados como dímeros e são extensamente discutidos na literatura (FAN, JING-SONG *et al.*, 2001; LIANG *et al.*, 1999; ROMES; TRIPATHY; SLEP, 2012).

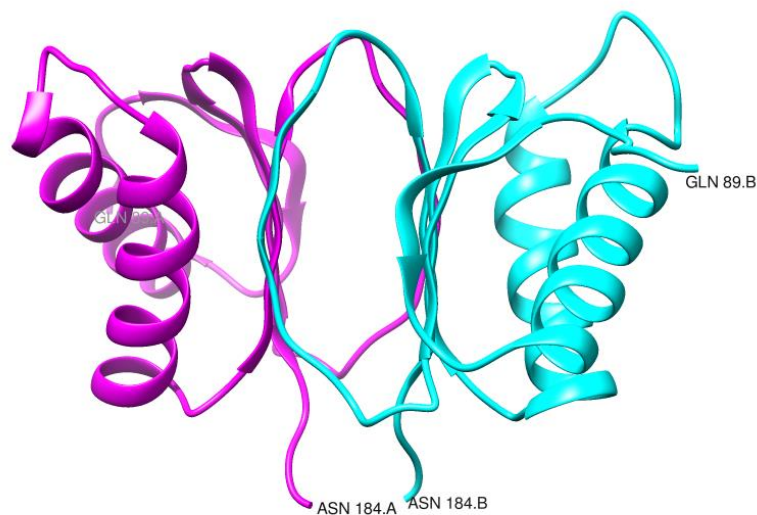


Figura 45: Estrutura tridimensional do homodímero de rSm21.7. Dados de difração de raios X dos cristais obtidos no NSLS foram coletados e a estrutura tridimensional foi determinada. Na imagem cada monômero é destacado em ciano e magenta. Observa-se que a face da molécula que contém as fitas- β formam a interface do dímero. A imagem foi gerada pelo programa UCSF Chimera (PETTERSEN *et al.*, 2004).

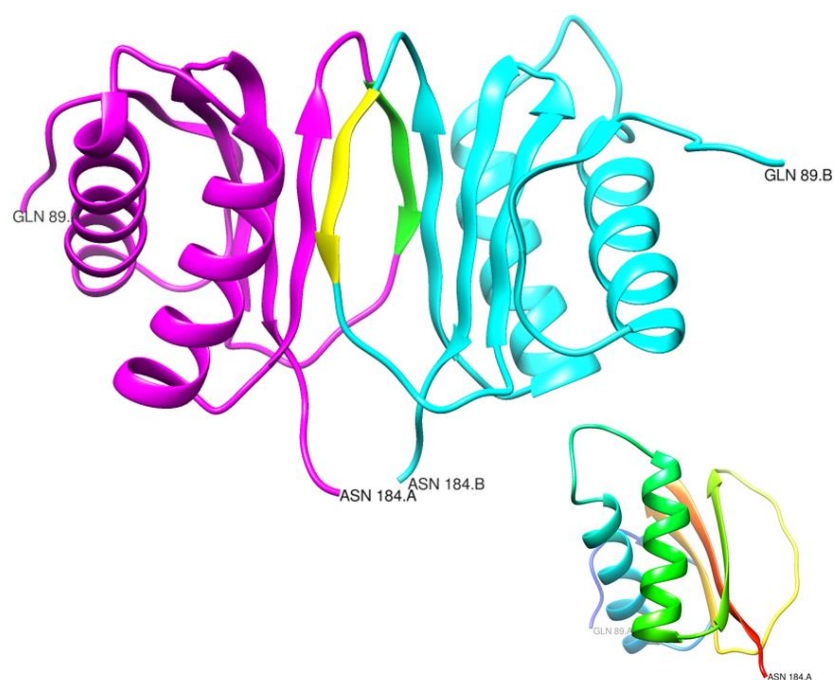


Figura 46: Fitas- β da interface do homodímero de rSm21.7. Nesta imagem são destacadas as fitas- β formadas pela dimerização da estrutura. Na figura menor observa-se o monômero da proteína rSm21.7 e a ausência da fita- β no *loop*. As imagens foram geradas pelo programa UCSF Chimera (PETTERSEN et al., 2004).

As fitas- β que estabilizam o dímero da proteína rSm21.7 interagem através dos aminoácidos 148, 150, 152, 154, 156 e 158 (em ambos os monômeros). As ligações de hidrogênio entre a amina de um resíduo e a carbonila de outro, com aproximadamente 3 Å de distância, podem ser visualizadas na figura 47.

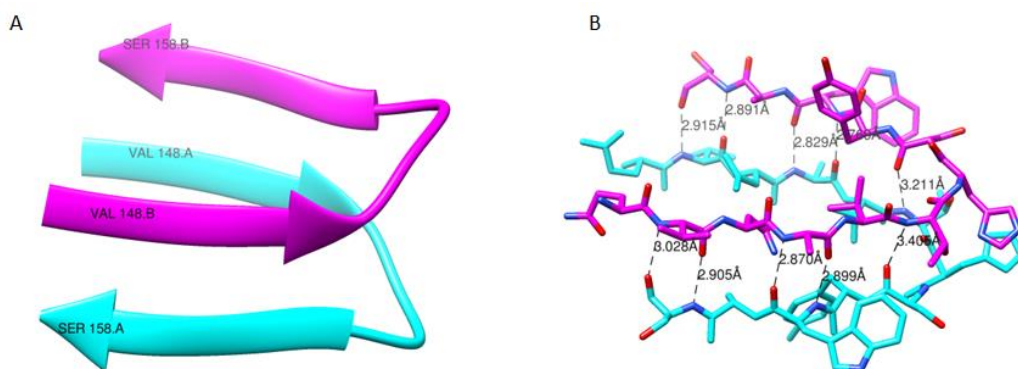


Figura 47: Interações entre os monômeros de rSm21.7. (A) A imagem mostra as fitas- β que interagem com o monômero oposto (Val148 a Ser158). (B) Ligações de hidrogênio entre grupos amina e carbonila das cadeias principais com as distâncias medidas. As imagens foram geradas pelo programa UCSF Chimera (PETTERSEN et al., 2004).

Uma análise quantitativa da estrutura cristalográfica do dímero DLC de rSm21.7 com os dados de CD obtidos para a proteína completa (200 aminoácidos) permite uma estimativa do conteúdo de elementos de estrutura secundária do domínio de mão EF (Tabela 8). Sabendo, pelos resultados de CD, que a proteína completa (200 aminoácidos) tem, em solução, uma porcentagem de α -hélices de 37% e que dos 96 resíduos C-terminais, 31 deles (32%) adotam esse tipo de estrutura secundária, então é possível estabelecer a seguinte relação:

$$\frac{31 \text{ (aa no DLC)} + X \text{ (aa no domínio mão EF)}}{200 \text{ (aa no total)}} = 0,37 \text{ (dado obtido no CD)}.$$

Esta relação sugere então que 43 resíduos de aminoácidos estariam em α -hélices no domínio de mão EF de rSm21.7. Por cálculos similares, o número de resíduos de aminoácidos em fitas- β e *random coil* neste domínio também foram estimados e são apresentados na tabela 8.

Tabela 8: Análise dos elementos de estrutura secundária da proteína rSm21.7. Nesta tabela foram utilizados os dados do dímero cristalográfico do DLC de rSm21.7 e os dados de CD da forma completa de rSm21.7 para o cálculo dos elementos de estrutura secundária do domínio de mão EF de rSm21.7.

	DLC cristalográfico	Dicroísmo Circular	Domínio mão EF
	# de resíduos em 96	# de resíduos em 200	# de resíduos em 104
α -hélice	31/96 = 32%	37% = 74/200	43/104 = 41%
fitas- β	28/96 = 29%	15% = 30/200	2/104 = 2%
<i>random coil</i> + volta β	37/96 = 39%	48% = 96/200	59/104 = 57%

De fato, os domínios de mão EF são constituídos de α -hélices e *loops*, como demonstrado nas figuras 9, 10 e 11 e como é extensamente apresentado na literatura (KRETSINGER; NOCKOLDS, 1973; STRYNADKA; JAMES, 1989; YAP, K. L. *et al.*, 1999). Das estruturas citadas na figura 11, a proteína S100 (PDB 1QLK) possui aproximadamente 58% de α -hélices e 42% de estruturas em *loop* e/ou *random coil*; a calpaína (PDB 1DVI) possui 47% de α -hélices, 3% de fitas- β e 50% de *loop* e/ou *random coil* enquanto a neurocalcina (PDB 1BJF) possui 55% de α -hélices, 5% de fitas- β e 40% de *loop* e/ou *random coil*. Assim, os cálculos que sugerem 41% de α -hélices no domínio de mão EF de rSm21.7 são bastante plausíveis.

Mesmo que o domínio de mão EF de rSm21.7 apresente estruturas helicoidais, como sugerido pelas análises de CD e pelos cálculos apresentados acima, não foi demonstrada a propriedade de ligação ao cálcio por essa proteína (FRANCIS; BICKLE, 1992). Como consequência, a estabilidade desse domínio poderia estar comprometida, facilitando a ação das proteases que o degradaram nos testes de proteólise limitada.

Novamente discutindo a dimerização, análises da superfície eletrostática da proteína rSm21.7 demonstraram que, para formação do dímero, as interfaces que interagem precisam possuir cargas opostas, possibilitando a interação entre os monômeros. Pode-se observar na figura 48 que a região do monômero apresentado em A que possui cargas negativas (inferior à esquerda), interage com a parte do monômero B que possui cargas positivas (inferior à direita); e vice versa. Deve-se ressaltar ainda que as análises de superfície do homodímero geradas pelo programa PISA ((KRISSINEL; HENRICK, 2007) <http://www.ebi.ac.uk/pdbe/pisa/>), sugerem ser este o dímero biológico, com uma área de contato de aproximadamente 1.100 Å².

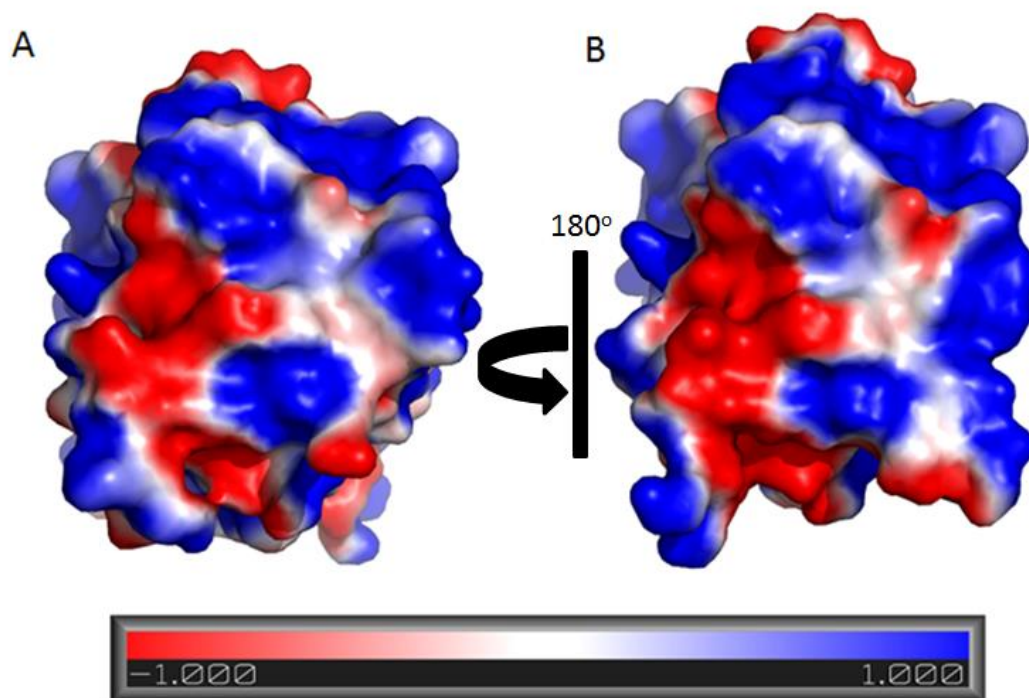


Figura 48: Análise da superfície eletrostática da interface do dímero de rSm21.7. Cada um dos monômeros, apresentado em A e B, respectivamente, foi posicionado de modo a apresentar sua superfície de contato. A cor azul representa carga positiva, enquanto a vermelha representa carga negativa e branca neutra. Pode-se observar que na região de interface dos dímeros as moléculas apresentam cargas opostas. O mapa da superfície foi obtido utilizando a ferramenta APBS no PYMOL. As imagens foram geradas no PYMOL (DELANO, W. L., 2007).

Em comparações entre as superfícies eletrostáticas dos dímeros de rSm21.7 e de Dyn2 de *S. cerevisiae* (PDB 4HT6), observa-se também algumas diferenças de carga (e forma), mesmo sendo o enovelamento de ambas semelhante (Figura 49). Nessa figura pode-se observar que na proteína Dyn2 existe uma fenda de maior profundidade (destacada pela seta vermelha) que é responsável pela ligação de um peptídeo (RAO, L. *et al.*, 2013). De fato, algumas cadeias leves de dineína já foram descritas na forma de complexos com peptídeos e proteínas, entretanto, as sequências destes últimos não apresentam similaridade óbvia, mas ligam-se a uma fenda similar à apresentada na figura 49 (LO *et al.*, 2001; RAO, L. *et al.*, 2013; ROMES; TRIPATHY; SLEP, 2012). Já na região equivalente a da fenda no DLC de rSm21.7, observa-se uma pequena fenda (Figura 49), sugerindo que Sm21.7 poderia interagir com outra molécula com especificidade diferente daquela observada para Dyn2.

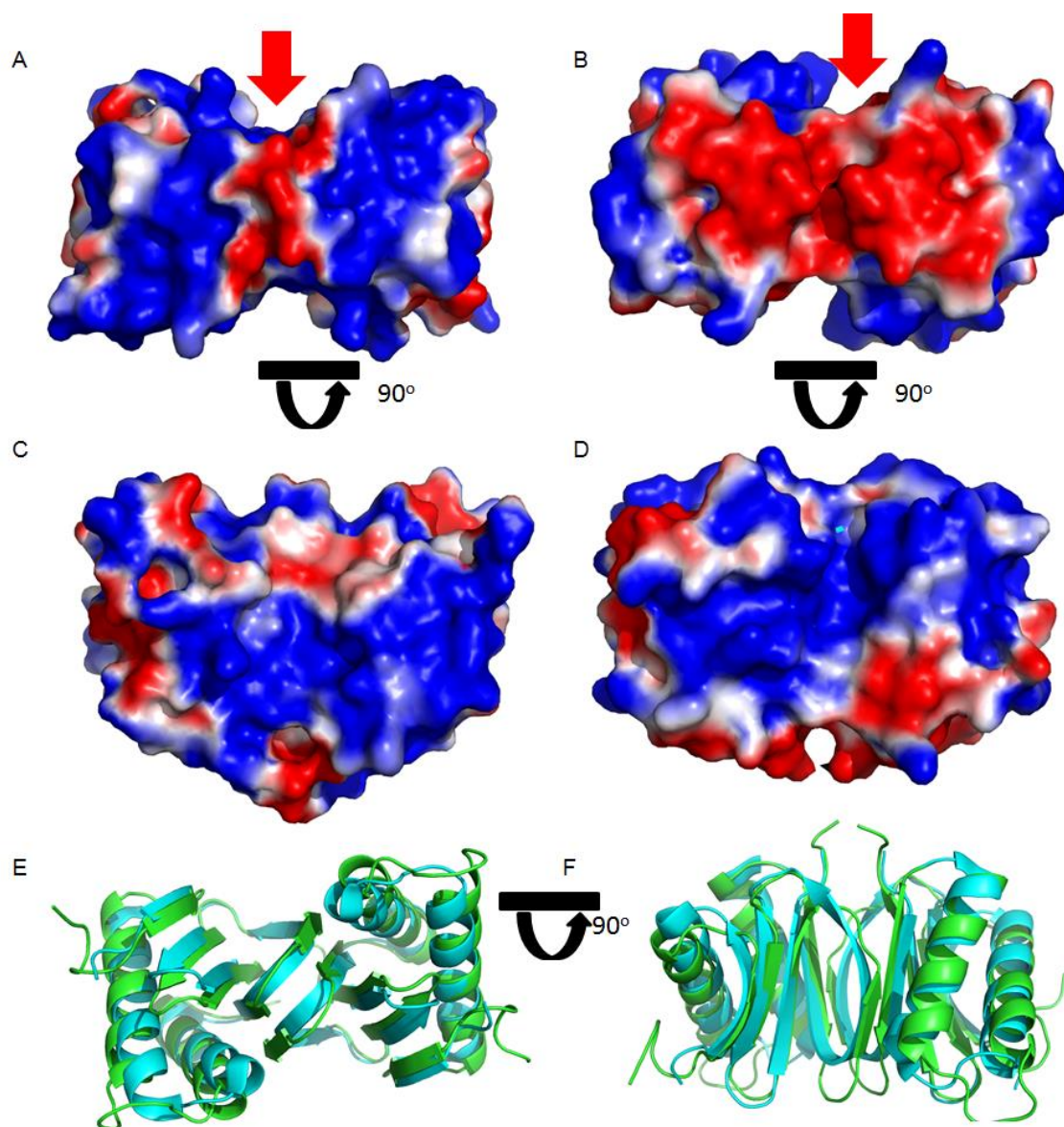


Figura 49: Análise de superfície eletrostática das proteínas rSm21.7 e Dyn2 (PDB 4HT6). (A) Superfície eletrostática da proteína rSm21.7. (B) Superfície eletrostática da proteína Dyn 2 (em orientação equivalente à imagem anterior). (C) Imagem A girada de 90°. (D) Imagem B girada de 90°. (E) Sobreposição de ambas as estruturas (Dyn2 em ciano e rSm21.7 em verde) na mesma orientação das imagens A e B. (F) Sobreposição de ambas as estruturas (Dyn2 em ciano e rSm21.7 em verde) na mesma orientação das imagens C e D. Os mapas das superfícies foram obtidos utilizando a ferramenta APBS no PYMOL. A imagem foi gerada no programa PYMOL (DELANO, W. L., 2007).

5.8 – Localização de epítomos reativos na estrutura do domínio de cadeia leve de dineína da proteína rSm21.7

Utilizando um programa de predição de epítomos (Bcepred - <http://www.imtech.res.in/raghava/bcepred/>), 16 sequências peptídicas contidas na proteína Sm21.7 foram sintetizadas em uma membrana de *spot* (Figura 50). Essa membrana foi incubada com soro de camundongos imunizados com a proteína recombinante rP22 (Sm21.7) expressa na construção pETDEST42 e alguns peptídeos reagiram de forma mais forte que outros (Rezende et al, em preparação). Tais ensaios foram realizados por Rezende e colaboradores no Laboratório de Imunologia Celular e Molecular do Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, UFMG.

Dos peptídeos mais reativos sintetizados na membrana (numerados como 1, 2, 3, 6, 10, 11, 13, 14 e 16 na figura 50), quase todos se encontravam no domínio de cadeia leve de dineína (Figura 51). Apenas uma pequena porção C-terminal do motivo de mão EF foi localizada em um peptídeo reativo (Rezende et al, em preparação).

```
MDSPMEKFIQTYLTLLRDGDETVETSKLSESCRKEKLDMDKQVNEWIALFDVDKDQKITFEEFC - Sm21.7 (63)
EKFQTYLTLLRDGDE - Peptídeo 7
      DETVETSKLSESCRKE - Peptídeo 4
            SESCOKEKLDMDKQVNE - Peptídeo 9
                  VNEWIALFDVDKDQKI - Peptídeo 8
                        KDQKITFEEFC - Peptídeo 15

RGLGLKQNEMRIERNHIKTVQSGREQSLPEGVSIIASTMPKPKQVEVTQLFKDIYNEVKKDPD - Sm21.7 (126)
RGLGL - Peptídeo 15
RGLGLKQNEMRIERNH - Peptídeo 1
      NEMRIERNHIKTVQSG - Peptídeo 14
            RNHIKTVQSGREQSLP - Peptídeo 13
                  SGREQSLPEGVSIIAS - Peptídeo 5
                        GVSIIASTMPKPKQVE - Peptídeo 3
                              PKQVEVTQLFKDIYNE - Peptídeo 16
                                    FKDIYNEVKKDPD - Peptídeo 2

MNKVVKTFKSELERRYGRVWQVNAVTHSYWASFSHEPFQSIQFQYDNKIILAWRTPSN - Sm21.7 (184)
MNK - Peptídeo 2
      KSELERRYGRVWQVNA - Peptídeo 6
            VNAVTHSYWASFSHEP - Peptídeo 12
                  YWASFSHEPFQSIQFQ - Peptídeo 11
                        FQSIQFQYDNKIILAW - Peptídeo 10
```

Figura 50: Sequência da proteína Sm21.7 contendo os peptídeos sintetizados na membrana e avaliados pela reatividade. Utilizando o programa Bcepred (<http://www.imtech.res.in/raghava/bcepred/>), 16 epítomos foram preditos e sintetizados em uma membrana de *spot*. Essa membrana foi, posteriormente, avaliada quanto à reatividade dos epítomos contra soro de camundongos imunizados com a proteína rP22 (Sm21.7).

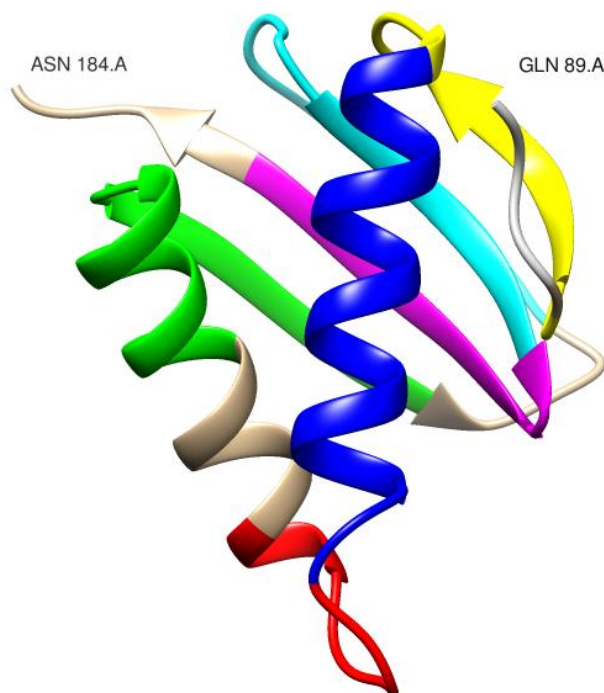


Figura 51: Peptídeos reativos localizados na estrutura do domínio de cadeia leve de dineína da proteína rSm21.7. Cada um dos peptídeos reativos presentes no domínio de cadeia leve de dineína foi destacado na estrutura de acordo com as cores: Peptídeo 2: vermelho, Peptídeo 3: amarelo, Peptídeo 6: verde, Peptídeo 10: magenta, Peptídeo 11: ciano, Peptídeo 13: cinza e Peptídeo 16: azul. Como se pode verificar na figura 50 ocorre sobreposição de alguns resíduos em peptídeos diferentes. A imagem foi gerada pelo programa UCSF Chimera (PETTERSEN et al., 2004).

O fato de quase todos os peptídeos reativos estarem localizados no domínio de cadeia leve de dineína pode ser um indicativo de que a clivagem da proteína ocorreu também no processo de imunização e que apenas esse domínio foi responsável por estimular o sistema imune e induzir a proteção dos camundongos vacinados através da redução do granuloma, da fibrose, da carga parasitária e do número de ovos no fígado (REZENDE *et al.*, 2011).

Devido às evidências de que o domínio de mão EF N-terminal de rSm21.7 é suscetível à ação de proteases e que os dados da literatura demonstram que o mesmo não se liga ao cálcio, sugerindo a existência de um domínio não funcional, pode-se sugerir que o domínio de cadeia leve de dineína é a unidade funcional dessa proteína.

Assim como outros domínios semelhantes, o domínio de cadeia leve de dineína de rSm21.7 poderia interagir com algum outro polipeptídeo do hospedeiro mamífero, como forma de adaptação e sobrevivência ou como forma de escape do sistema imune.

6 - Conclusões

- A proteína Sm21.7 foi subclonada em vetor de expressão pET28a(TEV);
- A expressão a 37°C produziu uma grande quantidade de proteína na fração solúvel;
- A purificação por cromatografias de afinidade e exclusão molecular produziu aproximadamente 30 mg de proteína com considerável grau de pureza por litro de expressão;
- Análises de CD demonstraram que a proteína rP22 encontra-se parcialmente enovelada;
- O estado oligomérico da proteína rSm21.7 não pode ser definido por Espalhamento Dinâmico da Luz, mas essa técnica demonstrou que a proteína encontrou-se monodispersa;
- Ensaios de *cross-linking* com glutaraldeído sugerem que a proteína completa adota a forma dimérica em solução;
- A proteína foi cristalizada em algumas condições diferentes, tanto pela técnica da gota suspensa quanto pela técnica da gota depositada;
- Apenas o domínio de cadeia leve de dineína foi cristalizado, uma vez que a proteína sofreu uma clivagem na gota de cristalização;
- A proteína rSm21.7 possui um domínio N-terminal suscetível a ação de proteases e um domínio C-terminal resistente à ação das mesmas;
- O domínio de cadeia leve de dineína é composto por 4 fitas- β , que formam uma folha- β antiparalela e duas α -hélices em uma estrutura semelhante a outras cadeias leves de dineína depositados no PDB;
- Os resultados de CD da proteína completa, junto com a estrutura dimérica do domínio de cadeia leve de dineína de rSm21.7 sugerem que o domínio N-terminal desta proteína seja parcialmente enovelado (com elementos estruturais helicoidais) mas pouco estável (não estabilizado) em solução;
- Os epítomos reativos responsáveis pela proteção fornecida pela imunização com rP22 (Sm21.7) são localizados em sua maioria no domínio de cadeia leve de dineína.

7 – Perspectivas

- Verificar se a degradação da proteína ocorre em extratos do parasita e no verme *in vivo*;
- Realizar ensaios de *cross-linking* apenas com o domínio de cadeia leve de dineína para confirmar sua forma dimérica;
- Realizar testes de ligação ao cálcio utilizando o mutante Q55G;
- Realizar co-imunoprecipitação para verificar se a proteína liga-se a outras moléculas.
- Concluir a escrita do artigo científico internacional que descreve a estrutura tridimensional desta proteína.

8 – Referências Bibliográficas

<http://www.ebi.ac.uk/pdbe/pisa/>
http://ekhidna.biocenter.helsinki.fi/dali_server
<http://www.imtech.res.in/raghava/bcepred/>
<http://www.expasy.org>
<http://validate.rcsb.org/>
<http://tools.neb.com/NEBcutter2/>
<http://www.csun.edu>
<http://www.rcsb.org/pdb/statistics/holdings.do>
<http://www.promega.com>
http://www.who.int/water_sanitation_health/diseases/schisto/en/
<http://www.who.int/schistosomiasis/en/>
<http://www.cdc.gov/parasites/schistosomiasis/biology.html>
<http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>

ADAMS, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D*, v. 66, n. 2, p. 213–221, 2010.

AHMED, H. M. *et al.* DNA Immunization with the Gene Encoding SM21.7 Protein Protects Mice Against *Schistosoma mansoni* Infections. *Journal of American Science*, v. 2, n. 4, p. 59–69, 2006.

AHMED, H. M. *et al.* Protection against *Schistosoma mansoni* infection with A recombinant schistosomula 21.7 kDa protein. *Arab J. Biotechnol.*, v. 4, p. 229–249, 2001.

AMONKAR, P.; MURALI, G.; KRISHNAMURTHY, S. *Schistosoma* induced squamous cell carcinoma of the bladder. *Indian Journal of Pathology & Microbiology*, PMID: 12024936, v. 44, n. 3, p. 363–364, jul. 2001.

ANAND, S.; HANSON, K. Disability-adjusted life years: a critical review. *Journal of Health Economics*, v. 16, n. 6, p. 685–702, dez. 1997.

ARTIMO, P. *et al.* ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research*, v. 40, n. W1, p. W597–W603, 1 jul. 2012.

- BAI, Y.; AUPERIN, T. C.; TONG, L. The use of in situ proteolysis in the crystallization of murine CstF-77. *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, PMID: 17277459 PMCID: PMC2330134, v. 63, n. Pt 2, p. 135–138, 1 fev. 2007.
- BANFORD, S. *et al.* FhCaBP3: A *Fasciola hepatica* calcium binding protein with EF-hand and dynein light chain domains. *Biochimie*, v. 95, n. 4, p. 751–758, abr. 2013.
- BARBAR, E. Dynein light chain LC8 is a dimerization hub essential in diverse protein networks. *Biochemistry*, PMID: 18092820, v. 47, n. 2, p. 503–508, 15 jan. 2008.
- BASKARAN, S. *et al.* Two-Site Recognition of Phosphatidylinositol 3-Phosphate by PROPPINs in Autophagy. *Molecular Cell*, v. 47, n. 3, p. 339–348, 10 ago. 2012.
- BECKWITH, S. M. *et al.* The “8-kD” cytoplasmic dynein light chain is required for nuclear migration and for dynein heavy chain localization in *Aspergillus nidulans*. *The Journal of Cell Biology*, PMID: 9832552 PMCID: PMC2133080, v. 143, n. 5, p. 1239–1247, 30 nov. 1998.
- BENASHSKI, S. E. *et al.* Dimerization of the highly conserved light chain shared by dynein and myosin V. *The Journal of Biological Chemistry*, PMID: 9252421, v. 272, n. 33, p. 20929–20935, 15 ago. 1997.
- BERG, J. M.; TYMOCZKO, J. L.; STRYER, Lubert. *Bioquímica*. 6. ed. [S.l.]: Guanabara Koogan, 2012.
- BERGQUIST, N. R.; LEONARDO, L. R.; MITCHELL, G. F. Vaccine-linked chemotherapy: can schistosomiasis control benefit from an integrated approach? *Trends in Parasitology*, PMID: 15734657, v. 21, n. 3, p. 112–117, mar. 2005.
- BERGQUIST, R.; UTZINGER, J.; MCMANUS, D. P. Trick or Treat: The Role of Vaccines in Integrated Schistosomiasis Control. *PLoS Neglected Tropical Diseases*, v. 2, n. 6, p. e244, 25 jun. 2008.
- BETHONY, J. M. *et al.* Vaccines to combat the neglected tropical diseases. *Immunological Reviews*, PMID: 21198676 PMCID: PMC3438653, v. 239, n. 1, p. 237–270, jan. 2011.
- BHATTACHARYA, S.; BUNICK, C. G.; CHAZIN, W. J. Target selectivity in EF-hand calcium binding proteins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, v. 1742, n. 1-3, p. 69–79, dez. 2004.
- BLANCHARD, H. *et al.* Structure of a calpain Ca(2+)-binding domain reveals a novel EF-hand and Ca(2+)-induced conformational changes. *Nature Structural Biology*, PMID: 9228945, v. 4, n. 7, p. 532–538, jul. 1997.
- BÖHM, G.; MUHR, R.; JAENICKE, R. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Engineering*, PMID: 1409538, v. 5, n. 3, p. 191–195, abr. 1992.

BOROS, D. L. Immunopathology of *Schistosoma mansoni* infection. *Clinical Microbiology Reviews*, PMID: 2504481 PMCID: PMC358119, v. 2, n. 3, p. 250–269, jul. 1989.

BOTROS, S. *et al.* Effect of praziquantel on the immature stages of *Schistosoma haematobium*. *International Journal for Parasitology*, v. 35, n. 13, p. 1453–1457, nov. 2005.

BOTTIEAU, E. *et al.* Imported Katayama fever: clinical and biological features at presentation and during treatment. *The Journal of Infection*, PMID: 16169593, v. 52, n. 5, p. 339–345, maio 2006.

BUNKER, R. D.; LOOMES, K. M.; BAKER, E. N. Purification, crystallization and preliminary crystallographic analysis of human dihydrodipicolinate synthase-like protein (DHDPsL). *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, PMID: 22232173 PMCID: PMC3253836, v. 68, n. Pt 1, p. 59–62, 1 jan. 2012.

CAPRON, A. *et al.* Vaccine strategies against schistosomiasis: from concepts to clinical trials. *International Archives of Allergy and Immunology*, PMID: 11306914, v. 124, n. 1-3, p. 9–15, mar. 2001.

CARNEIRO, F.R.G. *et al.* Spectroscopic characterization of the tumor antigen NY-REN-21 and identification of heterodimer formation with SCAND1. *Biochemical and Biophysical Research Communications*, v. 343, n. 1, p. 260–268, 28 abr. 2006.

CHANG, J. *et al.* Dynein light chain interaction with the peroxisomal import docking complex modulates peroxisome biogenesis in yeast. *Journal of Cell Science*, v. 126, n. 20, p. 4698–4706, 15 out. 2013.

CHEEVER, A. W.; HOFFMANN, K. F.; WYNN, T. A. Immunopathology of schistosomiasis mansoni in mice and men. *Immunology Today*, PMID: 10953099, v. 21, n. 9, p. 465–466, set. 2000.

CHEN, J. *et al.* Cloning and expression of 21.1-kDa tegumental protein of *Clonorchis sinensis* and human antibody response to it as a trematode-nematode pan-specific serodiagnosis antigen. *Parasitology Research*, PMID: 20865432, v. 108, n. 1, p. 161–168, jan. 2011.

CHIN, D.; MEANS, A. R. Calmodulin: a prototypical calcium sensor. *Trends in Cell Biology*, v. 10, n. 8, p. 322–328, ago. 2000.

CHIOU, C. *et al.* Crystallization and preliminary X-ray diffraction studies of the 51 kDa protein of the mosquito-larvicidal binary toxin from *Bacillus sphaericus*. *Acta Crystallographica. Section D, Biological Crystallography*, PMID: 10216313, v. 55, n. Pt 5, p. 1083–1085, maio 1999.

CIOLI, D. *et al.* Determination of ED50 values for praziquantel in praziquantel-resistant and -susceptible *Schistosoma mansoni* isolates. *International Journal for Parasitology*, v. 34, n. 8, p. 979–987, jul. 2004.

- COLLINS, J. H. Homology of myosin DTNB light chain with alkali light chains, troponin C and parvalbumin. *Nature*, v. 259, n. 5545, p. 699–700, 26 fev. 1976.
- COURA, Jr; AMARAL, Rs. Epidemiological and control aspects of schistosomiasis in Brazilian endemic areas. *Memórias do Instituto Oswaldo Cruz*, v. 99, ago. 2004.
- COWIE, R. H. Can snails ever be effective and safe biocontrol agents? *International Journal of Pest Management*, v. 47, n. 1, p. 23–40, jan. 2001.
- CURWEN, R.S. *et al.* The *Schistosoma mansoni* soluble proteome: a comparison across four life-cycle stages. *Molecular and Biochemical Parasitology*, v. 138, n. 1, p. 57–66, nov. 2004.
- DAUTER, Z.; DAUTER, M.; RAJASHANKAR, K. R. Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta Crystallographica Section D*, v. 56, n. 2, p. 232–237, 2000.
- DAVIS, A.; BILES, J. E.; ULRICH, A. M. Initial experiences with praziquantel in the treatment of human infections due to *Schistosoma haematobium*. *Bulletin of the World Health Organization*, PMID: 396053 PMCID: PMC2395866, v. 57, n. 5, p. 773–779, 1979.
- DELANO, W. L. *PyMOL: An Open-Source Molecular Graphics Tool*. [S.l: s.n.], 2007. v. DeLano Scientific San Carlos, California USA.
- DOENHOFF, M. J.; PICA-MATTOCCIA, L. Praziquantel for the treatment of schistosomiasis: its use for control in areas with endemic disease and prospects for drug resistance. *Expert Review of Anti-Infective Therapy*, PMID: 16597202, v. 4, n. 2, p. 199–210, abr. 2006.
- DORMITZER, P. R.; GRANDI, G.; RAPPUOLI, R. Structural vaccinology starts to deliver. *Nature Reviews. Microbiology*, PMID: 23154260, v. 10, n. 12, p. 807–813, dez. 2012.
- DOWNING, K. H; SUI, H. Structural insights into microtubule doublet interactions in axonemes. *Current Opinion in Structural Biology*, v. 17, n. 2, p. 253–259, abr. 2007.
- DROHAT, A. C. *et al.* Solution Structure of Calcium-Bound Rat S100B($\beta\beta$) As Determined by Nuclear Magnetic Resonance Spectroscopy^{†,‡}. *Biochemistry*, v. 37, n. 9, p. 2729–2740, mar. 1998.
- DUGUAY, D. *et al.* Dynein Light Chain Tctex-Type 1 Modulates Orexin Signaling through Its Interaction with Orexin 1 Receptor. *PLoS ONE*, v. 6, n. 10, p. e26430, 20 out. 2011.
- DUNSCH, A. K. *et al.* Dynein light chain 1 and a spindle-associated adaptor promote dynein asymmetry and spindle orientation. *The Journal of Cell Biology*, v. 198, n. 6, p. 1039–1054, 17 set. 2012.
- DYDA, F. *et al.* Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science (New York, N.Y.)*, PMID: 7801124, v. 266, n. 5193, p. 1981–1986, 23 dez. 1994.

ELBAZ, T; ESMAT, G. Hepatic and Intestinal Schistosomiasis: Review. *Journal of Advanced Research*, v. 4, n. 5, p. 445–452, set. 2013.

EL RIDI, R. A.F.; TALLIMA, H. A.-M. Novel Therapeutic and Prevention Approaches for Schistosomiasis: Review. *Journal of Advanced Research*, v. 4, n. 5, p. 467–478, set. 2013.

EMSLEY, P. *et al.* Features and development of Coot. *Acta Crystallographica Section D*, v. 66, n. 4, p. 486–501, 2010.

ESPINDOLA, F. S. *et al.* The light chain composition of chicken brain myosin-Va: calmodulin, myosin-II essential light chains, and 8-kDa dynein light chain/PIN. *Cell Motility and the Cytoskeleton*, PMID: 11093248, v. 47, n. 4, p. 269–281, dez. 2000.

FALLON, P. G.; DOENHOFF, M. J. Drug-resistant schistosomiasis: resistance to praziquantel and oxamniquine induced in *Schistosoma mansoni* in mice is drug specific. *The American Journal of Tropical Medicine and Hygiene*, PMID: 8059919, v. 51, n. 1, p. 83–88, jul. 1994.

FAN, J. *et al.* Structural basis of diverse sequence-dependent target recognition by the 8 kDa dynein light chain. *Journal of Molecular Biology*, v. 306, n. 1, p. 97–108, fev. 2001.

FAN, J. S. *et al.* Protein inhibitor of neuronal nitric-oxide synthase, PIN, binds to a 17-amino acid residue fragment of the enzyme. *The Journal of Biological Chemistry*, PMID: 9837926, v. 273, n. 50, p. 33472–33481, 11 dez. 1998.

FERRÉ-D'AMARÉ, A. R.; BURLEY, S. K. [10] Dynamic light scattering in evaluating crystallizability of macromolecules. *Methods in Enzymology*. [S.l.]: Elsevier, 1997. v. 276. p. 157–166.

FITZSIMMONS, C. M. *et al.* Human IgE response to the *Schistosoma haematobium* 22.6 kDa antigen. *Parasite Immunology*, v. 26, n. 8-9, p. 371–376, ago. 2004. Acesso em: 26 jul. 2014.

FITZSIMMONS, C. M. *et al.* Factors affecting human IgE and IgG responses to allergen-like *Schistosoma mansoni* antigens: Molecular structure and patterns of in vivo exposure. *International Archives of Allergy and Immunology*, PMID: 17019080, v. 142, n. 1, p. 40–50, 2007.

FONTANA, A. *et al.* Probing protein structure by limited proteolysis. *Acta Biochimica Polonica*, PMID: 15218531, v. 51, n. 2, p. 299–321, 2004.

FRANCIS, P.; BICKLE, Q. Cloning of a 21.7-kDa vaccine-dominant antigen gene of *Schistosoma mansoni* reveals an EF hand-like motif. *Molecular and Biochemical Parasitology*, PMID: 1371327, v. 50, n. 2, p. 215–224, fev. 1992.

FROHBERG, H. Results of toxicological studies on praziquantel. *Arzneimittel-Forschung*, PMID: 6542381, v. 34, n. 9B, p. 1137–1144, 1984.

GAUR, R. K. *et al.* Preliminary X-ray analysis of a human V_H fragment at 1.8 Å resolution. *Acta Crystallographica Section D Biological Crystallography*, v. 60, n. 5, p. 965–967, 1 maio 2004.

GAZZINELLI, G. *et al.* Immune responses during human Schistosomiasis mansoni. XI. Immunologic status of patients with acute infections and after treatment. *Journal of Immunology (Baltimore, Md.: 1950)*, PMID: 4020142, v. 135, n. 3, p. 2121–2127, set. 1985.

GELFAND, M. Relation between carcinoma of the bladder and infestation with *Schistosoma haematobium*. *The Lancet*, v. 289, n. 7502, p. 1249–1251, jun. 1967.

GENNERICH, A.; VALE, R. D. Walking the walk: how kinesin and dynein coordinate their steps. *Current Opinion in Cell Biology*, v. 21, n. 1, p. 59–67, fev. 2009.

GENTRY, H. R. *et al.* Structural and Biochemical Characterization of CIB1 Delineates a New Family of EF-hand-containing Proteins. *Journal of Biological Chemistry*, v. 280, n. 9, p. 8407–8415, 4 mar. 2005.

GOESSWEINER-MOHR, N. *et al.* Crystallization and first data collection of the putative transfer protein TraN from the Gram-positive conjugative plasmid pIP501. *Acta Crystallographica Section F Structural Biology and Crystallization Communications*, v. 68, n. 11, p. 1402–1405, 1 nov. 2012.

GÖNNERT, R.; ANDREWS, P. Praziquantel, a new broad-spectrum antischistosomal agent. *Zeitschrift für Parasitenkunde*, v. 52, n. 2, p. 129–150, 1 jan. 1977.

GOPALAKRISHNA, R.; ANDERSON, W. B. The effects of chemical modification of calmodulin on Ca²⁺-induced exposure of a hydrophobic region. Separation of active and inactive forms of calmodulin. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, v. 844, n. 2, p. 265–269, fev. 1985.

GRABAREK, Z. Structural Basis for Diversity of the EF-hand Calcium-binding Proteins. *Journal of Molecular Biology*, v. 359, n. 3, p. 509–525, jun. 2006.

GRAY, D. J *et al.* Schistosomiasis elimination: lessons from the past guide the future. *The Lancet Infectious Diseases*, v. 10, n. 10, p. 733–736, out. 2010.

GREEN, D. W.; INGRAM, V. M.; PERUTZ, M. F. The Structure of Haemoglobin. IV. Sign Determination by the Isomorphous Replacement Method. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 225, n. 1162, p. 287–307, 14 set. 1954.

GRYSEELS, B. *et al.* Human schistosomiasis. *The Lancet*, v. 368, n. 9541, p. 1106–1118, set. 2006.

GUIMARÃES, B. G. *et al.* The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS. *Journal of Synchrotron Radiation*, v. 16, n. 1, p. 69–75, 1 jan. 2009.

GURMU, D. *et al.* Expression, purification, crystallization and preliminary X-ray analysis of ORF60, the small subunit (R2) of ribonucleotide reductase from Kaposi's

sarcoma-associated herpesvirus (KSHV). *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, PMID: 20516613 PMCID: PMC2882783, v. 66, n. Pt 6, p. 734–737, 1 jun. 2010.

HAFALLA, J. C. R. *et al.* Molecular identification of a 21.7 kDa *Schistosoma japonicum* antigen as a target of the human IgE response. *Molecular and Biochemical Parasitology*, v. 98, n. 1, p. 157–161, jan. 1999.

HAGAN, P. *et al.* Human IgE, IgG4 and resistance to reinfection with *Schistosoma haematobium*. *Nature*, PMID: 1898985, v. 349, n. 6306, p. 243–245, 17 jan. 1991.

HAGAN, P. Reinfection, exposure and immunity in human schistosomiasis. *Parasitology Today*, v. 8, n. 1, p. 12–16, jan. 1992.

HANAHAN, D. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, v. 166, n. 4, p. 557–580, jun. 1983.

HIROKAWA, N. Kinesin and Dynein Superfamily Proteins and the Mechanism of Organelle Transport. *Science*, v. 279, n. 5350, p. 519–526, 23 jan. 1998.

HIRSCH, C. *et al.* Induction of protective immunity and modulation of granulomatous hypersensitivity in mice using PIII, an anionic fraction of *Schistosoma mansoni* adult worm. *Parasitology*, PMID: 9280892, v. 115 (Pt 1), p. 21–28, jul. 1997.

HIRSCH, C.; GOES, A. M. Characterization of fractionated *Schistosoma mansoni* soluble adult worm antigens that elicit human cell proliferation and granuloma formation in vitro. *Parasitology*, PMID: 8684827, v. 112 (Pt 6), p. 529–535, jun. 1996.

HOEFLICH, K. P.; IKURA, M. Calmodulin in Action. *Cell*, v. 108, n. 6, p. 739–742, mar. 2002.

HOFFMANN, K. F. Molecular Characterization of a 20.8-kDa *Schistosoma mansoni* Antigen. SEQUENCE SIMILARITY TO TEGUMENTAL ASSOCIATED ANTIGENS AND DYNEIN LIGHT CHAINS. *Journal of Biological Chemistry*, v. 272, n. 23, p. 14509–14515, 6 jun. 1997.

HOLM, L.; ROSENSTROM, P. Dali server: conservation mapping in 3D. *Nucleic Acids Research*, v. 38, n. Web Server, p. W545–W549, 1 jul. 2010.

HSIAO, Y. *et al.* Crystal structures of respiratory pathogen neuraminidases. *Biochemical and Biophysical Research Communications*, PMID: 19284989 PMCID: PMC3836282, v. 380, n. 3, p. 467–471, 13 mar. 2009.

HUANG, Y. *et al.* A novel tegumental protein 31.8 kDa of *Clonorchis sinensis*: sequence analysis, expression, and immunolocalization. *Parasitology Research*, v. 102, n. 1, p. 77–81, 18 out. 2007.

HUGHES, S. M. *et al.* Molecular analysis of a cytoplasmic dynein light intermediate chain reveals homology to a family of ATPases. *Journal of Cell Science*, PMID: 7738094, v. 108 (Pt 1), p. 17–24, jan. 1995.

HUNG, K.; HSU, C.; YU, C. Solution structure of human Ca²⁺-bound S100A12. *Journal of Biomolecular NMR*, v. 57, n. 3, p. 313–318, nov. 2013.

ISHIZAKI, T.; KAMO, E.; BOEHME, K. Double-blind studies of tolerance to praziquantel in Japanese patients with *Schistosoma japonicum* infections. *Bulletin of the World Health Organization*, PMID: 396055 PMCID: PMC2395874, v. 57, n. 5, p. 787–791, 1979.

ISMAIL, M. *et al.* Characterization of isolates of *Schistosoma mansoni* from Egyptian villagers that tolerate high doses of praziquantel. *The American Journal of Tropical Medicine and Hygiene*, PMID: 8780463, v. 55, n. 2, p. 214–218, ago. 1996.

ISMAIL, M. *et al.* Resistance to praziquantel: direct evidence from *Schistosoma mansoni* isolated from Egyptian villagers. *The American Journal of Tropical Medicine and Hygiene*, v. 60, n. 6, p. 932–935, 1 jun. 1999.

JACHIMSKA, B.; WASILEWSKA, M.; ADAMCZYK, Z. Characterization of globular protein solutions by dynamic light scattering, electrophoretic mobility, and viscosity measurements. *Langmuir: the ACS journal of surfaces and colloids*, PMID: 18512882, v. 24, n. 13, p. 6866–6872, 1 jun. 2008.

JAFFREY, S. R.; SNYDER, S. H. PIN: an associated protein inhibitor of neuronal nitric oxide synthase. *Science (New York, N.Y.)*, PMID: 8864115, v. 274, n. 5288, p. 774–777, 1 nov. 1996.

JEFFS, S. A. *et al.* Molecular cloning and characterisation of the 22-kilodalton adult *Schistosoma mansoni* antigen recognised by antibodies from mice protectively vaccinated with isolated tegumental surface membranes. *Molecular and Biochemical Parasitology*, v. 46, n. 1, p. 159–167, maio 1991.

JOHNSON, S. *et al.* Expression, limited proteolysis and preliminary crystallographic analysis of IpaD, a component of the *Shigella flexneri* type III secretion system. *Acta Crystallographica Section F Structural Biology and Crystallization Communications*, v. 62, n. 9, p. 865–868, 1 set. 2006.

KAISER, F. J. *et al.* Nuclear interaction of the dynein light chain LC8a with the TRPS1 transcription factor suppresses the transcriptional repression activity of TRPS1. *Human Molecular Genetics*, PMID: 12761050, v. 12, n. 11, p. 1349–1358, 1 jun. 2003.

KARKI, S.; HOLZBAUR, E. LF. Cytoplasmic dynein and dynactin in cell division and intracellular transport. *Current Opinion in Cell Biology*, v. 11, n. 1, p. 45–53, fev. 1999.

KATZ, N.; ROCHA, R. S.; CHAVES, A. Preliminary trials with praziquantel in human infections due to *Schistosoma mansoni*. *Bulletin of the World Health Organization*, PMID: 396054 PMCID: PMC2395863, v. 57, n. 5, p. 781–785, 1979.

KAWASAKI, H.; NAKAYAMA, S.; KRETSINGER, R.H. Classification and evolution of EF-hand proteins. *Biometals*, v. 11, n. 4, p. 277–295, 1 dez. 1998.

KENDREW, J. C.; PARRISH, R. G. The Crystal Structure of Myoglobin. III. Sperm-Whale Myoglobin. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 238, n. 1214, p. 305–324, 8 jan. 1957.

- KIM, K. Y. *et al.* Crystallization and preliminary X-ray crystallographic analysis of carboxyl-terminal region 4 of SigR from *Streptomyces coelicolor* A3(2). *Acta Crystallographica. Section F, Structural Biology Communications*, PMID: 24915084, v. 70, n. Pt 6, p. 747–749, jun. 2014.
- KIM, Y. *et al.* Identification and characterization of a novel 21.6-kDa tegumental protein from *Clonorchis sinensis*. *Parasitology Research*, v. 110, n. 5, p. 2061–2066, 1 maio 2012.
- KING, C. H. Toward the Elimination of Schistosomiasis. *New England Journal of Medicine*, v. 360, n. 2, p. 106–109, 8 jan. 2009.
- KING, S. M. AAA domains and organization of the dynein motor unit. *Journal of Cell Science*, PMID: 10862709, v. 113 (Pt 14), p. 2521–2526, jul. 2000.
- KNUDSEN, G. M. *et al.* Proteomic analysis of *Schistosoma mansoni* cercarial secretions. *Molecular & cellular proteomics: MCP*, PMID: 16112986, v. 4, n. 12, p. 1862–1875, dez. 2005.
- KRETSINGER, R. H.; NOCKOLDS, C. E. Carp muscle calcium-binding protein. II. Structure determination and general description. *The Journal of Biological Chemistry*, PMID: 4700463, v. 248, n. 9, p. 3313–3326, 10 maio 1973.
- KRICK, R. *et al.* Structural and functional characterization of the two phosphoinositide binding sites of PROPPINs, a -propeller protein family. *Proceedings of the National Academy of Sciences*, v. 109, n. 30, p. E2042–E2049, 24 jul. 2012.
- KRISSINEL, E.; HENRICK, K. Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology*, v. 372, n. 3, p. 774–797, set. 2007.
- LAEMMLI, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, v. 227, n. 5259, p. 680–685, 15 ago. 1970.
- LAMBERTUCCI, J. R. Acute schistosomiasis: clinical, diagnostic and therapeutic features. *Revista Do Instituto De Medicina Tropical De São Paulo*, PMID: 8115806, v. 35, n. 5, p. 399–404, out. 1993.
- LEWIT-BENTLEY, A; RÉTY, S. EF-hand calcium-binding proteins. *Current Opinion in Structural Biology*, v. 10, n. 6, p. 637–643, dez. 2000.
- LIANG, J. *et al.* Structure of the PIN/LC8 dimer with a bound peptide. *Nature Structural Biology*, PMID: 10426949, v. 6, n. 8, p. 735–740, ago. 1999.
- LI HSÜ, S.Y.; HSÜ, H.F.; BURMEISTER, L.F. *Schistosoma mansoni*: Vaccination of mice with highly X-irradiated cercariae. *Experimental Parasitology*, v. 52, n. 1, p. 91–104, ago. 1981.
- LI, M.; CHANG, W. Protein crystallization. *Photosynthesis Research*, v. 102, n. 2-3, p. 223–229, 1 dez. 2009.

LIN, G. *et al.* Crystal structure of calcium bound domain VI of calpain at 1.9 Å resolution and its role in enzyme assembly, regulation, and inhibitor binding. *Nature Structural Biology*, v. 4, n. 7, p. 539–547, jul. 1997.

LO, K. W.-H. *et al.* The 8-kDa Dynein Light Chain Binds to Its Targets via a Conserved (K/R)XTQT Motif. *Journal of Biological Chemistry*, v. 276, n. 17, p. 14059–14066, 27 abr. 2001.

LO, K. W.-H. *et al.* The 8-kDa dynein light chain binds to p53-binding protein 1 and mediates DNA damage-induced p53 nuclear accumulation. *The Journal of Biological Chemistry*, PMID: 15611139, v. 280, n. 9, p. 8172–8179, 4 mar. 2005.

LOPES, D. O. *et al.* Sm21.6 a novel EF-hand family protein member located on the surface of *Schistosoma mansoni* adult worm that failed to induce protection against challenge infection but reduced liver pathology. *Vaccine*, v. 27, n. 31, p. 4127–4135, jun. 2009.

MAKI, M. *et al.* Structures, functions and molecular evolution of the penta-EF-hand Ca²⁺-binding proteins. *Biochimica Et Biophysica Acta*, PMID: 12445459, v. 1600, n. 1-2, p. 51–60, 4 nov. 2002.

MAL, T. K. *et al.* Detecting Protein Kinase Recognition Modes of Calmodulin by Residual Dipolar Couplings in Solution NMR [†]. *Biochemistry*, v. 41, n. 43, p. 12899–12906, out. 2002.

MANDEL, C. R. *et al.* A serendipitous discovery that in situ proteolysis is essential for the crystallization of yeast CPSF-100 (Ydh1p). *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, PMID: 17012808 PMCID: PMC2225192, v. 62, n. Pt 10, p. 1041–1045, 1 out. 2006.

MASIGNANI, V. *et al.* Vaccination against *Neisseria meningitidis* Using Three Variants of the Lipoprotein GNA1870. *Journal of Experimental Medicine*, v. 197, n. 6, p. 789–799, 17 mar. 2003.

MATSUSHIMA, N. *et al.* Ca²⁺-bound calmodulin forms a compact globular structure on binding four trifluoperazine molecules in solution. *Biochem. J.*, v. 347, n. 1, p. 211–215, 2000.

MATTHEWS, B. W. Solvent content of protein crystals. *Journal of Molecular Biology*, PMID: 5700707, v. 33, n. 2, p. 491–497, 28 abr. 1968.

MCMANUS, D. P.; LOUKAS, A. Current status of vaccines for schistosomiasis. *Clinical Microbiology Reviews*, PMID: 18202444 PMCID: PMC2223839, v. 21, n. 1, p. 225–242, jan. 2008.

MCPHERSON, A.; DAY, J.; HARRIS, L. J. Crystals of the beta-subunit of bovine luteinizing hormone and indicators for the involvement of proteolysis in protein crystallization. *Acta Crystallographica. Section D, Biological Crystallography*, PMID: 15103132, v. 60, n. Pt 5, p. 872–877, maio 2004.

MICHIELS, J. *et al.* The functions of Ca(2+) in bacteria: a role for EF-hand proteins? *Trends in Microbiology*, PMID: 11827810, v. 10, n. 2, p. 87–93, fev. 2002.

MINARD, P. *et al.* Immunization of mice with cobalt-60 irradiated *Schistosoma mansoni* cercariae. *The American Journal of Tropical Medicine and Hygiene*, PMID: 626284, v. 27, n. 1 Pt 1, p. 76–86, jan. 1978.

MITCHELL, D. R.; ROSENBAUM, J. L. Protein-protein interactions in the 18S ATPase of *Chlamydomonas* outer dynein arms. *Cell Motility and the Cytoskeleton*, v. 6, n. 5, p. 510–520, 1986.

MOHAMED, M. M *et al.* Characterization of Sm20.8, a member of a family of schistosome tegumental antigens. *Molecular and Biochemical Parasitology*, v. 96, n. 1-2, p. 15–25, out. 1998.

MOSTAFA, M. H.; SHEWEITA, S. A.; O’CONNOR, P. J. Relationship between schistosomiasis and bladder cancer. *Clinical Microbiology Reviews*, PMID: 9880476 PMCID: PMC88908, v. 12, n. 1, p. 97–111, jan. 1999.

NAGEM, R. A. P.; DAUTER, Z.; POLIKARPOV, I. Protein crystal structure solution by fast incorporation of negatively and positively charged anomalous scatterers. *Acta Crystallographica Section D*, v. 57, n. 7, p. 996–1002, 2001.

NAGEM, R. A.P; POLIKARPOV, I.; DAUTER, Z. Phasing on Rapidly Soaked Ions. In: CHARLES W. CARTER, JR. AND ROBERT M. SWEET (Org.). . *Methods in Enzymology*. [S.l.]: Academic Press, 2003. v. Volume 374. p. 120–137.

NASH, T. E. *et al.* Treatment of *Schistosoma mekongi* with praziquantel: a double-blind study. *The American Journal of Tropical Medicine and Hygiene*, PMID: 6751116, v. 31, n. 5, p. 977–982, set. 1982.

NEVES, David P. *Parasitologia Humana*. 11. ed. São Paulo: Editora Atheneu, 2005.

NOVOTNÝ, J.; BRUCCOLERI, R. E. Correlation among sites of limited proteolysis, enzyme accessibility and segmental mobility. *FEBS letters*, PMID: 3542567, v. 211, n. 2, p. 185–189, 26 jan. 1987.

NUCCITELLI, A. *et al.* Structure-based approach to rationally design a chimeric protein for an effective vaccine against Group B *Streptococcus* infections. *Proceedings of the National Academy of Sciences*, 18 maio 2011.

ORR, R. *et al.* FhCaBP4: a *Fasciola hepatica* calcium-binding protein with EF-hand and dynein light chain domains. *Parasitology Research*, PMID: 22773043, v. 111, n. 4, p. 1707–1713, out. 2012.

OTWINOWSKI, Z.; MINOR, W. [20] Processing of X-ray diffraction data collected in oscillation mode. In: CHARLES W. CARTER, JR. (Org.). . *Methods in Enzymology*. [S.l.]: Academic Press, 1997. v. Volume 276. p. 307–326.

PASCHAL, B. M. *et al.* Homology of the 74-kD cytoplasmic dynein subunit with a flagellar dynein polypeptide suggests an intracellular targeting function. *The Journal of Cell Biology*, PMID: 1387402 PMCID: PMC2289596, v. 118, n. 5, p. 1133–1143, set. 1992.

- PAZOUR, G. J.; WILKERSON, C. G.; WITMAN, G. B. A dynein light chain is essential for the retrograde particle movement of intraflagellar transport (IFT). *The Journal of Cell Biology*, PMID: 9585416 PMCID: PMC2132779, v. 141, n. 4, p. 979–992, 18 maio 1998.
- PETTERSEN, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, PMID: 15264254, v. 25, n. 13, p. 1605–1612, out. 2004.
- PFISTER, K. K.; FAY, R. B.; WITMAN, G. B. Purification and polypeptide composition of dynein ATPases from *Chlamydomonas* flagella. *Cell Motility*, v. 2, n. 6, p. 525–547, 1982.
- PIPERNO, G.; LUCK, D. J. Axonemal adenosine triphosphatases from flagella of *Chlamydomonas reinhardtii*. Purification of two dyneins. *The Journal of Biological Chemistry*, PMID: 155062, v. 254, n. 8, p. 3084–3090, 25 abr. 1979.
- POTTS, B. C. M. *et al.* The structure of calyculin reveals a novel homodimeric fold for S100 Ca²⁺-binding proteins. *Nature Structural Biology*, v. 2, n. 9, p. 790–796, set. 1995.
- RAO, L. *et al.* The yeast dynein Dyn2-Pac11 complex is a dynein dimerization/processivity factor: structural and single-molecule characterization. *Molecular Biology of the Cell*, v. 24, n. 15, p. 2362–2377, 1 ago. 2013.
- RAO, S. T. *et al.* Structure of *Paramecium tetraurelia* calmodulin at 1.8 Å resolution. *Protein Science*, v. 2, n. 3, p. 436–447, 1 mar. 1993.
- RAUX, H.; FLAMAND, A.; BLONDEL, D. Interaction of the rabies virus P protein with the LC8 dynein light chain. *Journal of Virology*, PMID: 11024151 PMCID: PMC102061, v. 74, n. 21, p. 10212–10216, nov. 2000.
- RAVULAPALLI, R. *et al.* Homodimerization of calpain 3 penta-EF-hand domain. *The Biochemical Journal*, PMID: 15660530 PMCID: PMC1138966, v. 388, n. Pt 2, p. 585–591, 1 jun. 2005.
- RAYMENT, I. *et al.* Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science (New York, N.Y.)*, PMID: 8316857, v. 261, n. 5117, p. 50–58, 2 jul. 1993.
- REIS, R. A. G.; BORTOT, L. O.; CALIRI, A. In silico assessment of S100A12 monomer and dimer structural dynamics: implications for the understanding of its metal-induced conformational changes. *Journal of biological inorganic chemistry: JBIC: a publication of the Society of Biological Inorganic Chemistry*, PMID: 24944024, 19 jun. 2014.
- REZENDE, C.M.F. *et al.* Immunization with rP22 induces protective immunity against *Schistosoma mansoni*: Effects on granuloma down-modulation and cytokine production. *Immunology Letters*, v. 141, n. 1, p. 123–133, 30 dez. 2011.
- ROMES, E. M.; TRIPATHY, A.; SLEP, K. C. Structure of a yeast Dyn2-Nup159 complex and molecular basis for dynein light chain-nuclear pore interaction. *The*

Journal of Biological Chemistry, PMID: 22411995 PMCID: PMC3346085, v. 287, n. 19, p. 15862–15873, 4 maio 2012.

SABAH, A.A. *et al.* *Schistosoma mansoni*: Chemotherapy of infections of different ages. *Experimental Parasitology*, v. 61, n. 3, p. 294–303, jun. 1986.

SANTIAGO, M. L. *et al.* Identification of the *Schistosoma japonicum* 22.6-kDa antigen as a major target of the human IgE response: similarity of IgE-binding epitopes to allergen peptides. *International Archives of Allergy and Immunology*, PMID: 9784652, v. 117, n. 2, p. 94–104, out. 1998.

SANTOS, C. A. *et al.* Functional and structural studies of the disulfide isomerase DsbC from the plant pathogen *Xylella fastidiosa* reveals a redox-dependent oligomeric modulation *in vitro*. *FEBS Journal*, v. 279, n. 20, p. 3828–3843, out. 2012.

SCHNORRER, F.; BOHMANN, K.; NUSSLEIN-VOLHARD, C. The molecular motor dynein is involved in targeting Swallow and bicoid RNA to the anterior pole of *Drosophila* oocytes. *Nat Cell Biol*, v. 2, n. 4, p. 185–190, abr. 2000.

SCHOLTE, R.G.C. *et al.* Predictive risk mapping of schistosomiasis in Brazil using Bayesian geostatistical models. *Acta Tropica*, v. 132, p. 57–63, abr. 2014.

SCHROER, T. A. DYNACTIN. *Annual Review of Cell and Developmental Biology*, v. 20, n. 1, p. 759–779, nov. 2004.

SENAWONG, G. *et al.* Cloning, expression, and characterization of a novel *Opisthorchis viverrini* calcium-binding EF-hand protein. *Parasitology International*, v. 61, n. 1, p. 94–100, mar. 2012.

SHELDRIK, G. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallographica Section D*, v. 66, n. 4, p. 479–485, 2010.

SIDDIQUI, A. A.; SIDDIQUI, B. A.; GANLEY-LEAL, L. Schistosomiasis vaccines. *Human Vaccines*, PMID: 22048120 PMCID: PMC3323497, v. 7, n. 11, p. 1192–1197, nov. 2011.

SLEVIN, L. K. *et al.* The mechanism of dynein light chain LC8-mediated oligomerization of the Ana2 centriole duplication factor. *The Journal of Biological Chemistry*, PMID: 24920673 PMCID: PMC4110283, v. 289, n. 30, p. 20727–20739, 25 jul. 2014.

SMITH, J. J. *et al.* Expression and functional profiling reveal distinct gene classes involved in fatty acid metabolism. *Molecular Systems Biology*, v. 2, 21 fev. 2006.

SOUZA, C. P. De. Molluscicide control of snail vectors of schistosomiasis. *Memórias do Instituto Oswaldo Cruz*, v. 90, n. 2, p. 165–168, abr. 1995.

SPYRACOPOULOS, L. *et al.* Temperature Dependence of Dynamics and Thermodynamics of the Regulatory Domain of Human Cardiac Troponin C [†]. *Biochemistry*, v. 40, n. 42, p. 12541–12551, out. 2001.

SREERAMA, N.; WOODY, R. W. Computation and analysis of protein circular dichroism spectra. *Methods in Enzymology*, PMID: 15063656, v. 383, p. 318–351, 2004.

STEIN, L. D.; DAVID, J. R. Cloning of a developmentally regulated tegument antigen of *Schistosoma mansoni*. *Molecular and Biochemical Parasitology*, PMID: 2429181, v. 20, n. 3, p. 253–264, set. 1986.

STEINMANN, P. *et al.* Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk. *The Lancet Infectious Diseases*, v. 6, n. 7, p. 411–425, jul. 2006.

STELMA, F. F. *et al.* Efficacy and side effects of praziquantel in an epidemic focus of *Schistosoma mansoni*. *The American Journal of Tropical Medicine and Hygiene*, PMID: 7677219, v. 53, n. 2, p. 167–170, ago. 1995.

STRYNADKA, N C J; JAMES, M N G. Crystal Structures of the Helix-Loop-Helix Calcium-Binding Proteins. *Annual Review of Biochemistry*, v. 58, n. 1, p. 951–999, jun. 1989.

STUCHELL-BRERETON, M. D. *et al.* Functional interaction between dynein light chain and intermediate chain is required for mitotic spindle positioning. *Molecular Biology of the Cell*, v. 22, n. 15, p. 2690–2701, 1 ago. 2011.

SUBPIPATTANA, P.; GRAMS, R.; VICHASRI-GRAMS, S.. Analysis of a calcium-binding EF-hand protein family in *Fasciola gigantica*. *Experimental Parasitology*, v. 130, n. 4, p. 364–373, abr. 2012.

SWANSON, K. A. *et al.* Structural basis for immunization with postfusion respiratory syncytial virus fusion F glycoprotein (RSV F) to elicit high neutralizing antibody titers. *Proceedings of the National Academy of Sciences of the United States of America*, PMID: 21586636 PMCID: PMC3111287, v. 108, n. 23, p. 9619–9624, 7 jun. 2011.

TANAKA, M. *et al.* A comparison of the antischistosomal effect of levo- and dextro-praziquantel on *Schistosoma japonicum* and *S. mansoni* in mice. *The American Journal of Tropical Medicine and Hygiene*, PMID: 2505624, v. 41, n. 2, p. 198–203, ago. 1989.

TANEJA, B. *et al.* Structure of the N-terminal fragment of topoisomerase V reveals a new family of topoisomerases. *The EMBO Journal*, v. 25, n. 2, p. 398–408, 25 jan. 2006.

TANNER, N.; PRAG, G. Purification and crystallization of yeast Ent1 ENTH domain. *Acta Crystallographica Section F Structural Biology and Crystallization Communications*, v. 68, n. 7, p. 820–823, 1 jul. 2012.

TEIGELKAMP, S. *et al.* Calcium-dependent complex assembly of the myeloid differentiation proteins MRP-8 and MRP-14. *The Journal of Biological Chemistry*, PMID: 2071612, v. 266, n. 20, p. 13462–13467, 15 jul. 1991.

THÉTIOT-LAURENT, S. A.-L. *et al.* Schistosomiasis Chemotherapy. *Angewandte Chemie International Edition*, v. 52, n. 31, p. 7936–7956, 29 jul. 2013.

- TJANDRA, N. *et al.* Rotational Dynamics of Calcium-Free Calmodulin Studied by ¹⁵N-NMR Relaxation Measurements. *European Journal of Biochemistry*, v. 230, n. 3, p. 1014–1024, 28 jun. 2008.
- U OLVEDA, David. Bilharzia: Pathology, Diagnosis, Management and Control. *Tropical Medicine & Surgery*, v. 01, n. 04, 2013.
- VALE, R. D. The Molecular Motor Toolbox for Intracellular Transport. *Cell*, v. 112, n. 4, p. 467–480, fev. 2003.
- VALLEE, R. B.; SHEETZ, M. P. Targeting of motor proteins. *Science (New York, N.Y.)*, PMID: 8599110, v. 271, n. 5255, p. 1539–1544, 15 mar. 1996.
- VAN DER WERF, M. J. *et al.* Quantification of clinical morbidity associated with schistosome infection in sub-Saharan Africa. *Acta Tropica*, PMID: 12745133, v. 86, n. 2-3, p. 125–139, maio 2003.
- VICHASRI-GRAMS, S. *et al.* An analysis of the calcium-binding protein 1 of *Fasciola gigantica* with a comparison to its homologs in the phylum Platyhelminthes. *Molecular and Biochemical Parasitology*, v. 146, n. 1, p. 10–23, mar. 2006.
- VIJAY-KUMAR, S.; KUMAR, V. D. Crystal structure of recombinant bovine neurocalcin. *Nature Structural Biology*, PMID: 9886296, v. 6, n. 1, p. 80–88, jan. 1999.
- WANG, W. *et al.* Structure of the monomeric 8-kDa dynein light chain and mechanism of the domain-swapped dimer assembly. *The Journal of Biological Chemistry*, PMID: 12904292, v. 278, n. 42, p. 41491–41499, 17 out. 2003.
- WARREN, K. S. The pathology, pathobiology and pathogenesis of schistosomiasis. *Nature*, PMID: 351411, v. 273, n. 5664, p. 609–612, 22 jun. 1978.
- WATERHOUSE, A. M. *et al.* Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, v. 25, n. 9, p. 1189–1191, 1 maio 2009.
- WEBBE, G.; JAMES, C. A comparison of the susceptibility to praziquantel of *Schistosoma haematobium*, *S. japonicum*, *S. mansoni*, *S. intercalatum* and *S. mattheei* in hamsters. *Zeitschrift für Parasitenkunde*, v. 52, n. 2, p. 169–177, 1 jan. 1977.
- WERNIMONT, A.; EDWARDS, A. In Situ Proteolysis to Generate Crystals for Structure Determination: An Update. *PLoS ONE*, v. 4, n. 4, p. e5094, 7 abr. 2009.
- WHO. Report of the WHO informal consultation on the use of praziquantel during pregnancy/lactation and albendazole/ mebendazole in children under 24 months (WHO/CDS/CPE/ PVC/2002.4). In: : WORLD HEALTH ORGANIZATION, 2002, Geneva. *Anais...* Geneva: [s.n.], 2002.
- XIANG, S. *et al.* Crystal Structure of 1-Deoxy-D-xylulose 5-Phosphate Synthase, a Crucial Enzyme for Isoprenoids Biosynthesis. *Journal of Biological Chemistry*, v. 282, n. 4, p. 2676–2682, 26 jan. 2007.

XIAO, S.-h.; CATTO, B. A.; WEBSTER, L. T. Effects of Praziquantel on Different Developmental Stages of *Schistosoma mansoni* in Vitro and in Vivo. *Journal of Infectious Diseases*, v. 151, n. 6, p. 1130–1137, 1 jun. 1985.

YAP, K. L. *et al.* Diversity of conformational states and changes within the EF-hand protein superfamily. *Proteins*, PMID: 10591109, v. 37, n. 3, p. 499–507, 15 nov. 1999.

YAP, K. *et al.* Calmodulin Target Database. *Journal of Structural and Functional Genomics*, v. 1, n. 1, p. 8–14, 1 mar. 2000.

ZAYTSEVA, O. *et al.* The Novel Zinc Finger Protein dASCIZ Regulates Mitosis in *Drosophila* via an Essential Role in Dynein Light-Chain Expression. *Genetics*, v. 196, n. 2, p. 443–453, 1 fev. 2014.

ZHANG, M.; TANAKA, T.; IKURA, M. Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. *Nat Struct Mol Biol*, v. 2, n. 9, p. 758–767, set. 1995.

ZHANG, M.; YUAN, T. Molecular mechanisms of calmodulin's functional versatility. *Biochemistry and Cell Biology*, v. 76, n. 2-3, p. 313–323, maio 1998.

ZHOU, Y. *et al.* Prediction of EF-hand calcium-binding proteins and analysis of bacterial EF-hand proteins. *Proteins*, PMID: 16981205, v. 65, n. 3, p. 643–655, 15 nov. 2006.

ZHOU, Z. *et al.* Molecular cloning and identification of a novel *Clonorchis sinensis* gene encoding a tegumental protein. *Parasitology Research*, PMID: 17476530, v. 101, n. 3, p. 737–742, ago. 2007.

ZHOU, Z. *et al.* Oral administration of a *Bacillus subtilis* spore-based vaccine expressing *Clonorchis sinensis* tegumental protein 22.3 kDa confers protection against *Clonorchis sinensis*. *Vaccine*, PMID: 18329763, v. 26, n. 15, p. 1817–1825, 28 mar. 2008.

ZULAUF, M.; D'ARCY, A. Light scattering of proteins as a criterion for crystallization. *Journal of Crystal Growth*, v. 122, n. 1-4, p. 102–106, ago. 1992.

9 – Apêndice

CARACTERIZAÇÃO ESTRUTURAL E FUNCIONAL DE MUTANTES DA ENZIMA NahF DE *Pseudomonas putida* G7

O projeto inicial aprovado para este Doutorado no Departamento de Bioquímica e Imunologia, ICB, UFMG, foi intitulado “Clonagem, expressão e caracterização estrutural de UbiG e UbiE: metiltransferases envolvidas na via de biossíntese de ubiquinona em *E. coli*.” Durante dois anos, esse projeto foi desenvolvido no laboratório de Biologia Estrutural do referido departamento. A proteína UbiE apresentou-se insolúvel já nos primeiros experimentos, impossibilitando a continuidade dos trabalhos com a mesma para fins de determinação de estrutura. Já a proteína UbiG foi expressa e cristalizada com sucesso, entretanto, um grupo estrangeiro publicou um artigo científico (doi: 10.1107/S1744309111014278) com resultados idênticos aos obtidos nesse trabalho, o que tornaria muito difícil a publicação desses resultados para posterior defesa.

Assim, após dois anos de trabalho no projeto inicial, foi necessário abandoná-lo e iniciar novos trabalhos de pesquisa, mas com um tempo bastante reduzido para obtenção dos resultados. Dois projetos foram desenvolvidos nos dois anos restantes do Doutorado. O primeiro projeto intitulado “Caracterização estrutural e funcional de mutantes da enzima NahF de *Pseudomonas putida* G7” forneceu resultados interessantes, mas ainda insuficientes para uma publicação. Já o segundo, intitulado “Caracterização estrutural da proteína Sm21.7 recombinante de *Schistosoma mansoni*” forneceu resultados que foram publicados em um artigo e constituem a presente tese.

Como não foi possível ainda finalizar o projeto com os mutantes da enzima NahF, alguns resultados do mesmo são apresentados na forma de apêndice. Vale ressaltar que este trabalho foi realizado em parte em um doutorado sanduiche na Utah State University sob a supervisão do Dr. Alvan Hengge e com o apoio financeiro da CAPES.

1 – Introdução

1.1 – Hidrocarbonetos aromáticos policíclicos

Os hidrocarbonetos aromáticos policíclicos (HAPs) são uma classe de compostos orgânicos hidrofóbicos constituídos de dois ou mais anéis benzênicos fundidos (Revisado por HABE; OMORI, 2003). Esses compostos são amplamente distribuídos no ambiente, sendo naturalmente encontrados em combustíveis fósseis como carvão e petróleo. Também ocorre a formação de HAPs durante a combustão incompleta de matéria orgânica, tais como carvão, diesel, madeira e vegetação e em processos como erupções vulcânicas. Outras fontes de contaminação são o derramamento de petróleo, diesel e outros processos industriais como liquefação do carvão (Revisado por BAMFORTH; SINGLETON, 2005). Esses compostos são de difícil remoção e podem permanecer no ambiente por longos períodos devido à sua hidrofobicidade e baixa solubilidade em água (CERNIGLIA, 1992).

Um dos grandes interesses nos estudos com HAPs deve-se a suas propriedades mutagênicas e carcinogênicas. As primeiras descrições relacionando HAPs ao aparecimento de tumores datam do século XVIII, quando foi observado que o desenvolvimento de câncer em limpadores de chaminé poderia ser decorrente da exposição ocupacional (Revisado por BOSTRÖM *et al.*, 2002). Posteriormente, um grande número de pesquisas demonstraram os mecanismos pelos quais os HAPs atuam na carcinogênese. Assim, com base em sua abundância e toxicidade, 16 HAPs foram identificados pela EPA (*U.S. Environmental Protection Agency*) como poluentes prioritários nos programas de descontaminação ambiental (KEITH; TELLIARD, 1979).

No contexto de descontaminação ambiental, uma estratégia promissora é a biorremediação, que pode ser descrita como o processo no qual microrganismos são utilizados para reduzir a concentração e toxicidade de poluentes como petróleo, bifenis policlorados, compostos nitro aromáticos, solventes industriais, pesticidas e metais, além dos HAPs (Revisado por FERNANDEZ *et al.*, 2012).

Um exemplo da utilização da biorremediação ocorreu após um acidente com o petroleiro *Exxon Valdez*, em Prince William Sound, Alasca, que causou o vazamento de cerca de 41 milhões de litros de petróleo no oceano em 1989. Uma das abordagens utilizadas para descontaminação foi a adição de fertilizantes ao local contaminado para aumentar a quantidade de microrganismos presentes e, assim, a velocidade de

degradação dos contaminantes (ATLAS; HAZEN, 2011; LINDSTROM *et al.*, 1991). Após o tratamento, avaliações demonstraram a redução significativa dos poluentes e foi oficialmente declarada a conclusão da descontaminação (Revisado por ATLAS; HAZEN, 2011).

Nas últimas décadas uma grande variedade de microrganismos capazes de degradar HAPs foi isolada e as vias metabólicas responsáveis por essa atividade foram caracterizadas. Tais microrganismos consistem em bactérias, fungos lignofílicos, fungos não lignofílicos, algas e cianobactérias (BAMFORTH; SINGLETON, 2005; CERNIGLIA, 1992).

Assim, considerando a importância da biorremediação nos processos atuais de descontaminação, o conhecimento dos microrganismos, vias metabólicas e enzimas capazes de degradar esses compostos tóxicos é crucial para a utilização desse tipo de tratamento.

1.2 – Degradação do naftaleno por bactérias

O naftaleno é um hidrocarboneto aromático bicíclico (Figura 1) comumente encontrado no ambiente. Atualmente é utilizado como modelo no estudo de degradação de HAPs, uma vez que é o mais simples e solúvel desses compostos (EATON; CHAPMAN, 1992).

O metabolismo do naftaleno (Figura 1) é um processo bem estudado e caracterizado na espécie *Pseudomonas putida* G7, que possui um plasmídeo transmissível por conjugação denominado NAH7 (DUNN; GUNSALUS, 1973). Esse plasmídeo, de aproximadamente 83kb (SOTA *et al.*, 2006), contém os genes codificadores das enzimas da via de degradação do naftaleno, denominados genes *nah*, organizados em três operons. O primeiro operon contém os genes que codificam as enzimas que convertem naftaleno a salicilato (*nahAaAbAcAdBFCED*) e constituem a via superior de degradação (EATON; CHAPMAN, 1992; YEN; GUNSALUS, 1982). Já o segundo operon contém os genes codificadores de enzimas da via inferior de degradação na qual ocorre a transformação de salicilato aos intermediários do ciclo do ácido tricarboxílico, piruvato e acetaldeído (*nahGTHINLOMKJ*) (HABE; OMORI, 2003; YEN; GUNSALUS, 1982). O terceiro operon, é constituído por um gene que codifica para uma proteína regulatória denominada *nahR* (SCHELL; WENDER, 1986; YEN; GUNSALUS, 1982).

A última etapa da via superior de degradação do naftaleno é a conversão de 2-hidroxibenzilaldeído a salicilato (DAVIES, J. I.; EVANS, W. C, 1964; EATON; CHAPMAN, 1992) pela enzima NahF, uma aldeído desidrogenase (ALDH) (DAVIES, J. I.; EVANS, W. C, 1964).

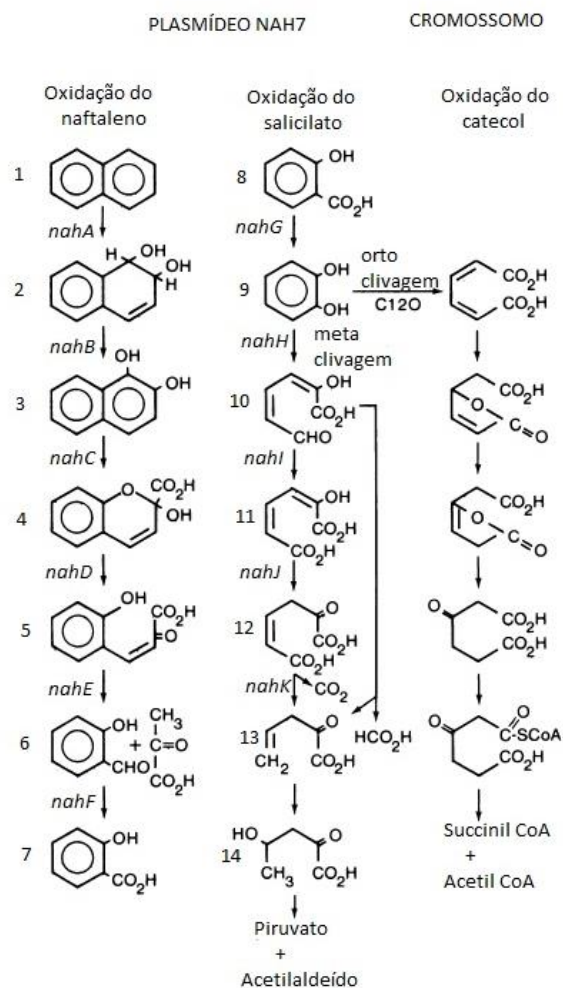


Figura 1: Via de degradação do naftaleno por *P. putida* G7. O esquema representa as vias superior e inferior de degradação do naftaleno. A via superior compreende a ação das enzimas NahA a NahF (naftaleno (1) a salicilato (7)) e a via inferior transforma salicilato (8) a intermediários do ciclo do ácido cítrico pela oxidação do catecol via meta-clivagem (NahG a NahK). (Adaptado de YEN; GUNSALUS, 1985).

1.3 – As enzimas da família Aldeído Desidrogenases

As aldeídos desidrogenases são um grupo de enzimas que catalisam a conversão de aldeídos aos ácidos carboxílicos correspondentes utilizando NAD(P)^+ como cofator (Revisado por YOSHIDA *et al.*, 1998).

O mecanismo catalítico dessas enzimas é amplamente estudado e envolve as seguintes etapas: 1) ataque nucleofílico ao aldeído pelo grupo tiolato da cisteína catalítica (284 em NahF), levando à formação de um tio-hemiacetal tetraédrico intermediário covalentemente ligado à enzima; 2) transferência de um hidreto para o anel piridina do NAD(P)^+ (oxidação), que produz um tioéster intermediário e 3) hidrólise do tioéster intermediário resultante (desacilação), que resulta em um segundo intermediário tetraédrico com a liberação do ácido carboxílico e dissociação do cofator reduzido (Figura 2) (Revisado por KOPPAKA *et al.*, 2012).

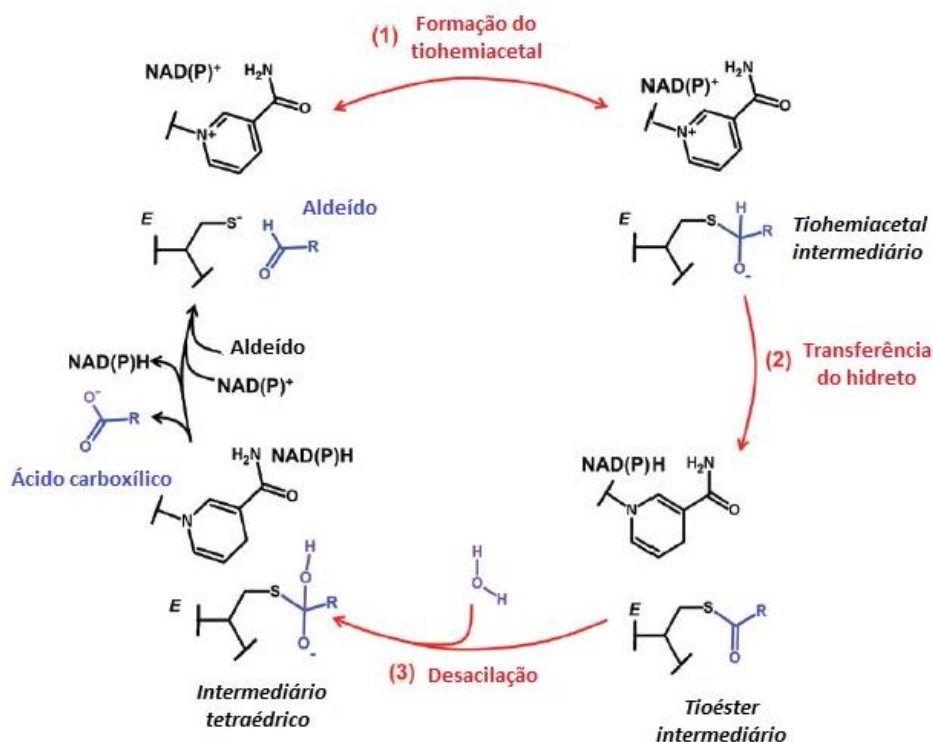


Figura 2: Mecanismo catalítico das aldeído desidrogenases. As etapas da catálise dessas enzimas descritas na imagem são: 1) ataque nucleofílico ao aldeído eletrofílico pelo grupo tiolato da cisteína ativada formando um tio-hemiacetal tetraédrico intermediário 2) transferência de um hidreto para o anel piridina do NAD(P)^+ com formação do tioéster intermediário 3) hidrólise do tioéster intermediário resultante e dissociação do cofator reduzido e subsequente regeneração da enzima (Adaptado de MUÑOZ-CLARES; GONZÁLEZ-SEGURA; DÍAZ-SÁNCHEZ, 2011).

Além da cisteína catalítica descrita, análises de modificação química e mutagênese sítio dirigida identificaram o outro resíduo catalítico, um glutamato (250 em NahF) (ABRIOLA *et al.*, 1987; ABRIOLA; MACKERELL; PIETRUSZKO, 1990; LEE *et al.*, 2011; MARCHAL; RAHUEL-CLERMONT; BRANLANT, 2000; VEDADI; MEIGHEN, 1997; WANG; WEINER, 1995; WENZL *et al.*, 2009, p.). Esse glutamato foi identificado como a base geral que interage com a cisteína do sítio ativo para ativação inicial da mesma (LEE *et al.*, 2011; STEINMETZ *et al.*, 1997) ou com uma molécula de água que atua na hidrólise para liberação do produto durante a catálise (TSYBOVSKY *et al.*, 2013; WANG; WEINER, 1995).

Outro resíduo importante é uma asparagina (149 em NahF), que foi identificado como parte do *oxyanion hole* envolvido na estabilização dos intermediários da reação que estão covalentemente ligados à enzima (COBESSI *et al.*, 2000; STEINMETZ *et al.*, 1997) e na ligação e posicionamento do substrato para o ataque nucleofílico (MUÑOZ-CLARES; GONZÁLEZ-SEGURA; DÍAZ-SÁNCHEZ, 2011).

Na primeira estrutura de ALDH resolvida por cristalografia demonstrou-se que a mesma é uma molécula dimérica que possui um domínio de ligação ao NAD(P)⁺, um domínio catalítico e um domínio de oligomerização (LIU *et al.*, 1997) (Figura 3).

Muitas enzimas utilizam dinucleotídeos, NAD e FAD, como cofatores e a maioria dos domínios que ligam dinucleotídeos tem uma estrutura típica de *Rossmann fold*. O domínio de ligação ao NAD(H)⁺ das ALDH possui uma variação do tradicional *Rossmann fold*, que, nessas enzimas, é constituído de 5 fitas- β conectadas por 4 α -hélices, em vez de seis fitas- β usualmente observadas nas desidrogenases dependentes de NAD(P)⁺ (LIU *et al.*, 1997; ROSSMANN; MORAS; OLSEN, 1974).

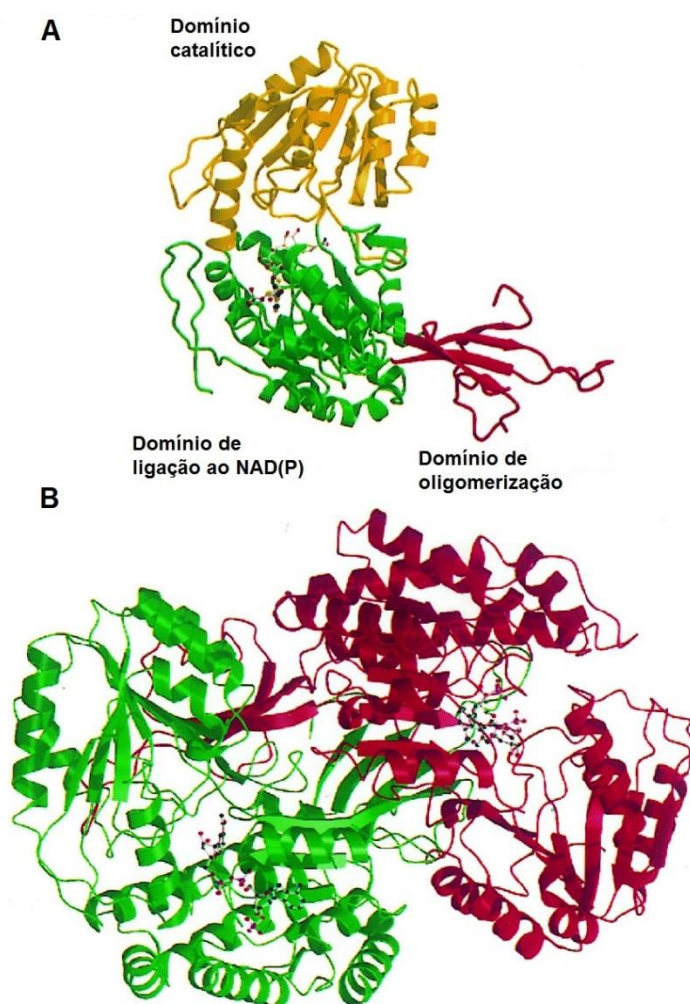


Figura 3: Estrutura de uma aldeído desidrogenase de *Vibrio harveyi*. (A) A imagem destaca os domínios de ligação ao NAD(P)⁺ (verde), de oligomerização (vermelho) e catalítico (amarelo) característicos dessa classe de enzimas. (B) Forma dimérica da ALDH de *Vibrio harveyi* com cada monômero destacado em uma cor, verde ou vermelho (Adaptado de AHVAZI et al., 2000).

Na via de degradação do naftaleno por *P. putida* G7, uma ALDH converte salicialdeído a salicilato. Essa enzima, denominada NahF (salicilaldeído desidrogenase), é codificada pelo gene *nahF* (DAVIES, J. I.; EVANS, W. C., 1964).

A enzima NahF recombinante foi caracterizada recentemente por nosso grupo e a proteína foi cristalizada na presença do substrato salicilaldeído (Figura 4). A estrutura tridimensional da mesma em uma resolução de 2,4Å demonstrou a presença dos domínios catalítico, de oligomerização e de ligação ao NAD⁺ (PDB 4JZ6) (COITINHO et al., 2012).

Baseado nos dados estruturais de NahF, além da cisteína (284) e glutamato (250) catalíticos e da asparagina (149) que estabiliza o *oxyanion hole* (Figura 4), observou-se

que alguns outros resíduos de aminoácidos podem estar envolvidos na interação com o substrato nativo, o salicilaldeído. Um desses resíduos, arginina 157, foi selecionado para avaliação através de mutagênese sítio dirigida. Esse não é um resíduo conservado, mas é positivamente carregado e aparentemente, está envolvido de alguma forma na catálise devido a sua posição no sítio ativo (figura 4). Assim, esse resíduo foi substituído, por alanina, um resíduo sem carga e por lisina, um resíduo também positivamente carregado e com tamanho semelhante à arginina (R157A, R157K).

Outra mutação em resíduos não catalíticos foi realizada na glicina 150 (Figura 4). Essa mutação foi baseada na avaliação da atividade enzimática de NahF utilizando-se várias moléculas como substrato, demonstrando que a enzima catalisa com maior eficiência substratos aromáticos (Coitinho et al, em preparação). Adicionalmente, a comparação estrutural das ALDH NahF e NahI (também determinada em nosso laboratório) demonstrou que na posição correspondente à glicina 150 em NahF, existe uma leucina em NahI. Tentativas de realizar testes de atividade enzimática com NahI não obtiveram sucesso com a utilização de substratos aromáticos. Nossa hipótese é que a presença da glicina forma um bolsão catalítico maior em NahF permitindo a entrada de substratos maiores e consequente catálise. Assim, foram realizadas as seguintes mutações: G150A, G150V, G150L e G150F com o intuito de verificar se o aumento do resíduo nessa posição tornaria o bolsão inacessível aos substratos aromáticos.

Assim, as mutações podem auxiliar no entendimento do mecanismo de ação da enzima NahF e a preferência por substratos aromáticos, em oposição à outra ALDH existente na via de degradação do naftaleno, a NahI. Em adição às mutações já citadas, foram feitas algumas nos resíduos catalíticos: C284A, C284S, E250A, E250Q e N149A.

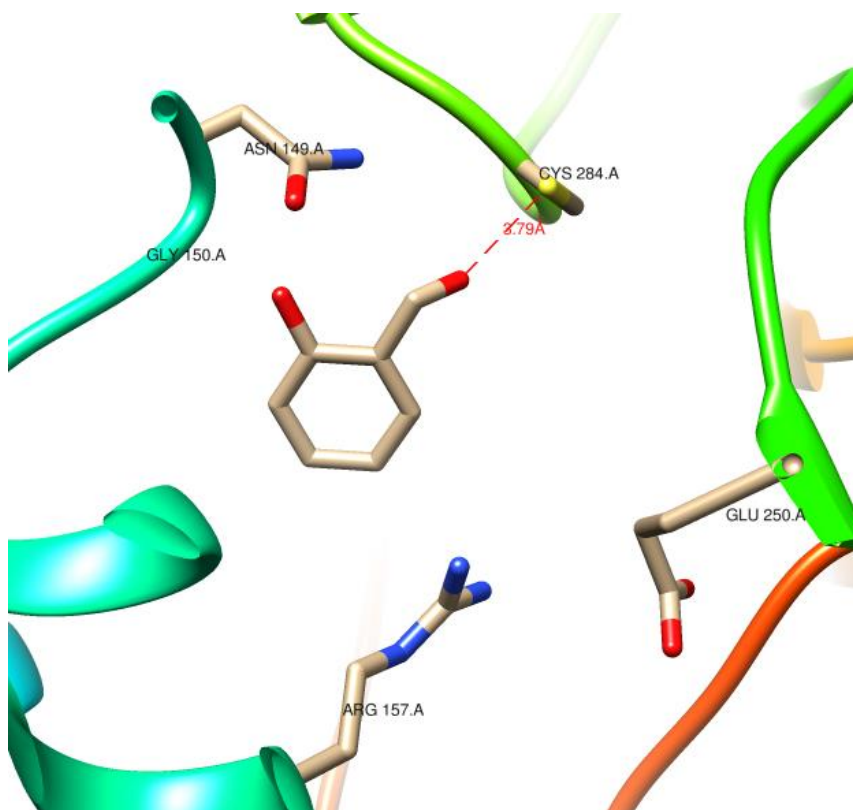


Figura 4: Posição de alguns resíduos e do substrato no sítio ativo da enzima NahF (PDB 4JZ6). Após determinação da estrutura tridimensional dessa enzima, foi possível observar o substrato salicilaldeído no sítio ativo. Também são visualizados na imagem os resíduos catalíticos Cys284 e Glu250, Asn149 que estabiliza o *oxyanion hole*, além de Arg157 e Gly150, resíduos que foram alterados por mutagênese sítio dirigida. A distância entre o átomo S da cisteína reativa e o aldeído é de 3,794Å. A imagem foi gerada pelo programa UCSF Chimera (PETTERSEN et al., 2004).

2 – Materiais e métodos

2.1 – Mutações sítio dirigidas na sequência codificadora da enzima NahF

A proteína NahF nativa foi previamente clonada no vetor pET28a(TEV) (COITINHO *et al.*, 2012), que possui uma sequência codificadora para uma cauda de histidinas N-terminal. As mutações foram inseridas no DNA nativo utilizando o kit *QuikChange Lightning Site-Directed Mutagenesis* (Stratagene). Basicamente, o procedimento utiliza o vetor recombinante nahF-pET28a(TEV) e iniciadores direto e reverso contendo cada mutação a ser realizada (Tabela 1). A enzima *PfuUltra HF DNA polymerase* é utilizada para a síntese do plasmídeo contendo a mutação a partir dos iniciadores e, posteriormente, a enzima *DpnI* é utilizada para digerir o DNA inicial. A

síntese dos plasmídeos contendo as mutações foi realizada segundo o protocolo do fabricante.

Após sequenciamento dos plasmídeos para confirmar a presença da mutação, cada um foi transformado em bactérias *E. coli* Arctic Express (DE3) (item 4.1.4) e as colônias foram selecionadas em placas de meio LB-ágar contendo gentamicina (20µg/mL) e canamicina (50µg/mL).

Tabela 1: Mutações realizadas na enzima NahF e os iniciadores utilizados no processo de mutagênese (o códon alterado está destacado)

Mutação		Iniciadores
N149A	Direto	5' GCATCGTTCCATGGG <u>CCG</u> GCACCGCAGTGCTG 3'
	Reverso	5' CAGCACTGCGGTG <u>CCG</u> CCCATGGAACGATGC 3'
C284A	Direto	5' CTATTCCAAGGTCAGATCG <u>CC</u> ATGTCCACTGAGCGCTTGG 3'
	Reverso	5' CCAAGCGCTCAGTGGACATGGCGATCTGACCTTGAATAG 3'
C284S	Direto	5' CTATTCCAAGGTCAGATCAG <u>C</u> ATGTCCACTGAGCGCTTGG 3'
	Reverso	5' CAAGCGCTCAGTGGACATGCTGATCTGACCTTGAATAG 3'
E250A	Direto	5' GCTGCCTACTGG <u>CG</u> CTCGGCGGCAA 3'
	Reverso	5' TTGCCGCCGAGCG <u>CC</u> CAGTAGGCAGC 3'
E250Q	Direto	5' GCTGCCTACTGCAGCTCGGCGGC 3'
	Reverso	5' GCCGCCGAGCTGCAGTAGGCAGC 3'
G150A	Direto	5' CGTTCCATGGAACGCCACCGCAGTGCTGG 3'
	Reverso	5' CCAGCACTGCGGTGGCGTTCCATGGAACG 3'
G150F	Direto	5' GCATCGTTCCATGGAAC <u>TT</u> CACCGCAGTGCTGGCGG 3'
	Reverso	5' CCGCCAGCACTGCGGTGAAGTTCCATGGAACGATGC 3'
G150V	Direto	5' CGTTCCATGGAACGTCACCGCAGTGCTGGC 3'
	Reverso	5' GCCAGCACTGCGGTGACGTTCCATGGAACG 3'
G150L	Direto	5' CATCGTTCCATGGAACCTCACCAGTGCTGGCGG 3'
	Reverso	5' CCGCCAGCACTGCGGTGAGGTTCCATGGAACGATG 3'
R157A	Direto	5' CGCAGTGCTGGCGGCAGCAGCCATCGCTT 3'
	Reverso	5' AAGCGATGGCTGCTGCCGCCAGCACTGCG 3'
R157K	Direto	5' CCGCAGTGCTGGCGGCAAAAGCCATCGCTTA 3'
	Reverso	5' TAAGCGATGGCTTTTGCCGCCAGCACTGCGG 3'

2.2 – Expressão em sistema procariótico, lise bacteriana e purificação dos mutantes de NahF

A expressão e lise bacteriana dos mutantes de NahF foram realizadas de acordo com Coitinho e colaboradores, 2012 (COITINHO *et al.*, 2012). A expressão das proteínas mutantes e sua presença na fração solúvel foram verificadas através de SDS-PAGE. Apesar da síntese dos 11 plasmídeos mutantes e transformação dos mesmos em bactérias, apenas os mutantes NahF-R157A e NahF-R157K foram expressos e purificados até o momento.

Após expressão e lise, os mutantes de NahF foram purificados como citado em 4.3. Posteriormente, as proteínas foram concentradas (Vivaspin 20 MWCO 10 kDa - GE Healthcare) para 15 mg/mL para realização dos testes de cristalização. Já para os testes de cinética, as proteínas foram concentradas para 5mg/mL e posteriormente congeladas a -80°C com 30% (v/v) de glicerol.

2.3 – Determinação da estrutura dos mutantes de NahF por cristalografia de raios X

2.3.1 – Cristalização dos mutantes de NahF

Os mutantes de NahF foram cristalizados pela técnica da gota suspensa (item 4.8.1). As condições de cristalização para a NahF foram estabelecidas anteriormente (COITINHO *et al.*, 2012) e os mutantes foram cristalizados em condições similares. Para capturar o substrato no sítio ativo, a proteína foi incubada em banho de gelo com o salicilaldeído na concentração de 5mM (pelo menos 10 vezes mais que a concentração molar da proteína) por períodos de tempo variando entre 30 minutos e 2 horas. Posteriormente essa mesma proteína foi utilizada nos testes de cristalização.

2.3.2 – Coleta de dados de difração de raios X dos mutantes de NahF, determinação das fases e refinamento das estruturas

Os cristais dos mutantes de NahF foram submetidos a experimentos de difração de raios-X na linha W01B-MX2 (GUIMARÃES *et al.*, 2009) no LNLS. Antes da exposição dos cristais aos raios X, os mesmos foram congelados como citado em 4.8.2.

O cristal do mutante NahF-R157A foi exposto aos raios X por 20 segundos para coleta das imagens. O mesmo foi girado de 0,5 graus para coleta de cada imagem, totalizando 180 graus de rotação e 360 imagens. Já o cristal do mutante NahF-R157K foi exposto por 15 segundos para coleta de cada imagem e girado de 0,5 graus, totalizando 360 imagens e 180 graus de rotação. As imagens coletadas foram gravadas em um detector Rayonix MarMosaic 225 CCD.

Após processamento dos dados como citado em 4.8.3, o método de Substituição Molecular foi usado para solucionar o problema das fases utilizando-se o programa PHENIX (ADAMS *et al.*, 2010) e a estrutura da enzima NahF *wild type* como modelo (Coitinho *et al.*, em preparação). Posteriormente, as estruturas foram refinadas como descrito em 4.8.4.

2.4 – Testes de atividade enzimática dos mutantes de NahF

Os testes de cinética com os mutantes da enzima NahF foram realizados em um espectrofotômetro Cary50 Bio utilizando salicilaldeído como substrato. A mistura da reação consistiu em tampão Hepes pH 8,5, 1mM EDTA e força iônica de 0,2M, ajustada pela adição de NaCl. O substrato foi dissolvido em acetonitrila antes de sua adição à mistura da reação em uma cubeta de quartzo. Posteriormente, o cofator NAD⁺ foi adicionado em uma concentração final de 200µM. A reação foi iniciada pela adição da enzima em uma concentração previamente determinada. As medidas de velocidade inicial para a reação enzimática foram monitoradas por espectrofotometria a 340nm.

Em um gráfico de absorbância em função do tempo, a inclinação da curva pode ser determinada, correspondendo à velocidade inicial da reação. Posteriormente, um gráfico de Michaelis-Menten (velocidade inicial (*v*) *versus* concentração de substrato [*S*]) foi obtido. Os parâmetros do estado estacionário foram determinados por regressão não linear desses dados utilizando o programa PRISMA.

3 – Resultados preliminares e discussão

3.1 – Mutagênese sítio dirigida, expressão e purificação das proteínas recombinantes

Utilizando o kit *QuikChange Lightning Site-Directed Mutagenesis* (Stratagene) 11 mutantes da enzima NahF foram sintetizados (N149A, E250A, E250Q, C284A, C284S, R157A, R157K, G150A, G150V, G150L e G150F). Como já citado, apesar da síntese desses plasmídeos mutados, apenas as proteínas NahF-R157A e NahF-R157K foram expressas e purificadas até o presente momento.

Após a expressão em bactérias *E. coli* Arctic Express (DE3), as proteínas foram purificadas por cromatografias de afinidade e exclusão molecular e avaliadas por SDS-PAGE (Figuras 5 e 6). A segunda etapa de purificação (exclusão molecular) foi utilizada como uma forma de polimento da amostra. O cromatograma obtido apresentou um pico principal de eluição em um volume de 79 ml, o que corresponde a um dímero com 100 kDa, de acordo com a curva de calibração da coluna com proteínas de massa molecular específica (ver figura 24).

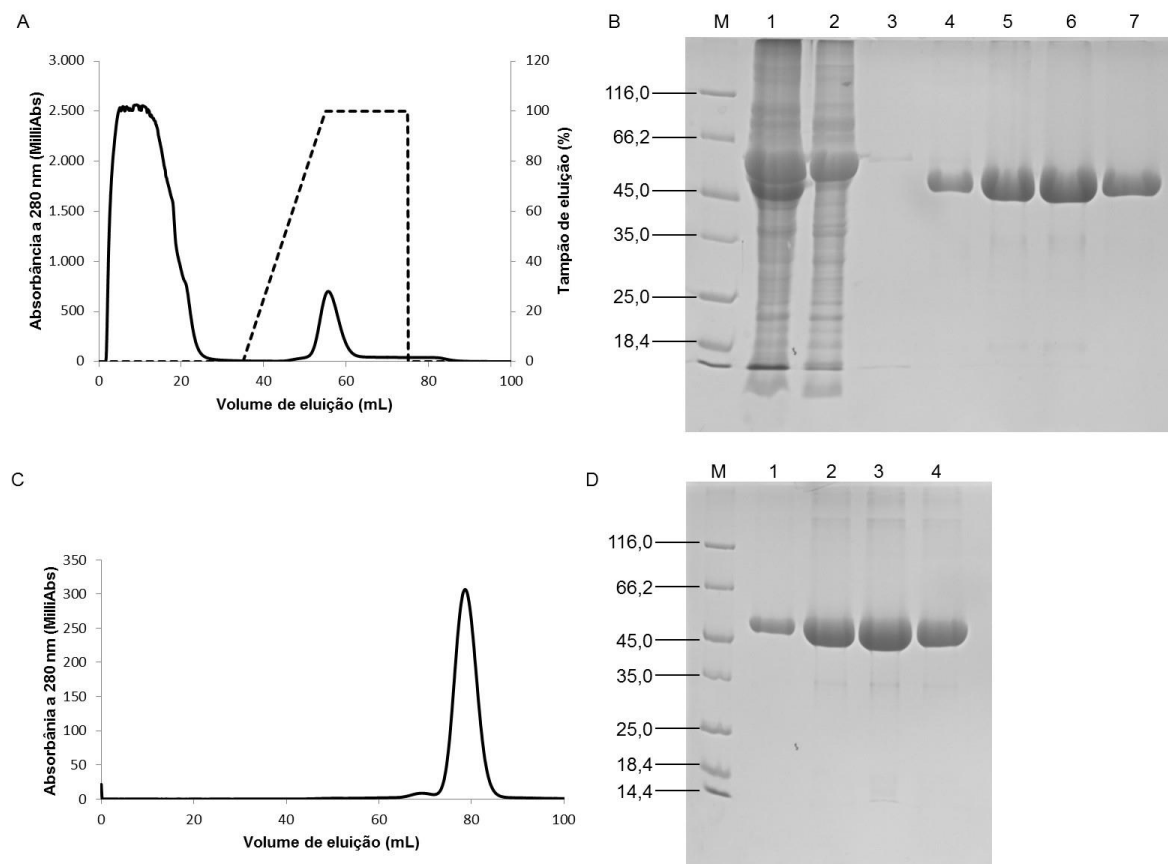


Figura 5: Comatografias de afinidade e exclusão molecular do mutante NahF-R157A. Os cromatogramas obtidos são apresentados em A (Afinidade) e C (Exclusão Molecular). A pureza das amostras purificadas foi analisada por SDS-PAGE corado com *Coomassie Brilliant Blue R250*. B – Análise da cromatografia de afinidade por SDS-PAGE. M – marcador de massa molecular em kDa (*Unstained Protein Molecular Weight Marker* – Fermentas), 1 – fração solúvel, 2 – frações referentes ao *flow through*, 3 – frações referentes à lavagem, 4 a 7 – frações correspondentes à proteína de interesse eluída. D – Análise da cromatografia de exclusão molecular por SDS-PAGE. M – marcador de massa molecular em kDa (*Unstained Protein Molecular Weight Marker* – Fermentas), 1 a 4 – frações eluídas da coluna.

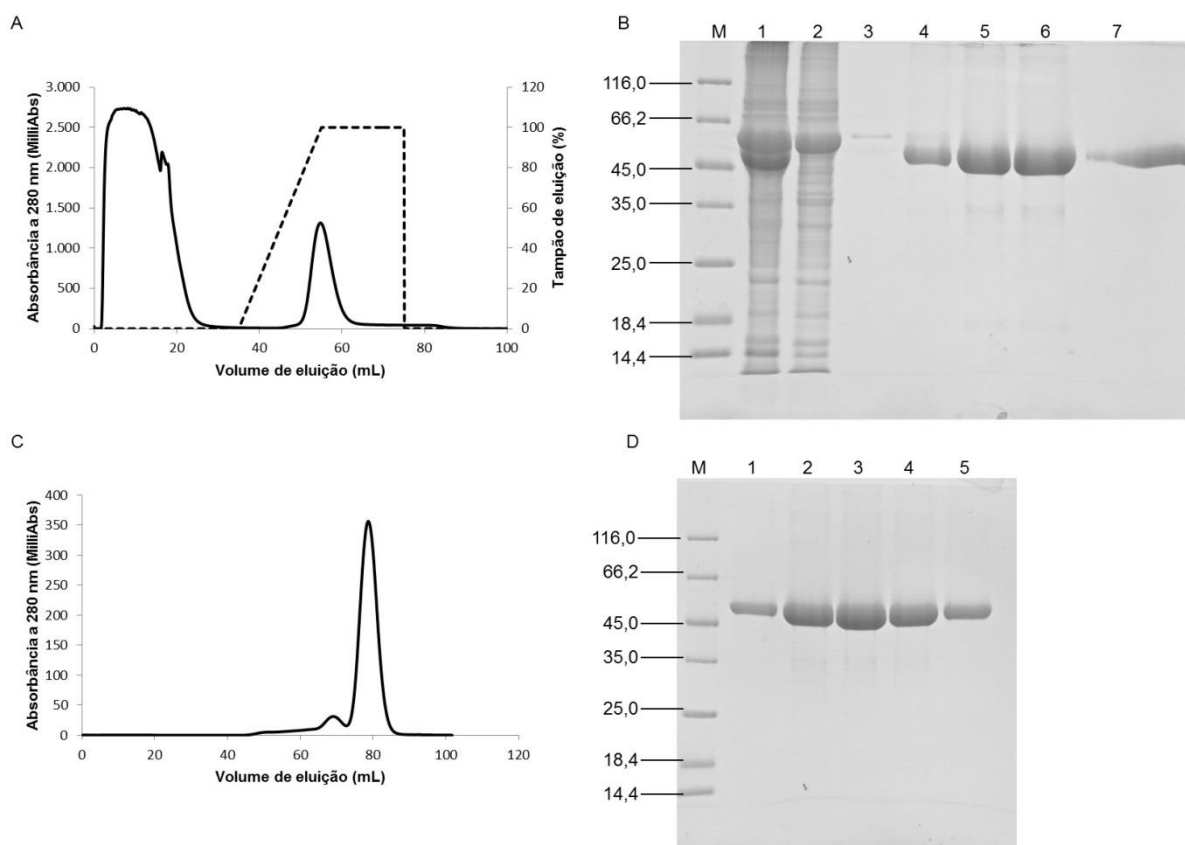


Figura 6: Cromatografias de afinidade e exclusão molecular do mutante NahF-R157K. Os cromatogramas obtidos são apresentados em A (Afinidade) e C (Exclusão Molecular). A pureza das amostras purificadas foi analisada por SDS-PAGE. B – Análise da cromatografia de afinidade por SDS-PAGE corado com *Coomassie Brilliant Blue R250*. M – marcador de massa molecular em kDa (*Unstained Protein Molecular Weight Marker* – Fermentas), 1 – fração solúvel, 2 – frações referentes ao *flow through*, 3 – frações referentes à lavagem, 4 a 7 – frações correspondentes à proteína de interesse eluída. D – Análise da cromatografia de exclusão molecular por SDS-PAGE. M – marcador de massa molecular em kDa (*Unstained Protein Molecular Weight Marker* – Fermentas), 1 a 5 – frações eluídas da coluna.

3.2 – Cristalização dos mutantes de NahF, coleta de dados de difração de raios X e processamento dos dados

Os mutantes de NahF foram cristalizados em condições contendo tampão acetato de sódio, sulfato de amônio e isopropanol (Figura 7). De todos os cristais testados, os que alcançaram maior resolução foram obtidos nas seguintes condições (Tabela 2):

Tabela 2: Condições de cristalização dos mutantes NahF-R157A e NahF-R157K. Essas são variações da condição inicial da proteína *wild type*: 0,1M acetato de sódio pH 5,0, 1,5M sulfato de amônio, 5% (v/v) isopropanol.

Condição de cristalização	
NahF-R157A	0,1M acetato de sódio pH 4,0, 1M sulfato de amônio, 5% (v/v) isopropanol na presença de 5mM de NAD ⁺
NahF-R157K	0,1M acetato de sódio pH 5,0, 1,2M sulfato de amônio, 5% (v/v) isopropanol na presença de 5mM de salicilaldeído

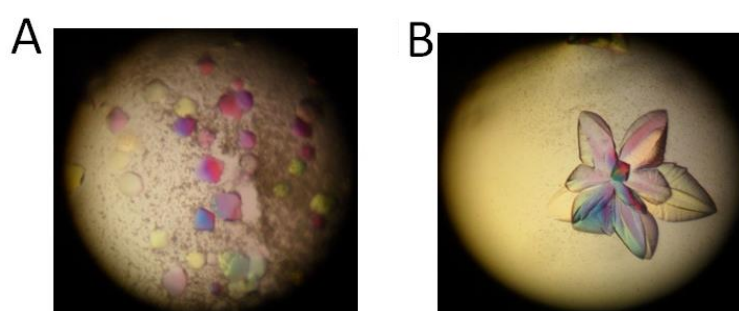


Figura 7: Cristais dos mutantes NahF-R157A e NahF-R157K obtidos após a otimização da condição inicial. Todos os cristais das proteínas foram obtidos em condições de cristalização contendo tampão acetato, sulfato de amônio e isopropanol. A metodologia utilizada foi a da gota suspensa. Nessa imagem os cristais de **NahF-R157A** (A) e **NahF-R157K** (B) são apresentados.

As principais estatísticas de coleta de dados e processamento dos mesmos são apresentadas na tabela 3.

Tabela 3: Estatísticas de coleta de dados e processamento dos cristais dos mutantes NahF-R157A e NahF-R157K. Os dados entre parênteses indicam os valores da menor e maior faixas de resolução, respectivamente.

	NahF-R157A	NahF-R157K
Comprimento de onda (Å)	1,459	1,459
Temperatura (K)	100	100
Distância cristal-detector (mm)	160	145
Rotação por imagem (°)	0,5	0,5
Tempo de exposição por imagem (s)	20	10
Grupo espacial	P6 ₄ 22	P6 ₄ 22
Parâmetros de célula unitária (Å)	a=b=170,53, c=158,23	a=b=168,95, c=157,20
Resolução (Å)	50,00-2,80 (50,00-6,03/2,90-2,80)	50,00-2,15 (50,00-4,63/2,23-2,15)
Nº de observações	350.168	561.794
Nº. of reflexes únicas	33.283 (3.576/3.260)	70.858 (7.404/6.856)
Completaza (%)	99,9 (99,5/100,0)	98,7(97,0/97,2)
(I/ σ(I))	18,1 (30,4/4,6)	16,4 (33,4/3,4)
Multiplicidade	10,5 (9,9/10,6)	8,0 (8,3/7,2)
R _{merge} ^a (%)	11,7 (5,3/51,7)	10(4,3/49,3)
Mosaicidade (°)	0,55	0,56
Coeficiente Matthews (Å ³ /Da)	6,1	5,9
Solvent content (%)	79,8	79,0

$$^a R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)} \times 100$$

3.3 – Determinação da estrutura tridimensional dos mutantes de NahF

Após a obtenção dos primeiros modelos através de Substituição Molecular, os mesmos foram refinados. A estrutura do mutante NahF-R157K foi refinada até valores de R_{work} e R_{free} de 18,1% e 19,9%, respectivamente. Já o refinamento da estrutura do mutante NahF-R157A não foi concluído até o momento e os valores de R_{work} e R_{free} estão atualmente em 21% e 24%, respectivamente. De todo o modo as estatísticas de refinamento de NahF-R157K e NahF-R157A podem ser observadas na tabela 4.

Tabela 4: Estatísticas de refinamento da estrutura dos mutantes NahF-R157A e NahF-R157K. O refinamento da estrutura do mutante NahF-R157A não foi concluído até o momento, por isto os valores apresentados para este mutante devem sofrer modificação até o fim do refinamento.

	NahF-R157A	NahF-R157K
$R_{work}^a(\%)$	21,1	18,1
$R_{free}^b(\%)$	24,39	19,9
Nº de átomos:		
Proteína	3.668	3.705
Água	55	394
Etilenoglicol	-	36
SO ₄	-	10
Salicilaldeído	-	10
Sódio	-	6
R.m.s.d:		
Comprimento de ligação(Å)	0,009	0,006
Ângulo de ligação(°)	1,2	0,99
Planaridade	0,005	0,004
Diedro	15,7	14,4
Gráfico de Ramachandran:		
Favoráveis (%)	95	97
Permitidos (%)	4	3
Outliers (%)	1	-

$$^a R_{work} = \sum_{hkl} |F_{obs}| - |F_{calc}| / \sum_{hkl} |F_{obs}|$$

^b R_{free} é o valor de R_{work} calculado para 5% do conjunto de dados não inclusos no refinamento

Mesmo não sendo possível finalizar até o presente momento o refinamento da estrutura do mutante NahF-R157A, um modelo parcial foi construído. Esse modelo apresentou o enovelamento característico da família das aldeído desidrogenases, constituído por um domínio catalítico, um domínio de ligação ao NAD⁺ e um domínio de oligomerização (Figura 8). Este mutante também se apresenta na forma de um dímero, apesar da presença de apenas uma molécula de NahF-R157A por unidade assimétrica.



Figura 8: Estrutura do mutante NahF-R157A e os domínios constituintes. Essa estrutura do mutante NahF-R157A a 2,8Å de resolução representa um monômero constituído dos domínios catalítico, em magenta (resíduos 251 a 458), domínio de ligação ao NAD⁺, em ciano (resíduos 1 a 116, 137 a 250 e 459 a 466) e domínio de oligomerização, em vermelho (resíduos 467 a 483). A figura foi gerada pelo programa UCSF Chimera (PETTERSEN et al., 2004).

Mesmo com um modelo incompleto da estrutura do mutante NahF-R157A, foi possível observar uma particularidade nesse mutante. Foi observado que a α -hélice constituída pelos resíduos de valina 153 a valina 164 sofre um pequeno desvio devido a uma alteração na orientação da leucina 154 (Figura 9). Observa-se na figura 9 que, quando a arginina 157 foi mutada para alanina, a leucina 154 sofre um deslocamento. Esse posicionamento da leucina é impossibilitado na enzima *wild type* por um provável choque estérico (Figura 9). Esse pequeno desvio na α -hélice poderia influenciar no correto posicionamento dos resíduos no sítio ativo, consequentemente reduzindo a eficiência catalítica da enzima, como será discutido posteriormente.

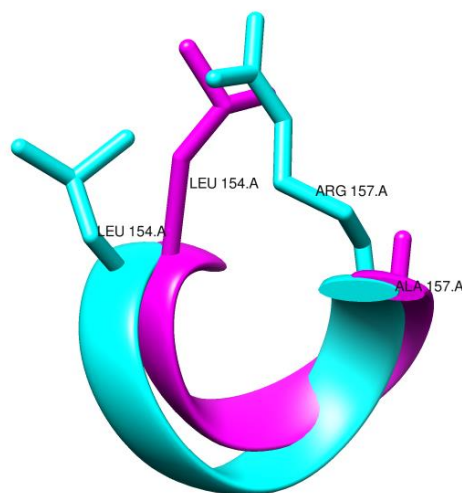


Figura 9: Sobreposição das estruturas da enzima NahF *wild type* e do mutante NahF-R157A, destacando os resíduos 154 e 157. Comparando-se ambas as estruturas, pode-se observar que na enzima contendo a mutação de arginina (ciano) para alanina (magenta) ocorre um deslocamento da leucina na posição 154. A figura foi gerada pelo programa UCSF Chimera (PETTERSEN et al., 2004).

A estrutura do mutante nahF-R157K foi resolvida e refinada. Ela também possui os 3 domínios característicos e uma organização dimérica. Após a determinação da estrutura, observou-se a existência de uma densidade eletrônica positiva no sítio ativo da enzima em formato semelhante ao que seria seu substrato (salicilaldeído) (Figura 10).

Após alguns ciclos de refinamento foi observado que a densidade eletrônica existente no sítio ativo, assemelha-se mais ao produto da reação catalisada pela NahF, o ácido salicílico (Figura 10) do que ao salicilaldeído. Desta forma, é possível supor que este mutante foi capaz de oxidar o salicilaldeído a ácido carboxílico; possivelmente devido a algum resquício de cofator presente na solução de proteína.

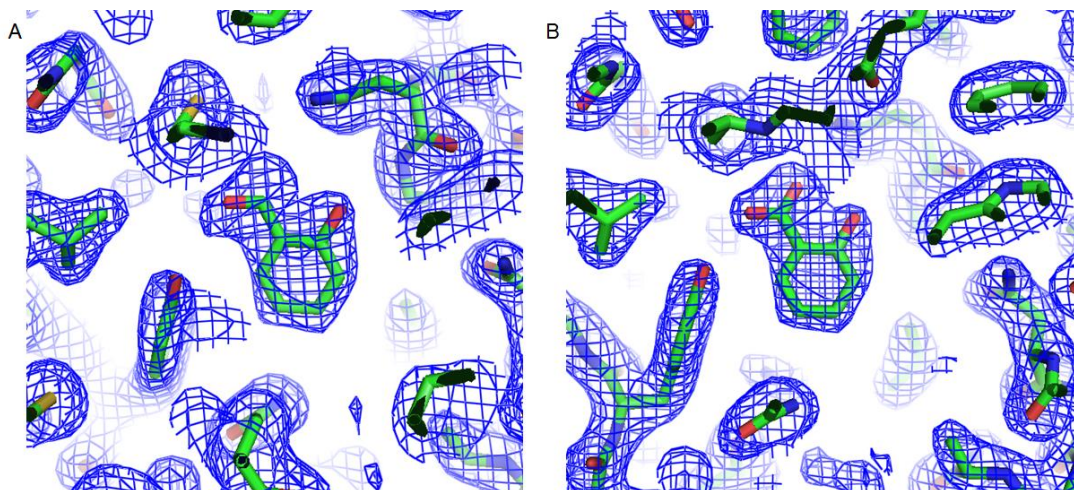


Figura 10: Região do sítio ativo da enzima NahF-R157K sugerindo a presença de um ligante. (A) A enzima mutante foi co-cristalizada com o salicilaldeído e observou-se uma densidade eletrônica no sítio ativo que poderia corresponder à molécula desse substrato. (B) O refinamento da estrutura sugeriu, entretanto, que o ligante mais provável de estar presente no sítio ativo seria o produto da reação (salicilato). Os mapas ($2mF_{\text{obs}} - DF_{\text{calc}}$, Φ_{calc}) foram contornados a $1,0 \sigma$. A imagem foi gerada pelo programa Pymol (DELANO, W. L., 2007).

Sobrepondo a estrutura desse mutante com a enzima *wild type*, observa-se um pequeno deslocamento da lisina 157 (mutante) quando comparada com a arginina correspondente na molécula *wild type* (Figura 11). Esse deslocamento aumenta a distância entre o aminoácido nessa posição e o substrato, podendo ser responsável pela redução na eficiência catalítica, que será discutida posteriormente.

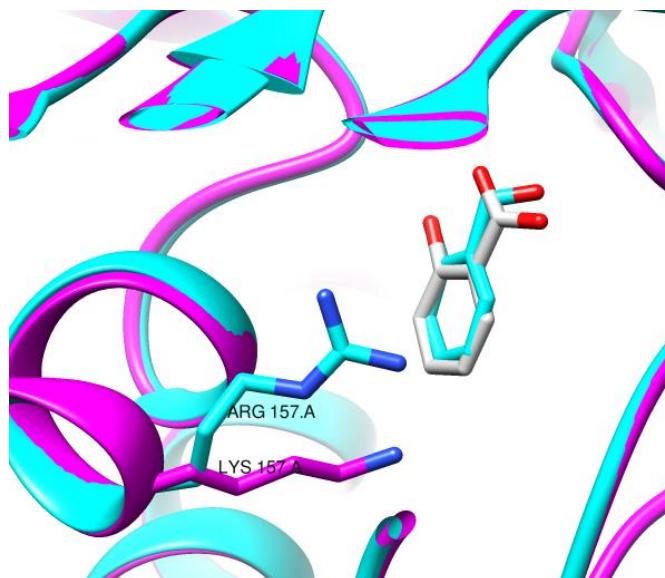


Figura 11: Sobreposição da enzima NahF *wild type* e do mutante NahF-R157K na região da mutação. Nesta figura o posicionamento do resíduo na posição 157 em cada uma das estruturas está destacado. A arginina da enzima *wild type* está apresentada em ciano e a lisina do mutante em magenta. A figura foi gerada pelo programa UCSF Chimera (PETTERSEN et al., 2004).

3.4 – Cinética enzimática dos mutantes de NahF

O testes de cinética demonstraram que os mutantes da enzima NahF seguem a equação de Michaelis-Menten (Figura 12), assim como a enzima *wild type*. Os valores de k_{cat} e K_M para a NahF foram determinados anteriormente (Coitinho et al, em preparação) e comparados aos parâmetros cinéticos dos mutantes (Tabela 5).

Foi observado que a eficiência catalítica do mutante NahF-R157A foi reduzida cerca de 320 vezes, enquanto a do mutante NahF-R157K foi 3 vezes menor que a da enzima *wild type* (Tabela 5).

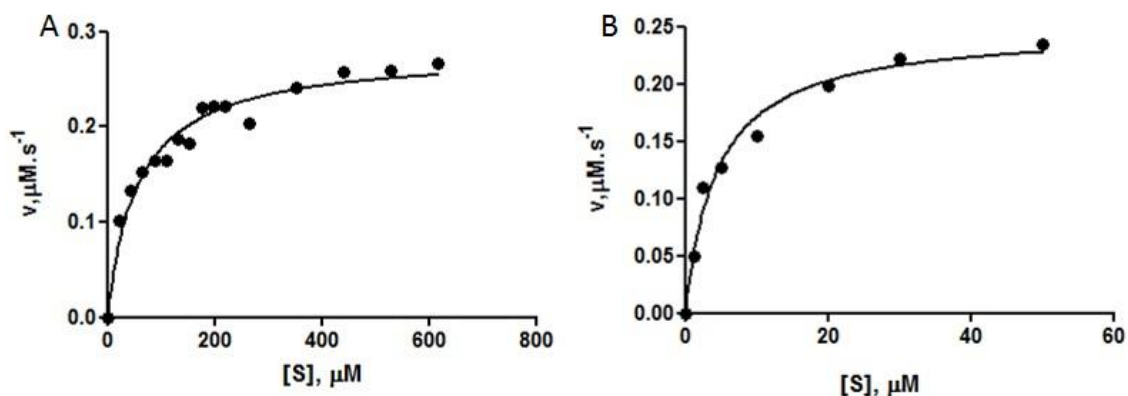


Figura 12: Gráficos de Michaelis-Menten para os mutantes NahF-R157A e NahF-R157K. Os testes de cinética foram realizados usando salicilaldeído como substrato e NAD⁺ como cofator. Pode-se observar as curvas de Michaelis-Menten para os mutantes NahF-R157A (A) e NahF-R157K (B).

Tabela 9: Parâmetros cinéticos dos mutantes de NahF comparados à enzima *wild type*. Os parâmetros foram calculados para o substrato salicilaldeído.

Enzima	K_M (μM)	k_{cat} (s ⁻¹)	k_{cat}/K_M
NahF <i>wild type</i>	2,48	12,7	5.132
NahF-R157A	56,96	0,9	16
NahF-R157K	4,47	8,3	1.851

Esses dados de cinética demonstram a importância da arginina na posição 157. Quando foi realizada a mutação para outro resíduo longo e também carregado positivamente (lisina), a enzima teve sua eficiência reduzida, mas não tão drasticamente quanto à da alteração para alanina.

Na enzima succínico semialdeído desidrogenase da cianobactéria do gênero *Synechococcus*, mutantes da arginina 139 (correspondente a Arg157 em NahF) apresentaram comportamento semelhante. Na mutação R139A, a atividade enzimática não chega a 30% da atividade da enzima *wild type*. Já a mutação R139K exibe mais de 80% da atividade, comparado com a *wild type*, demonstrando a necessidade de um resíduo positivamente carregado nessa posição (PARK; RHEE, 2013), assim como para NahF.

4 – Conclusões

- Os mutantes da enzima NahF foram sintetizados com sucesso utilizando-se o kit *QuikChange Lightning Site-Directed Mutagenesis* (Stratagene);
- Os mutantes NahF-R157A e NahF-R157K foram produzidos em sistema heterólogo, purificados e utilizados para testes de cristalização e cinética;
- O refinamento da estrutura do mutante NahF-R157A não foi concluído até o momento mas já foi possível observar uma alteração entre ela e a estrutura *wild type*;
- O refinamento da estrutura do mutante NahF-R157K foi concluído e não apresentou alterações significativas quando comparada à estrutura *wild type*, exceto pela localização da lisina 157 que apresentou-se mais afastada do substrato;
- O mutante NahF-R157A teve sua eficiência catalítica reduzida aproximadamente 320 vezes, enquanto a do mutante nahF-R157K foi reduzida apenas 3 vezes;
- A presença de um resíduo positivamente carregado na posição 157 da enzima NahF é importante para sua função catalítica.

5 – Perspectivas

- Concluir o refinamento da estrutura do mutante R157A;
- Expressar, purificar, cristalizar e determinar os parâmetros cinéticos dos outros mutantes sintetizados;
- Avaliar a atividade enzimática dos mutantes contra outros substratos aromáticos.

6- Referências Bibliográficas

ABRIOLA, D. P. *et al.* Active site of human liver aldehyde dehydrogenase. *Biochemistry*, PMID: 3676276, v. 26, n. 18, p. 5679–5684, 8 set. 1987.

ABRIOLA, D. P.; MACKERELL, A. D.; PIETRUSZKO, R. Correlation of loss of activity of human aldehyde dehydrogenase with reaction of bromoacetophenone with glutamic acid-268 and cysteine-302 residues. Partial-sites reactivity of aldehyde dehydrogenase. *The Biochemical Journal*, PMID: 1968743 PMCID: PMC1131112, v. 266, n. 1, p. 179–187, 15 fev. 1990.

ADAMS, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D*, v. 66, n. 2, p. 213–221, 2010.

AHVAZI, B. *et al.* Crystal structure of the NADP⁺-dependent aldehyde dehydrogenase from *Vibrio harveyi*: structural implications for cofactor specificity and affinity. *The Biochemical Journal*, PMID: 10903148 PMCID: PMC1221214, v. 349 Pt 3, p. 853–861, 1 ago. 2000.

ATLAS, R. M.; HAZEN, T. C. Oil Biodegradation and Bioremediation: A Tale of the Two Worst Spills in U.S. History. *Environmental Science & Technology*, v. 45, n. 16, p. 6709–6715, 15 ago. 2011.

BAMFORTH, S. M; SINGLETON, Ian. Bioremediation of polycyclic aromatic hydrocarbons: current knowledge and future directions. *Journal of Chemical Technology & Biotechnology*, v. 80, n. 7, p. 723–736, jul. 2005.

BOSTRÖM, C. *et al.* Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. *Environmental Health Perspectives*, PMID: 12060843 PMCID: PMC1241197, v. 110 Suppl 3, p. 451–488, jun. 2002.

CERNIGLIA, C. Biodegradation of polycyclic aromatic hydrocarbons. *Biodegradation*, v. 3, n. 2-3, p. 351–368, 1 jun. 1992.

COBESSI, D. *et al.* Structural and biochemical investigations of the catalytic mechanism of an NADP-dependent aldehyde dehydrogenase from *Streptococcus mutans*. *Journal of Molecular Biology*, PMID: 10864505, v. 300, n. 1, p. 141–152, 30 jun. 2000.

COITINHO, J. B. *et al.* Expression, purification and preliminary crystallographic studies of NahF, a salicylaldehyde dehydrogenase from *Pseudomonas putida* G7 involved in naphthalene degradation. *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, PMID: 22232182 PMCID: PMC3253845, v. 68, n. Pt 1, p. 93–97, 1 jan. 2012.

DAVIES, J. I.; EVANS, W. C. Oxidative metabolism of naphthalene by soil pseudomonads. The ring-fission mechanism. *Biochem J*, v. 91, p. 251–261, 1964.

DELANO, W. L. *PyMOL: An Open-Source Molecular Graphics Tool*. [S.l: s.n.], 2007. v. DeLano Scientific San Carlos, California USA.

DUNN, N. W.; GUNSALUS, I. C. Transmissible plasmid coding early enzymes of naphthalene oxidation in *Pseudomonas putida*. *Journal of Bacteriology*, PMID: 4712575 PMCID: PMC285353, v. 114, n. 3, p. 974–979, jun. 1973.

EATON, R. W.; CHAPMAN, P. J. Bacterial metabolism of naphthalene: construction and use of recombinant bacteria to study ring cleavage of 1,2-dihydroxynaphthalene and subsequent reactions. *Journal of Bacteriology*, PMID: 1447127 PMCID: PMC207464, v. 174, n. 23, p. 7542–7554, dez. 1992.

FERNANDEZ, M. *et al.* Enhanced Tolerance to Naphthalene and Enhanced Rhizoremediation Performance for *Pseudomonas putida* KT2440 via the NAH7 Catabolic Plasmid. *Applied and Environmental Microbiology*, v. 78, n. 15, p. 5104–5110, 1 ago. 2012.

GUIMARÃES, B. G. *et al.* The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS. *Journal of Synchrotron Radiation*, v. 16, n. 1, p. 69–75, 1 jan. 2009.

HABE, H.; OMORI, T. Genetics of Polycyclic Aromatic Hydrocarbon Metabolism in Diverse Aerobic Bacteria. *Bioscience, Biotechnology and Biochemistry*, v. 67, n. 2, p. 225–243, jan. 2003.

KEITH, L.; TELLIARD, W. ES&T Special Report: Priority pollutants: I-a perspective view. *Environmental Science & Technology*, v. 13, n. 4, p. 416–423, abr. 1979.

KOPPAKA, V. *et al.* Aldehyde dehydrogenase inhibitors: a comprehensive review of the pharmacology, mechanism of action, substrate specificity, and clinical application. *Pharmacological Reviews*, PMID: 22544865 PMCID: PMC3400832, v. 64, n. 3, p. 520–539, jul. 2012.

LEE, Y. *et al.* Contribution of conserved Glu255 and Cys289 residues to catalytic activity of recombinant aldehyde dehydrogenase from *Bacillus licheniformis*. *Biochemistry. Biokhimiia*, PMID: 22117550, v. 76, n. 11, p. 1233–1241, nov. 2011.

LINDSTROM, J. E. *et al.* Microbial populations and hydrocarbon biodegradation potentials in fertilized shoreline sediments affected by the T/V Exxon Valdez oil spill. *Applied and Environmental Microbiology*, PMID: 1662935 PMCID: PMC183612, v. 57, n. 9, p. 2514–2522, set. 1991.

LIU, Z. J. *et al.* The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD and the Rossmann fold. *Nature Structural Biology*, PMID: 9095201, v. 4, n. 4, p. 317–326, abr. 1997.

MARCHAL, S.; RAHUEL-CLERMONT, S.; BRANLANT, G. Role of glutamate-268 in the catalytic mechanism of nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase from *Streptococcus mutans*. *Biochemistry*, PMID: 10727225, v. 39, n. 12, p. 3327–3335, 28 mar. 2000.

MUÑOZ-CLARES, R. A.; GONZÁLEZ-SEGURA, L.; DÍAZ-SÁNCHEZ, A. G. Crystallographic evidence for active-site dynamics in the hydrolytic aldehyde dehydrogenases. Implications for the deacylation step of the catalyzed reaction.

Chemico-Biological Interactions, PMID: 21195066, v. 191, n. 1-3, p. 137–146, 30 maio 2011.

PARK, J.; RHEE, S. Structural Basis for a Cofactor-dependent Oxidation Protection and Catalysis of Cyanobacterial Succinic Semialdehyde Dehydrogenase. *Journal of Biological Chemistry*, v. 288, n. 22, p. 15760–15770, 31 maio 2013.

PETTERSEN, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, PMID: 15264254, v. 25, n. 13, p. 1605–1612, out. 2004.

ROSSMANN, M. G.; MORAS, D.; OLSEN, K. W. Chemical and biological evolution of a nucleotide-binding protein. *Nature*, v. 250, n. 5463, p. 194–199, 19 jul. 1974. Acesso em: 1 set. 2014.

SCHELL, M. A.; WENDER, P. E. Identification of the nahR gene product and nucleotide sequences required for its activation of the sal operon. *Journal of Bacteriology*, PMID: 3007442 PMCID: PMC214548, v. 166, n. 1, p. 9–14, abr. 1986.

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. *Journal of Bacteriology*, v. 188, n. 11, p. 4057–4067, 1 jun. 2006.

STEINMETZ, C. G. *et al.* Structure of mitochondrial aldehyde dehydrogenase: the genetic component of ethanol aversion. *Structure (London, England: 1993)*, PMID: 9195888, v. 5, n. 5, p. 701–711, 15 maio 1997.

TSYBOVSKY, Y. *et al.* The mechanism of discrimination between oxidized and reduced coenzyme in the aldehyde dehydrogenase domain of Aldh1l1. *Chemico-Biological Interactions*, PMID: 23295222 PMCID: PMC3602205, v. 202, n. 1-3, p. 62–69, 25 fev. 2013.

VEDADI, M.; MEIGHEN, E. Critical glutamic acid residues affecting the mechanism and nucleotide specificity of *Vibrio harveyi* aldehyde dehydrogenase. *European journal of biochemistry / FEBS*, PMID: 9219528, v. 246, n. 3, p. 698–704, 15 jun. 1997.

WANG, X.; WEINER, H. Involvement of glutamate 268 in the active site of human liver mitochondrial (class 2) aldehyde dehydrogenase as probed by site-directed mutagenesis. *Biochemistry*, PMID: 7819202, v. 34, n. 1, p. 237–243, 10 jan. 1995.

WENZL, M. V. *et al.* Role of the General Base Glu-268 in Nitroglycerin Bioactivation and Superoxide Formation by Aldehyde Dehydrogenase-2. *Journal of Biological Chemistry*, v. 284, n. 30, p. 19878–19886, 24 jul. 2009.

YEN, K M; GUNSALUS, I C. Plasmid gene organization: naphthalene/salicylate oxidation. *Proceedings of the National Academy of Sciences*, v. 79, n. 3, p. 874–878, 1 fev. 1982.

YEN, K. M.; GUNSALUS, I. C. Regulation of naphthalene catabolic genes of plasmid NAH7. *Journal of Bacteriology*, PMID: 3997772 PMCID: PMC215876, v. 162, n. 3, p. 1008–1013, jun. 1985.

YOSHIDA, A. *et al.* Human aldehyde dehydrogenase gene family. *European journal of biochemistry / FEBS*, PMID: 9490025, v. 251, n. 3, p. 549–557, 1 fev. 1998.

Preliminary crystallographic studies of a *Schistosoma mansoni* antigen (Sm21.7) dynein light-chain (DLC) domain

M. A. F. Costa, F. T. G. Rodrigues, B. C. A. Chagas, C. M. F. Rezende, A. M. Goes and R. A. P. Nagem*

Departamento de Bioquímica e Imunologia,
Instituto de Ciências Biológicas, Universidade
Federal de Minas Gerais, 31270-901 Belo
Horizonte-MG, Brazil

Correspondence e-mail: nagem@icb.ufmg.br

Received 20 February 2014
Accepted 24 April 2014

Schistosomiasis is an inflammatory chronic disease that represents a major health problem in tropical and subtropical countries. The drug of choice for treatment, praziquantel, is effective in killing adult worms but fails to kill immature forms and prevent reinfection. One prominent antigen candidate for an anti-schistosomiasis vaccine is the protein Sm21.7 (184 amino-acid residues) from *Schistosoma mansoni*, a tegumental protein capable of reducing the worm burden in a murine immunization model. In the present work, the Sm21.7 gene was cloned and expressed in *Escherichia coli* and the full-length protein was purified to homogeneity. Crystals of recombinant Sm21.7 suitable for X-ray diffraction were obtained using PEG monomethyl ether 2000 as a precipitant. X-ray diffraction images of a native crystal (at 2.05 Å resolution) and a quick-cryosoaked NaI derivative (at 1.95 Å resolution) were collected on the W01B-MX2 beamline at the Laboratório Nacional de Luz Síncrotron (LNLS, Brazilian Synchrotron Light Laboratory/MCT). Both crystals belonged to the hexagonal space group $P6_122$, with similar unit-cell parameters $a = b = 108.5$, $c = 55.8$ Å. SIRAS-derived phases were used to generate the first electron-density map, from which a partial three-dimensional model of Sm21.7 (from Gln89 to Asn184) was automatically constructed. Analysis of dissolved crystals by SDS-PAGE confirmed that the protein was cleaved in the crystallization drop and only the Sm21.7 C-terminal domain was crystallized. The structure of the Sm21.7 C-terminal domain will help in the localization of the epitopes responsible for its protective immune responses, constituting important progress in the development of an anti-schistosomiasis vaccine.

1. Introduction

Schistosomiasis is an inflammatory chronic disease caused by worms of the *Schistosoma* genus that represents a major health problem in tropical and subtropical countries (Gryseels *et al.*, 2006). The main schistosomes infecting humans are *S. haematobium* in Africa, *S. mansoni* in Africa and South America, *S. intercalatum* in West Africa and *S. japonicum* and *S. mekongi* in Asia. The former causes fibrosis, structuring and calcification of the urinary tract, while the others cause chronic liver and intestinal fibrosis (Ross *et al.*, 2013; Thétiot-Laurent *et al.*, 2013).

Estimates indicate that over 200 million individuals worldwide are currently infected (King, 2009), with more than 250 000 deaths per year (van der Werf *et al.*, 2003). Moreover, the impact of the severe and debilitating effects of schistosomiasis account for the loss of 4.5 million disability-adjusted life years (DALY; Steinmann *et al.*, 2006).

Current schistosomiasis control strategies are predominantly based on the treatment of infected individuals with safe and effective drugs (Gryseels *et al.*, 2006). The drug of choice, praziquantel, is effective in killing adult worms but fails to kill immature forms and to prevent reinfection (Gray *et al.*, 2010). Besides, there is evidence supporting the development of drug-resistant parasites (Fallon & Doenhoff, 1994; Ismail *et al.*, 1996). These problems limit the success of control strategies based on chemotherapy only. Therefore, a protective anti-



schistosomiasis vaccine would be valuable to the current disease-control program.

Recently, a significant effort has been made to develop a vaccine against schistosomiasis, and several vaccine candidates have been identified. One prominent vaccine candidate is the protein Sm21.7 (184 residues; accession No. XP_002569898) from *S. mansoni*. It was identified by screening a sporocyst cDNA expression library using irradiated cercaria-vaccinated rabbit serum. Sequence analysis demonstrated that the antigen is composed of an N-terminal EF-hand-like motif (Arg17–Gly67) and an assigned C-terminal dynein light-chain (DLC) domain (Lys81–Thr181). Surprisingly, ^{45}Ca -binding experiments with recombinant (GST-fusion protein) and parasite-derived antigen showed no evidence of calcium binding, suggesting the presence of a nonfunctional EF-hand motif (Francis & Bickle, 1992).

Sm21.7 was tested as an antigen in recombinant vaccines (Ahmed *et al.*, 2001) and DNA vaccines (Ahmed *et al.*, 2006), conferring a significant level of protection in mice against *S. mansoni* challenge infection (Ahmed *et al.*, 2001, 2006). Northern blot analysis demonstrated that mRNA for Sm21.7 is present in the sporocyst, schistosomula and adult stages (Francis & Bickle, 1992), and RT-PCR showed that transcription occurred in eggs, cercaria and adult worms (Fitzsimmons *et al.*, 2007). The antigen has been observed in schistosomula (Curwen *et al.*, 2004), eggs, cercariae (Fitzsimmons *et al.*, 2007) and adult antigen preparation (Curwen *et al.*, 2004; Fitzsimmons *et al.*, 2007).

Recently, our group has demonstrated that an adult worm protein preparation (fraction PIII) containing a tegument protein named P22 was able to engender protective and immunomodulatory effects against murine schistosomiasis. A cDNA clone encoding P22 was isolated from an *S. mansoni* adult worm cDNA library using anti-PIII rabbit serum and showed 100% identity to Sm21.7 (Rezende *et al.*, 2011).

Mice immunized with recombinant Sm21.7 (rP22; C-terminally 6×His-tagged fusion protein) exhibited a 51 and 22% decrease in adult worm burden and in hepatic eggs, respectively. Additionally, a recombinant Sm21.7 vaccine produced a reduction of 60% in liver granuloma size and a 71% reduction in fibrosis in mice, suggesting that Sm21.7 could contribute to downmodulate granulomatous hypersensitivity to *S. mansoni* eggs (Rezende *et al.*, 2011).

Even though relevant biological information about Sm21.7 has been obtained in recent years, there is a lack of structural data for both of its domains. The primary structure of the Sm21.7 amino-terminal domain (EF-hand-like motif) shares 52% similarity with calmodulin, but does not bind calcium as reported previously (Francis & Bickle, 1992). The carboxy-terminal domain (dynein light-chain domain) shares only 24% identity (*E*-value of 0.29) with a fragment of a *Kluyveromyces lactis* Hsv2p protein for which the three-dimensional structure was recently determined (Baskaran *et al.*, 2012; Krick *et al.*, 2012). Apart from this, tertiary structures for DLC domains are available for proteins sharing only 19% identity to Sm21.7 (Romes *et al.*, 2012; Gallego *et al.*, 2013; Rao *et al.*, 2013).

In order to better characterize Sm21.7, compare its structure with the most similar proteins available and localize the epitopes responsible for its protective immune responses, we aimed to clone, express and purify the full-length antigen. Crystallization conditions were also determined, but despite all efforts, X-ray diffraction results showed that only the C-terminal domain of this protein was crystallized. Owing to the importance of schistosomiasis and the promising results of a vaccine composed of recombinant Sm21.7, we aimed to determine the initial phases for the crystalline structure of the Sm21.7 carboxy-terminal domain by the heavy-atom method.

2. Experimental methods

2.1. Cloning, expression and purification

The gene sequence encoding full-length Sm21.7 (Gene ID 8347700) was amplified by PCR from a previously constructed plasmid (pETDEST42-rP22; Rezende *et al.*, 2011). The primers Sm21.7Fw (5'-CAT **ATG** GAT AGT CCA ATG GAA AAA TTT ATT C-3') and Sm21.7Rv (5'-C **TCG AGT** CTA GTT ACT TGG TGT ACG CCA AG-3'), containing *Nde*I and *Xho*I restriction sites (bold sequences), respectively, were used to amplify the Sm21.7 coding sequence. The PCR-amplified fragment was inserted into pGEM-T vector (Promega) and the recombinant plasmids pGEM-Sm21.7 were transformed into *Escherichia coli* DH5α cells. pGEM-Sm21.7 plasmids were then double-digested with *Nde*I and *Xho*I and the released inserts were ligated into pET-28a(TEV) vector (Carneiro *et al.*, 2006) previously digested with the same enzymes. The recombinant plasmids pET-28a(TEV)-Sm21.7 were transformed into *E. coli* BL21(DE3) and selected on kanamycin plates.

Large-scale expression cultures of the recombinant 6×His-Sm21.7 (rSm21.7) were carried out by inoculating 20 ml of an overnight rSm21.7 culture into 1 l Luria–Bertani broth medium containing 50 µg ml⁻¹ kanamycin. The culture was incubated at 310.15 K and 200 rev min⁻¹ until a cell density of between 0.4 and 0.8 (OD_{600 nm}) was achieved. Expression of the recombinant protein was then induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM followed by incubation of the culture for 4 h at 310.15 K and 200 rev min⁻¹.

Cells were harvested by centrifugation at 7000g for 15 min at 277.15 K and resuspended in lysis buffer [50 mM Tris–HCl pH 7.5, 1% (w/v) sucrose, 1% (v/v) glycerol, 1% (v/v) Tween 20, 1 mM phenylmethylsulfonyl fluoride (PMSF)]. The lysis process was performed by three passes through an EmulsiFlex-C3 homogenizer (Avestin). The lysate was clarified by centrifugation at 10 000g for 1 h at 277.15 K. In order to purify the rSm21.7, imidazole was added to the supernatant to a final concentration of 30 mM and the mixture was loaded onto a 5 ml HisTrap HP Ni²⁺-affinity column (GE Healthcare) connected to an ÄKTAprime plus system (GE Healthcare). The affinity column had previously been equilibrated with a binding buffer solution (20 mM sodium phosphate pH 7.5, 500 mM NaCl, 30 mM imidazole, 1 mM PMSF). After washing the column with four volumes of the binding buffer, the rSm21.7 antigen was eluted using a gradient of imidazole. For this purpose, an elution buffer solution (20 mM sodium phosphate pH 7.5, 500 mM NaCl, 500 mM imidazole, 1 mM PMSF) was automatically mixed with the binding buffer solution within the system during the chromatographic run. The fractions containing rSm21.7 were analyzed by SDS–PAGE, pooled, concentrated to 5 ml using a sample concentrator (Vivaspin 20, MWCO 10 kDa; GE Healthcare) and loaded onto a size-exclusion column (HiLoad 16/60 Superdex 200 prep grade; GE Healthcare). The column was previously equilibrated with a buffer solution consisting of 50 mM sodium phosphate pH 7.5, 50 mM NaCl, 1 mM PMSF. The same ÄKTAprime plus system (GE Healthcare) was used. Eluted peak fractions were pooled and the homogeneity of the purified samples was verified by SDS–PAGE and mass spectrometry.

Following size-exclusion chromatography, crystallization tests were carried out with freshly concentrated protein at approximately 20 mg ml⁻¹ (Vivaspin 20, MWCO 10 kDa; GE Healthcare). The concentration of purified rSm21.7 was estimated by measuring the absorbance at 280 nm and employing its theoretical molar extinction coefficient as calculated using the *ProtParam* software (Gasteiger *et al.*, 2005) available on the ExPASy Proteomics Server (<http://www.expasy.org>; Artimo *et al.*, 2012).

2.2. Crystallization, data collection and processing

All crystallization trials were accomplished at 291.15 K using the hanging-drop vapour-diffusion method in 24-well plates.

Initial screening was performed using in-house-prepared solutions based on the Crystal Screen and Crystal Screen 2 formulations (Hampton Research). Drops consisting of 1 μ l protein solution (20 mg ml⁻¹ rSm21.7 in 50 mM sodium phosphate buffer pH 7.5, 50 mM NaCl, 1 mM PMSF) and an equal volume of crystallization solution were equilibrated against 1 ml reservoir solution (the same crystallization solution) in a manual setup.

Of the 98 conditions tested two hits were identified: (i) 200 mM ammonium sulfate, 100 mM sodium acetate trihydrate pH 4.6, 30% (w/v) PEG monomethyl ether 2000 and (ii) 200 mM ammonium sulfate, 100 mM sodium cacodylate trihydrate pH 6.5, 30% (w/v) PEG 8000. These two promising conditions were optimized by varying the pH and the precipitant and salt concentrations. Moreover, in each new condition, drops of different volumes and protein:precipitant solution ratios (2 μ l:2 μ l, 1 μ l:2 μ l and 2 μ l:1 μ l) were tested, totalling 288 drops.

X-ray diffraction data sets were collected on the W01B-MX2 beamline (Guimarães *et al.*, 2009) at the Laboratório Nacional de Luz Síncrotron (LNLS, Brazilian Synchrotron Light Laboratory/MCT; Campinas-SP, Brazil). Before data acquisition of a native sample, a single rSm21.7 crystal was transferred for a few seconds to a 10 μ l drop of a cryogenic solution formed by the addition of 1 μ l ethylene glycol to 9 μ l reservoir solution. The crystal was then flash-cooled to 100 K in a nitrogen stream and used for data acquisition. An iodide derivative was obtained using the quick-cryosoaking approach (Dauter *et al.*, 2000; Nagem *et al.*, 2001, 2003). Essentially, a native crystal was soaked for 30 s in the same cryogenic solution additionally containing sodium iodide (at a final concentration of 0.4 M) and then flash-cooled in the nitrogen stream. In both cases, X-ray diffraction images were recorded on a Rayonix MarMosaic 225 CCD detector and were processed with *HKL-2000* (Otwinowski & Minor, 1997).

The phase problem was solved by using the SIRAS (single isomorphous replacement with anomalous scattering) method within *PHENIX* (Adams *et al.*, 2010). Iodide sites were obtained using the anomalous scattering contribution present in the derivative data set, but both isomorphous and anomalous contributions were used for phase determination. SIRAS-derived phases were used to generate the first electron-density map, enabling the construction of a preliminary model of rSm21.7. Restrained refinement and visual inspection of the model were performed with *PHENIX* and *Coot* (Emsley *et al.*, 2010), respectively.

3. Results and discussion

The 555 bp gene encoding rSm21.7 was successfully amplified and cloned into the expression vector pET-28a(TEV). This vector adds an N-terminal 6 $\times$ His tag and a TEV protease cleavage site to the protein (MGHHHHHHENLYFQGH—target protein; Carneiro *et al.*, 2006).

After expression in *E. coli* BL21(DE3) and cell lysis (Fig. 1a), rSm21.7 was purified by Ni²⁺-affinity and size-exclusion chromatography (Fig. 1b). In the last purification step, the protein eluted with an apparent molecular weight of 33 kDa, making the correct assignment of its oligomeric state in the conditions tested impossible. The purified protein showed a single band of approximately 25 kDa on SDS-PAGE (Fig. 1a), which is in good agreement with the calculated (23 688.8 Da) and mass-spectrometry-derived (23 564.4 Da; data not shown) molecular weight of rSm21.7. The final

yield of purified protein was approximately 30 mg per litre of cell culture, allowing the preparation of approximately 1.5 ml of a 20.0 mg ml⁻¹ protein sample for crystallization assays per expression batch.

Owing to the fact that purified rSm21.7 is susceptible to degradation with time, a protease inhibitor (PMSF) was added to the lysis and purification buffers. Recombinant Sm21.7 samples also show signs of precipitation when stored at 277.15 K or thawed. Consequently, all experiments were performed with freshly purified protein.

Following the initial crystallization screening, hits were optimized by testing 288 new crystallization conditions with varying pH values (in the range 4.0–5.5 using sodium acetate buffer or 5.5–7.4 using sodium cacodylate buffer) and precipitant and salt concentrations [PEG 8000 or PEG monomethyl ether 2000 from 20 to 40% (w/v) and

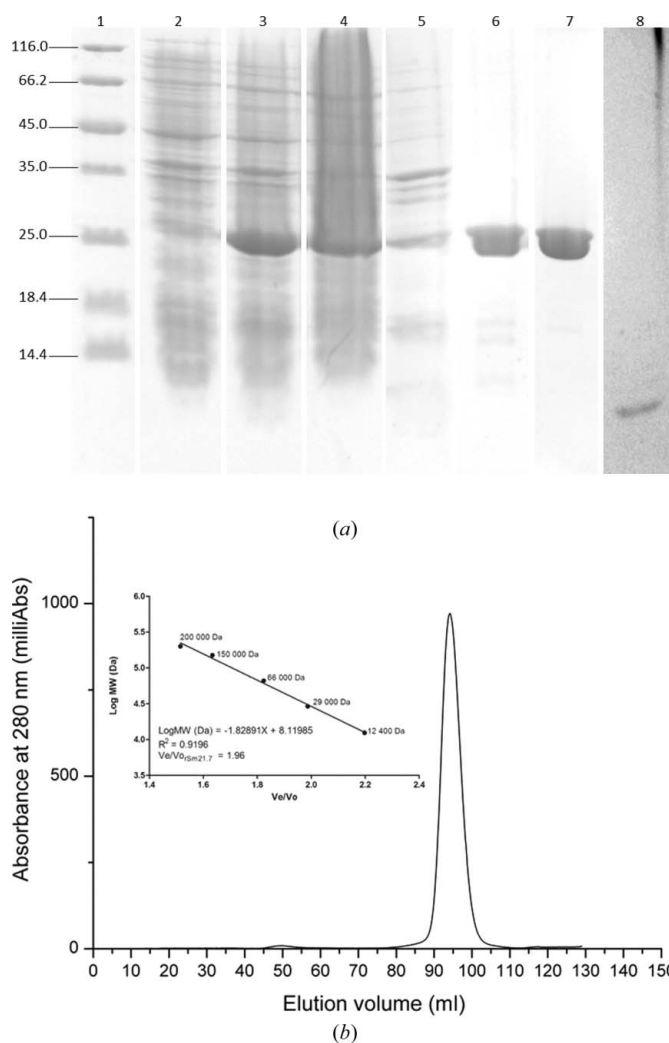


Figure 1 Quality assessment of rSm21.7 expression and purification. (a) SDS-PAGE of rSm21.7 samples during protein expression and purification. Lane 1, molecular-weight marker (labelled in kDa); lane 2, non-induced cells; lane 3, 4 h IPTG-induced cells; lane 4, soluble fraction after cell lysis; lane 5, insoluble fraction after cell lysis; lane 6, sample obtained after Ni²⁺-affinity chromatography; lane 7, sample obtained after size-exclusion chromatography; lane 8, sample obtained from crystals used for structure determination. (b) Size-exclusion chromatography profile of rSm21.7. The protein eluted at a volume close to 94 ml, with an apparent molecular weight of 33 kDa, consistent with either a monomer or a dimer in solution. The column was calibrated with Sigma Gel Filtration Molecular Weight Markers (catalogue No. MWGF200) comprising the following standard molecules: cytochrome *c* from horse heart (12.4 kDa), carbonic anhydrase from bovine erythrocytes (29 kDa), albumin from bovine serum (66 kDa), alcohol dehydrogenase from yeast (150 kDa) and β -amylase from sweet potato (200 kDa).

ammonium sulfate from 100 to 300 mM, respectively]. Different crystallization drop volumes and protein:precipitant solution ratios (2 μ l:2 μ l, 1 μ l:2 μ l and 2 μ l:1 μ l) were also tested.

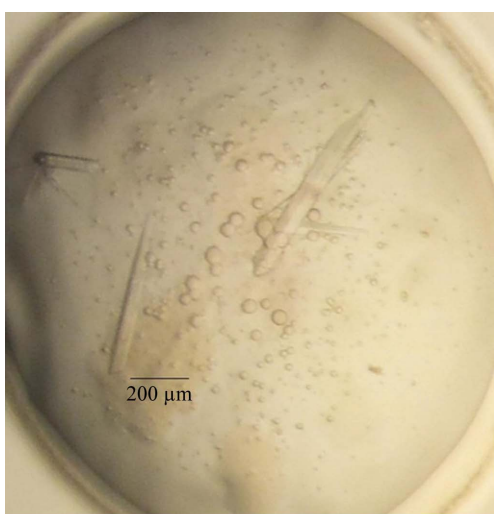
The best diffracting plate-like crystals with area dimensions of 0.5 $\times$ 0.1 mm were obtained after 30 d using the following optimized reservoir conditions: (i) 100 mM ammonium sulfate, 25%(w/v) PEG monomethyl ether 2000, 100 mM sodium acetate buffer pH 5.5 (1 μ l protein solution and 2 μ l reservoir solution) and (ii) 200 mM ammonium sulfate, 25%(w/v) PEG 8000, 100 mM sodium cacodylate buffer pH 5.5 (2 μ l protein solution and 2 μ l reservoir solution). Both conditions resulted in the formation of a protein precipitate from which crystals displaying the same macroscopic morphology grew (Fig. 2).

A single native rSm21.7 crystal grown in PEG monomethyl ether 2000 was submitted to an X-ray diffraction experiment and diffracted to 2.05 Å resolution (Fig. 3a). The crystal belonged to the hexagonal space group $P6_122$ as determined during phasing and model building (Table 1). The calculated Matthews coefficient (V_M ; Matthews, 1968) of 2.00 Å³ Da⁻¹ with a solvent content of 38.4% suggests the

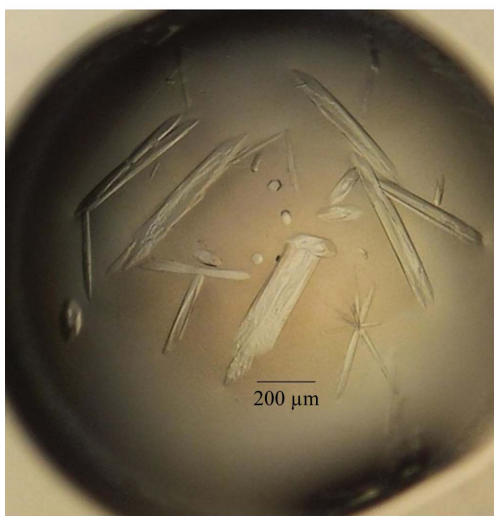
presence of one full-length rSm21.7 molecule in the asymmetric unit of the crystal. However, all attempts to solve the rSm21.7 crystal structure by the molecular-replacement method using the EF-hand-like motif of *Paramecium tetraurelia* calmodulin (PDB entry 1clm; Rao *et al.*, 1993) and the dynein light-chain domain of *Kluyveromyces lactis* Hsv2p protein (PDB entries 4exv and 4av9; Baskaran *et al.*, 2012; Krick *et al.*, 2012) as search models were unsuccessful. As a consequence, the heavy-atom method was used for phasing and structure determination.

An iodide derivative was prepared by soaking a native crystal (grown in PEG 8000) in a cryogenic solution containing 0.4 M sodium iodide for 30 s. This crystal diffracted to 1.95 Å resolution (Fig. 3b) and was shown to be isomorphous to the native crystal (Table 1). An almost 6% anomalous difference observed in Friedel mates suggests systematic incorporation of iodide anions into the crystallographic structure.

Indeed, two iodide anion sites with high occupancies (1.00 and 0.95) were found in the crystallographic structure of the derivative using *SHELXD* (Sheldrick, 2010). These sites, together with both



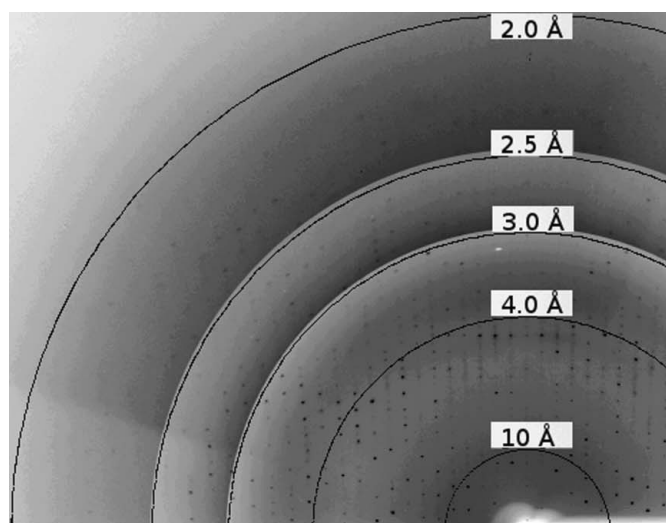
(a)



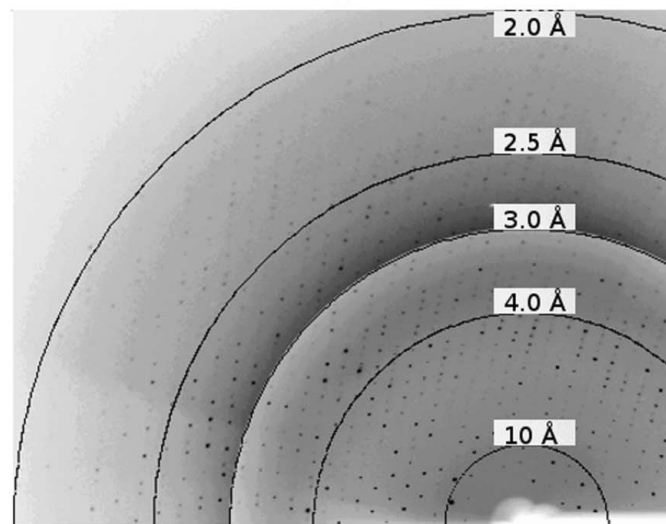
(b)

Figure 2

Crystals of rSm21.7 from *S. mansoni*. (a) Crystals grown in 100 mM ammonium sulfate, 25%(w/v) PEG monomethyl ether 2000, 100 mM sodium acetate buffer pH 5.5. (b) Crystals obtained in 200 mM ammonium sulfate, 25%(w/v) PEG 8000, 100 mM sodium cacodylate buffer pH 5.5.



(a)



(b)

Figure 3

X-ray diffraction images of rSm21.7. Image obtained from (a) a native crystal and (b) an iodide-derivative crystal using a Rayonix MarMosaic 225 CCD detector. The brightness and/or contrast of the displayed images were adjusted by resolution shells to facilitate spot visualization.

Table 1

Data-collection and processing statistics for native and iodide-derivative rSm21.7 crystals.

Values in parentheses are for the lowest/highest resolution shells.

	Native	Iodide derivative
Wavelength (Å)	1.459	1.459
Temperature (K)	100	100
Crystal-to-detector distance (mm)	120	105
Rotation range per image (°)	0.5	0.5
Exposure time per image (s)	16	6
Total rotation range (°)	130	150
Space group	<i>P</i> 6 ₁ 22	<i>P</i> 6 ₁ 22
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 108.5, <i>c</i> = 55.8	<i>a</i> = <i>b</i> = 109.1, <i>c</i> = 56.4
Resolution range (Å)	50.00–2.05 (50.00–4.69/ 2.11–2.05)	50.00–1.95 (50.00–4.46/ 2.01–1.95)
No. of observations	169511	183507
No. of unique reflections	22353 (1899/1778)	27084 (2284/2092)
Data completeness (%)	96.7 (97.8/92.9)	98.9 (99.0/92.0)
$\langle I/\sigma(I) \rangle$	26.9 (47.5/2.7)	16.9 (52.4/2.1)
Multiplicity	7.6 (8.3/4.9)	6.8 (9.4/3.1)
$R_{\text{merge}}^{\dagger}$ (%)	6.8 (3.4/50.9)	10.7 (4.3/52.1)
Average mosaicity (°)	0.81	1.18
Matthews coefficient ‡ (Å ³ Da ^{−1})	4.20	4.30
Solvent content ‡ (%)	70.7	71.4
$R_{\text{anom}}^{\S}$ (%)	—	5.98 ¶
$R_{\text{isom}}^{\dagger\dagger}$ (%)	—	21.6 ¶
FOM ‡†	—	0.43 (0.69)

$^{\dagger} R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl) \times 100$, where $I_i(hkl)$ is the *i*th intensity measurement of reflection *hkl* and $\langle I(hkl) \rangle$ is its average. ‡ Numbers refer to suggested values calculated by assuming the presence of 96 carboxy-terminal residues of rSm21.7 per asymmetric unit (Gln89–Asn184). $^{\S} R_{\text{anom}} = \sum_{hkl} |I(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} I(hkl) \times 100$. ¶ The signal was computed with reflections up to 3.0 Å resolution. $^{\dagger\dagger} R_{\text{isom}} = \sum_{hkl} ||F_{\text{deriv}}(hkl)| - |F_{\text{nat}}(hkl)|| / \sum_{hkl} |F_{\text{nat}}(hkl)| \times 100$. ‡† The values refer to FOM prior to and after (in parentheses) density modification computed for all reflections.

data sets (native and derivative), were used for SIRAS phasing within PHENIX. Density modification followed by automatic model building against the native data set resulted in a partial rSm21.7 model with an *R* value and *R*_{free} of 0.23 and 0.25, respectively. Surprisingly, this partial model contained 96 residues from the carboxy-terminal dynein light-chain domain of rSm21.7 (from Gln89 to Asn184), with a theoretical molecular weight of 11 277 Da. As confirmed by SDS–PAGE prepared with dissolved crystals (Fig. 1a), only the C-terminal domain of rSm21.7 was crystallized. The observed Matthews coefficient (*V*_M) of this crystal is 4.20 Å³ Da^{−1}, with a solvent content of 70.7% (Table 1). The cleavage of Sm21.7 is also inferred from two-dimensional gel electrophoresis/mass-spectrometry results (Curwen *et al.*, 2004), suggesting that a similar *in vivo* post-translation modification might occur. Visual model inspection/correction and automated maximum-likelihood refinement are currently under way. The detailed structure of the C-terminal dynein light-chain domain of Sm21.7 will be published elsewhere.

This work was supported by the Fundação de Desenvolvimento à Pesquisa do Estado de Minas Gerais (FAPEMIG; Project Rede-170/08) and the Laboratório Nacional de Luz Síncrotron (LNLS, Brazilian Synchrotron Light Laboratory/MCT; Project-9248/ROBOLAB-14983/MX1-14663). MAFC, FTGR and CMFR are recipients of scholarships from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). BCAC is the recipient of a scholarship from

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Ahmed, H. M., Romeih, M. H., Sherif, S. A., Fahim, F. A. & Saber, M. A. (2001). *Arab J. Biotechnol.* **4**, 229–249.
- Ahmed, H. M., Romeih, M. H., Shousha, T. S. A., Dabaa, E. E. & Saber, M. A. (2006). *J. Am. Sci.* **2**, 59–69.
- Artimo, P. *et al.* (2012). *Nucleic Acids Res.* **40**, W597–W603.
- Baskaran, S., Ragusa, M. J., Boura, E. & Hurley, J. H. (2012). *Mol. Cell*, **47**, 339–348.
- Carneiro, F. R. G., Silva, T. C. L., Alves, A. C., Haline-Vaz, T., Gozzo, F. C. & Zanchin, N. I. T. (2006). *Biochem. Biophys. Res. Commun.* **343**, 260–268.
- Curwen, R. S., Ashton, P. D., Johnston, D. A. & Wilson, R. A. (2004). *Mol. Biochem. Parasitol.* **138**, 57–66.
- Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* **D56**, 232–237.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Fallon, P. G. & Doenhoff, M. J. (1994). *Am. J. Trop. Med. Hyg.* **51**, 83–88.
- Fitzsimmons, C. M., McBeath, R., Joseph, S., Jones, F. M., Walter, K., Hoffmann, K. F., Kariuki, H. C., Mwatha, J. K., Kimani, G., Kabatereine, N. B., Vennervald, B. J., Ouma, J. H. & Dunne, D. W. (2007). *Int. Arch. Allergy Immunol.* **142**, 40–50.
- Francis, P. & Bickle, Q. (1992). *Mol. Biochem. Parasitol.* **50**, 215–224.
- Gallego, P., Velazquez-Campoy, A., Regué, L., Roig, J. & Reverter, D. (2013). *J. Biol. Chem.* **288**, 12283–12294.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. Walker, pp. 571–607. Totowa: Humana Press.
- Gray, D. J., McManus, D. P., Li, Y., Williams, G. M., Bergquist, R. & Ross, A. G. (2010). *Lancet Infect. Dis.* **10**, 733–736.
- Gryseels, B., Polman, K., Clerinx, J. & Kestens, L. (2006). *Lancet*, **368**, 1106–1118.
- Guimarães, B. G., Sanfelici, L., Neuenschwander, R. T., Rodrigues, F., Grizolli, W. C., Raulik, M. A., Piton, J. R., Meyer, B. C., Nascimento, A. S. & Polikarpov, I. (2009). *J. Synchrotron Rad.* **16**, 69–75.
- Ismail, M., Metwally, A., Farghaly, A., Bruce, J., Tao, L. F. & Bennett, J. L. (1996). *Am. J. Trop. Med. Hyg.* **55**, 214–218.
- King, C. H. (2009). *N. Engl. J. Med.* **360**, 106–109.
- Krick, R., Busse, R. A., Scacioc, A., Stephan, M., Janshoff, A., Thumm, M. & Kühnel, K. (2012). *Proc. Natl Acad. Sci. USA*, **109**, E2042–E2049.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* **D57**, 996–1002.
- Nagem, R. A. P., Polikarpov, I. & Dauter, Z. (2003). *Methods Enzymol.* **374**, 120–137.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Rao, L., Romes, E. M., Nicholas, M. P., Brenner, S., Tripathy, A., Gennerich, A. & Slep, K. C. (2013). *Mol. Biol. Cell*, **24**, 2362–2377.
- Rao, S. T., Wu, S., Satyshur, K. A., Ling, K.-Y., Kung, C. & Sundaralingam, M. (1993). *Protein Sci.* **2**, 436–447.
- Rezende, C. M. F., Silva, M. R., Santos, I. G. D., Silva, G. A. B., Gomes, D. A. & Goes, A. M. (2011). *Immunol. Lett.* **141**, 123–133.
- Romes, E. M., Tripathy, A. & Slep, K. C. (2012). *J. Biol. Chem.* **287**, 15862–15873.
- Ross, A. G. P., Olveda, R. M., Acosta, L., Harn, D. A., Chy, D., Li, Y., Gray, D. J., Gordon, C. A., McManus, D. P. & Williams, G. M. (2013). *Microbes Infect.* **15**, 858–865.
- Sheldrick, G. M. (2010). *Acta Cryst.* **D66**, 479–485.
- Steinmann, P., Keiser, J., Bos, R., Tanner, M. & Utzinger, J. (2006). *Lancet Infect. Dis.* **6**, 411–425.
- Thétiot-Laurent, S. A. L., Boissier, J., Robert, A. & Meunier, B. (2013). *Angew. Chem. Int. Ed.* **52**, 7936–7956.
- Werf, M. J. van der, de Vlas, S. J., Brooker, S., Looman, C. W., Nagelkerke, N. J. D., Habbema, J. D. F. & Engels, D. (2003). *Acta Trop.* **86**, 125–139.