

UNIVERSIDADE ESTADUAL DE PONTA GROSSA
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA APLICADA

RENATO FERRAS PENTEADO

ESTUDOS ESTRUTURAIS POR CRISTALOGRAFIA E MODELAGEM
COMPUTACIONAL DA LIPASE DE PINHÃO MANSO (*Jatropha curcas*) E DA
TRIOSE FOSFATO ISOMERASE DE *Naegleria gruberi*

PONTA GROSSA
2016

RENATO FERRAS PENTEADO

ESTUDOS ESTRUTURAIS POR CRISTALOGRAFIA E MODELAGEM
COMPUTACIONAL DA LIPASE DE PINHÃO MANSO (*Jatropha curcas*) E DA
TRIOSE FOSFATO ISOMERASE DE *Naegleria gruberi*

Dissertação apresentada para a obtenção
do título de Mestre em Química Aplicada no
Programa de Pós-Graduação em Química
Aplicada da Universidade Estadual de
Ponta Grossa.

Orientador: Prof. Dr. Jorge Iulek
Coorientadora: Profa. Dra. Viviane Paula
Martini

PONTA GROSSA
2016

Ficha Catalográfica
Elaborada pelo Setor de Tratamento da Informação BICEN/UEPG

P419 Penteado, Renato Ferras
Estudos estruturais por cristalografia e modelagem computacional da lipase de Pinhão manso (*Jatropha curcas*) e da triose fosfato isomerase de *Naegleria gruberi*/ Renato Ferras Penteado. Ponta Grossa, 2016.
92f.

Dissertação (Mestrado em Química Aplicada - Área de Concentração: Química), Universidade Estadual de Ponta Grossa.
Orientador: Prof. Dr. Jorge Iulek.
Coorientadora: Prof^a Dr^a Viviane Paula Martini.

1.*Jatropha curcas*. 2.*Naegleria gruberi*. 3.Cristalografia. 4.Lipase. 5.Triose fosfato isomerase. I.Iulek, Jorge. II. Martini, Viviane Paula. III. Universidade Estadual de Ponta Grossa. Mestrado em Química Aplicada. IV. T.

CDD: 548

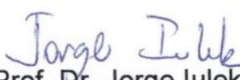
TERMO DE APROVAÇÃO

RENATO FERRAS PENTEADO


**“ESTUDOS ESTRUTURAIS POR CRISTALOGRAFIA E
MODELAGEM COMPUTACIONAL DA LIPASE DE PINHÃO
MANSO (*Jatropha curcas*) E DA TRIOSE FOSFATO
ISOMERASE DE *Naegleria gruberi*”.**

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Química Aplicada da Universidade Estadual de Ponta Grossa, pela seguinte banca examinadora.

Orientador(a):


Prof. Dr. Jorge Iulek
UEPG/PR


Prof. Dra. Nádia Krieger
UFPR/PR


Prof. Dra. Romajana Picada Pereira
UEPG/PR

Ponta Grossa, 09 de agosto de 2016.

AGRADECIMENTOS

À minha mãe pelo apoio incondicional durante toda a vida.

Ao meu orientador Prof. Dr. Jorge Iulek, por ter-me concedido a oportunidade de realizar este trabalho, mostrar o poderoso mundo do Linux, seus *scripts*.

À Prof^a. Dr^a. Viviane Paula Martini, pela amizade e disposição para esclarecer as dúvidas iniciais sobre produção de proteínas.

À Agnes Thiane Pereira Machado pela introdução ao laboratório L11 e operação dos equipamentos.

À Universidade Estadual de Ponta Grossa e ao Programa de Pós Graduação em Química Aplicada, pela oportunidade de realização deste projeto.

Aos órgãos financiadores do projeto: Fundação Araucária e Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pelas bolsas, INBEQMeDI (Instituto Nacional de Biotecnologia Estrutural e Química Medicinal em Doenças Infecciosas) e CNPq pelos recursos para aquisição de equipamentos e reagentes, (573607/2008-7 e 08/57910-00) e LNLS pelo apoio para uso de suas instalações.

RESUMO

O conhecimento da estrutura de proteínas é de grande importância, uma vez que esta informação permite o entendimento dos mecanismos pelos quais elas desempenham suas funções biológicas. Lipases constituem uma família enzimática capaz de realizar a síntese ou hidrólise de ligações éster de substratos triacilgliceróis (TAGs) contendo ácidos graxos de cadeia longa. São alvo de muitos estudos dadas suas potencialidades em um grande número de aplicações envolvendo o grupo funcional éster, por exemplo, em química orgânica sintética. Já o conhecimento estrutural de algumas enzimas é importante para o desenvolvimento de novas drogas terapêuticas ou mesmo contribuir para o entendimento de aspectos evolutivos estruturais, como daquelas pertencentes a vias metabólicas. Neste trabalho foram realizadas a modelagem por homologia da estrutura lipase da planta *Jatropha curcas* e a determinação experimental da estrutura da triose fosfato isomerase do microrganismo *Naegleria gruberi* a partir de três conjuntos de imagens de difração de Raios X. Das três estruturas experimentais obtidas, duas pertencem ao grupo de espaço $C2$, com células unitárias diferentes, e uma ao grupo de espaço $P4_122$. As fases iniciais foram obtidas com o procedimento de substituição molecular utilizando o programa PHASER e todas as estruturas foram refinadas iterativamente com o auxílio dos programas COOT e PHENIX. Em uma das estruturas foi possível modelar três moléculas do agente precipitante Jeffamine[®] presente na condição de cristalização e uma molécula do tampão Tris (no sítio ativo do monômero B). Comparações estruturais foram realizadas entre o modelo refinado e validado e algumas das proteínas homólogas, tendo em vista diferenças observadas no alinhamento baseado em estrutura entre elas e características notadas durante o procedimento de refinamento. Experimentos de dicroísmo circular mostraram que a desnaturação térmica é irreversível para esta proteína.

Palavras-chaves: *Jatropha curcas*, *Naegleria gruberi*, Cristalografia, Lipase, Triose fosfato isomerase, Estrutura tridimensional

ABSTRACT

Knowledge of protein structures is of huge importance, since this information allows to understand the mechanisms through which proteins carry out their biological functions. Lipases constitute an enzymatic family capable to perform synthesis or hydrolysis of ester bonds of triacyl glycerols (TAGs) with long chain fatty acids. These enzymes are the theme of many investigations given their potential to be used in a wide variety of applications involving chemicals with the ester functional group, e.g., in organic synthesis. On the other hand, structural knowledge of some enzymes is important for the development of new therapeutic drugs or even to contribute for the understanding of structural evolutionary features, like those belonging to metabolic pathways. In this work were accomplished the homology modeling of the lipase from *Jatropha curcas* and the structure determination of the triosephosphate isomerase from *Naegleria gruberi* from three X ray diffraction data sets. Among three experimental structures obtained, two belong to $C2$ space group, with different unit cells, and one to $P4_122$ space group. Initial phases were obtained by molecular replacement procedure using the Phaser program and all structures were refined interactively with Coot and Phenix programs. In one structure it was possible to model three molecules of the precipitant agent Jeffamine[®] present in the crystallization solution and one molecule of Tris buffer (placed at the active site). Structural comparisons were performed among the refined and validated model and some of its homologues, taking into account the differences observed in the structural-based alignment among them and characteristics noticed during the refinement procedure. Circular dichroism experiments have shown that thermal denaturation is irreversible to triosephosphate isomerase of *Naegleria gruberi*.

Keywords: *Jatropha curcas*, *Naegleria gruberi*, Crystallography, Lipase, Triosephosphate isomerase, Three dimensional structure

LISTA DE FIGURAS

Figura 1	Esquema geral de reação catalisada por lipase.....	18
Figura 2	Topologia da estrutura secundária canônica do dobramento α/β hidrolase. Hélices α e fitas β são representadas por cilindros e setas, respectivamente. A localização dos resíduos da tríade catalítica está indicada pelos triângulos em preto na ordem nucleófilo – ácido – histidina.....	20
Figura 3	Mecanismo para a reação de hidrólise de um éster genérico, mostrando como atuam os resíduos da tríade catalítica do sítio ativo de uma lipase.....	21
Figura 4	(a) Árvore <i>Jatropha curcas</i> e (b) suas sementes.....	22
Figura 5	Mecanismo para a reação de isomerização da DHAP em D-GAP por uma enzima triose fosfato isomerase.....	25
Figura 6	Estágios assumidos pelo microrganismo <i>Naegleria gruberi</i> . Na parte superior encontra-se a forma ameboide (com destaque para um pseudópodo); na parte inferior esquerda, a forma cística (com destaque para os poros presentes no cisto) e, na parte inferior direita, a forma flagelada (com destaque para seus dois flagelos). As possíveis interconversões entre as formas estão mostradas com setas de duas pontas.....	26
Figura 7	Esquema da via glicolítica de <i>Naegleria gruberi</i> mostrando as principais etapas. A ação da TIM se dá na interconversão indicada pelo número 10.....	28
Figura 8	Esquema da montagem para experimento de cristalização pelo método da gota suspensa.....	33
Figura 9	Caminho percorrido pelos Raios X ao incidirem sobre um material cristalino na condição de difração.....	34
Figura 10	Efeito de absorção diferencial das componentes anti-horária (L) e horária (R) da radiação eletromagnética. A linha tracejada indica o trajeto do vetor (radiação) resultante elipticamente polarizada.....	38
Figura 11	Alinhamento entre a sequência da lipase de <i>Jatropha curcas</i> e as homólogas usadas como molde para a modelagem.....	50
Figura 12	Gráfico de Ramachandran para o modelo selecionado da enzima lipase de <i>Jatropha curcas</i> . Em vermelho, regiões favoráveis, em amarelo, regiões adicionalmente permitidas, em bege, regiões generosamente permitidas e em branco, regiões proibidas.....	51
Figura 13	Estrutura final obtida por modelagem por homologia para a lipase de <i>Jatropha curcas</i>	52
Figura 14	(a) Cromatograma de afinidade mostrando o perfil de eluição da NgTIM no gradiente de imidazol e (b) SDS-PAGE das frações 51 a 63 (na raia 1 encontram-se os padrões de massa molecular, MM; nas raias 2 a 7, as frações 51-56 e nas raias 8 a 14 as frações 57-63, onde a enzima apresenta-se mais pura, correspondendo ao	53

segundo pico no gradiente do cromatograma). Os padrões de massa molecular são: albumina sérica bovina, 66 kDa; β -amilase de batata doce, 56 kDa; anidrase carbônica de eritrócito bovino, 29 kDa e citocromo c de coração de cavalo, 12,4 kDa.....

- Figura 15 (a) Cromatograma de exclusão de tamanho das frações parcialmente puras da primeira etapa de purificação; a enzima *NgTIM* foi eluída no primeiro pico. (b) Imagem do SDS-PAGE correspondente ao primeiro pico do cromatograma. Na raia 1 encontram-se os padrões de massa molecular, MM; nas raias de 2 a 15 encontram-se as frações correspondentes aos volumes de 8,5 a 15 mL do cromatograma. Os padrões de massa molecular são: albumina sérica bovina, 66 kDa; β -Amilase de batata doce, 56 kDa; Anidrase carbônica de eritrócito bovino, 29 kDa e Citocromo c de coração de cavalo, 12,4 kDa..... 54
- Figura 16 Cristais obtidos nas condições descritas anteriormente em (a) A, (b) B e (c) C. A seta vermelha indica o cristal usado na coleta das imagens de difração de Raios X..... 55
- Figura 17 Cristais obtidos nas condições descritas acima (a) A, (b) B e (c) C na segunda produção da triose fosfato isomerase de *Naegleria gruberi*. A seta vermelha indica o cristal usado na coleta das imagens de difração de Raios X em cada condição..... 56
- Figura 18 Alinhamento das sequências de aminoácidos das proteínas *AfTIM* e *NgTIM*. Acima da sequência de *AfTIM* estão representadas as suas estruturas secundárias como calculadas pelo programa DSSP (KABSCH; SANDER, 1983): os cilindros correspondem a hélices α , as setas correspondem a fitas β e a espiral corresponde a uma hélice irregular..... 60
- Figura 19 Região do resíduo Ile²⁰⁶ no início do refinamento em uma região favorável. Em azul, mapa de densidade eletrônica com contorno em 1σ , em verde, mapa Fourier-diferença com contorno em $+3\sigma$ e -3σ 62
- Figura 20 (a) Região próxima à Val³⁸ no início refinamento e (b) a mesma região após a construção manual do modelo no sexto ciclo de refinamento. Em azul, mapas de densidade eletrônica com contorno em 1σ , em verde e vermelho, mapas Fourier-diferença com contorno em $+3\sigma$ e -3σ , respectivamente..... 62
- Figura 21 Fórmula geral da molécula do agente precipitante polieteramina, Jeffamine[®]-M2070..... 65
- Figura 22 Moléculas de Jeffamine[®] modeladas na estrutura obtida a partir do conjunto de dados *NgTIM_2B*. (a) Disposição geral das moléculas de Jeffamine[®] e Tris ao redor do dímero de triose fosfato isomerase de *N. gruberi*; (b) Jeff_A; (c) Jeff_B e (d) Jeff_C. Em azul, mapa de densidade eletrônica com contorno em 1σ , em verde, mapa Fourier-diferença com contorno em $+3\sigma$ e -3σ 66
- Figura 23 Molécula de Tris (Tris(hidroximetil)aminoetano) modelada no sítio 67

	ativo do monômero B. Em azul, mapa de densidade eletrônica com contorno em 1σ , em verde, mapa Fourier-diferença com contorno em $+3\sigma$ e -3σ	
Figura 24	Validações geométricas. (a) Rotâmeros <i>outliers</i> (i) Phe ¹⁵¹ e (ii) Glu ²¹⁸ . Em azul, mapa de densidade eletrônica com contorno em 1σ , em verde, mapa Fourier-diferença com contorno em $+3\sigma$ e -3σ . (b) Gráfico de Ramachandran da estrutura NgTIM_1A. Em vermelho, regiões favoráveis, em amarelo, regiões adicionalmente permitidas, em bege, regiões generosamente permitidas e em branco, regiões proibidas.....	69
Figura 25	Ajuste na densidade eletrônica, índices RSCC (azul) e RSR (vermelho) por resíduo para a estrutura obtida do conjunto de dados NgTIM_1A.....	71
Figura 26	Gráfico de Ramachandran da estrutura NgTIM_2B. Em vermelho, regiões favoráveis, em amarelo, regiões adicionalmente permitidas, em bege, regiões generosamente permitidas e em branco, regiões proibidas.....	72
Figura 27	Ajuste na densidade eletrônica, índices RSCC (azul) e RSR (vermelho) por resíduo para a estrutura obtida do conjunto de dados NgTIM_2B. (a) monômero A e (b) monômero B.....	73
Figura 28	Gráfico de Ramachandran para a estrutura NgTIM_2C. Em vermelho, regiões favoráveis, em amarelo, regiões adicionalmente permitidas, em bege, regiões generosamente permitidas e em branco, regiões proibidas.....	74
Figura 29	Ajuste na densidade eletrônica, índices RSCC (azul) e RSR (vermelho) por resíduo para a estrutura obtida do conjunto de dados NgTIM_2C. (a) monômero A e (b) monômero B.....	75
Figura 30	(a) Regiões de contato entre os monômeros: em roxo o monômero A e em vermelho o monômero B. (b) Diagrama da topologia adotada pela triose fosfato isomerase de <i>N. gruberi</i>	77
Figura 31	Alinhamento baseado em estrutura construído para ambos os monômeros de NgTIM e para algumas homólogas representativas. As estruturas secundárias (estimadas pelo programa DSSP) hélice α , fitas β e hélices irregulares são representadas cilindros, setas e espirais, respectivamente. As estrelas em azul denotam os resíduos componentes do sítio ativo e as estrelas em amarelo denotam os resíduos que apresentam interação com o substrato. <i>A. thaliana</i> _CL (<i>AfTIM</i> do cloroplasto) e <i>A. thaliana</i> _CI (<i>AfTIM</i> do citossol).....	79
Figura 32	A alça do monômero B (verde) projeta-se em direção ao solvente em uma extensão maior que o monômero A (azul), possivelmente devido à molécula de Tris (Tris(hidroximetil)aminoetano) presente no sítio ativo da primeira (representado aqui na forma de bastões).....	81
Figura 33	Superposição dos monômeros B de NgTIM (verde) e TbTIM (azul)	82

complexada com G3P. Representação na forma de bastões dos resíduos Arg¹³⁵, Trp¹⁶⁷ e Glu¹³⁰ de NgTIM; Glu¹²⁹, Arg¹³⁴ e Trp¹⁷⁰ de TbTIM e do ligante G3P.....

- Figura 34 Superposição das triose fosfato isomerases de *N. gruberi* (monômero B em verde), *A. thaliana* (código do PDB 4OBT) e *P. falciparum* (ambas monômeros A, em roxo e rosa, respectivamente)..... 83
- Figura 35 Espectro de dicroísmo circular da enzima triose fosfato isomerase de *N. gruberi*. A curva corresponde aos dados experimentais para o intervalo de comprimento de onda de 195 a 260 nm..... 85
- Figura 36 Perfil de desnaturação térmica da enzima triose fosfato isomerase de *N. gruberi* obtido por dicroísmo circular a 222 nm para aquecimento de 20 a 95 °C e posterior resfriamento de 95 a 20 °C.. 85

LISTA DE TABELAS

Tabela 1	Dados referentes aos parâmetros de coleta e resultados do processamento das imagens dos padrões de difração para os três conjuntos de dados utilizados para determinação de estrutura. Os conjunto de dados <i>NgTIM_1A</i> , <i>NgTIM_2B</i> e <i>NgTIM_2C</i> foram obtidos a partir do cristal B mostrado na Figura 16 e dos cristais A e C mostrados na Figura 17, respectivamente.....	58
Tabela 2	Resultados da substituição molecular para os conjuntos de dados <i>NgTIM_1A</i> , <i>NgTIM_2B</i> , <i>NgTIM_2C</i> . Os ângulos de rotação estão em graus e as translações em coordenadas fracionárias.....	61
Tabela 3	Estatísticas finais dos refinamentos para as estruturas obtidas dos conjuntos de dados <i>NgTIM-1A</i> , <i>NgTIM-2B</i> e <i>NgTIM-2C</i>	64
Tabela 4	Valores dos índices RSCC e RSR para as três moléculas do agente precipitante Jeffamine [®] e do tampão Tris - Tris(hidroximetil)aminoetano – modelados na estrutura <i>NgTIM_2B</i>	68

LISTA DE SIGLAS

AfTIM	Triose fosfato isomerase de <i>Arabidopsis thaliana</i>
BLAST	<i>Basic Local Alignment Search Tool</i> ou Ferramenta de Busca e Alinhamento Local Básica
BSA	<i>Bovine Serum Albumin</i> ou Albumina Sérica Bovina
CCP4	Pacote de programas (<i>Collaborative Computational Project</i> No. 4 ou Projeto Computacional Colaborativo No. 4)
DOPE	<i>Discrete Optimized Protein Energy</i> ou Energia de Proteína Discretamente Otimizada
EO	<i>Ethylene Oxide</i> ou Óxido de Etileno
IPTG	Isopropil- β -D-tiogalactosídeo
JCSG	<i>Joint Center for Structural Genomics</i> ou Centro Associado de Genômica Estrutural
LNLS	Laboratório Nacional de Luz Síncrotron
NgTIM	Triose fosfato isomerase de <i>Naegleria gruberi</i>
NgTIM_1A	Código para conjunto de dados de difração da NgTIM, cristal 1A
NgTIM_2B	Código para conjunto de dados de difração da NgTIM, cristal 2B
NgTIM_2C	Código para conjunto de dados de difração da NgTIM, cristal 2C
PDB	<i>Protein Data Bank</i> ou Banco de Dados de Proteínas
PEG	Polietilenoglicol
PO	<i>Propylene Oxide</i> ou Óxido de Propileno
R_{factor}	Índice residual, usado para estimar o ajuste de uma estrutura de proteína aos dados de difração
R_{free}	Índice residual livre, calculado a partir de 5 a 10% das reflexões de um conjunto de dados que não é usado para refinamento
RFZ	Score (pontuação) da função de rotação em uma substituição molecular
RSCC	<i>Real Space Correlation Coefficient</i> ou Coeficiente de Correlação no Espaço Real
RSR	<i>Real Space Residual</i> ou Residual no Espaço Real
SDS-PAGE	<i>Sodium Dodecyl Sulphate – Polyacrylamide Gel Electrophoresis</i> ou Eletroforese em Gel de Poliacrilamida Contendo Dodecilsulfato de Sódio
SCOP	<i>Structural Classification of Protein</i> ou Classificação Estrutural de Proteínas
TFZ	Score (pontuação) da função de translação em uma substituição molecular
TLS	<i>Translation – Libration – Screw</i> ou Translação – Libração - Parafuso

Tris

Tris-(hidroximetil)aminometano

SUMÁRIO

1	INTRODUÇÃO.....	17
1.1	Justificativa.....	17
1.2	Revisão Bibliográfica.....	18
1.2.1	Lipases.....	18
1.2.2	<i>Jatropha curcas</i>	21
1.2.2.1	Lipase de <i>Jatropha curcas</i>	22
1.2.3	Triose fosfato isomerase.....	24
1.2.4	<i>Naegleria gruberi</i>	25
1.2.4.1	Triose fosfato isomerase de <i>Naegleria gruberi</i>	27
2	TÉCNICAS UTILIZADAS.....	29
2.1	Obtenção das proteínas por expressão heteróloga.....	29
2.2	Cromatografia de afinidade e por exclusão por tamanho.....	30
2.2.1	Cromatografia de afinidade por Ni ²⁺	30
2.2.2	Cromatografia por exclusão por tamanho.....	31
2.3	Eletroforese.....	31
2.4	Dosagem de proteínas.....	32
2.5	Cristalização de proteínas.....	32
2.6	Difração de Raios X.....	33
2.6.1	Lei de Bragg.....	34
2.6.2	Fator de estrutura.....	35
2.6.3	Problema das fases.....	35
2.6.4	Processamento de imagens.....	36
2.6.5	Substituição Molecular.....	36
2.6.6	Refinamento.....	36
2.6.7	Validação.....	37
2.7	Dicroísmo Circular (CD).....	38
3	OBJETIVOS.....	39
3.1	Objetivos gerais.....	39
3.2	Objetivos específicos.....	39
4	MATERIAIS E MÉTODOS.....	40
4.1	Meios de cultivo e aditivos.....	40
4.2	Programas computacionais.....	40
4.3	Produção e expressão das enzimas.....	40
4.4	Extração da enzima lipase de <i>Jatropha curcas</i>	42

4.5	Modelo por homologia.....	42
4.6	Extração da enzima triose fosfato isomerase de <i>Naegleria gruberi</i>	42
4.7	Purificação por cromatografia de afinidade.....	43
4.8	Dosagem proteica.....	43
4.9	Ensaio de cristalização.....	43
4.10	Coleta e processamento dos conjuntos de dados de difração de Raios X.....	44
4.11	Substituição molecular.....	45
4.12	Refinamentos estruturais.....	45
4.13	Validações e comparações estruturais.....	46
4.14	Dicroísmo Circular.....	46
5	RESULTADOS E DISCUSSÃO.....	48
5.1	Lipase de <i>Jatropha curcas</i>	48
5.1.1	Modelo por homologia.....	48
5.2	Triose fosfato isomerase de <i>Naegleria gruberi</i>	52
5.2.1	Produção e purificação da triose fosfato isomerase de <i>Naegleria gruberi</i>	52
5.2.2	Ensaio de cristalização da triose fosfato isomerase de <i>Naegleria gruberi</i> (NgTIM).....	54
5.2.3	Processamento das imagens de difração de Raios X.....	56
5.2.4	Substituição molecular.....	59
5.2.5	Refinamentos das estruturas.....	61
5.2.6	Ligantes na estrutura da triose fosfato isomerase de <i>Naegleria gruberi</i> obtida do conjunto de dados NgTIM_2B.....	65
5.2.7	Validações.....	68
5.2.7.1	Validação da estrutura da triose fosfato isomerase de <i>Naegleria gruberi</i> obtida do conjunto de dados NgTIM_1A.....	68
5.2.7.2	Validação da estrutura da triose fosfato isomerase de <i>Naegleria gruberi</i> obtida do conjunto de dados NgTIM_2B.....	71
5.2.7.3	Validação da estrutura da triose fosfato isomerase de <i>Naegleria gruberi</i> obtida do conjunto de dados NgTIM_2C.....	74
5.2.8	Análises e comparações estruturais para a triose fosfato isomerase de <i>Naegleria gruberi</i> obtida do conjunto de dados NgTIM_2B.....	76
5.2.9	Dicroísmo Circular (CD) da triose fosfato isomerase de <i>Naegleria gruberi</i> : estabilidade térmica.....	84
6	CONCLUSÃO.....	86

1 INTRODUÇÃO

1.1 Justificativa

As proteínas são compostos de origem biológica que são encontrados em todas as formas de vida. São formadas a partir da reação de condensação de moléculas de aminoácidos com a liberação de uma molécula de água por ligação formada, sendo a forma enantiomérica mais comum a L, embora alguns organismos apresentem D-aminoácidos na parede celular (TORTORA; FUNKE; CASE, 2012). As proteínas desempenham diversas funções nos organismos vivos, como por exemplo, suporte estrutural, catálise de reações, transporte de substâncias e proteção do organismo. A surpreendente diversidade funcional das proteínas provém da ordem com a qual os diferentes L-aminoácidos estão conectados (sua estrutura primária), sendo possível correlacionar os níveis organizacionais superiores (estrutura secundária, terciária e quaternária – esta última, quando existente) e fazer inferências do tipo de estrutura tridimensional adotada com base na estrutura primária. (NELSON; COX, 2002).

Desta forma, o conhecimento da sequência de aminoácidos e da estrutura tridimensional é muito importante para o entendimento da respectiva função e, a partir destas informações, torna-se possível, por exemplo, o desenvolvimento de inibidores enzimáticos, ou mutações estruturais com a finalidade de melhorar as características da proteína.

Não há relatos da determinação da estrutura tridimensional da lipase putativa de pinhão manso, *Jatropha curcas* (*J. curcas*), que será denominada aqui de *JcLip*, e da triose fosfato isomerase (TIM) do organismo *Naegleria gruberi* (*N. gruberi*), que será denominada *NgTIM*. Assim, estudos por cristalografia de Raios X podem se tornar importantes ao permitirem a caracterização de suas estruturas tridimensionais.

As duas enzimas são interessantes do ponto de vista biotecnológico. A primeira, para química orgânica de síntese, uma vez que se trata de uma enzima lipolítica utilizada em biocatálise para obtenção de biodiesel. Já a segunda, trata-se de uma enzima relacionada à via glicolítica do protozoário *N. gruberi*, ao qual uma espécie correlata, *Naegleria fowleri*, causa uma doença grave e geralmente fatal em humanos, a Meningoencefalite Amebiana Primária. Os estudos com a enzima triose

fosfato isomerase (TIM) de *N. gruberi* poderão ser importantes considerando-se que se pode utilizar a estrutura desta enzima como modelo para a similar da cepa patogênica *Naegleria fowleri*.

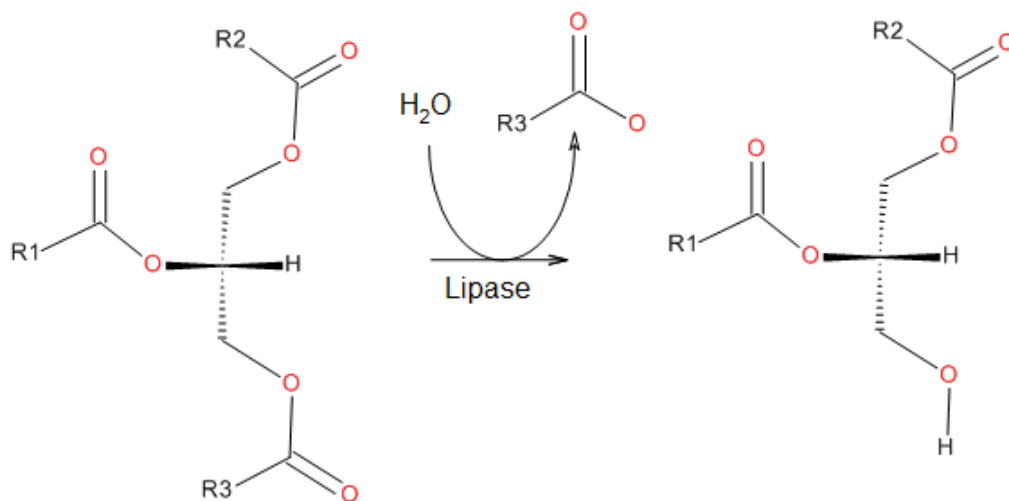
1.2 Revisão Bibliográfica

1.2.1 Lipases

Estima-se que a maioria das enzimas utilizadas industrialmente ou comercializadas são de origem microbiana. Isso se deve basicamente à facilidade de obtenção, cultivo, produção, extração e purificação (SHARMA; CHISTI; BANERJEE, 2001).

Lipases (E.C. 3.1.1.3, triacilglicerol hidrolases) são enzimas que catalisam a hidrólise de ligações éster de triacilgliceróis (TAG's) em ambientes aquosos (Figura 1). Elas podem também realizar a síntese de TAG's em um meio orgânico (SHARMA; CHISTI; BANERJEE, 2001). Devido a isso, são interessantes do ponto de vista biotecnológico para serem usadas em um grande número de reações catalisadas, tais como: esterificação, transesterificação, acilação regioseletiva de gliceróis e mentóis, síntese de peptídeos, entre outros produtos químicos (SHARMA; CHISTI; BANERJEE, 2001).

Figura 1. Esquema geral de reação catalisada por lipase.

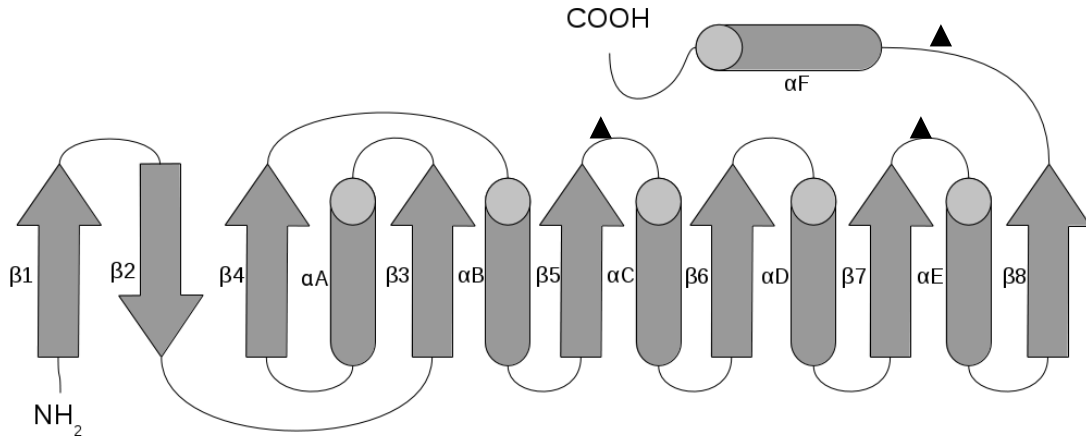


Fonte: O autor.

As lipases apresentam diversas aplicações: em formulação de detergentes (com cerca de 32% das vendas de mercado, entretanto, para esta finalidade precisam ser termoestáveis e resistir ao ambiente alcalino), em síntese de biosurfactantes, em indústria óleo-química (como na hidrólise e modificação de óleos e gorduras), em agroquímica, em manufatura de papel, em nutrição (como o realce de sabor em processamento de alimentos), em cosméticos e processamento de fármacos (resolução de misturas racêmicas) (SHARMA; CHISTI; BANERJEE, 2001).

Do ponto de vista estrutural, as lipases, bem como as esterases, apresentam grande similaridade na sequência de aminoácidos, embora possam diferir radicalmente na especificidade e função da sua atividade catalítica. A comparação entre as estruturas tridimensionais destas enzimas mostra uma característica estrutural comum conhecida como dobramento α/β hidrolase. Essencialmente, este tipo de estrutura é constituído por um arranjo central de fitas β rodeado por um número variável de hélices α e acomoda a chamada tríade catalítica, que é geralmente composta por resíduos de serina, aspartato ou glutamato, e histidina. A topologia canônica do tipo α/β hidrolase é mostrada na Figura 2. A serina nucleofílica encontra-se em uma alça entre a fita $\beta 5$ e a hélice α seguinte com a sequência Gly-X-Ser-X-Gly, onde 'X' pode ser qualquer aminoácido, exceto prolina. A sequência dos motivos estruturais fita $\beta 5$ -alça-hélice α é denominada “cotovelo nucleofílico” e é responsável pelo posicionamento ótimo da serina nucleofílica para que possa, por um lado, interagir com a cadeia lateral da histidina do sítio ativo e, por outro lado, interagir com o substrato. A ordem dos resíduos da tríade catalítica na família α/β hidrolase é serina – resíduo ácido – histidina, diferindo da ordem serina – histidina – ácido aspártico encontrada em proteases (GUILHAM; LEHNER, 2005). O resíduo ácido catalítico fica localizado em uma alça após a fita $\beta 7$, embora possa apresentar certa variabilidade na sua posição, como foi observado na estrutura da lipase pancreática humana (após a fita $\beta 6$), não obstante a histidina catalítica ainda pertença à alça que segue a última fita β .

Figura 2. Topologia da estrutura secundária canônica do dobramento α/β hidrolase. Hélices α e fitas β são representadas por cilindros e setas, respectivamente. A localização dos resíduos da tríade catalítica está indicada pelos triângulos em preto na ordem nucleófilo - ácido - histidina.



Fonte: O autor.

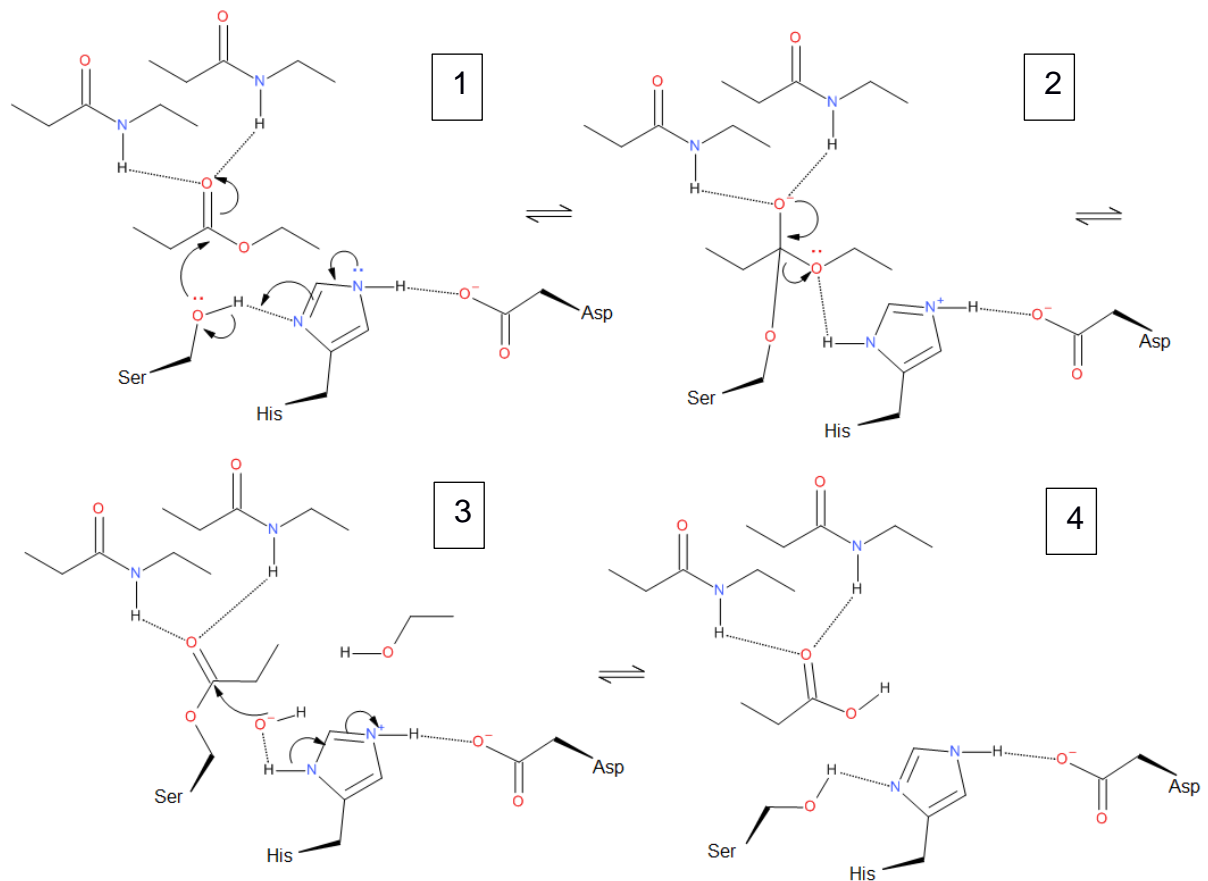
Muitas lipases, assim como esterases, possuem um elemento de estrutura secundária que cobre o sítio ativo (conhecido como “lid”) que deve sofrer alterações conformacionais para permitir o acesso do substrato ao mesmo (GUILHAM; LEHNER, 2005). Em outros casos, entretanto, as enzimas não possuem tal estrutura.

A diferenciação entre lipases e esterases ainda não é bem definida. Inicialmente poder-se-ia fazê-la através do fato de as lipases apresentarem o fenômeno da ativação interfacial, como proposto por Sarda e Desnuelle (1958). Entretanto, esta característica não é encontrada em todas as lipases e, em trabalhos mais recentes, a denominação destas é feita como carboxilesterases que hidrolisam acilgliceróis contendo ácidos graxos de cadeia longa (aqueles com cadeia carbônica com dez ou mais átomos de carbono), enquanto que aquelas capazes de hidrolisar ésteres de glicerol cujos grupos acila possuem menos de dez carbonos são denominadas esterases (JAEGER; DIJKSTRA; REETZ, 1999; JAEGER et al., 1994).

A reação de hidrólise da ligação éster do substrato se processa possivelmente segundo o mecanismo descrito na Figura 3 que envolve quatro passos básicos. Primeiramente, ocorre o ataque nucleofílico pelo átomo de oxigênio da hidroxila serínica ao carbono carbonílico da ligação éster (passo 1), formando um intermediário tetraédrico (passo 2), cuja carga negativa é estabilizada por ligações de hidrogênio com grupos NH da cadeia principal e pelo momento de dipolo da hélice αC (Figura 2). Possivelmente, este efeito eletrostático estabilizante justifica a

elevada conservação da referida hélice em α/β hidrolases (KNOWLES, 1991). Após a liberação da porção alcóxi do substrato, a porção carboxílica permanece covalentemente ligada à lipase por uma ligação éster com o resíduo de serina (passo 3). A renovação da enzima ocorre pelo ataque nucleofílico de uma molécula de água com seu potencial nucleofílico aumentado pela interação com o resíduo de histidina da tríade catalítica, liberando a porção carboxílica para o solvente (passo 4) (JAEGER; DIJKSTRA; REETZ, 1999).

Figura 3. Mecanismo para a reação de hidrólise de um éster genérico, mostrando como atuam os resíduos da tríade catalítica do sítio ativo de uma lipase.



Fonte: O autor.

1.2.2 *Jatropha curcas*

Jatropha curcas, conhecida popularmente como pinhão manso, pertence à família Euphorbiaceae e é comumente encontrada em regiões da América do Sul, África e Ásia. A planta é classificada como uma árvore pequena ou um arbusto

grande (Figura 4a) que apresenta geralmente de 3 a 6 m de altura (embora em condições favoráveis possa alcançar de 8 a 10 m de altura) e passou a ter importância econômica em áreas de clima e solo extremos devido à sua resistência à seca (KUMAR; SHARMA, 2006). Praticamente todas as partes da planta são usadas na medicina natural, por exemplo, em algumas regiões da África e da Malásia as folhas são usadas como purgativo; nas Filipinas contra diarreia; em Camarões como remédio contra reumatismo, icterícia, entre outros. O óleo obtido das sementes (Figura 4b) é utilizado na fabricação de sabões e como fonte de energia (STAUBMANN et al., 1999). Além disso, a presença de substâncias tóxicas como ésteres de forbol é apontada como uma vantagem no uso desta planta na produção de biodiesel, uma vez que não entraria em conflito com produções destinadas à indústria de alimentos (GU et al., 2010). Atualmente, o interesse em investimentos na plantação de *J. curcas* deve-se às suas potencialidades energéticas. Suas sementes podem ser processadas para extração de óleo que pode ser usado para produção de biodiesel (SHARMA; DHAMIJA; PARASHAR, 2012).

Figura 4. (a) Árvore *Jatropha curcas* e (b) suas sementes.



Fonte: <http://shop.theplantattraction.com/Jatropha-Curcas-Tree-Bio-Diesel-Fuel-Source-Physic-Nut-5-Seeds-S-Jatropha-Curcas.html>

1.2.2.1 Lipase de *Jatropha curcas*

Lipases de eucariotos, como de plantas, estão envolvidas em muitos estágios do

metabolismo de lipídeos, incluindo digestão, absorção, restituição de gorduras, além de papel no metabolismo de lipoproteínas (BALASHEV; JENSEN; BJORNHOLM, 2001).

Nas plantas, as lipases são encontradas em tecidos que servem como reserva de energia. O mecanismo de interação entre lipases e lipídeos ainda não está totalmente claro e é objeto de intensa investigação, como descrito por Balashev; Jensen e Bjornholm (2001).

As pesquisas com esta classe enzimática estão relacionadas principalmente em caracterizações estruturais, elucidação do mecanismo de ação, determinação de parâmetros cinéticos, sequenciamento e clonagem dos seus genes (SHARMA et al., 2001) e tentativas no sentido de melhorar as propriedades das lipases já conhecidas, por exemplo, usando evolução direcionada (POWELL et al., 2001).

Uma lipase de planta de grande interesse é a de pinhão manso, *Jatropha curcas*. Ela foi relativamente pouco estudada, sendo encontrada nas sementes dormentes e após terem sido germinadas (ABRIGOR et al., 2002), embora Staubmann et al. (1999) tenham encontrado atividade desta enzima apenas em sementes germinadas. Ambos os autores conseguiram realizar apenas a purificação parcial desta enzima. Nos experimentos realizados por Staubmann et al. (1999), a atividade lipolítica durante a etapa de hidrólise dos substratos testados tem uma ligeira preferência pela hidrólise do ácido gama-linoleico. No que diz respeito às reações de transesterificação, a eficiência deste processo foi maior em uma atividade de água reduzida, quando comparada com a do preparado comercial denominado Lipozyme. Isto possivelmente está relacionado à semelhança da atividade da água no meio reacional destes experimentos com aquela encontrada nas sementes durante o processo de germinação. Abrigor et al. (2002) relatam que as sementes formam germinadas sob luz ambiente a aproximadamente 25 °C por seis dias. Diferentemente do encontrado nos primeiros experimentos descritos acima, neste caso a atividade lipolítica foi maior para triacilgliceróis com ácidos graxos de cadeia longa. Além disso, a velocidade da reação foi maior quando substratos monoacil foram usados em relação àqueles diacil e triacilglicerol. Outros estudos realizados com esta lipase destinados ao desenvolvimento de novos processos de produção de biodiesel estão relacionados à sua imobilização em procedimentos mistos enzimático/químico (SOUSA et al., 2010). Neste caso, os autores encontraram uma

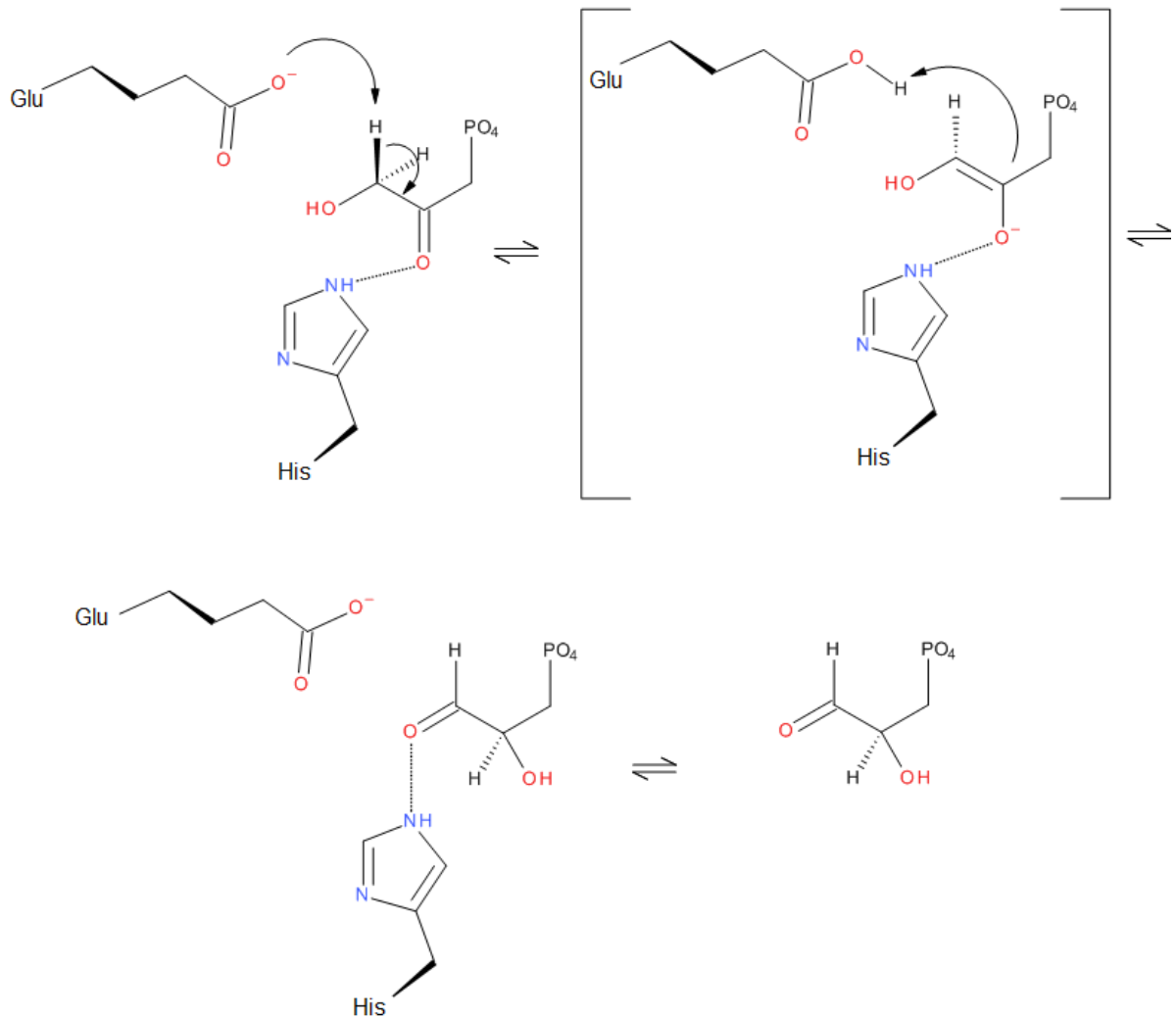
excelente capacidade de hidrólise de diferentes matérias-primas para produção de biodiesel. Além disso, o glicerol resultante da etapa de hidrólise, que constitui um resíduo em processos de produção de biodiesel, apresentou grau de pureza elevado (geralmente este resíduo apresenta metanol como contaminante e pH alcalino proveniente dos catalisadores usados em procedimentos convencionais), tornando possível sua utilização em alimentos, cosméticos, fármacos, entre outros (SOUSA et al., 2010). Motivados pelas dificuldades encontradas em processos convencionais de produção de biodiesel, Gu et al., (2010) utilizaram a tecnologia de extração reativa auto catalisada, onde as sementes de *J. curcas* germinadas servem como a fonte de óleo e de catalisador (lipase). Desta forma, as complicações decorrentes da extração e purificação são reduzidas, bem como o custo global da produção.

1.2.3 Triose fosfato isomerase

Triose fosfato isomerase (E.C. 5.3.1.1; TIM) é uma enzima encontrada na via glicolítica responsável por realizar a conversão estereoespecífica da di-hidroxiacetona fosfato (DHAP) em D-gliceraldeído-3-fosfato (D-GAP). A reação se processa sem a intermediação de cofatores ou íons metálicos. Nesta reação, uma ligação C-H precisa ser quebrada. Assim como em muitas outras reações biológicas, o referido átomo de carbono encontra-se adjacente a um grupamento carbonila (neste caso, do grupo cetona da DHAP) (WIERENGA et al., 2010). O mecanismo para a reação de isomerização de DHAP em D-GAP é mostrado na Figura 5. A partir dos estudos de Rose e O'Connell, (1969), pôde-se concluir que a base capaz de realizar a abstração do próton da DHAP deveria ser um resíduo de aminoácido de cadeia lateral acídica. De La Mare et al. (1972), comprovaram, a partir de estudos realizados com a TIM de músculo de galinha, que este resíduo é o Glu¹⁶⁷. Além disso, dois fatores importantes em sua estrutura são a possibilidade de estabilização eletrostática dos ânions gerados e de exclusão do solvente (GO, 2010; LOLIS, 1990). A manutenção da atividade da triose fosfato isomerase é de importância crítica para o funcionamento apropriado das células humanas, resultando em uma doença grave quando esta apresenta decréscimo de atividade (a deficiência de triose fosfato isomerase), como pôde ser observado nos estudos realizados por Orosz et al. (2009). Por isso, torna-se importante o estudo da estrutura da enzima de *N. guberi* a

fim de procurar diferenças e/ou similaridades com outras TIMs de diferentes organismos, que possam servir para esclarecer aspectos evolutivos das espécies.

Figura 5. Mecanismo para a reação de isomerização da DHAP em D-GAP por uma enzima triose fosfato isomerase.



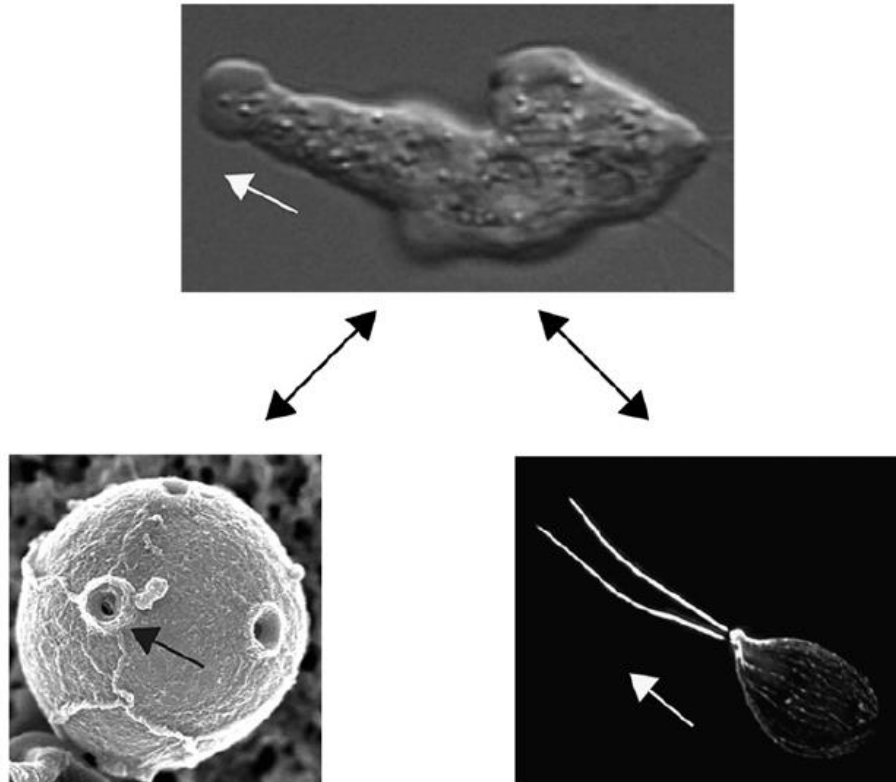
Fonte: O autor.

1.2.4 *Naegleria gruberi*

O microrganismo *Naegleria gruberi* é um protista heterotrófico de vida livre comumente encontrado em ambientes aeróbicos e microaeróbicos em água potável e solos úmidos. Sua forma predominante é a amebóide que pode se reproduzir a cada 1,6 horas quando se alimenta de bactérias, além de apresentar também as formas flagelada e cística (FULTON, 1993; FULTON, 1970). Os referidos estágios

morfológicos são mostrados na Figura 6.

Figura 6. Estágios assumidos pelo microrganismo *Naegleria gruberi*. Na parte superior encontra-se a forma ameboide (com destaque para um pseudópodo); na parte inferior esquerda, a forma cística (com destaque para os poros presentes no cisto) e, na parte inferior direita, a forma flagelada (com destaque para seus dois flagelos). As possíveis interconversões entre as formas estão mostradas com setas de duas pontas.



Fonte: Fritz-Laylin et al. (2011).

N. gruberi não apresenta risco para a saúde humana, entretanto, o gênero *Naegleria* inclui o microrganismo *N. fowleri*, que, embora viva em fontes de água aquecidas e em ambientes de forma semelhante à *N. gruberi*, pode também agir como um patógeno oportunista em humanos, causando uma doença quase sempre fatal, a Meningoencefalite Amebiana Primária (MAP) (VISVESVARA; MOURA; SCHUSTER, 2007). Os sintomas iniciais aparecem cerca de cinco dias após a infecção e incluem dores de cabeça, febre, náusea e vômito. Entre os sintomas tardios apresentados pelo infectado estão: enrijecimento do pescoço, confusão mental, falta de atenção nas pessoas e nos arredores, perda de equilíbrio, convulsões e alucinação (CDC, 2016). Após o aparecimento dos sintomas iniciais, a

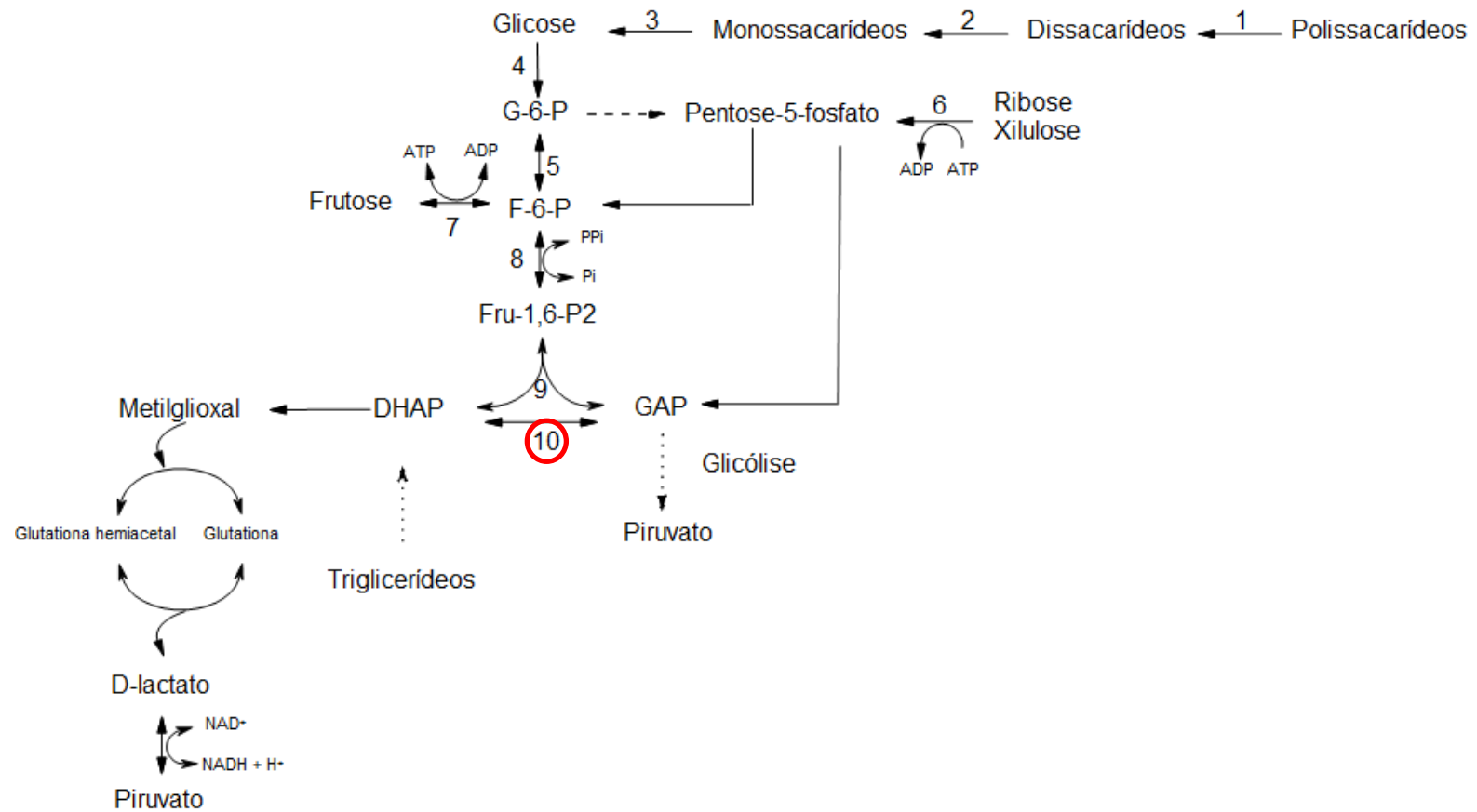
enfermidade evolui rapidamente e o infectado entra em óbito em cerca de cinco dias (podendo variar de um a doze dias).

Há poucos anos, foi realizado o sequenciamento completo do genoma de *N. gruberi*, gerando a possibilidade de pesquisas no desenvolvimento de drogas com ação contra a espécie patogênica relacionada *N. fowleri* (OPPERDOES; DE JONCKHEERE; TIELENS, 2011). Neste sentido, a resolução de estrutura tridimensional desta proteína pode ser de grande valia para o conhecimento de elementos envolvidos no seu processo catalítico.

1.2.4.1 Triose fosfato isomerase de *Naegleria gruberi*

Até o momento, não há estudos de caracterização da triose fosfato isomerase de *N. gruberi*, tendo em vista que seu gene codificante ainda é considerado como preditivo da proteína. No que diz respeito ao seu papel no metabolismo deste organismo, as informações disponíveis que se encontram no trabalho de Opperdoes; De Jonckheere e Tielens (2011), que é baseado no genoma de *N. gruberi* sequenciado por Fritz-Laylin et al. (2011), são apenas de cunho especulativo e necessitam confirmação experimental. Um esquema da parte citossólica da via glicolítica de *N. gruberi* é reproduzida na Figura 7.

Figura 7. Esquema da via glicolítica de *Naegleria gruberi* mostrando as principais etapas. A ação da Triose fosfato isomerase se dá na interconversão indicada pelo número 10.



Fonte: O autor, baseado na via glicolítica proposta por Opperdoes; De Jonckheere e Tielens (2011).

2 Técnicas utilizadas

2.1 Obtenção das proteínas por meio de expressão heteróloga

Denomina-se expressão heteróloga a técnica utilizada para expressar uma proteína em um organismo diferente daquele de onde a mesma tem origem. Dentre os sistemas mais comumente utilizados para expressão heteróloga, destacam-se as células bacterianas de *Escherichia coli*. Este sistema é caracterizado por apresentar maior rendimento, rapidez de biossíntese, baixo custo e a possibilidade de realizar produções desde pequena a grande escala (LEITE et al., 2001). Por outro lado, Twyman et al. (2003), afirmam que podem existir desvantagens nestes sistemas devido à necessidade de modificações pós-traducionais complexas que necessitem da manutenção de suas estruturas originais íntegras, sem adição e remoção de resíduos de aminoácidos, que possam comprometer sua função e especificidade naturais.

Para podermos utilizar a expressão heteróloga, o gene que codifica a proteína de interesse deve ser inserido em um fragmento de DNA denominado de plasmídeo (isto é, um tipo de DNA encontrado em células bacterianas que atua independentemente do DNA genômico, conferindo ao organismo que o detém alguma característica que aumente sua capacidade de sobrevivência em relação aos demais, por exemplo, resistência a antibióticos). Os plasmídeos podem ser modificados geneticamente para incorporar, além de um gene de resistência a antibiótico que lhes confere capacidade de seleção, uma região de policlonagem com sítios de enzimas de restrição, para inserir DNAs de qualquer origem, e uma origem de replicação para que ele possa ser replicado dentro da célula. A seleção dos organismos contendo o plasmídeo que carrega o gene de interesse é feita por meio da adição do antibiótico específico ao meio de cultura, em uma concentração que permita a diferenciação entre as células transformadas (que contem o plasmídeo, então resistentes ao antibiótico) das não transformadas (sensíveis ao antibiótico). Alguns plasmídeos podem conter promotores de transcrição fortes para expressar proteínas dentro de outros organismos, sendo chamados assim de plasmídeos de expressão (SILVA, 2001).

2.2 Cromatografia de afinidade e por exclusão por tamanho

Após a determinação da forma de obtenção da proteína a ser estudada (de origem natural ou de expressão heteróloga), é necessário separá-la das demais e para isso existem diversas técnicas de separação e purificação. Duas técnicas de grande importância para a purificação de proteínas são as cromatografias de afinidade e por exclusão de tamanho (ou filtração em gel). As técnicas cromatográficas, de forma geral, exploram características das proteínas com diferença de tamanho, carga elétrica, afinidade de ligação, entre outras. Basicamente, compreende um material sólido poroso, denominado fase estacionária, que possui propriedades químicas adequadas e é mantido encerrado em uma coluna. Faz-se percolar um solvente/solução pela fase estacionária e, em seguida, a solução de proteínas. A migração de cada proteína será diferente e dependerá das propriedades de cada uma no que diz respeito à interação com a fase estacionária (STRYER; TYMOCZKO; BERG, 2004).

2.2.1 Cromatografia de afinidade por íons Ni^{2+}

Neste método, as proteínas são separadas levando-se em consideração sua especificidade de ligação pela fase estacionária que contém íons níquel(II) complexados a um polímero. A ligação torna-se possível devido à presença de resíduos de histidina presentes na proteína. Átomos de nitrogênio do anel da cadeia lateral da histidina completam a esfera de coordenação dos íons Ni^{2+} por complexação, retendo a proteína de interesse enquanto as demais, por não possuírem afinidade tão grande pela fase estacionária, são eluídas da coluna. A diferença de afinidade entre a proteína alvo e as demais é conseguida na etapa de clonagem, onde utiliza-se um plasmídeo que codifica uma região com várias histidinas (tipicamente seis resíduos), além da sequência de aminoácidos de interesse. Em seguida, a proteína alvo é eluída da coluna usando-se uma solução contendo imidazol (STRYER; TYMOCZKO; BERG, 2004).

2.2.2 Cromatografia por exclusão por tamanho

Esta técnica, também conhecida como filtração em gel, proporciona a separação das proteínas com base em seus diferentes tamanhos. A fase estacionária é composta por um polímero contendo ligações cruzadas de forma a produzir poros de tamanho específico. As proteínas maiores migram mais rapidamente pela fase estacionária por serem incapazes de penetrar nos poros das esferas da fase estacionária; ocorre o oposto com as menores, o seu movimento sofre um retardo a medida que seguem por um caminho mais tortuoso ao penetrar nos poros da fase estacionária. Dessa forma, as proteínas maiores são eluídas nas frações iniciais (STRYER; TYMOCZKO e BERG, 2004).

2.3 Eletroforese

Uma importante técnica de separação e identificação de proteínas, por meio de sua migração em campo elétrico, é a eletroforese. Constitui-se em um método analítico especialmente útil que permite inclusive a determinação do peso molecular aproximado de proteínas. A eletroforese de proteínas é geralmente executada em géis de um polímero com ligações cruzadas, por exemplo, a poli(acrilamida). O gel age como uma peneira molecular, dificultando a migração da proteína na proporção aproximada da razão entre carga e massa ou refletindo somente a massa da proteína. Em alguns casos, a migração também é afetada pela forma da proteína.

A velocidade de migração no gel v é dada pela equação abaixo (Equação 1):

$$v = \frac{Ez}{f} \quad \text{Equação 1}$$

Onde:

v é a velocidade de migração, E é a intensidade do campo elétrico, z é a carga total da proteína e f é o coeficiente de atrito.

Em um dos métodos eletroforéticos, usado para determinar a pureza e o peso molecular, utiliza-se o detergente dodecilsulfato de sódio (SDS). O SDS liga-se às proteínas em quantidades aproximadamente proporcionais ao peso molecular da

proteína, de forma que o detergente contribui com uma carga líquida negativa grande em relação à carga da proteína. A razão carga/massa para a maioria das proteínas torna-se a mesma, fazendo com que elas possam ser separadas quase exclusivamente com base em suas massas, com aquelas de menor massa migrando mais rapidamente (STRYER; TYMOCZKO; BERG, 2004).

2.4 Dosagem de proteínas

A quantificação do teor de proteínas em solução pode ser feita empregando-se diferentes métodos, um dos mais usados é o método de Bradford (BRADFORD, 1976).

A base do método é a capacidade de ligação do corante Coomassie G-250 a resíduos de aminoácido contendo grupos funcionais básicos ou aromáticos. A ligação à proteína ocorre em cerca de dois minutos e dura por duas horas. Na condição da reação, a interação entre proteína e corante ocorre de forma a deslocar o equilíbrio de dissociação do corante para sua forma aniônica, que absorve radiação majoritariamente no comprimento de onda de 595 nm, sendo este o comprimento de onda então usado nas leituras de absorção das amostras (ZAIA; ZAIA; LICHTIG, 1998).

2.5 Cristalização de Proteínas

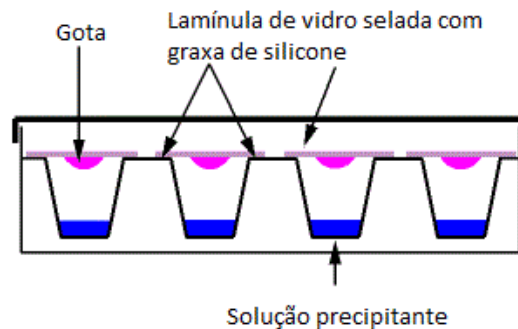
Para que seja possível o estudo estrutural usando a técnica de difração de Raios X é necessário cristalizar a proteína alvo.

A formação e o crescimento de um cristal de proteína demanda alcançar uma condição de supersaturação na solução que se encontra. Isso pode ocorrer quando faz-se o solvente evaporar lentamente ou pela alteração de parâmetros como: temperatura, pH, força iônica ou presença de aditivos que auxiliem a agregação da proteína na forma cristalina. Uma macromolécula biológica segue as mesmas leis termodinâmicas de qualquer outro tipo de molécula inorgânica ou orgânica no que se refere à supersaturação, nucleação e crescimento de cristais (BOISTELLE; ASTIER, 1988).

Existem várias possibilidades de levar a solução da proteína à saturação, uma

delas é a técnica de difusão de vapor. Na sua forma mais comumente utilizada, a da difusão de vapor em gota suspensa (Figura 8), alguns poucos microlitros da solução da proteína são misturados com igual volume de uma solução precipitante presente no reservatório. Esta gota é montada na superfície de uma lamínula de vidro previamente siliconizada que cobre e sela o reservatório. A mistura na gota possui uma concentração menor que a da solução no reservatório, portanto, a água evapora da gota para o reservatório. Como resultado, as concentrações da proteína e do precipitante aumentam até o limite de solubilidade da proteína ser excedido; a solução torna-se supersaturada, nucleação e separação de fase começam a ocorrer e os cristais podem ser formados (RUPP, 2010).

Figura 8: Esquema da montagem para experimento de cristalização pelo método da gota suspensa.



Fonte: O autor.

2.6 Difração de Raios X

O procedimento experimental mais comum de se obter uma imagem detalhada da estrutura de uma macromolécula, permitindo uma resolução dos seus átomos individuais, é interpretar a difração sofrida por Raios X a partir de muitas moléculas idênticas em um arranjo ordenado como em um cristal. Este método é conhecido como Cristalografia de Raios X de Monocristal (RHODES, 2000).

O cristal é colocado entre uma fonte de Raios X e um detector apropriado, ficando no caminho de um feixe estreito de Raios X provenientes da fonte. A forma mais simples de detector é um filme que, quando revelado, exibe pequenas áreas enegrecidas (*spots*), correspondentes aos pontos onde os feixes de Raios X que emergiram do cristal atingiram-no e o impressionaram. Estes *spots* são chamados de reflexões porque os feixes emergem do cristal como se tivessem sido refletidos por planos de átomos que o constituem, e o conjunto destes *spots* é chamado de

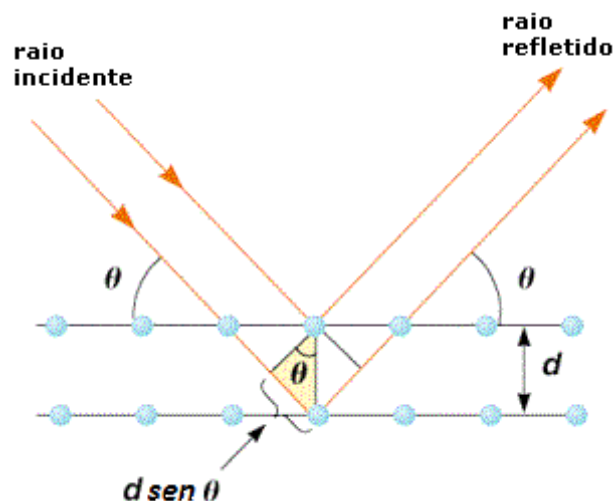
padrão de difração.

Uma vez que o espaçamento da rede real (a rede cristalina) é inversamente proporcional ao das reflexões, as dimensões da cela unitária podem ser calculadas a partir do espaçamento observado na rede recíproca, ou seja, na imagem obtida por difração de Raios X (BLOW, 2005).

2.6.1 Lei de Bragg

A explicação do fenômeno de difração pode ser feita pela Lei de Bragg. Segundo esta lei, o ângulo com o qual um feixe de Raios X emerge do cristal (raio difratado) pode ser determinado tratando o fenômeno como uma reflexão por meio de um conjunto de planos atômicos equivalentes paralelos entre si. W. L. Bragg demonstrou que a difração ocorre quando a diferença da distância percorrida pelos raios incidentes é igual a um múltiplo do comprimento de onda da radiação incidente. A diferença da distância depende do ângulo de reflexão e do espaçamento entre os planos, segundo a equação da Lei de Bragg (Equação 2) (BLOW, 2005; RHODES, 2000) (Figura 9):

Figura 9: Caminho percorrido pelos Raios X ao incidirem sobre um material cristalino na condição de difração.



Fonte: wiki.stoa.usp.br/Usuário:Clovisdsn.

$$2d \sin \theta = n\lambda$$

Equação 2

2.6.2 Fator de estrutura

No padrão de difração, cada *spot* da imagem gerada é produzido por um feixe difratado (que é a soma de todas as ondas “refletidas” por um grupo de planos de átomos) que possui uma amplitude e fase.

A expressão que define o fator de estrutura para cada reflexão é dada pela Equação 3:

$$F_{hkl} = \sum_{j=1}^n f_j e^{2\pi i(hx_j + ky_j + lz_j)} \quad \text{Equação 3}$$

Nesta expressão, f_j é o fator de espalhamento atômico e a soma é realizada sobre todos os átomos da célula unitária. A variável fator de espalhamento atômico nos fornece a quantidade de radiação será espalhada por um átomo em uma dada direção. Quando são conhecidas as fases e amplitudes de todos os fatores de estrutura, é possível calcular a transformada de Fourier inversa, fornecendo a densidade eletrônica do cristal, como mostrado pela equação abaixo (Equação 4) (BLOW, 2005):

$$\rho(x, y, z) = \frac{1}{V} \sum_{hkl} F(hkl) e^{-2\pi i(hx + ky + zl)} \quad \text{Equação 4}$$

2.6.3 O problema das fases

As imagens coletadas durante o experimento de difração fornecem as intensidades dos raios difratados, mas não apresentam qualquer informação a respeito das fases das ondas que compõem estes raios. Isso impossibilita-nos de calcular a densidade eletrônica diretamente. Existem métodos para resolver este problema como a Substituição Isomórfica (onde são produzidos os chamados derivados isomórficos por meio da introdução de metais de elevada massa atômica no cristal de proteína) e a Substituição Molecular (um dos mais utilizados atualmente e brevemente discutido a seguir) (RUPP, 2010).

2.6.4 Processamento das imagens

Encerrado o procedimento de coleta dos dados de difração de Raios X, o conjunto de imagens obtido guarda a informação do módulo dos fatores de estrutura em cada um dos *spots* que formam o padrão de difração sob a forma da intensidade de cada ponto. Esta intensidade é dada pelo grau de enegrecimento do *spot*. A obtenção das intensidades a partir de cada reflexão é possível pelo emprego de programas computacionais como MOSFLM e XDS, responsáveis por fazer a integração dos *spots*. Em seguida, realiza-se a indexação, onde são determinadas informações como os parâmetros da célula unitária, grupo de espaço, mosaicidade, etc. A terceira operação constitui-se do escalonamento, feito através de programas como XSCALE. Esta etapa é necessária porque cada imagem apresenta intensidades em diferentes escalas devido a variações da radiação incidente, portanto, deve-se atribuir um fator de escala para cada uma delas a fim de tornar possível a comparação entre os fatores de estrutura obtidos.

2.6.5 Substituição Molecular

Neste método, utiliza-se a estrutura de uma proteína conhecida, com a maior identidade possível no que diz respeito à sequência de aminoácidos e conseqüentemente provável estrutura parecida, como modelo de partida, para que fases aproximadas possam ser calculadas para o padrão de difração da proteína desconhecida. Essa estrutura, modelo inicial, é então colocada dentro da célula unitária obtida experimentalmente para a proteína alvo, de modo que as fases possam ser calculadas. A estimativa da orientação e posição do modelo inicial é decomposta em dois componentes: a rotação tridimensional da molécula na orientação correta e translação no espaço tridimensional da mesma orientada na posição correta (BLOW, 2005).

2.6.6 Refinamento

No processo de refinamento realiza-se o ajuste do modelo à densidade eletrônica obtida experimentalmente. Neste procedimento, um dos objetivos é

melhorar a concordância entre os fatores de estrutura calculados (F_{calc}) e os fatores de estrutura observados (F_{obs}), por meio da variação de parâmetros como posições atômicas, fatores de temperatura, ângulos de ligação, entre outros. Uma maneira de acompanhar o avanço do refinamento é pelo cálculo dos índices R_{factor} e R_{free} ; abaixo está descrita a forma como estes índices são obtidos (Equação 5).

$$R_{factor} = \frac{\sum_h (|F_{obs}| - |F_{calc}|)}{\sum_h |F_{obs}|} \quad \text{Equação 5}$$

O R_{free} é calculado da mesma maneira que R_{factor} mas, para esse índice, é separado um conjunto com cerca de 5 a 10% de todos os fatores de estrutura, que não é usado no processo de refinamento; desta forma o progresso do refinamento pode ser julgado sem haver tendenciosidades nas fases das reflexões, densidades eletrônicas após cada ciclo de refinamento.

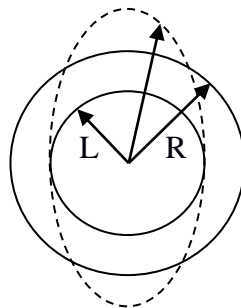
2.6.7 Validação

A validação é a etapa na qual se faz a avaliação da qualidade do modelo final como um todo; busca-se estimar erros ao olhar para regiões locais, levando em consideração diversos parâmetros estereoquímicos. Uma das ferramentas utilizadas para a validação é o programa PROCHECK (LASKOWSKI et. al., 1993), onde são avaliados parâmetros como ângulos planos, comprimento das ligações, a planaridade dos anéis da cadeia lateral, a quiralidade, as conformações das cadeias laterais, impedimento estérico e avalia a qualidade do gráfico de Ramachandran. O gráfico de Ramachandran permite determinar se os aminoácidos encontram-se em posições energeticamente favoráveis ou não. Além disso, deve-se verificar a interação do modelo com moléculas ou íons que são constituintes do meio em que o modelo está inserido; esta análise pode ser feita com o programa WHATCHECK (FILHO et al., 2003).

2.7 Dicroísmo Circular (CD)

A técnica de dicroísmo circular se baseia na absorção diferencial de radiação eletromagnética circularmente polarizada. A luz polarizada no plano pode ser concebida como sendo constituída por duas componentes circularmente polarizadas de iguais intensidades (uma com rotação no sentido anti-horário, denominada L, e outra com rotação no sentido horário, denominada R). Ao passar pela amostra, o feixe de radiação teria a mesma polarização se suas componentes fossem absorvidas na mesma intensidade ou se elas não sofressem absorção. Entretanto, se uma das suas componentes é absorvida em extensão diferente da outra, diz-se que a radiação é elipticamente polarizada. Quando uma substância é composta por moléculas quirais, um sinal de dicroísmo circular poderá ser medido por umas das seguintes razões: (a) o cromóforo é intrinsecamente quiral, como um átomo de carbono com quatro substituintes diferentes; (b) está covalentemente ligado ao centro quiral em uma molécula ou (c) está disposto em um ambiente assimétrico em virtude do arranjo tridimensional adotado pela molécula. Este sinal é (obtido em um espectropolarímetro) a medida da diferença na absorção entre as componentes L e R ($\Delta A = A_L - A_R$). Geralmente se expressa o sinal em termos da elipticidade (em graus) $\theta = \tan^{-1}(b/a)$, onde b é o menor vetor (L) e a é o maior vetor (R), na Figura 10. Um espectro de dicroísmo circular é obtido quando o sinal de dicroísmo é medido em função do comprimento de onda (KELLY; JESS; PRICE, 2005).

Figura 10. Efeito de absorção diferencial das componentes anti-horária (L) e horária (R) da radiação eletromagnética. A linha tracejada indica o trajeto do vetor (radiação) resultante (elipticamente polarizada).



Fonte: O autor.

3 Objetivos

3.1 Objetivo Geral:

- Realizar estudos estruturais da lipase de pinhão manso, *Jatropha curcas*, e da triose fosfato isomerase de *Naegleria gruberi*.

3.2 Objetivos Específicos:

- Obter via expressão heteróloga a lipase de pinhão manso, *Jatropha curcas*;
- Obter via expressão heteróloga a triose fosfato isomerase de *Naegleria gruberi*;
- Superexpressar, purificar, e determinar a estrutura tridimensional da enzima lipase de *Jatropha curcas*;
- Superexpressar, purificar e determinar a estrutura tridimensional da enzima triose fosfato isomerase de *Naegleria gruberi*.

4 Materiais e Métodos

4.1 Meios de cultivo e aditivos

Os meios para cultivo das bactérias foram preparados segundo Sambrook; Fritsch e Maniatis, (1989) e apresentavam as seguintes composições:

- 2xYT: 16 g L⁻¹ de peptona, 10 g L⁻¹ de extrato de levedura e 5 g L⁻¹ de NaCl;
- Ampicilina: dissolvida em água ultrapura para uma concentração estoque de 100 mg mL⁻¹;
- Isopropil-β-D-tiogalactosídeo (IPTG): 1 g de IPTG em 5 mL de água ultrapura para concentração estoque de 0,8 mmol L⁻¹;
- Meio sólido: 2xYT contendo 15 g L⁻¹ de ágar;
- SOC pH 7,0 (1 L): 20 g de triptona, 5 g de extrato de levedura, 0,5 g de NaCl, 10 mL de solução 250 mmol L⁻¹ de KCl, 5 mL de solução 2 mol L⁻¹ de MgCl₂ e 20 mmol L⁻¹ de glicose.

4.2 Programas computacionais

Os programas e servidores utilizados foram: a) TCOFFEE (NOTREDAME; HIGGINS e HERINGA, 2000) e BLAST (*Basic Local Alignment Search Tool*) (PARK et al., 2012), para alinhamento da sequência de amino ácidos; b) para o trabalho cristalográfico, XDS (*X-ray Detector Software*) (KABSCH, 2009), CHAINSAW no modo MAXI (STEIN, 2008) para gerar um modelo com a maior conservação de átomos possível entre da NgTIM e AfTIM para a substituição molecular, PHASER (McCOY et al., 2007), Phenix.refine (AFONINE et al., 2005), COOT (EMSLEY et al., 2004); c) para predição de massa molecular das enzimas, ProtParam (GASTEIGER et al., 2005).

4.3 Produção e expressão das enzimas

Após a realização de uma busca no site do *National Center for Biotechnology Information* (NCBI), verificou-se a disponibilidade do gene codificador da lipase da

espécie *Jatropha curcas*; verificou-se também que não existia a estrutura tridimensional desta enzima depositada no *Protein Data Bank* (PDB). A sequência de amino ácidos da lipase foi submetida à análise pelo servidor XTAIPred (SLABINSKI et al., 2007) que indicou uma reduzida probabilidade de cristalização da lipase devido às porções N e C-terminal. Diante disto, 26 resíduos iniciais e 8 finais foram deletados da sequência. Em seguida, a síntese do gene codificante da nova sequência truncada foi encomendada à empresa Genscript (a clonagem foi realizada no plasmídeo pET15b) em uma construção que permite obter a enzima com seis resíduos de histidina na região N-terminal. O plasmídeo pET15b, recombinante, contendo o gene codificante da lipase de *Jatropha curcas* (*JcLip*), foi inserido em células de bactérias *E. coli* BL21-DE3, por eletroporação (técnica de transformação onde as células são submetidas a um campo elétrico para produzir poros em suas membranas e permitir a inserção de plasmídeos). O mesmo procedimento foi adotado para a triose fosfato isomerase de *N. gruberi*, entretanto, neste caso não foi necessário excluir nenhum amino ácido da sequência ao se planejar o gene a ser sintetizado.

As células transformadas por eletroporação foram deixadas sob agitação a 250 rpm e 37 °C por 1 h em 1 mL de meio SOC; em seguida, as células foram inoculadas em meio sólido suplementado com ampicilina 5 µg mL⁻¹ por 16 h a 37 °C.

Foram selecionadas cinco colônias e suas células foram propagadas em 5 mL de meio 2xYT suplementado com ampicilina a uma concentração final de 5 µg mL⁻¹ e mantido sob agitação a 185 rpm, 37 °C por 16 h. As células foram armazenadas em 50% de glicerol a temperatura de -20 °C para uso posterior.

Foi realizado teste de expressão para as cinco colônias escolhidas anteriormente, usando temperaturas de 20 °C e 37 °C, por 4 h e 16 h; para cada combinação de tempo e temperatura o indutor IPTG foi adicionado nas seguintes concentrações: 0,1 mmol L⁻¹; 0,2 mmol L⁻¹; 0,3 mmol L⁻¹; 0,4 mmol L⁻¹; 0,5 mmol L⁻¹ e 1,0 mmol L⁻¹. Não foram realizados testes de expressão para a triose fosfato isomerase; para sua superexpressão utilizou-se temperatura de 37 °C por 16 h e o indutor IPTG a uma concentração de 0,1 mmol L⁻¹. Foram utilizados 2 L de meio de 2xYT suplementados com ampicilina a uma concentração final de 5 µg mL⁻¹.

Para a expressão da enzima *JcLip*, foi preparado novo inóculo em 2 L do meio 2xYT, a partir do resultado do teste de expressão.

4.4 Extração da enzima lipase de *Jatropha curcas*

Os meios de cultura num volume de 2 L contendo as células após indução foram centrifugados a $7000 \times g$ por 20 minutos a $4\text{ }^{\circ}\text{C}$. O sobrenadante foi descartado e o pellet de células foi ressuscitado em 40 mL da solução tampão de lise (Tris-HCl a 30 mmol L^{-1} , NaCl a 150 mmol L^{-1} , pH 7,9). Esta suspensão foi lisada por sonicação por 10 vezes de 30 s, entremeadas por repouso por 30 s. O procedimento de lise foi realizado a temperatura de próxima de $0\text{ }^{\circ}\text{C}$. Posteriormente, o lisado foi centrifugado a $15000 \times g$ por 1 h a $4\text{ }^{\circ}\text{C}$ e o sobrenadante foi utilizado para cromatografia.

4.5 Modelagem por homologia

A busca por sequências similares à de JcLip foi realizada utilizando-se a ferramenta BLAST, por apenas as existentes no PDB (*Protein Data Bank*). As sequências foram então alinhadas com o programa CLUSTALW (CHENNA, 2003) e são representadas com o programa ALINE (BOND; WOLFGANG; SCHUTTELKOPF; 2009). Este alinhamento, juntamente com os arquivos de coordenadas do PDB das homólogas, foi usado pelo programa MODELLER (SALI; BLUNDELL, 1993) para construção de 1000 modelos iniciais, otimizados internamente por dinâmica molecular. Os modelos foram avaliados pelos métodos DOPE (SALI; SHEN, 2006), GA341 (BINO; SALI, 2003) e pelo gráfico de Ramachandran gerado usando o programa PROCHECK (LASKOWSKI et al, 1993).

4.6 Extração da enzima triose fosfato isomerase de *Naegleria gruberi*

Terminado o tempo de indução, o volume de 2 L de meio 2xYT foi centrifugado a $7000 \times g$ por 20 minutos a $4\text{ }^{\circ}\text{C}$. O sobrenadante foi descartado e o *pellet* de células foi ressuscitado em 40 mL da solução tampão de lise, contendo lisozima a uma concentração final de $2,5\text{ mg mL}^{-1}$. A mistura foi mantida sob agitação a temperatura de $4\text{ }^{\circ}\text{C}$ por 40 min. Em seguida, esta suspensão foi submetida a 10 sucessivos processos de congelamento e descongelamento em nitrogênio líquido. Após a lise celular, a amostra foi centrifugada a $15000 \times g$ por 1 h a $4\text{ }^{\circ}\text{C}$. O sobrenadante foi

usado para a etapa de purificação cromatográfica.

4.7 Purificação por cromatografia de afinidade

Alíquotas de 10 mL do sobrenadante contendo as proteínas solúveis obtido na etapa anterior foram injetadas em uma coluna cromatográfica com fase estacionária de resina com íons Ni^{2+} . Uma coluna com volume de 5 mL foi previamente equilibrada com 15 mL da solução tampão de lise a um fluxo de 1 mL min^{-1} (este fluxo mantido durante as demais etapas). A etapa de lavagem para retirada das proteínas com ligação fraca à resina foi realizada com 15 mL do tampão de lavagem (Tris-HCl a 30 mmol L^{-1} , NaCl a 150 mmol L^{-1} , imidazol 10 mmol L^{-1} pH 7,9). A eluição da enzima recombinante foi realizada usando um gradiente do tampão de lavagem e do tampão de eluição (Tris-HCl a 30 mmol L^{-1} , NaCl a 150 mmol L^{-1} , imidazol 500 mmol L^{-1} pH 7,9).

As frações provenientes da cromatografia foram analisadas por eletroforese em géis de poli(acrilamida) a 12% (SDS-PAGE) em condições desnaturantes (LAEMMLI, 1970), corridas sob tensão constante de 180 V. Foram tomadas alíquotas de 20 μL das frações eluídas que foram então dissolvidas em tampão de amostra para SDS-PAGE na proporção de 1:1 e aquecidas a aproximadamente $98 \text{ }^\circ\text{C}$ por 10 min. A coloração do gel foi realizada com o corante Coomassie R-250 (10% (V/V) de ácido acético glacial, 0,25% (m/V) de azul de Coomassie R-250 e 45% (V/V) de metanol).

4.8 Dosagem proteica

A medida da concentração das enzimas foi realizada por medida de absorvância em comprimento de onda 595 nm através do método de Bradford (BRADFORD, 1976). A curva de calibração foi construída utilizando-se albumina sérica bovina (BSA) como padrão na faixa de concentrações de $0,002 \text{ mg mL}^{-1}$ a $0,020 \text{ mg mL}^{-1}$.

4.9 Ensaio de cristalização

Inicialmente, a solução de proteína purificada foi concentrada a 7 mg mL^{-1} em concentradores Centricon (com corte de massa molecular de 10 kDa) por

centrifugação a $3800 \times g$ a $4 \text{ }^\circ\text{C}$.

Os ensaios de cristalização foram conduzidos usando a técnica de difusão de vapor em gota suspensa (RUPP, 2010) em placas de cristalização de 24 poços. Cada condição de cristalização foi construída utilizando-se $500 \text{ } \mu\text{L}$ da solução para o reservatório; cada gota foi montada manualmente na Universidade Estadual de Ponta Grossa (UEPG) e continha $2 \text{ } \mu\text{L}$ da solução do reservatório (poço) e $2 \text{ } \mu\text{L}$ da solução de proteína. As condições (soluções precipitantes) foram obtidas comercialmente sob a forma de *kits*: Morpheus; Midas; JCSG Screen; Clear Strategy Screen; PGA Screen, todos do fabricante Molecular Dimensions (EUA, Flórida, Orlando). Além destes ensaios manuais, foram realizados ensaios de cristalização automatizados pelo robô de cristalização (HoneyBee 963, Digilab Global) do LNLS. Nos experimentos conduzidos pelo robô, o método usado foi o da gota sentada em placas com 96 poços; cada gota com um volume total de $2 \text{ } \mu\text{L}$ ($1 \text{ } \mu\text{L}$ de solução precipitante e $1 \text{ } \mu\text{L}$ de solução de proteína). As concentrações de proteína usadas nestes experimentos foram de 5 mg mL^{-1} , 7 mg mL^{-1} e 10 mg mL^{-1} , gerando 3 gotas para cada poço (uma para cada concentração). Os *kits* usados pelo robô foram: “JCSG”, “PACT Suite”, “WizardScreen I e II” e “PrecipitantSynergy”.

4.10 Coleta e processamento dos conjuntos de dados de difração de Raios X

Em uma primeira tentativa de obter um conjunto de dados de difração, um cristal obtido em uma condição montada manualmente foi submetido a experimento no difratômetro Rigaku MicroMax-007 com fonte de ânodo rotatório e detector R-AXIS IV++ do Instituto de Física da São Carlos – USP. Foram submetidos ao experimento de difração para coleta de dados outros cristais, obtidos mais tarde em ensaios realizados na Universidade Estadual de Ponta Grossa (UEPG) e no robô de cristalização, na estação W01B-MX2 do LNLS. O processamento das imagens de cada conjunto de dados foi realizado usando o programa XDS e XSCALE (KABSCH, 1994).

Realizaram-se estudos de corte de resolução para todos os conjuntos de dados obtidos e que foram usados para determinação de estrutura levando-se em consideração os valores dos índices completeza, relação sinal/ruído ($\langle I/\sigma(I) \rangle$), R_{symm} , R_{meas} e principalmente $CC_{1/2}$ (EVANS; GARIB, 2013).

4.11 Substituição molecular

Para tornar possível a solução da estrutura pelo método de substituição molecular, realizou-se uma busca com a ferramenta BLAST (*Basic Local Alignment Search Tool*) disponível no site do NCBI (*National Center for Biotechnology Information*), para identificar proteínas homólogas com estrutura depositada no PDB que pudessem ser usadas como modelo inicial com base nos valores de identidade e “*E-value*”. Após a definição daquela de maior identidade, sua sequência foi alinhada com a sequência de triose fosfato isomerase de *N. gruberi* com o programa TCOFFEE (NOTREDAME; HIGGINS; HERINGA, 2000). Este alinhamento foi usado como entrada no programa CHAINSAW (STEIN, 2008) para preparar o modelo para determinação das funções de rotação e translação com o programa PHASER (McCOY et al., 2007).

4.12 Refinamentos estruturais

O processo de refinamento consistiu de etapas iterativas de construções e ajustes manuais do modelo no espaço real utilizando o programa COOT (EMSLEY; COWTAN, 2004) e refinamentos no espaço recíproco com o programa PHENIX (por meio do módulo phenix.refine). Os mapas de densidade eletrônica utilizados foram calculados pelo método FEM (*Feature-Enhanced Map*) (AFONINE et al., 2015).

Refinamentos incluindo TLS (*Translation-Libration-Screw*) foram testados para todas as estruturas. Este tipo de refinamento pode conduzir a melhoras nos residuais R_{factor} e R_{free} do modelo e mesmo indicar alguma característica dinâmica da estrutura. Basicamente, este procedimento consiste em permitir que certos grupos de átomos do modelo em refinamento sejam tratados como corpos rígidos. A escolha de quais átomos constituem um domínio de TLS foi realizada tendo como base as divisões da cadeia polipeptídica em corpos rígidos propostas pelo servidor TLSMD (PAINTER; MERRITT, 2006 - TLSMD web server for the generation of multi-group TLS models). Todas as possibilidades apontadas pelo servidor foram testadas nos refinamentos e foram mantidas aquelas em que foi observada melhora nos índices R_{factor} e R_{free} .

4.13 Validações e comparações estruturais

Os modelos refinados foram validados com os programas PROCHECK (LASKOWSKI et al., 1993), MOLPROBITY (CHEN, et al., 2010) e MOLEMAN2 (KLEYWEGT, 1996). Também foram calculados o coeficiente de correlação no espaço real (RSCC – *Real Space Correlation Coefficient*) e residual no espaço real (RSR – *Real Space Residual*), ambos por aminoácido para todos os modelos (a definição destes dois índices é dada nas equações 6 e 7 abaixo). O modelo final da estrutura foi comparado com algumas de suas homólogas em relação a características estruturais existentes em cada uma como reflexo das diferenças e/ou semelhanças observadas nos alinhamentos de suas sequências.

$$RSCC = \frac{\sum_r (\rho(r)_{obs} - \overline{(\rho(r)_{obs})}) \cdot (\rho(r)_{calc} - \overline{(\rho(r)_{calc})})}{(\sum_r (\rho(r)_{obs} - \overline{(\rho(r)_{obs})})^2 \cdot \sum_r (\rho(r)_{calc} - \overline{(\rho(r)_{calc})})^2)^{\frac{1}{2}}} \quad \text{Equação 6}$$

$$RSR = \frac{\sum \rho_{obs} - \rho_{calc}}{\sum \rho_{obs} + \rho_{calc}} \quad \text{Equação 7}$$

4.14 Dicroísmo Circular

Com o intuito de obter dados de porcentagem de estrutura secundária e acompanhar o efeito da temperatura na desnaturação da enzima, foram realizadas medidas de dicroísmo circular em espectropolarímetro Jasco J-810 no LNLS. Para este procedimento, utilizou-se uma amostra de 200 μL da proteína a 0,11 mg mL^{-1} em tampão Tris-HCl 30 mmol L^{-1} pH 7,9. As medidas foram realizadas no intervalo de comprimento de onda de 260 nm a 195 nm para obtenção de informação de estrutura secundária e 222 nm na faixa de temperatura de 20 $^{\circ}\text{C}$ a 95 $^{\circ}\text{C}$, com intervalo de temperatura de 5 $^{\circ}\text{C}$, para seguir o comportamento de desnaturação e eventual renaturação. As análises dos dados foram realizadas com o pacote CDPro, composto pelos programas SELCON3, CONTINLL e CDSSTR, que permitem estimar a porcentagem de estrutura secundária através da comparação dos dados experimentais com espectros de um grupo de 43 proteínas de referência (SREERAMA; WOODY, 2000). O servidor CAPITO (WIEDEMANN, 2013) também

foi usado para o mesmo fim.

5 Resultados e Discussão

5.1 Lipase de *Jatropha curcas*

Foram realizadas duas transformações de células de *E. coli* por eletroporação. Após a primeira, as condições para a produção da enzima lipase de *Jatropha curcas* (JcLip) foram testadas e, pela avaliação do SDS-PAGE de cada combinação, não foi observada diferença significativa na concentração de proteínas nas bandas, de forma geral, entre as duas temperaturas, os dois tempos de indução e as seis concentrações do indutor usados. Entretanto, decidiu-se induzir a expressão da enzima apenas na temperatura de 37 °C por 16 h com 0,3 mmol L⁻¹ de IPTG, pois foi a condição que apresentou bandas de maior destaque no gel. A ausência de bandas de superexpressão nos géis e a grande semelhança entre estes e o padrão observado em géis de eletroforese das proteínas por células não transformadas foram indicativos iniciais de que não ocorreu expressão da enzima.

Mesmo assim, ainda realizou-se produção com dois litros de meio de cultura, para a qual tanto a cromatografia de afinidade quanto os seus respectivos géis de eletroforese não foram novamente conclusivos no que se refere à expressão. Diante disso, algumas suspeitas foram levantadas: (a) o vetor escolhido é do tipo pET que usa o promotor T7 e assim, se alguma quantidade elevada de proteína pudesse ser produzida, ocorreria a produção de grandes quantidades de tRNA, levando à morte celular; (b) a proteína pode apresentar alguma toxicidade para as células (TERPE, 2006).

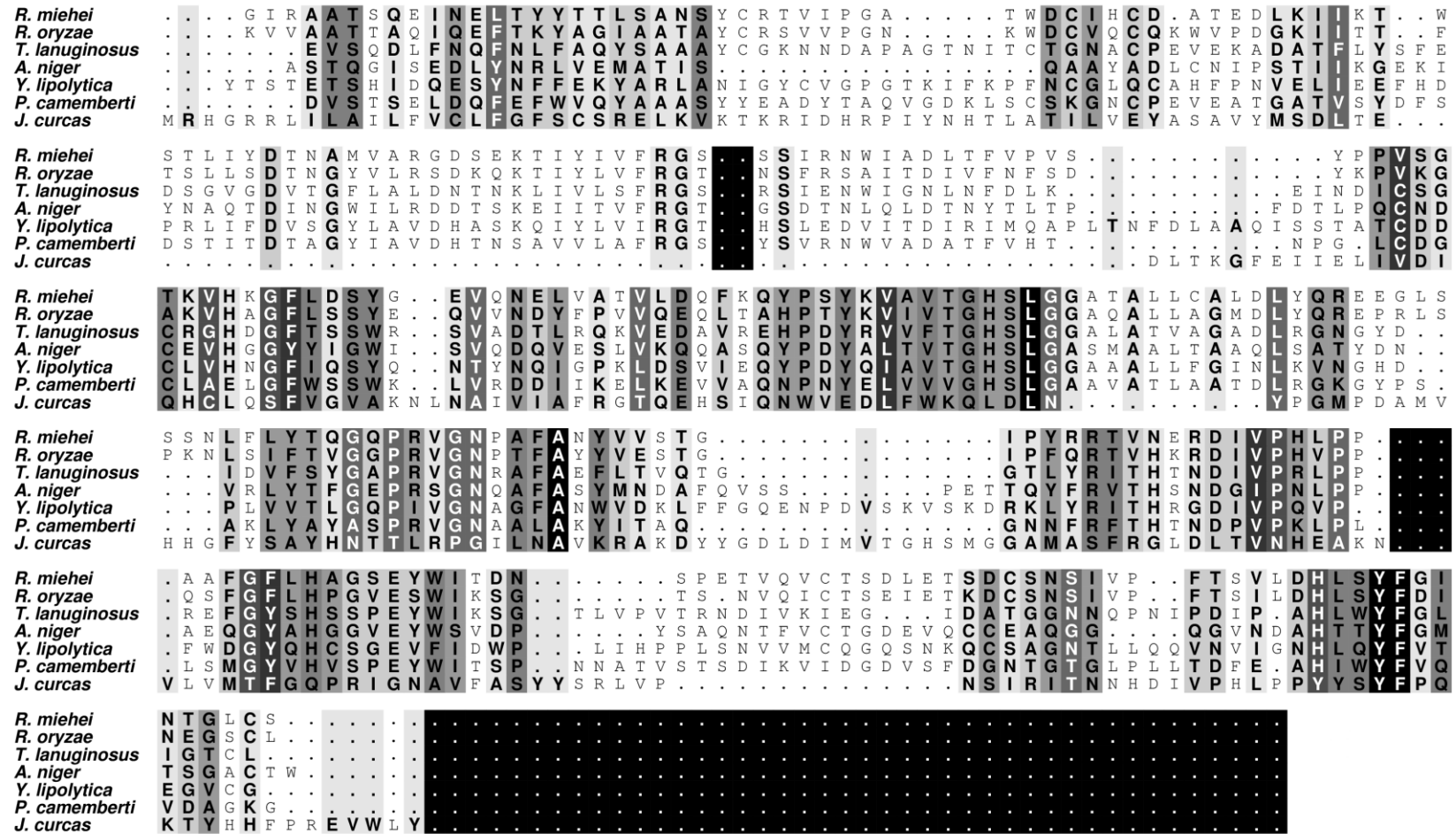
Uma segunda tentativa de transformação foi realizada da mesma forma e, novamente, sem indicativo de sucesso.

5.1.1 Modelo por homologia

Devido ao insucesso na expressão da JcLip, procurou-se, ao menos, modelar sua estrutura por homologia, para o que o alinhamento utilizado está mostrado na Figura 11. A estrutura de maior identidade é a de *Rhizopus oryzae*, código PDB 1TIC, com 34%. Dada a indisponibilidade de modelo para a região entre os resíduos 1 a 26 e 348 a 356, modelou-se apenas do resíduo 27 ao 347. A partir de 1000

modelos gerados, foi selecionado como representativo aquele que apresentou valor global normalizado de energia DOPE mais baixo, 1,293 (SHEN; SALI, 2006).

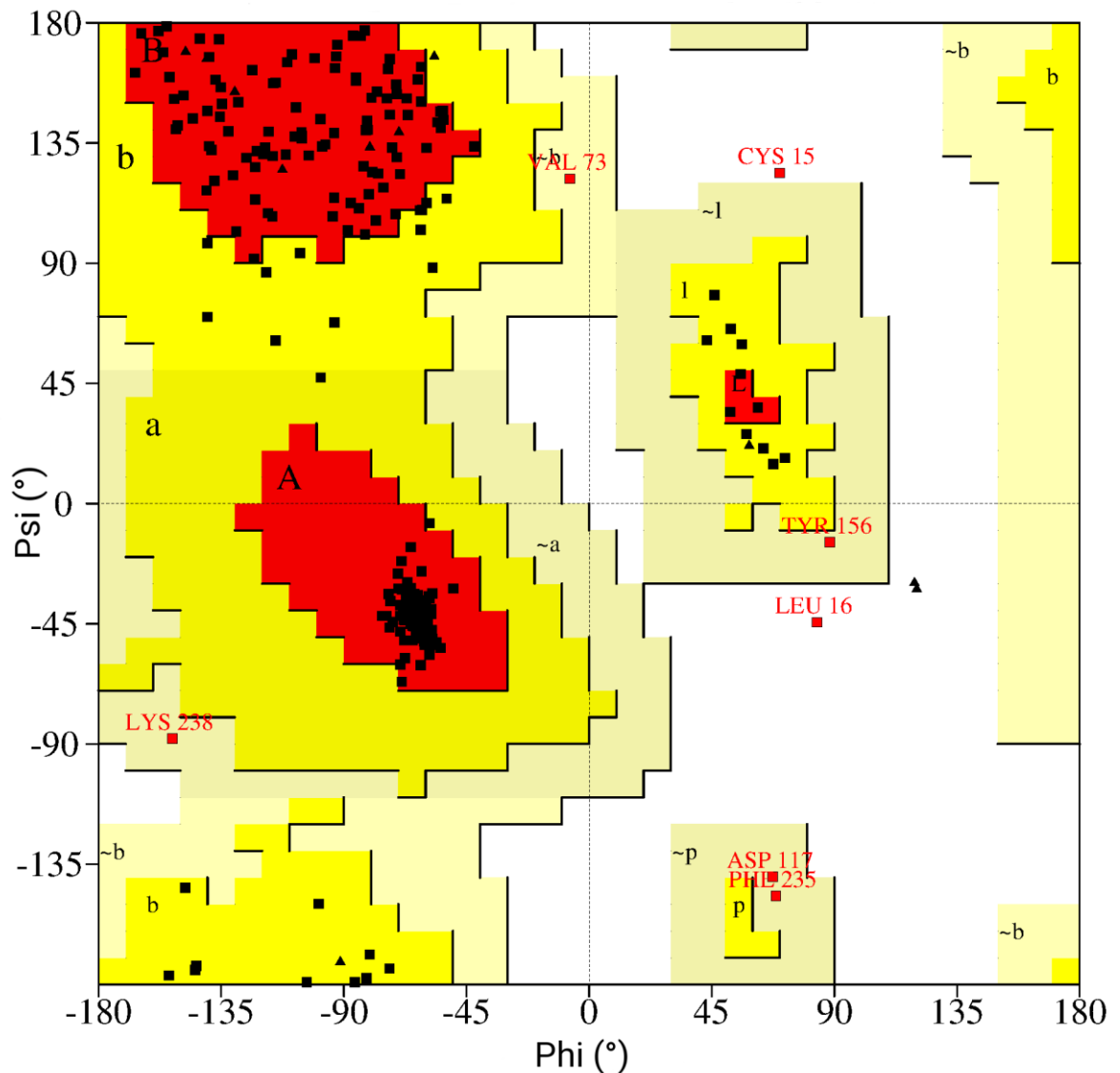
Figura 11. Alinhamento entre a sequência da lipase de *Jatropha curcas* e as homólogas usadas como molde para a modelagem.



Fonte: O autor.

O gráfico de Ramachandran (Figura 12) do modelo selecionado indica que 83,8% dos resíduos estão nas regiões mais favoráveis, 13,1% nas regiões adicionalmente permitidas, 2,3% nas regiões generosamente permitidas e 0,9% em regiões proibidas.

Figura 12. Gráfico de Ramachandran para o modelo selecionado da enzima lipase de *Jatropha curcas*. Em vermelho, regiões favoráveis, em amarelo, regiões adicionalmente permitidas, em bege, regiões generosamente permitidas e em branco, regiões proibidas.



Fonte: O autor.

O modelo apresentou uma organização estrutural contendo hélices α (em vermelho), folhas β (em amarelo) e alças (em verde), como demonstrado na Figura

13, aproximando-se do arranjo típico de lipases.

Figura 13. Estrutura final obtida por modelagem por homologia para a lipase de *Jatropha curcas* .



Fonte: O autor.

5.2 Triose fosfato isomerase de *Naegleria gruberi*

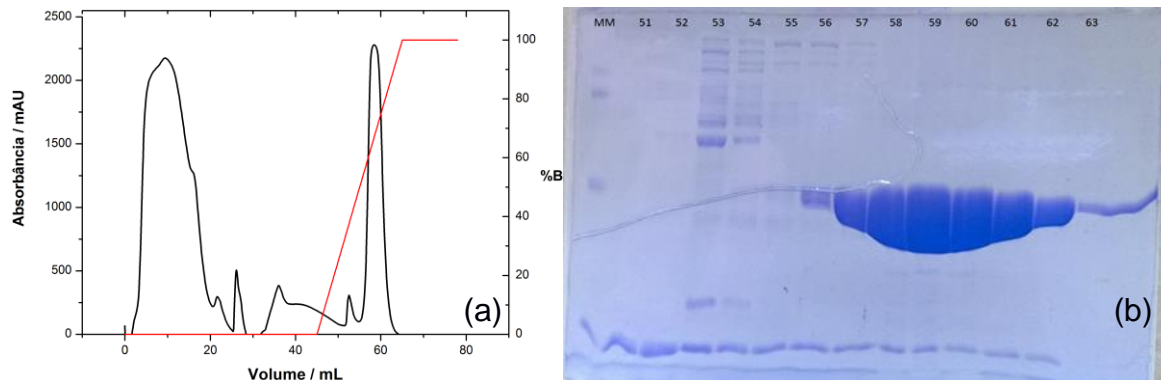
5.2.1 Produção e purificação da triose fosfato isomerase de *Naegleria gruberi*

A enzima recombinante triose fosfato isomerase de *N. gruberi* (NgTIM) foi produzida em grandes quantidades logo na primeira tentativa, por isso, não foi realizado teste otimização de expressão; todos os experimentos de cristalização e de dicroísmo circular puderam ser realizados a partir desta primeira produção. Cada alíquota da amostra bruta (sobrenadante resultante do procedimento de lise) necessitou, basicamente, de somente duas etapas de purificação. Na primeira, através de cromatografia de afinidade, foi possível perceber que as frações eluídas no gradiente, a uma concentração de imidazol de 279,5 mmol L⁻¹ a 427,0 mmol L⁻¹, apresentavam elevada concentração da NgTIM, como pode ser visualizado no cromatograma da Figura 14a e na imagem do respectivo gel de eletroforese da

Figura 14b.

A sequência de aminoácidos da enzima *NgTIM* foi submetida à análise pelo servidor *ProtParam*, que calculou uma massa molecular de 29054,7 Da. As bandas dos géis de eletroforese mostram concordância com este valor, uma vez que ficaram perto do marcador de massa molecular correspondente a 29 kDa. Importante destacar que esta massa molecular calculada leva em consideração os seis resíduos de histidina (*HisTag*) e seu conectores expressos para permitir a purificação da proteína recombinante (o valor sem *HisTag* é 27390,9 Da). Convém lembrar que os experimentos de cristalização e de dicroísmo circular foram realizados sem a retirada da *HisTag*.

Figura 14: (a) Cromatograma de afinidade mostrando o perfil de eluição da *NgTIM* no gradiente de imidazol e (b) SDS-PAGE das frações 51 a 63 (na raia 1 encontram-se os padrões de massa molecular, MM; nas raias 2 a 7, as frações 51-56 e nas raias 8 a 14 as frações 57-63, onde a enzima apresenta-se mais pura, correspondendo ao segundo pico no gradiente do cromatograma). Os padrões de massa molecular são: albumina sérica bovina, 66 kDa; β -amilase de batata doce, 56 kDa; anidrase carbônica de eritrócito bovino, 29 kDa e citocromo c de coração de cavalo, 12,4 kDa.



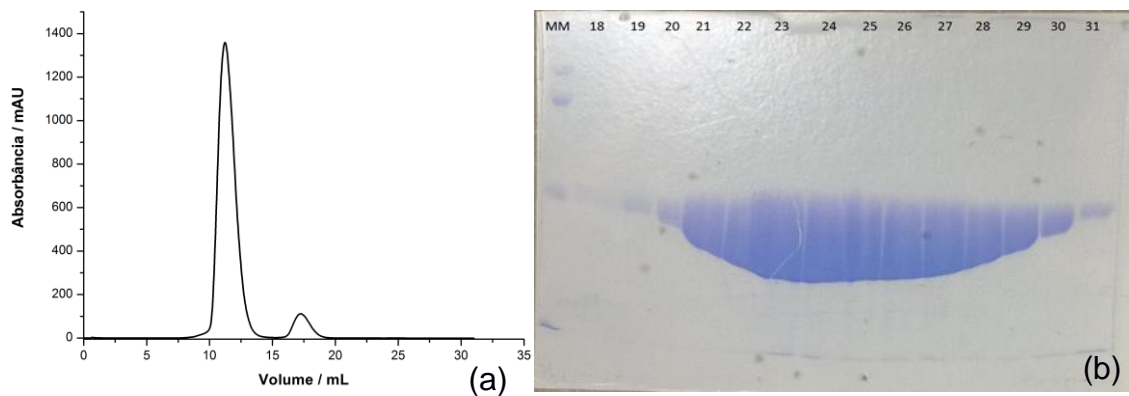
Fonte: O autor.

As frações 1 a 50 foram amostradas para análise por SDS-PAGE e nenhuma delas apresentou banda com massa molecular compatível com a da enzima *NgTIM* (géis não mostrados).

Uma segunda etapa de purificação foi necessária para eliminar as impurezas de menor massa molecular observadas na análise por SDS-PAGE da etapa anterior. A Figura 15a traz o cromatograma por exclusão de tamanho das frações 57 a 63 da cromatografia de afinidade combinadas. A Figura 15b mostra o resultado da análise por SDS-PAGE, evidenciando o quão eficiente foi a purificação da proteína após esta segunda cromatografia. Os critérios utilizados para decidir sobre o grau de

pureza adequado foram a qualidade visual do SDS-PAGE, a adequada separação entre os picos do cromatograma e o perfil do pico correspondente.

Figura 15: (a) Cromatograma de exclusão de tamanho das frações parcialmente puras da primeira etapa de purificação; a enzima NgTIM foi eluída no primeiro pico. (b) Imagem do SDS-PAGE correspondente ao primeiro pico do cromatograma. Na raia 1 encontram-se os padrões de massa molecular, MM; nas raia de 2 a 15 encontram-se as frações correspondentes aos volumes de 8,5 a 15 mL do cromatograma. Os padrões de massa molecular são: albumina sérica bovina, 66 kDa; β -Amilase de batata doce, 56 kDa; Anidrase carbônica de eritrócito bovino, 29 kDa e Citocromo c de coração de cavalo, 12,4 kDa.



Fonte: O autor.

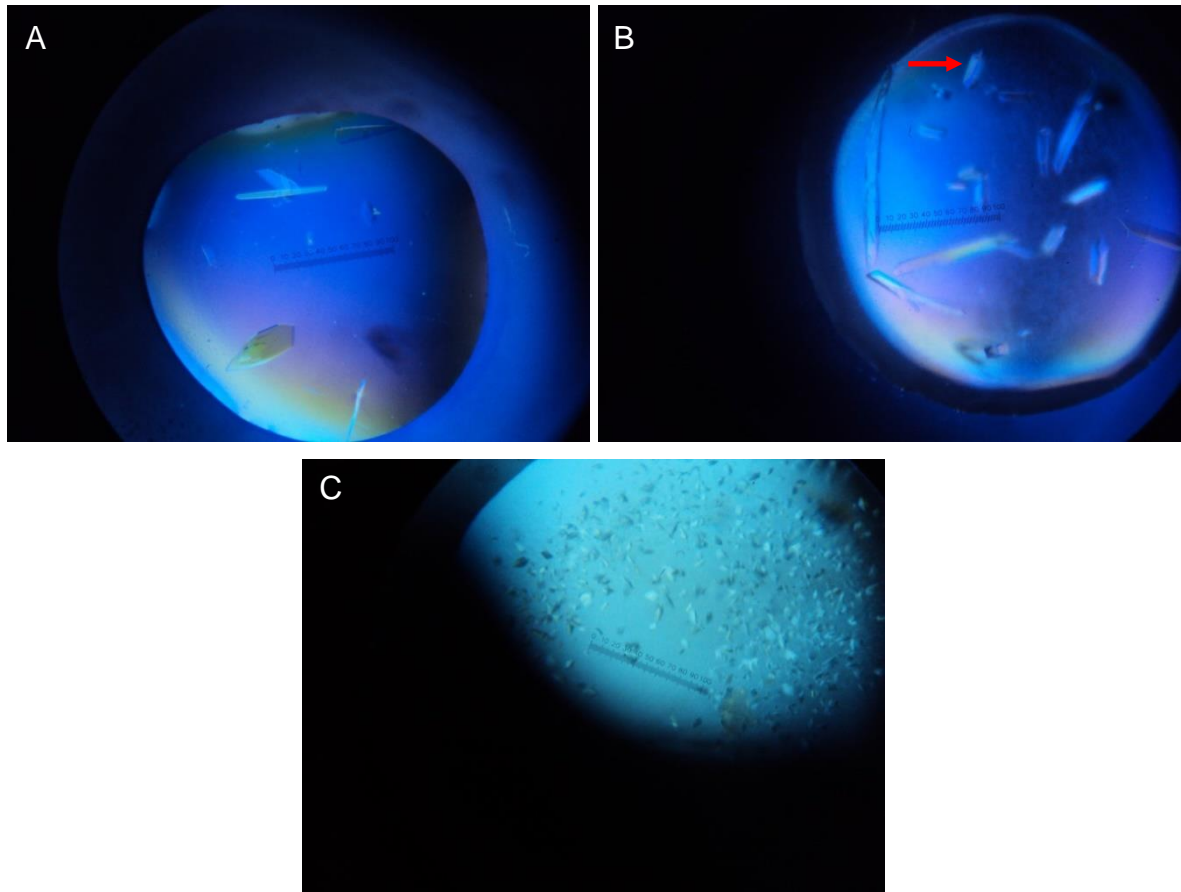
O perfil do gel de eletroforese mostrado na Figura 15b foi usado para decidir quais as frações poderiam ser combinadas; desta forma, fez-se a junção das frações 22 a 26 (correspondendo aos volumes de 11 a 13 mL no cromatograma, uma vez que cada fração contém 0,5 mL) e, em seguida, estas foram dialisadas para retirada do imidazol e a amostra resultante foi concentrada até atingir 7 mg mL^{-1} .

5.2.2 Ensaios de cristalização da triose fosfato isomerase de *Naegleria gruberi* (NgTIM)

Os experimentos realizados manualmente na UEPG geraram cristais em algumas condições, em três das quais cristais aparentemente melhores, todas do *kit* Midas (duas na parte I e uma na parte II do referido *kit*), após sete dias da montagem das gotas. As composições destas condições são: (a) $0,2 \text{ mol L}^{-1}$ de NaCl, tampão MES-NaOH $0,1 \text{ mol L}^{-1}$ pH 6,0 e Jeffamine[®] ED2003 30% (V/V); (b) etanol 15% (V/V) e propoxilato de pentaeritritol 40% (V/V) e (c) $0,2 \text{ mol L}^{-1}$ de KCl, tampão glicina $0,1 \text{ mol L}^{-1}$ pH 9,5 e etoxilato de pentaeritritol 20% (V/V). Os cristais destas três condições foram os únicos que apresentaram um padrão de difração de Raios X

com qualidade suficiente para permitir uma coleta de dados completa, entretanto, apenas um, da condição (b), foi usado para efetuar o processamento e determinação estrutural, pois foi o único que apresentou imagens com qualidade adequada para isso. Os cristais das três condições estão mostrados na Figura 16.

Figura 16. Cristais obtidos nas condições descritas anteriormente em (a) A, (b) B e (c) C. A seta vermelha indica o cristal usado na coleta das imagens de difração de Raios X.

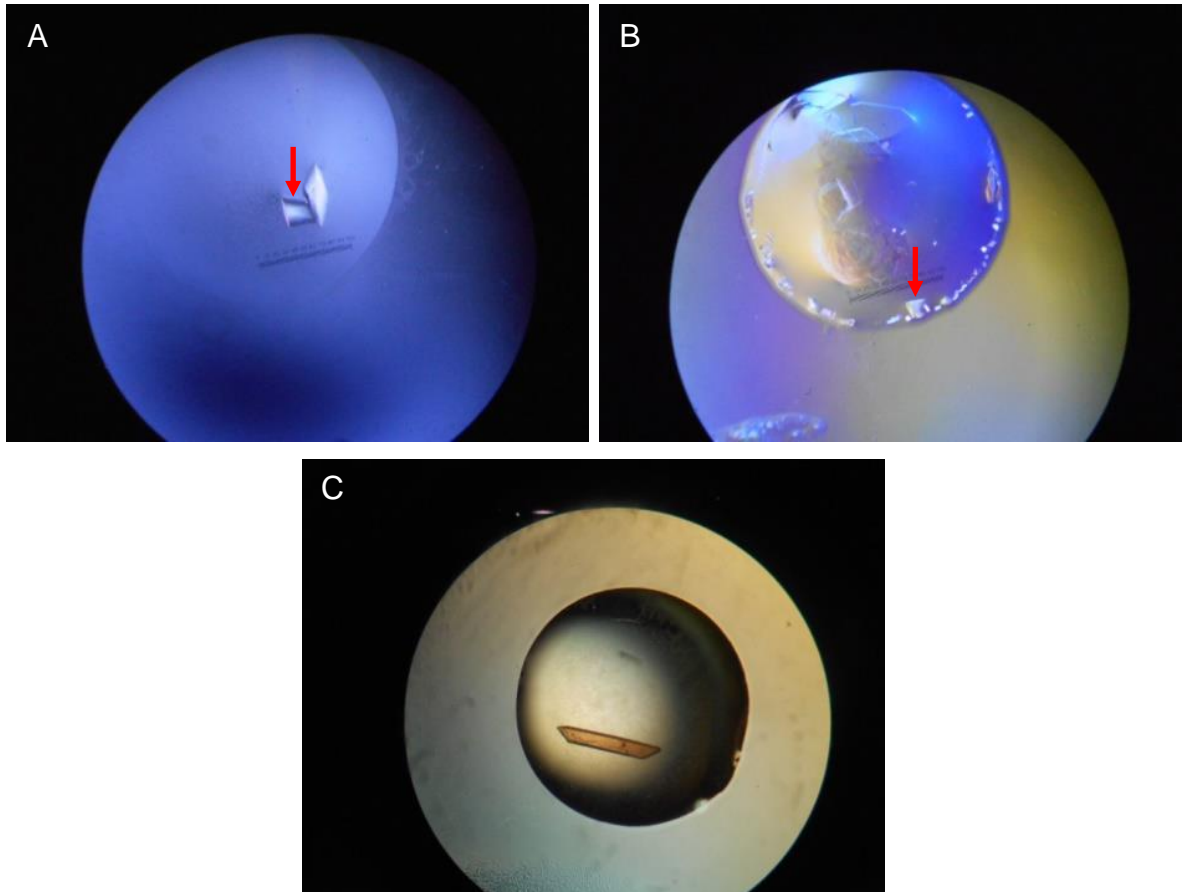


Fonte: O autor.

Uma segunda produção da NgTIM foi realizada com o intuito de obter cristais que fornecessem padrões de difração de Raios X a uma resolução maior. As tentativas de cristalização da segunda produção da enzima que resultaram em cristais cujos padrões de difração são adequados para determinação de estrutura foram nas seguintes condições: do *kit* Midas (a) Jeffamine[®] M-2070 30% (v/v), KCl 0,2 mol L⁻¹ e tampão Tris-HCl 0,1 mol L⁻¹ pH 8,0; (b) 45% (v/v) de propileno glicol 400 e 10% (v/v) etanol; do *kit* Morpheus (c) mistura de ácidos carboxílicos 0,1 mol L⁻¹ (formato de sódio, acetato de amônio, citrato de sódio, racemato de tartarato de

sódio e potássio, oxamato de sódio), *buffer system* I 0,1 mol L⁻¹ (imidazol e ácido 2-(N-morpholino)-etano sulfônico) pH 6,5, PEG 1000, PEG 3350 e 2-metil-2,4-pentanodiol. Os cristais obtidos em cada condição são mostrados na Figura 17.

Figura 17. Cristais obtidos nas condições descritas acima (a) A, (b) B e (c) C na segunda produção da triose fosfato isomerase de *Naegleria gruberi*. A seta vermelha indica o cristal usado na coleta das imagens de difração de Raios X em cada condição.



Fonte: O autor.

5.2.3 Processamento das imagens de difração de Raios X

A Tabela 1 mostra os valores dos índices provenientes dos processamentos das imagens de difração dos conjuntos de dados *NgTIM_1A* (Figura 16 (b)), *NgTIM_2B* e *NgTIM_2C* (Figura 17 (a) e (c)).

O procedimento de indexação mostrou que dois dos conjuntos de dados apresentam grupo de espaço *P4₁22* e os outros dois pertencem ao grupo de espaço *C2*, todos com celas unitárias diferentes.

O programa MATTHEWS (MATTHEWS, 1968) do CCP4 (WINN, et al., 2011) foi utilizado para determinação do conteúdo de solvente dos cristais e indicou uma molécula da enzima na unidade assimétrica do cristal $P4_122$ NgTIM_1A e conteúdo de solvente em volume de 54,15%. Por outro lado, a mesma análise para os cristais NgTIM_2B e NgTIM_2C revelou duas moléculas na unidade assimétrica e os respectivos conteúdos de solvente estão mostrados na Tabela 1. Esta informação do número de moléculas na unidade assimétrica é importante, uma vez que ela é usada durante o procedimento de substituição molecular na determinação das funções de rotação e translação.

Tabela 1: Dados referentes aos parâmetros de coleta e resultados do processamento das imagens dos padrões de difração para os três conjuntos de dados utilizados para determinação de estrutura. Os conjunto de dados NgTIM_1A, NgTIM_2B e NgTIM_2C foram obtidos a partir do cristal B mostrado na Figura 16 e dos cristais A e C mostrados na Figura 17, respectivamente.

	NgTIM_1A	NgTIM_2B	NgTIM_2C
Fonte de Raios X	W01B-MX2, LNLS	W01B-MX2, LNLS	W01B-MX2, LNLS
Comprimento de onda (Å)	1,45866	1,45864	1,45864
Temperatura (K)	100	100	100
Detector	PILATUS 2M	PILATUS 2M	PILATUS 2M
Distância Cristal-Detector (mm)	130	120	105
Intervalo de rotação por imagem (°)	1,0	0,5	0,5
Intervalo total de rotação (°)	534	178	200
Tempo de exposição por imagem (s)	10	10	10
Grupo de espaço	<i>P</i> 4 ₁ 2 2	<i>C</i> 2	<i>C</i> 2
<i>a</i> , <i>b</i> , <i>c</i> (Å)	79,6980; 79,6980; 98,1100	201,809; 53,395; 46,542	196,666; 52,893; 44,430
Mosaicidade (°)	0,199	0,226	0,566
Faixa de resolução (Å)	98,0 - 2,64 (2,72-2,64)	250,0 - 1,74 (3,98-1,74)	250,0 - 1,56 (3,57-1,56)
Número total de reflexões	367311 (30675)	129604 (5543)	175271 (7737)
Número de reflexões únicas	9781 (804)	47641 (3336)	59256 (4103)
Completeza (%)	99,9 (100,0)	95,5 (82,4)	91,2 (70,3)
Multiplicidade	37,55 (38,15)	2,72 (1,66)	2,95 (1,88)
$\langle I/\sigma(I) \rangle$	35,68 (6,93)	16,67 (1,38)	15,71 (1,16)
R_{symm} (%)	10,2 (73,3)	4,1 (53,1)	3,0 (75,9)
R_{meas} (%)	10,4 (74,3)	4,9 (75,1)	4,0 (102,7)
$CC_{1/2}$ (%)	99,9 (97,8)	99,8 (60,1)	99,9 (55,6)
Fator <i>B</i> global do Wilson plot (Å ²)	61,92	26,00	26,33
Moléculas proteicas na unidade assimétrica	1	2	2
V_M (Å ³ Da ⁻¹)	2,68	2,17	2,08
Porcentagem de solvente (%)	54,15	43,29	40,99

Fonte: O autor.

5.2.4 Substituição Molecular

Devido à identidade de 58% com *NgTIM*, decidiu-se usar como modelo para a substituição molecular a triose fosfato isomerase de *Arabidopsis thaliana* (*AfTIM*) para o primeiro conjunto de dados, *NgTIM_1A*. A Figura 18 mostra o alinhamento entre as sequências de *AfTIM* e *NgTIM*. A Tabela 2 traz os resultados da substituição molecular para os três conjuntos de dados, notando-se que, para os conjuntos *NgTIM_2B* e *NgTIM_2C*, utilizou-se como modelo a estrutura da própria *NgTIM_1A* já parcialmente refinada.

Figura 18. Alinhamento das seqüências de aminoácidos das proteínas *AtTIM* e *NgTIM*. Acima da seqüência de *AtTIM* estão representadas as suas estruturas secundárias como calculadas pelo programa DSSP (KABSCH; SANDER, 1983): os cilindros correspondem a hélices α , as setas correspondem a fitas- β e a espiral corresponde a uma hélice irregular.



Fonte: O autor.

Tabela 2. Resultados da substituição molecular para os conjuntos de dados NgTIM_1A, NgTIM_2B, NgTIM_2C. Os ângulos de rotação estão em graus e as translações em coordenadas fracionárias.

	Ângulos de rotação (α, β, γ) / °	Z Score	Translações (x,y,z)	Z Score
NgTIM_1A				
A	308,569 / 75,818 / 4,301	4,1	-0,65441 / -0,57011 / -0,13292	7,5
NgTIM_2B				
A	162,663 / 21,929 / 291,778	26,3	0,02756 / -0,53094 / -0,42470	25,6
B	344,157 / 157,642 / 340,424		-0,02903 / -0,38475 / 1,07946	
NgTIM_2C				
A	162,865 / 22,096 / 297,640	38,6	0,05802 / -0,50715 / -0,43732	34,3
B	347,093 / 156,908 / 335,916		-0,05285 / -0,36079 / 1,10390	

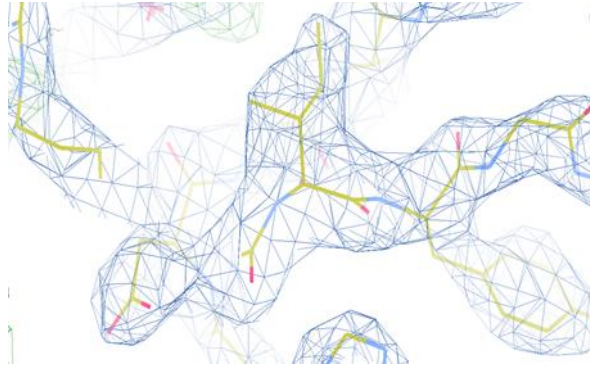
Fonte: O autor.

5.2.5 Refinamento das estruturas

Após a substituição molecular da estrutura NgTIM_1A, o modelo foi inicialmente submetido a um processo de construção automatizada, que resultou em R_{work} de 29,19% e R_{free} de 33,68%. A Figura 19 mostra o adequado ajuste do modelo à densidade eletrônica num trecho favorável, com destaque para a Ile²⁰⁶. Por outro lado, a Figura 20a mostra parte que precisou ser trabalhada em ciclo de reconstrução e refinamento, correspondente à região próxima à Val³⁸. No procedimento de reconstrução automática, algumas moléculas de água foram inseridas onde a sequência de aminoácidos do modelo feito pelo programa CHAINSAW apresentava uma descontinuidade. Após dezesseis ciclos de refinamento e reconstrução com restrições, obteve-se R_{work} 20,18% e R_{free} de 26,02%. A Figura 20b mostra a mesma região próxima à Val³⁸ após a sequência ter

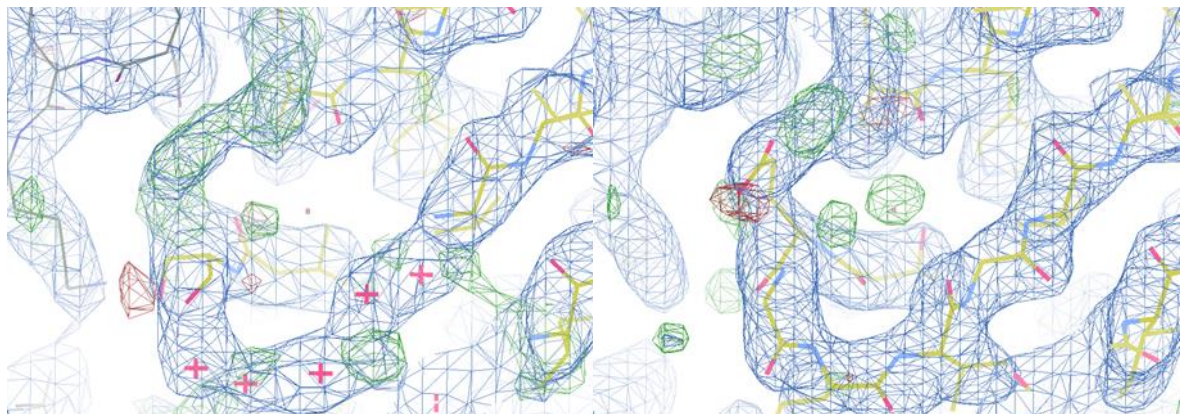
seu refinamento, tendo sido completada com os resíduos faltantes.

Figura 19: Região do resíduo Ile²⁰⁶ no início do refinamento em uma região favorável. Em azul, mapa de densidade eletrônica com contorno em 1σ , em verde, mapa Fourier-diferença com contorno em $+3\sigma$ e -3σ .



Fonte: O autor.

Figura 20: (a) Região próxima à Val³⁸ no início do refinamento e (b) a mesma região após a construção manual do modelo no décimo-sexto ciclo de refinamento. Em azul, mapas de densidade eletrônica com contorno em 1σ , em verde e vermelho, mapas Fourier-diferença com contorno em $+3\sigma$ e -3σ , respectivamente.



(a)

(b)

Fonte: O autor.

Diferentes grupos para TLS foram testados de acordo com o indicado pelo servidor TLSMD (PAINTER, et al., 2006). Mediante este estudo, decidiu-se utilizar dois grupos de TLS para a cadeia polipeptídica, a saber, um para os resíduos 1 a 134 e outro para os resíduos 135 a 252.

Após o sucesso na obtenção de cristais da segunda produção da NgTIM que geraram padrões de difração de mais alta resolução, optou-se por iniciar o refinamento destas em preferência ao parcialmente realizado, da NgTIM_1A, com

base na hipótese de que a uma maior resolução esta tarefa seria mais fácil e poderia ajudar no refinamento da primeira. Entretanto, durante o refinamento da estrutura de maior resolução, *NgTIM_2C*, percebeu-se que a região dos resíduos Ala¹⁶⁸ a Thr¹⁷⁴ no monômero A apresenta razoável desordem, dificultando a correta modelagem da cadeia principal. Não obstante, algumas tentativas de refinar esta região foram realizadas, nas quais os resíduos Gly¹⁷⁰, Thr¹⁷¹ e Gly¹⁷² não foram modelados, mas mesmo assim nenhuma melhora na densidade foi observada.

Já o monômero B da estrutura *NgTIM_2C* apresenta duas regiões de grande dificuldade para modelagem: Gly¹²⁹ a Glu¹³⁴ e Thr¹⁴⁰ a Val¹⁴³. De forma geral, a dificuldade deste monômero é ainda maior que a do monômero A. Estas dificuldades motivaram a busca por referências nas outras estruturas. Assim, verificou-se que o conjunto de dados *NgTIM_2B* provê uma estrutura de maior facilidade para a modelagem, o que nos levou a refiná-la prioritariamente, isso devido principalmente ao método de cálculo dos mapas de densidade eletrônica, FEM, (AFONINE et al., 2015) que permite a amplificação dos “sinais autênticos” e eliminação parcial de ruídos. De forma geral, o ajuste do modelo à densidade eletrônica neste caso está bastante claro, apesar da resolução ser algo menor, 1,74 Å. A única exceção é a região N-terminal, na qual alguns dos resíduos pertencentes à His-tag (necessária para a purificação) estão razoavelmente desordenados, com alguma dificuldade para sua modelagem.

Estudos de refinamento com TLS também foram realizados para as estruturas provenientes destes dois outros conjuntos de dados. Particularmente para *NgTIM_2C*, a partir da observação de que os resíduos no intervalo Ser¹⁰⁶-Val¹¹⁶ do monômero B, correspondentes a uma região de hélice α , se encontravam sistematicamente deslocados no sentido N-terminal em relação ao monômero A, decidiu-se realizar um teste para TLS específico para esta região, entretanto, nenhuma melhora na qualidade da densidade foi observada. Diante disso, a partir dos grupos sugeridos pelo servidor TLSMD (PAINTER, et al., 2006), encontrou-se como melhor usar um único grupo por monômero. Já no tocante à estrutura *NgTIM_2B*, os testes de TLS mostraram que o uso de três grupos por monômero levava à significativa melhoria nos valores de R_{work} e R_{free} , enquanto que somente um número muito maior de grupos levava à melhorias posteriores, todavia, marginais. Assim, prosseguiu-se com três grupos por monômero, a saber:

monômero A 1 a 89; 90 a 172; 173 a 252 e o monômero B, 1 a 60; 61 a 179; 180 a 252. A Tabela 3 mostra as estatísticas finais de refinamento das três estruturas.

Tabela 3. Estatísticas finais dos refinamentos das estruturas obtidas dos conjuntos de dados NgTIM-1A, NgTIM-2B e NgTIM-2C.

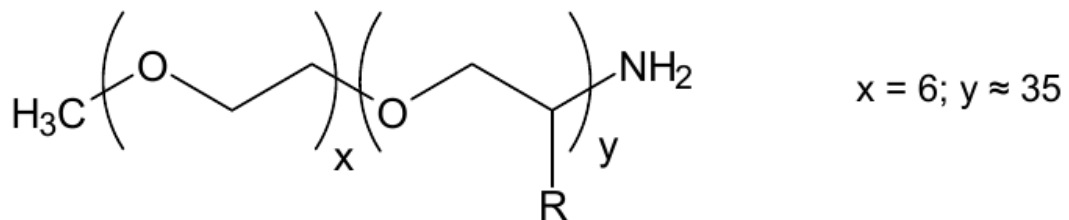
	NgTIM-1A	NgTIM-2B	NgTIM-2C
No. de reflexões no refinamento	9806	49829	69690
Completeza (%)	99,9	95,5	91,2
No. de reflexões usadas para o cálculo de R_{free}	980	1982	2146
R_{factor} (%)	20,18	15,50	18,25
R_{free} (%)	26,02	20,45	19,79
Átomos	1919	3824	3561
Moléculas de água	151	548	208
Fator B para átomos proteicos (Å^2)	53,466	30,426	32,296
<i>Desvios da idealidade RMS</i>			
Comprimento de ligação (Å)	0,001	0,008	0,013
Ângulos de ligação ($^\circ$)	0,448	1,064	1,238
<i>Gráfico de Ramachandran</i>			
Regiões favoráveis (%)	90,0	93,2	91,1
Regiões permitidas (%)	10,0	6,4	8,4
Regiões generosamente permitidas (%)	0,0	0,5	0,5
Regiões proibidas (%)	0,0	0,0	0,0

Fonte: O autor.

5.2.6 Ligantes na estrutura da triose fosfato isomerase de *Naegleria gruberi* obtida do conjunto de dados NgTIM_2B

O cristal submetido ao experimento de difração de Raios X foi gerado em uma condição/gota contendo o precipitante Jeffamine-M2070[®] (uma polieteramina sintetizada a partir de um copolímero de óxido de etileno/óxido de propileno, cuja fórmula geral está representada na Figura 21), KCl, NaCl e tampão Tris-HCl. Em algumas regiões foram observadas densidades eletrônicas que poderiam corresponder a algumas destas espécies químicas, principalmente Jeffamine[®] (três moléculas) e Tris (uma molécula). No caso da Jeffamine[®], há uma complicação extra na sua modelagem que reside no fato de sua síntese não ser estereoespecífica e, intrinsecamente, a polimerização entre os monômeros de EO e PO poder ocorrer de forma a gerar uma grande variedade de compostos. Diante disto, alguns modelos para Jeffamine[®] foram gerados e tentou-se refiná-los. A Figura 22 mostra as moléculas de Jeffamine[®] melhor ajustadas e suas respectivas densidades eletrônicas. A disposição geral de cada uma destas moléculas nas proximidades do dímero da triose fosfato isomerase de *N. gruberi* é mostrada na Figura 22a. A molécula Jeff_A (Figura 22b) encontra-se em uma cavidade próxima à interface dos monômeros, a molécula Jeff_B (Figura 22c) encontra-se próxima ao resíduo Thr¹⁰⁸ do monômero A, enquanto que Jeff_C (Figura 22d) está disposta em outra de cavidade próxima aos resíduos Val¹⁰², do monômero A, e Glu¹⁰⁵ do monômero B.

Figura 21. Fórmula geral da molécula do agente precipitante polieteramina, Jeffamine[®]-M2070.



R = H para (EO), ou CH₃ para (PO)

Fonte: O autor.

Figura 22. Moléculas de Jeffamine[®] modeladas na estrutura obtida a partir do conjunto de dados NgTIM_2B. (a) Disposição geral das moléculas de Jeffamine[®] e Tris ao redor do dímero de triose fosfato isomerase de *N. gruberi*; (b) Jeff_A; (c) Jeff_B e (d) Jeff_C. Em azul, mapa de densidade eletrônica com contorno em 1σ , em verde, mapa Fourier-diferença com contorno em $+3\sigma$ e -3σ . Continua.

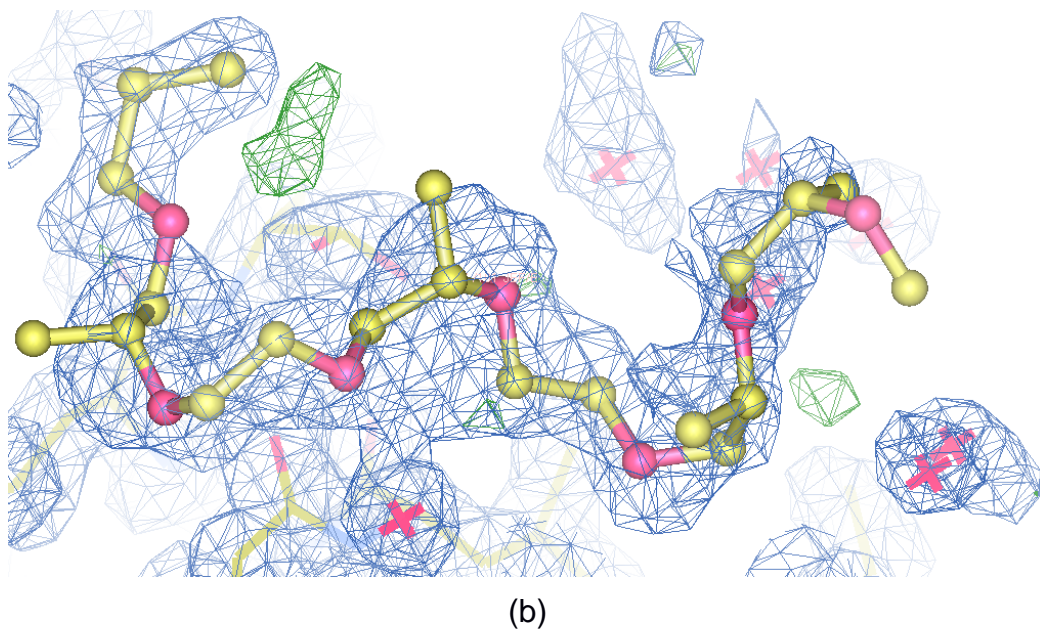
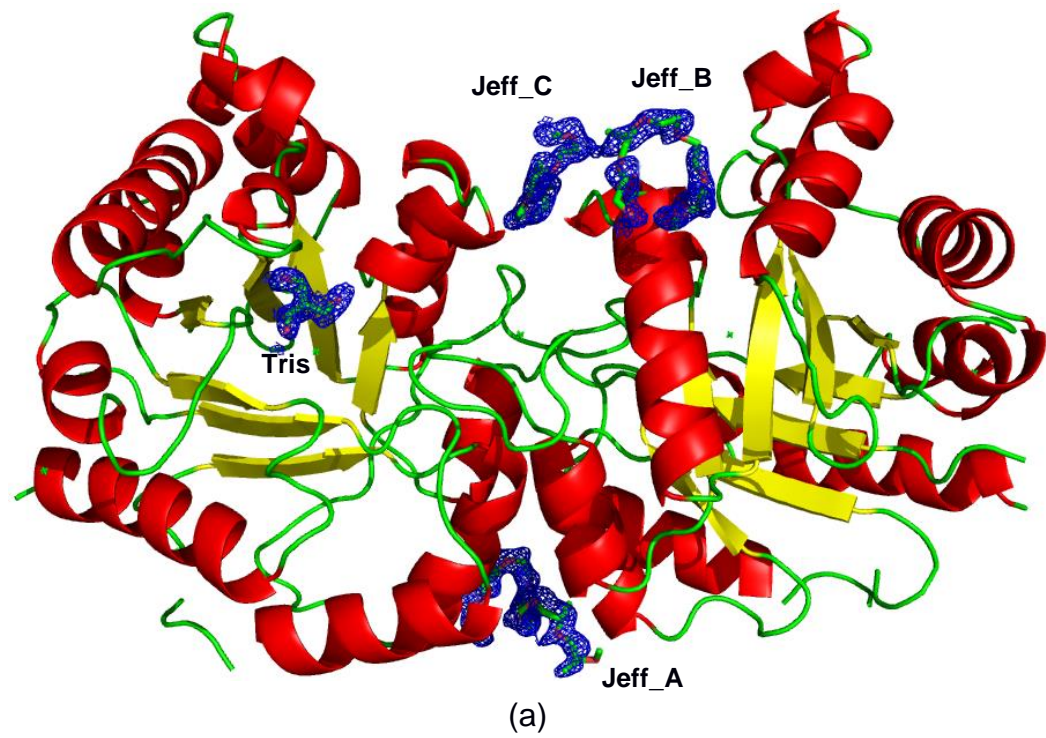
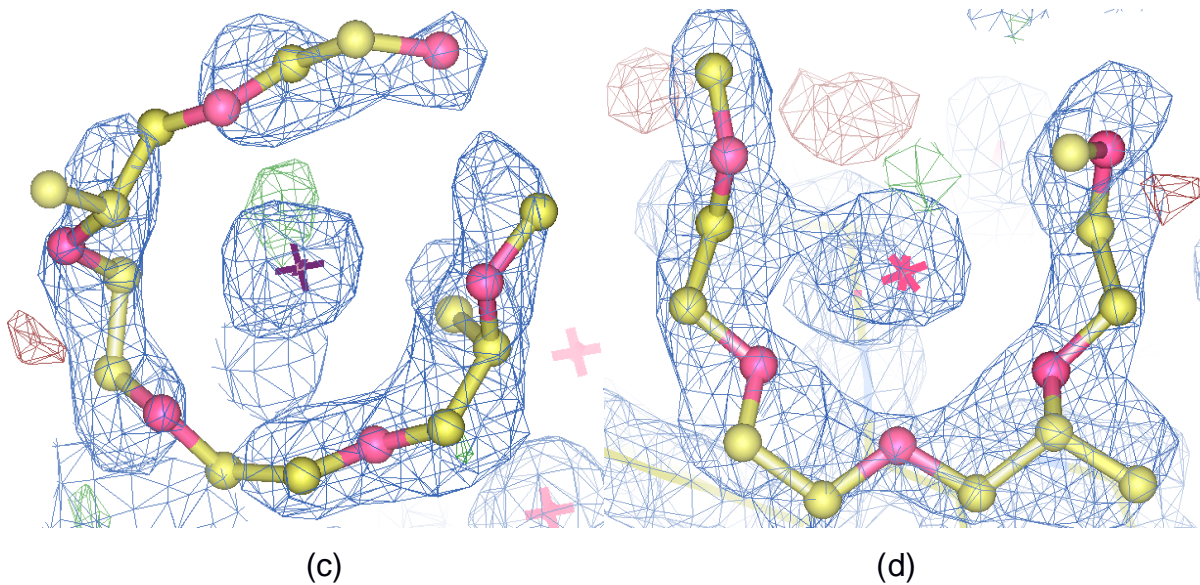


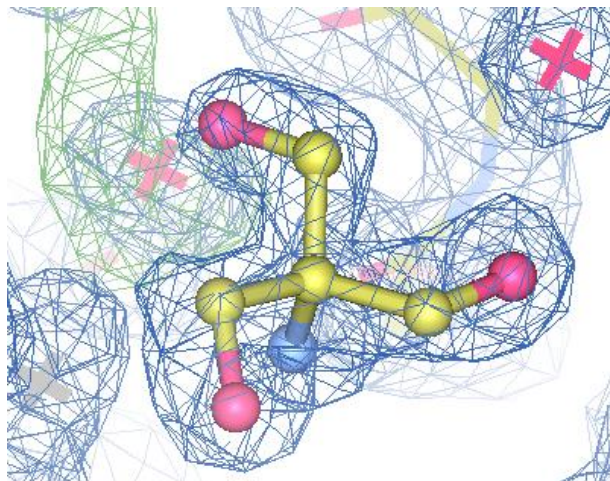
Figura 22. Moléculas de Jeffamine[®] modeladas na estrutura obtida a partir do conjunto de dados NgTIM_2B. (a) Disposição geral das moléculas de Jeffamine[®] e Tris ao redor do dímero de triose fosfato isomerase de *N. gruberi*; (b) Jeff_A; (c) Jeff_B e (d) Jeff_C. Em azul, mapa de densidade eletrônica com contorno em 1σ , em verde, mapa Fourier-diferença com contorno em $+3\sigma$ e -3σ . Conclusão.



Fonte: O autor.

Na cavidade do sítio ativo foi modelada uma molécula do tampão Tris - Tris(hidroximetil)aminoetano - (Figura 23), à qual foi atribuída a diferença observada no posicionamento da cadeia principal próximo a esta região.

Figura 23. Molécula de Tris (Tris(hidroximetil)aminoetano) modelada no sítio ativo do monômero B. Em azul, mapa de densidade eletrônica com contorno em 1σ , em verde, mapa Fourier-diferença com contorno em $+3\sigma$ e -3σ .



Fonte: O autor.

Os valores dos índices RSCC e RSR para as moléculas do agente precipitante e do tampão são mostradas na Tabela 4 e evidenciam a boa qualidade do ajuste dos modelos à densidade eletrônica.

Tabela 4. Valores dos índices RSCC e RSR para as três moléculas do agente precipitante Jeffamine[®] e do tampão Tris - Tris(hidroximetil)aminoetano – modelados na estrutura NgTIM_2B.

Molécula	RSCC	RSR
Jeff_A	0,719	0,332
Jeff_B	0,837	0,256
Jeff_C	0,819	0,289
Tris	0,881	0,225

Fonte: O autor.

5.2.7 Validações

5.2.7.1 Validação da estrutura da triose fosfato isomerase de *Naegleria gruberi* obtida do conjunto de dados NgTIM_1A

A validação pelo MOLPROBITY apontou características geométricas incomuns em apenas os rotâmeros dos resíduos Phe¹⁵¹ e Glu²¹⁸ (este na conformação alternada A). Todavia, a Figura 24 a-i e a-ii mostra o adequado ajuste destes resíduos à densidade eletrônica. O gráfico de Ramachandran (Figura 24b) desta estrutura gerada pelo PROCHECK indica que não existem resíduos com ângulos phi/psi em regiões não permitidas ou apenas generosamente permitidas.

Figura 24. Validações geométricas. (a) Rotâmeros *outliers* (i) Phe¹⁵¹ e (ii) Glu²¹⁸. Em azul, mapa de densidade eletrônica com contorno em 1σ , em verde, mapa Fourier-diferença com contorno em $+3\sigma$ e -3σ . (b) Gráfico de Ramachandran da estrutura NgTIM_1A. Em vermelho, regiões favoráveis, em amarelo, regiões adicionalmente permitidas, em bege, regiões generosamente permitidas e em branco, regiões proibidas. Continua.

(a)

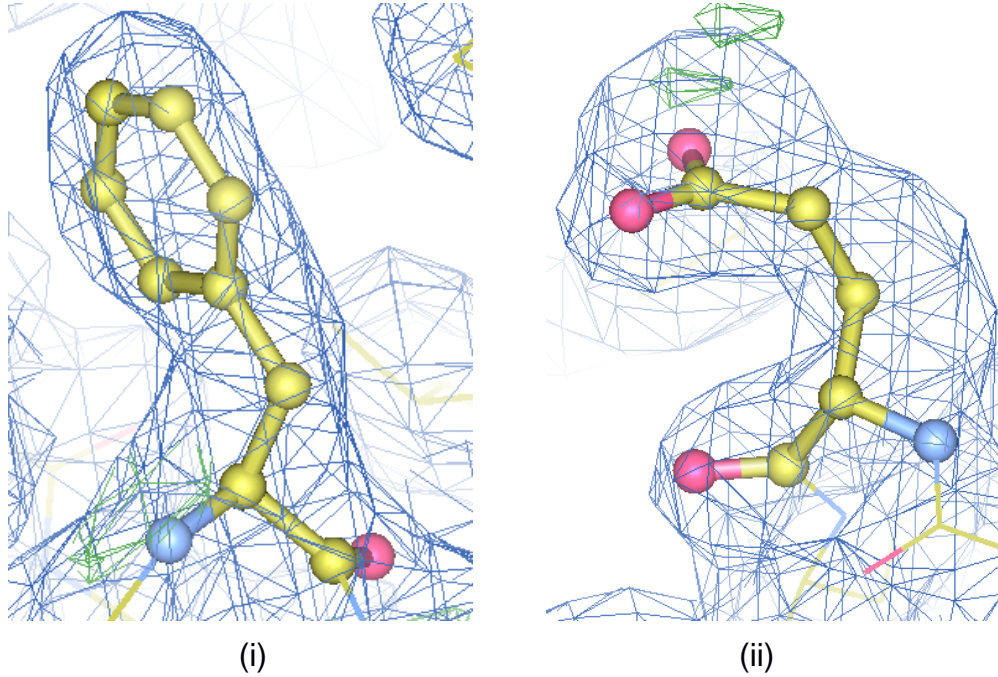
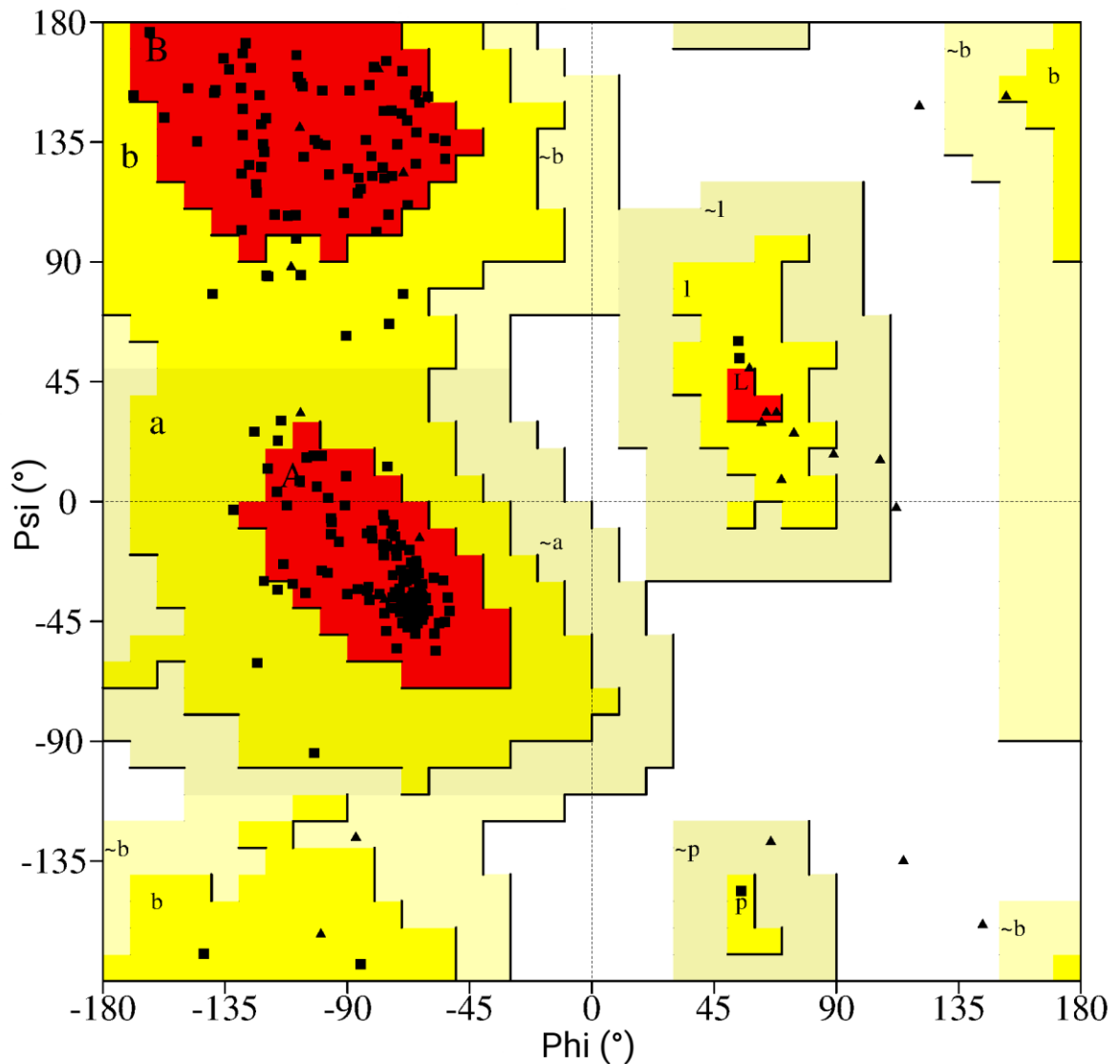


Figura 24. Validações geométricas. (a) Rotâmeros *outliers* (I) Phe¹⁵¹ e (II) Glu²¹⁸. Em azul, mapa de densidade eletrônica com contorno em 1σ , em verde, mapa Fourier-diferença com contorno em $+3\sigma$ e -3σ . (b) Gráfico de Ramachandran da estrutura NgTIM_1A. Em vermelho, regiões favoráveis, em amarelo, regiões adicionalmente permitidas, em bege, regiões generosamente permitidas e em branco, regiões proibidas. Conclusão.

(b)

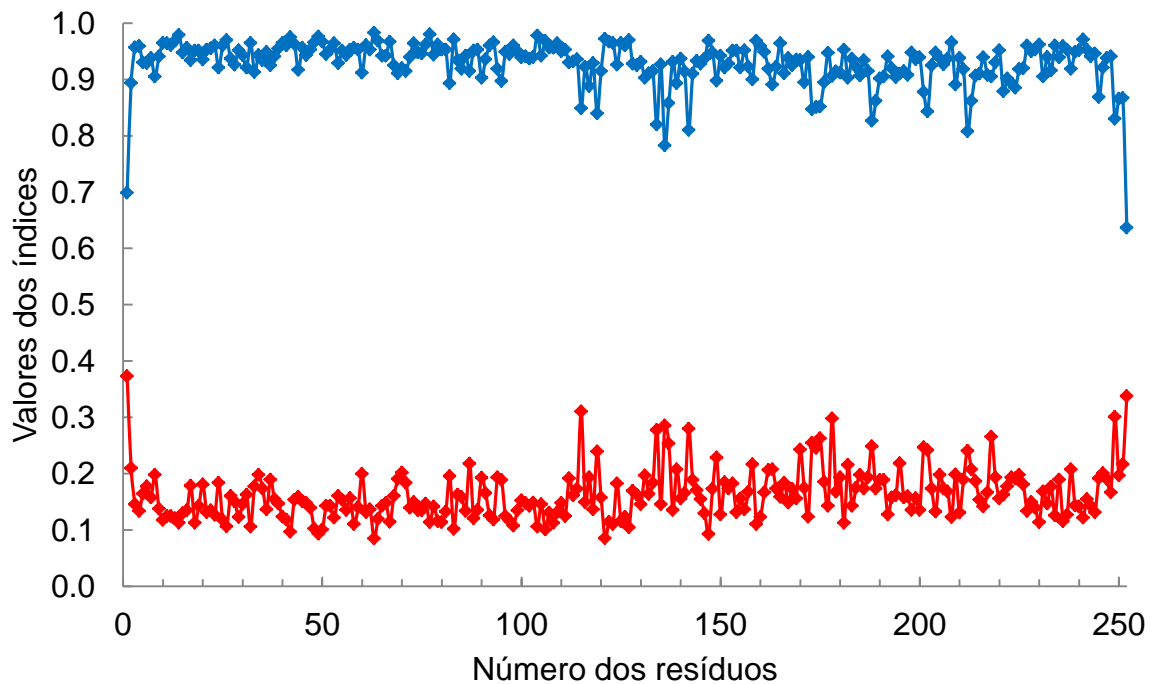


(b)

Fonte: O autor.

A Figura 25 traz gráficos dos índices RSCC e RSR calculados para cada resíduo pelo programa MAPMAN (KLEYWEGT et al., 1996). Os valores globais para a parte proteica são RSCC 0,9285 e RSR 0,1627, mostrando o adequado ajuste entre o modelo e a densidade eletrônica.

Figura 25. Ajuste na densidade eletrônica, índices RSCC (azul) e RSR (vermelho) por resíduo para a estrutura obtida do conjunto de dados NgTIM_1A.



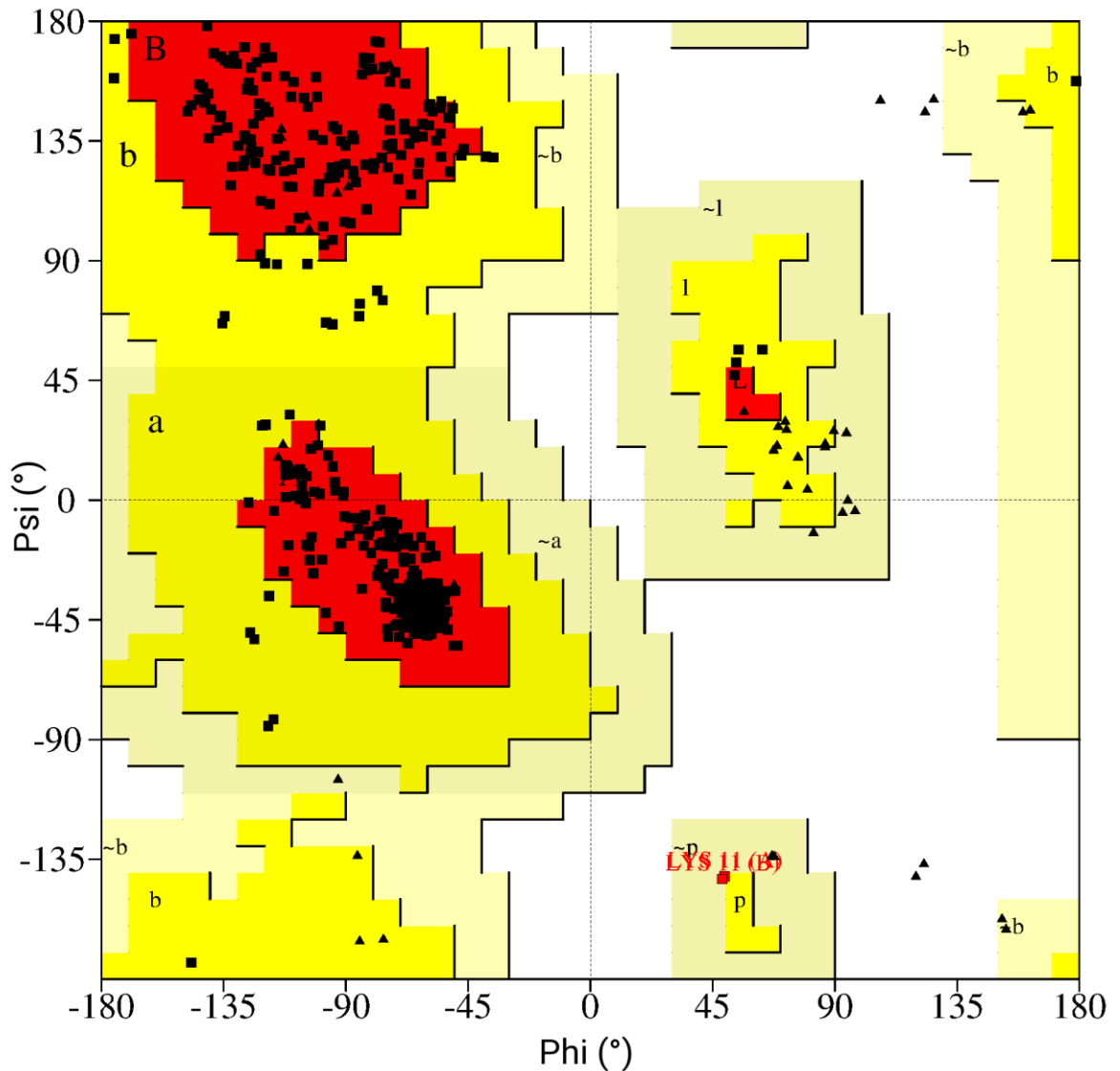
Fonte: O autor.

5.2.7.2 Validação da estrutura da triose fosfato isomerase de *Naegleria gruberi* obtida do conjunto de dados NgTIM_2B

A validação realizada pelo MOLPROBITY mostrou que o modelo possui desvios em alguns poucos parâmetros geométricos, mas todos perfeitamente justificáveis pela densidade eletrônica, a saber: a) ângulo ω do monômero B entre Phe²²⁸ e Leu²²⁹ (157,89°); b) rotâmeros dos resíduos Thr¹⁷⁴, Arg²⁰⁴, Asp²³⁹ e Glu²⁵² do monômero A, Gln²⁰¹ (conformações alternadas A e B), Asp²³⁹ e Asn²⁴⁵ do monômero B.

A Figura 26 mostra o gráfico de Ramachandran gerado pelo programa PROCHECK. As porcentagens de resíduos nas regiões mais favoráveis, adicionalmente permitidas, generosamente permitidas e proibidas são, respectivamente, 93,2%, 6,4%, 0,5% e 0,0%. Os únicos resíduos que se apresentam em região generosamente permitida são os resíduos de Lys¹¹ de ambos os monômeros e, assim para os outros resíduos apontados antes, a verificação do seu correto/adequado posicionamento na densidade eletrônica corrobora a boa modelagem.

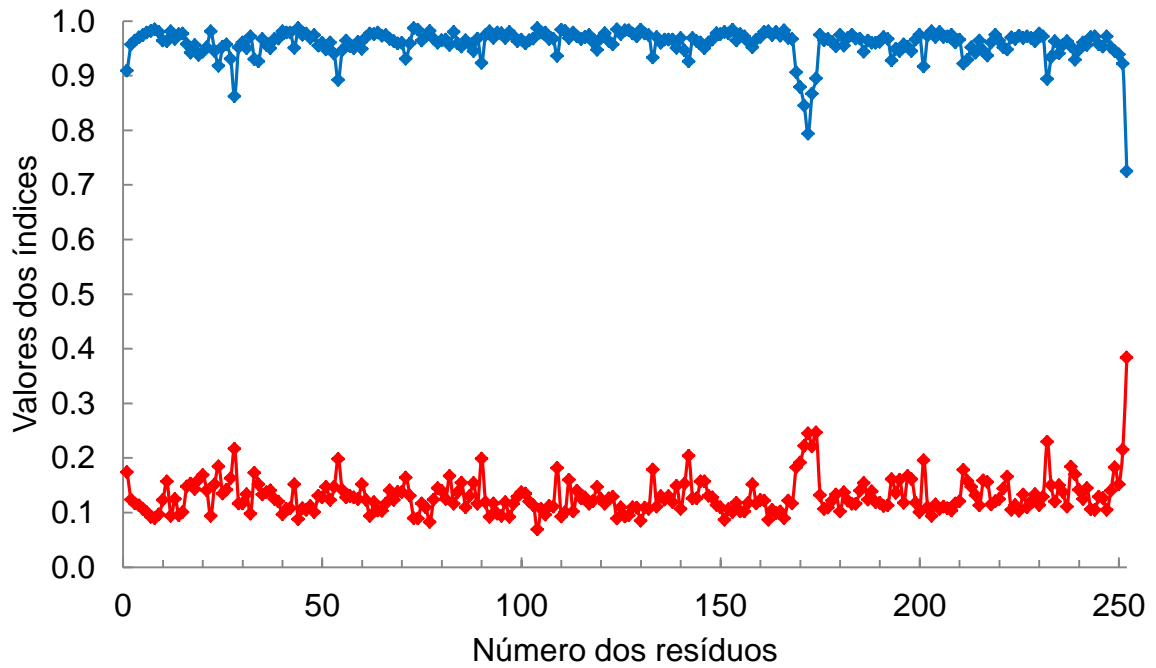
Figura 26. Gráfico de Ramachandran da estrutura NgTIM_2B. Em vermelho, regiões favoráveis, em amarelo, regiões adicionalmente permitidas, em bege, regiões generosamente permitidas e em branco, regiões proibidas.



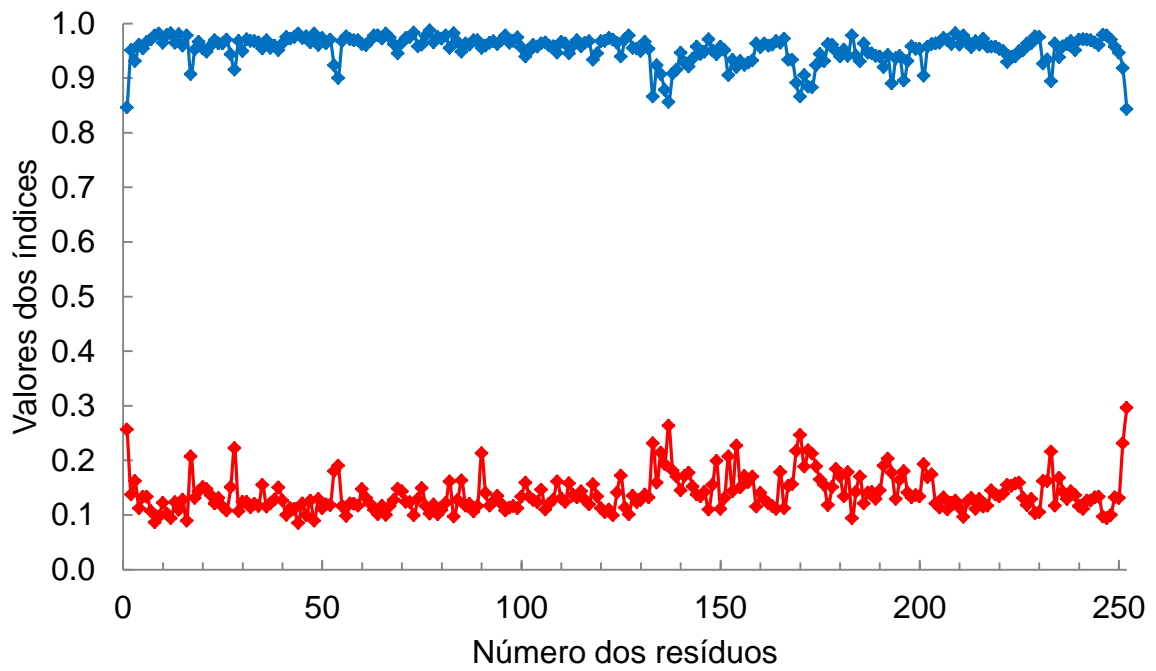
Fonte: O autor.

A qualidade dos modelos de ambas as cadeias também foi avaliada pelos índices RSCC e RSR, como demonstrado na Figura 27. Os valores globais de RSCC e RSR para a estrutura são: para monômero A, 0,9588 e 0,1292 e para o monômero B, 0,9525 e 0,1378, respectivamente.

Figura 27. Ajuste na densidade eletrônica, índices RSCC (azul) e RSR (vermelho) por resíduo para a estrutura obtida do conjunto de dados *NgTIM_2B*. (a) monômero A e (b) monômero B.



(a)



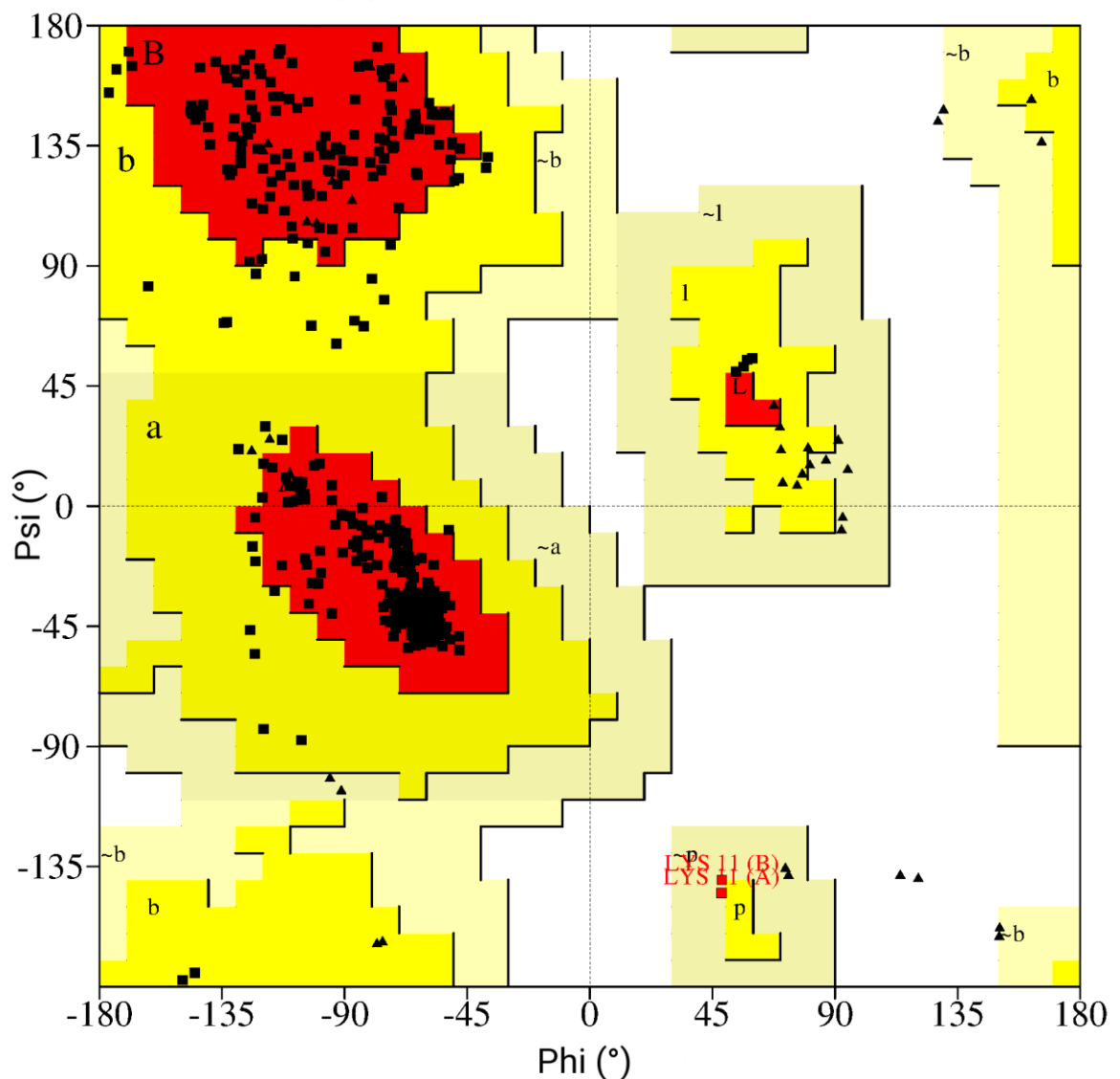
(b)

Fonte: O autor.

5.2.7.3 Validação da estrutura da triose fosfato isomerase de *Naegleria gruberi* obtida do conjunto de dados NgTIM_2C

Esta estrutura não apresentou nenhum desvio no tocante à conformação das cadeias laterais. O gráfico de Ramachandran é mostrado na Figura 28.

Figura 28. Gráfico de Ramachandran para a estrutura NgTIM_2C. Em vermelho, regiões favoráveis, em amarelo, regiões adicionalmente permitidas, em bege, regiões generosamente permitidas e em branco, regiões proibidas.

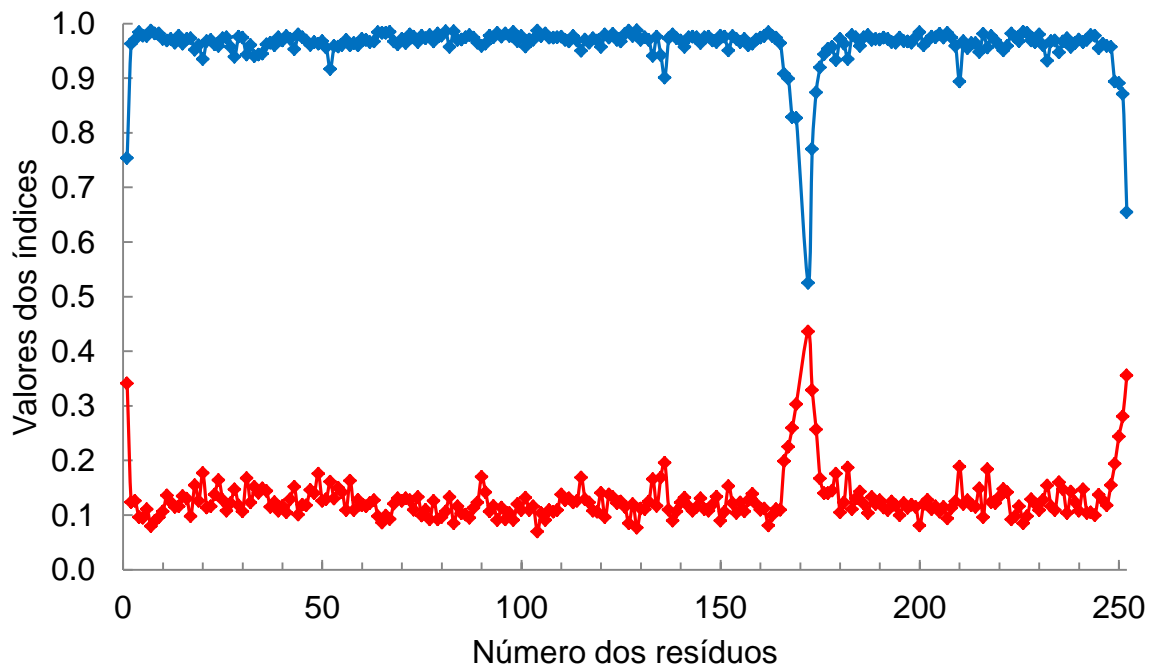


Fonte: O autor.

Os resíduos Lys¹¹ de ambos os monômeros, assim como nas outras estruturas, não deixam dúvida quanto à correta modelagem dado o bom ajuste na densidade

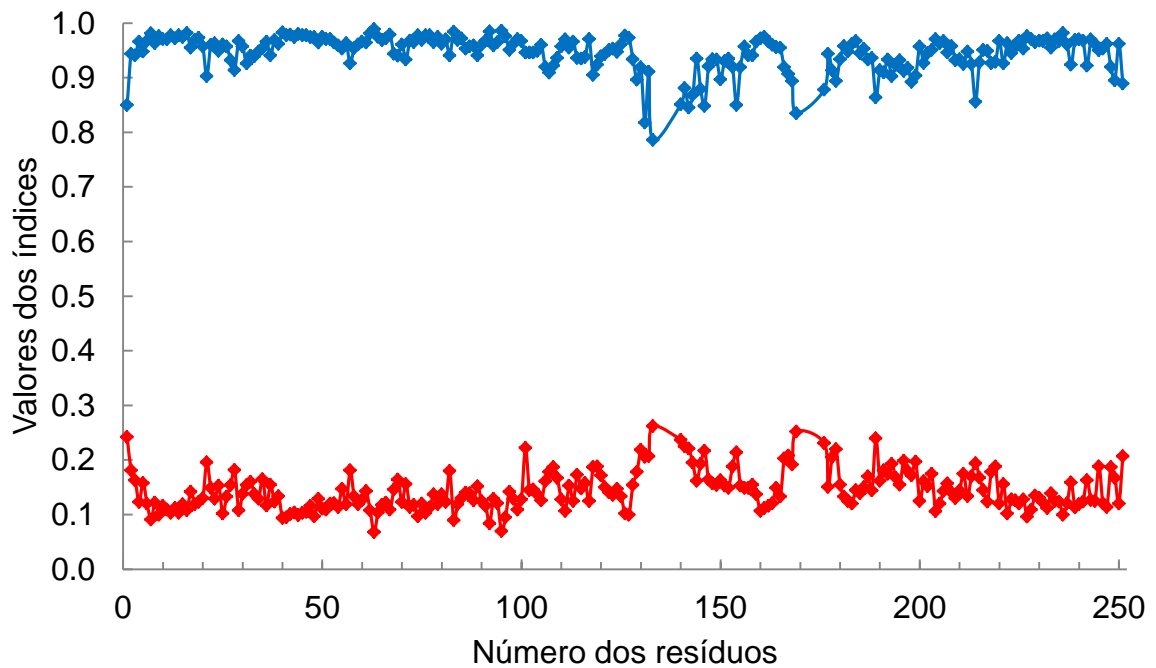
eletrônica. Para este modelo, também foram calculados os índices RSCC e RSR (Figura 29). Os respectivos valores globais são: para o monômero A, 0,9602 e 0,1292 e, para o monômero B, 0,9455 e 0,1439. A baixa qualidade destes índices na região entre os resíduos 170 a 175 do monômero A, 132 a 141 e 170 a 175 do monômero B é uma consequência da desordem elevada desta região, o que levou a deixar de se modelar alguns deles.

Figura 29. Ajuste na densidade eletrônica, índices RSCC (azul) e RSR (vermelho) por resíduo para a estrutura obtida do conjunto de dados *NgTIM_2C*. (a) monômero A e (b) monômero B. Continua.



(a)

Figura 29. Ajuste na densidade eletrônica, índices RSCC (azul) e RSR (vermelho) por resíduo para a estrutura obtida do conjunto de dados *NgTIM_2C*. (a) monômero A e (b) monômero B. Conclusão.



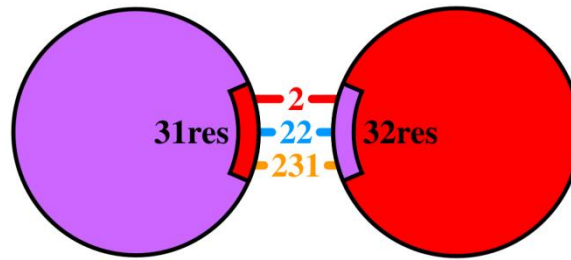
(b)

Fonte: O autor.

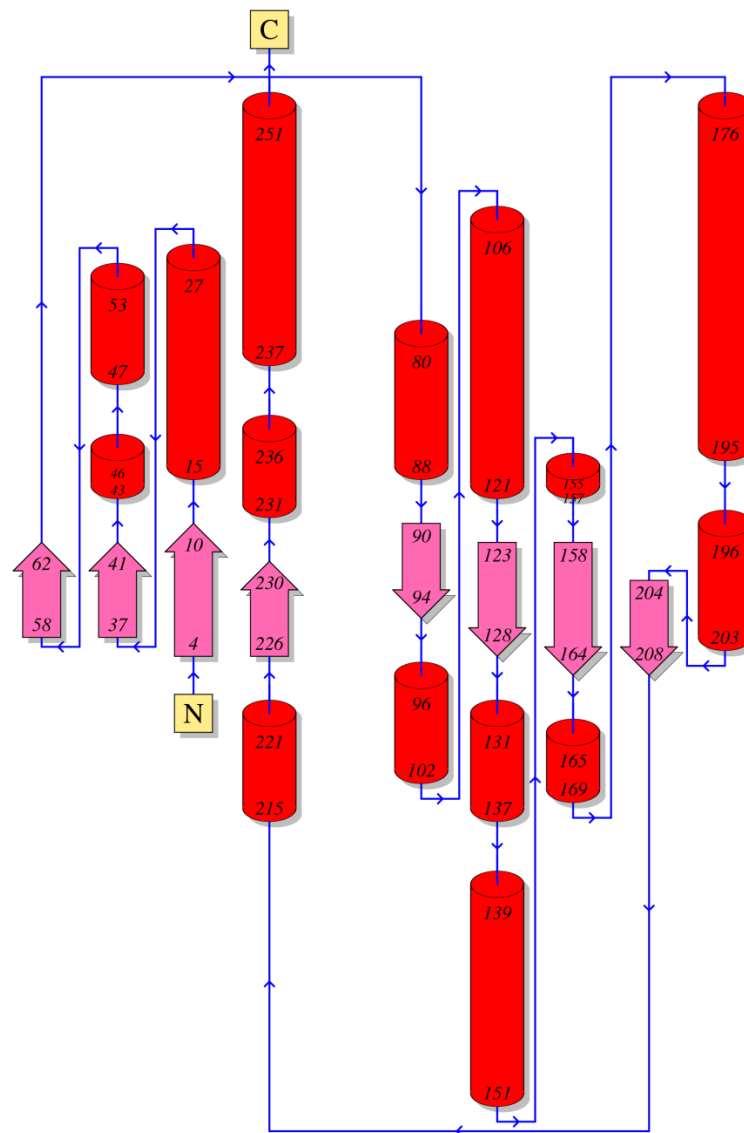
5.2.8 Análises e comparações estruturais para a triose fosfato isomerase de *Naegleria gruberi* obtida do conjunto de dados *NgTIM_2B*

A Figura 30 traz a representação pictórica das regiões de contato entre os monômeros do modelo e a topologia adotada pelo protômero (representativo monômero A) obtidas a partir da análise realizada pela ferramenta PDBsum (LASKOWSKI, 2001). As duas moléculas refinadas na unidade assimétrica devem corresponder ao dímero biológico. A área da região de interação é de 1440 \AA^2 e envolve 31 resíduos do monômero A e 32 do monômero B, a extensão destas interações de cada monômero é representada pela pequena região de mesma cor no contramonômero, como mostrado na Figura 30a. As interações desta região são mantidas por 2 pontes salinas, 22 ligações de hidrogênio e 231 contatos não ligados. A Figura 30b mostra que *NgTIM* pertence à classe estrutural alfa/beta (α/β) e possui enovelamento barril TIM β/α , segundo a classificação SCoP (CONTE, et al., 2002).

Figura 30. (a) Regiões de contato entre os monômeros: em roxo o monômero A e em vermelho o monômero B. (b) Diagrama da topologia adotada pela triose fosfato isomerase de *N. gruberi*.



(a)



(b)

Fonte: O autor.

Uma característica presente em outras estruturas de triose fosfato isomerasas depositadas no PDB e também na *NgTIM*, é a região da alça Pro¹⁶⁵-Gly¹⁷²,

altamente conservada como pode ser visto no alinhamento da Figura 31, realizado a partir de sobreposição de estruturas, de sete triose fosfato isomerases representativas de vários organismos. Nesta figura, os resíduos participantes do sítio ativo são destacados pelas estrelas em azul e os resíduos que apresentam contatos com o substrato estão destacados pelas estrelas em amarelo.

Figura 31. Alinhamento baseado em estrutura construído para ambos os monômeros de NgTIM e para algumas homólogas representativas. As estruturas secundárias (estimadas pelo programa DSSP) hélice α , fitas β e hélices irregulares são representadas cilindros, setas e espirais, respectivamente. As estrelas em azul denotam os resíduos componentes do sítio ativo e as estrelas em amarelo denotam os resíduos que apresentam interação com o substrato. *A. thaliana_CL* (AfTIM do cloroplasto) e *A. thaliana_CI* (AfTIM do citossol). Continua.

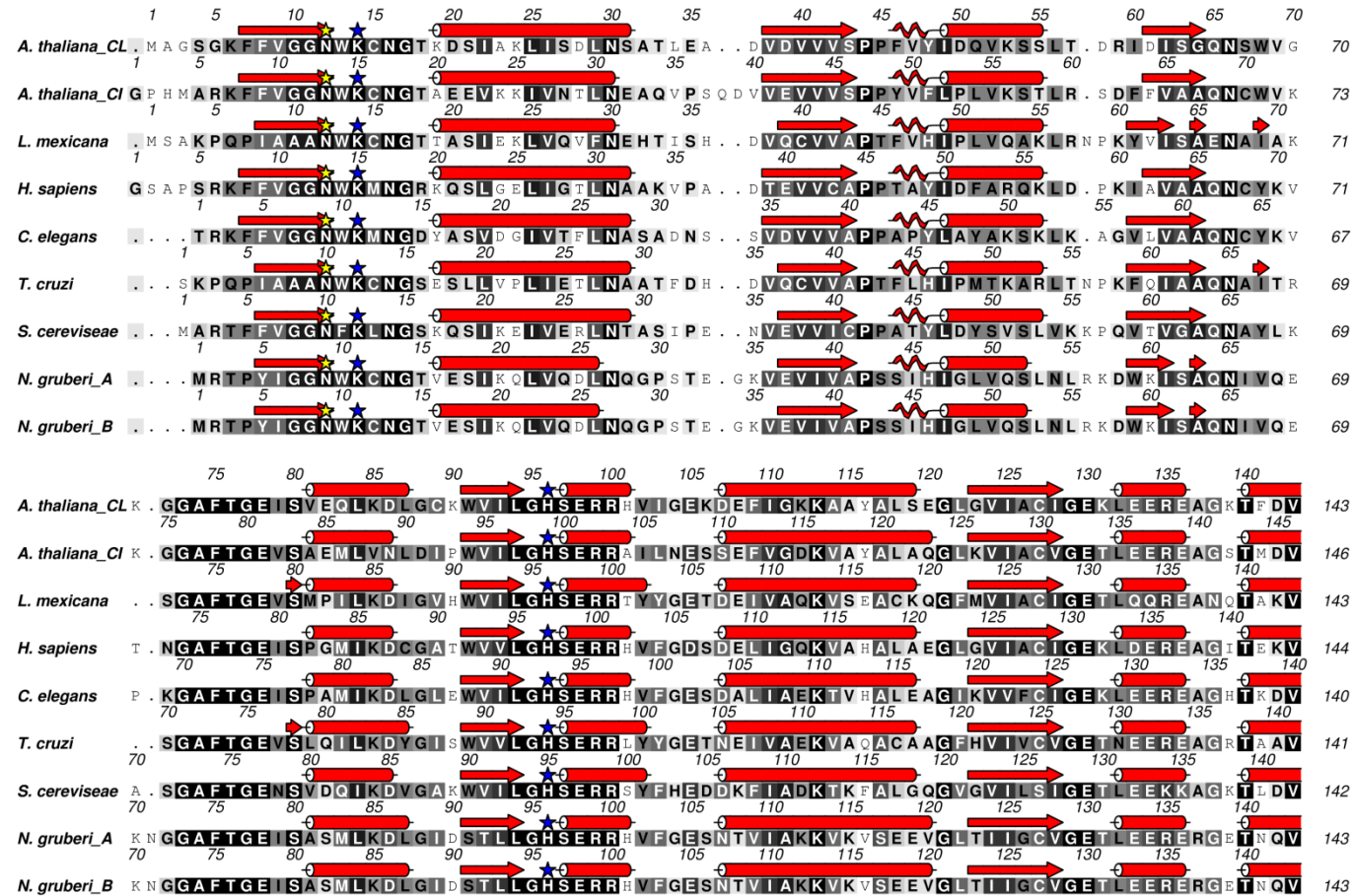
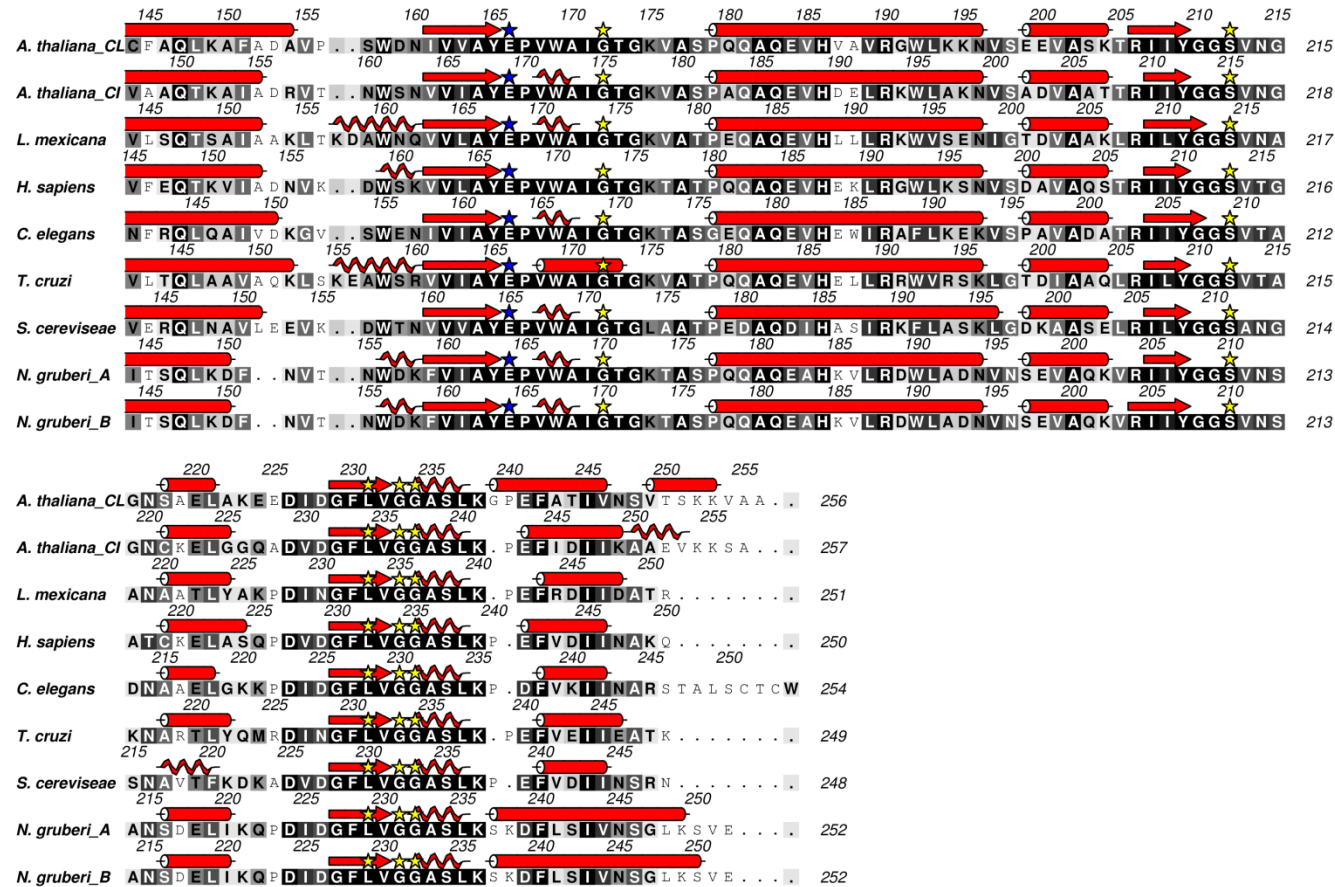


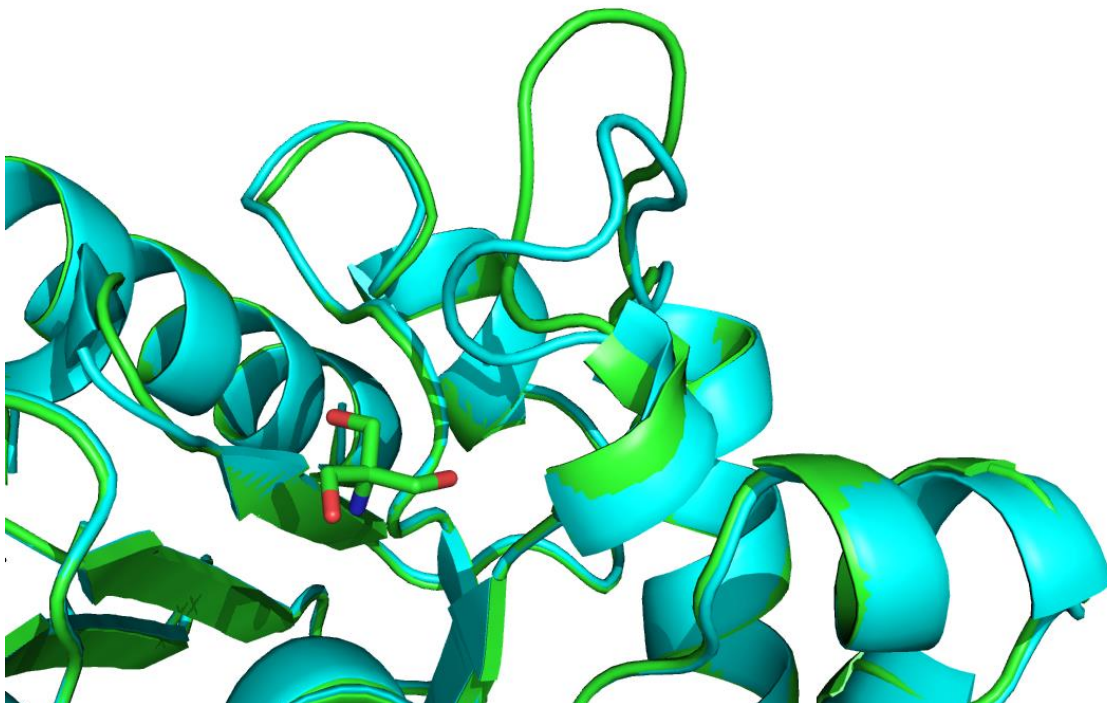
Figura 31. Alinhamento baseado em estrutura construído para ambos os monômeros de NgTIM e para algumas homólogas representativas. As estruturas secundárias (estimadas pelo programa DSSP) hélice α , fitas β e hélices irregulares são representadas cilindros, setas e espirais, respectivamente. As estrelas em azul denotam os resíduos componentes do sítio ativo e as estrelas em amarelo denotam os resíduos que apresentam interação com o substrato. *A. thaliana_CL* (AfTIM do cloroplasto) e *A. thaliana_CI* (AfTIM do citossol). Conclusão.



Fonte: O autor.

A sequência AIGTG (resíduos 168 a 172) é deslocada sobre a cavidade do sítio ativo (Figura 32) e desempenha um papel importante ao potencializar a capacidade do Glu¹⁶⁴ em abstrair o próton do substrato (GAP ou DHAP, dependendo do sentido considerado da reação) durante a reação de isomerização (MALABANAN; AMYES; RICHARD, 2010). O Glu¹⁶⁴ em todas estruturas de NgTIM refinadas encontra-se em uma posição considerada como um indício de que não há substrato no sítio ativo. Na estrutura NgTIM_2B, ambos os monômero A e B apresentam esta alça em uma conformação aberta. Entretanto, há no sítio ativo do monômero B uma molécula de Tris, cuja posição pode ser a causa do maior distanciamento da alça deste monômero em relação ao do outro monômero (Figura 32). Isso poderia levar à percepção errônea de que a conformação no monômero A está fechada (na forma usual da denominação, ou seja, com deslocamento angular na “dobradiça” do seu N-terminal).

Figura 32. A alça do monômero B (verde) projeta-se em direção ao solvente em uma extensão maior que o monômero A (azul), possivelmente devido à molécula de Tris (Tris(hidroximetil)aminoetano) presente no sítio ativo da primeira (representado aqui na forma de bastões).

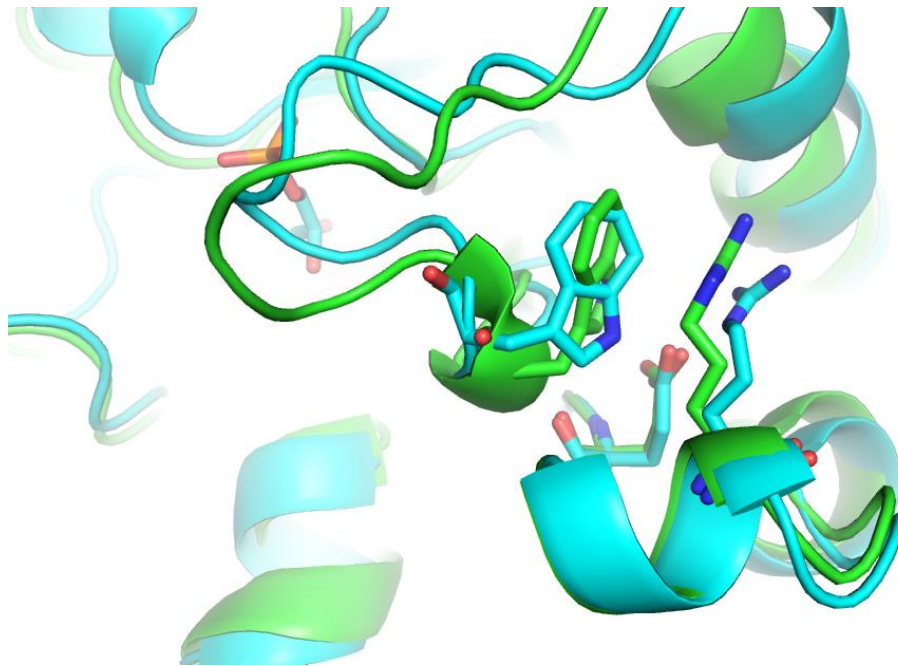


Fonte: O autor.

Relacionada ao fechamento do sítio ativo por esta alça, a sequência de resíduos

Pro¹⁶⁵Val¹⁶⁶Trp¹⁶⁷ (também altamente conservada), que constitui uma das dobradiças da alça, poderia contribuir para mantê-la aberta através da interação π -stacking (CHAKRABARTI; BHATTACHARYYA, 2007) entre Arg¹³⁵ e Trp¹⁶⁷. Após o fechamento da alça, devido à presença do substrato, o triptofano correspondente se desloca, mas a cadeia lateral da Arg¹³⁵ não sofre deslocamento, o que poderia ser devido à interação com o Glu¹³⁰, como pode ser observado, por exemplo, nos aminoácidos correspondentes na estrutura da triose fosfato isomerase de *T. brucei* (*TbTIM*, código PDB 6TIM). A Figura 33 mostra uma sobreposição entre os monômeros B de *NgTIM* e 6TIM.

Figura 33. Superposição dos monômeros B de *NgTIM* (verde) e *TbTIM* (azul) complexada com G3P. Representação na forma de bastões dos resíduos Arg¹³⁵, Trp¹⁶⁷ e Glu¹³⁰ de *NgTIM*; Glu¹²⁹, Arg¹³⁴ e Trp¹⁷⁰ de *TbTIM* e do ligante G3P.



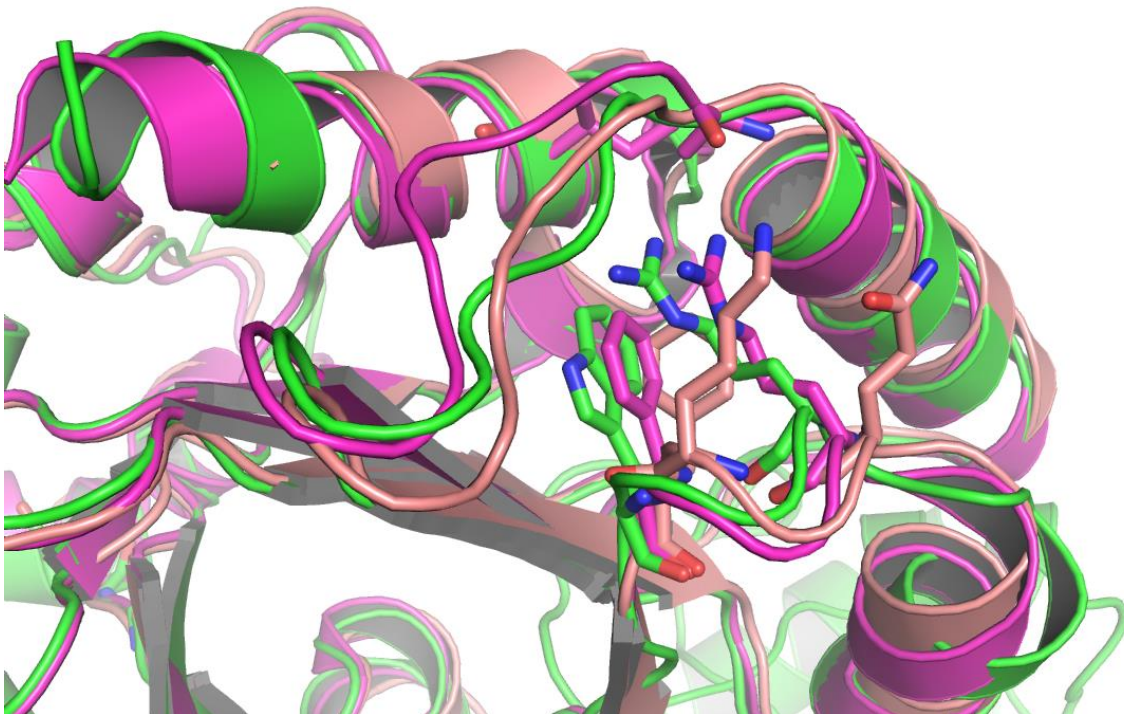
Fonte: O autor.

Visando destacar características particulares existentes na estrutura da *NgTIM*, um alinhamento estrutural foi realizado contra várias de suas homólogas (dados não mostrados). De modo geral, *NgTIM* apresenta elevado grau de conservação em todos os níveis estruturais (estrutura primária, secundária, terciária e quaternária). Em algumas regiões do alinhamento, é possível observar algumas poucas inserções, deleções e variações, mais especificamente: Gly⁷³ é uma inserção, Glu³³ está

ausente em todas as estruturas (embora na homóloga 1YDV (*Plasmodium falciparum*) o glutamato é substituído por prolina e em 4OBT por glutamina) e Ser²³⁵ encontra-se na posição de resíduos Ala, Glu, Pro, Val e Gly de algumas estruturas.

Durante as etapas iterativas de refinamento, chamou a atenção a interação entre os resíduos Arg⁵⁶ e Trp⁵⁹, pertencentes ao segmento **RKDW** (nas homólogas, esta região possui geralmente seis resíduos). Em 4OBT (*Arabidopsis thaliana*), o triptofano é substituído por fenilalanina (**RSDF**) (e talvez a substituição da Gly²⁹ por Ala³⁰ seja responsável por expulsar água junto com I²⁴² e impedir a arginina de efetuar uma interação XH- π nesta); em 1M6J (*Entamoeba histolytica*) o segmento correspondente é (**AGEANGANI**); em 1YDV, (**QSKF**). A superposição das estruturas de NgTIM, 4OBT e 1YDV mostra o posicionamento dos resíduos do segmento referido anteriormente (Figura 34).

Figura 34. Superposição das triose fosfato isomerases de *N. gruberi* (monômero B em verde), *A. thaliana* (código do PDB 4OBT) e *P. falciparum* (ambos monômeros A, em roxo e rosa, respectivamente).



Fonte: O autor.

No caso de NgTIM e as demais que apresentam arginina ou glutamina, as

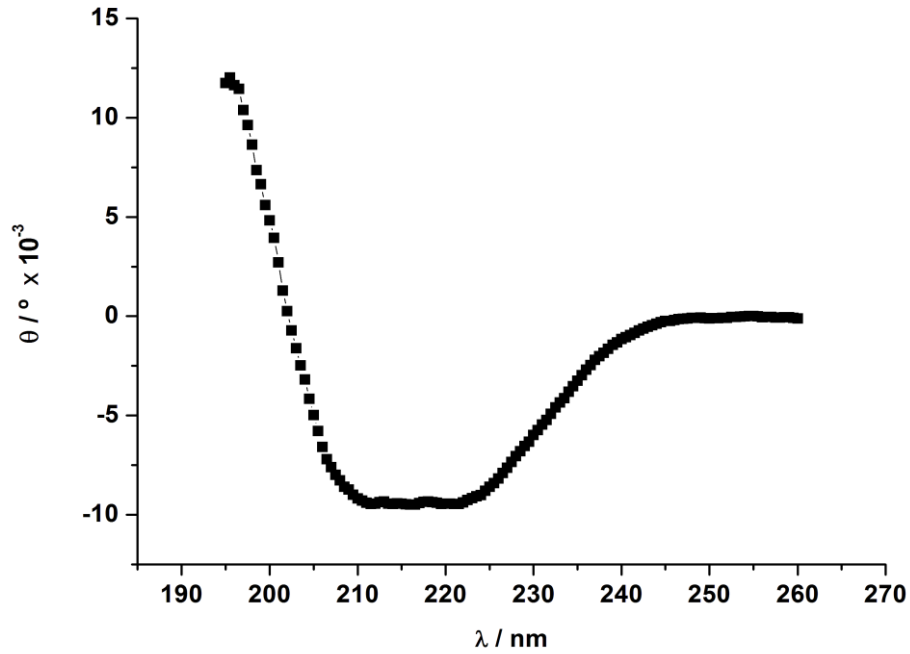
interações das cadeias laterais entre estes resíduos são do tipo XH- π com um resíduo aromático (embora em 1YDV, estas interações não ocorram devido à grande distância entre glutamina e fenilalanina, isso por sua vez pela presença de cadeia lateral de uma lisina interpondo-se entre estes dois resíduos – na NgTIM a cadeia lateral do glutamato na mesma posição reforça o posicionamento da arginina próximo ao triptofano), enquanto que nas outras estruturas estas interações podem ser entre cadeias laterais hidrofóbicas de isoleucinas e leucinas, em ambas as porções terminais do referido segmento, além das ligações de hidrogênio convencionais da cadeia principal.

Em relação às deleções, NgTIM não possui dois resíduos entre Asn¹⁵² e Val¹⁵³, além de outros dois resíduos entre Val¹⁵³ e Thr¹⁵⁴ (embora, neste caso, as homólogas apresentem de um a três resíduos) o que implica em hélices α mais curtas em relação às homólogas.

5.2.9 Dicroísmo Circular (CD) da triose fosfato isomerase de *Naegleria gruberi*: estabilidade térmica

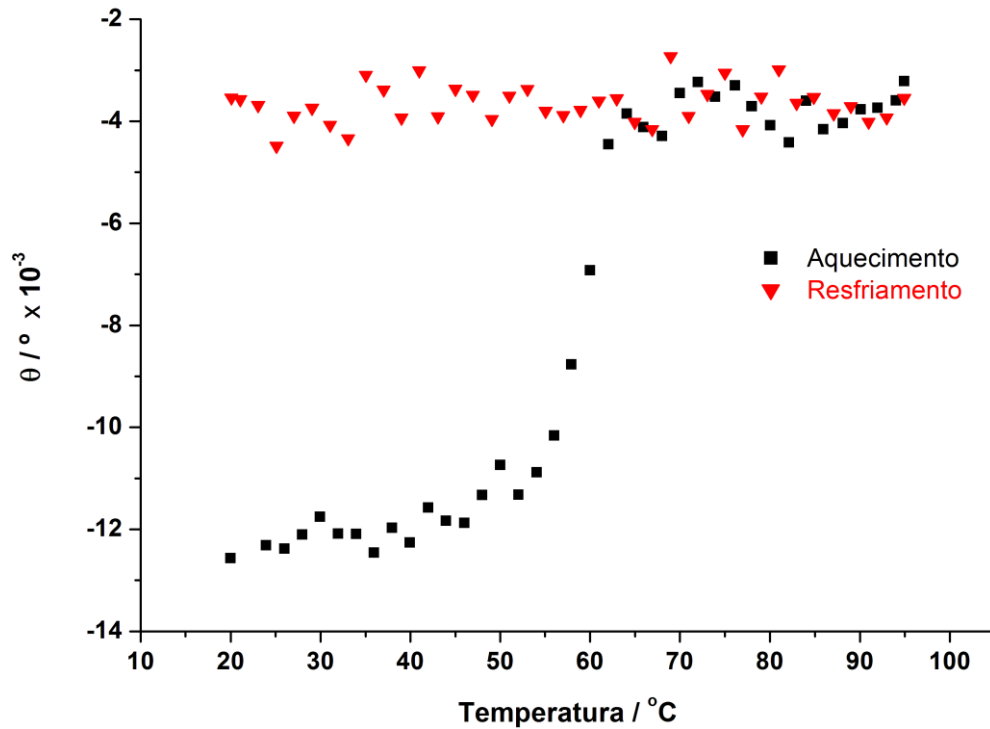
Ambos o pacote CD-Pro e o servidor CAPITO não proveram resultados corretos com o espectro medido (Figura 35), embora a forma da curva seja semelhante àquela apontada como característica para estruturas de triose fosfato isomerase. Novos experimentos deveriam ser planejados, todavia, dado que a estrutura terciária foi determinada por difração de Raios X, esta foi priorizada. Por outro lado, a informação do comportamento da enzima quando submetida ao aquecimento foi obtida. O gráfico da Figura 36 mostra que aproximadamente na faixa entre 50 °C (onde os pontos experimentais de elipticidade passam a se distanciar da tendência aproximadamente constante e horizontal na parte inferior da curva em preto) e 65 °C (onde os valores de elipticidade passam novamente a apresentar certa constância em relação à temperatura) a proteína sofre desnaturação. Após a perda completa da estrutura secundária, o resfriamento da amostra (curva em vermelho) não levou ao reenovelamento, uma vez que o padrão sigmoidal da curva não foi retomado, portanto, a desnaturação térmica é irreversível para a NgTIM, nas condições do experimento.

Figura 35: Espectro de dicroísmo circular da enzima triose fosfato isomerase de *N. gruberi*. A curva corresponde aos dados experimentais para o intervalo de comprimento de onda de 195 a 260 nm.



Fonte: O autor.

Figura 36: Perfil de desnaturação térmica da enzima triose fosfato isomerase de *N. gruberi* obtido por dicroísmo circular a 222 nm para aquecimento de 20 a 95 °C e posterior resfriamento de 95 a 20 °C.



Fonte: O autor.

6 Conclusão

Este trabalho resultou na modelagem por homologia da lipase de *Jatropha curcas* devido ao insucesso das tentativas de expressá-la. Uma suspeita é que a construção para expressá-la não está correta porque as bactérias submetidas ao processo de transformação desenvolveram colônias em meios seletivos, mas não apresentaram qualquer desenvolvimento de colônias em meios preparados como amostras controle (sem transformação). Outras possibilidades são a destruição dos ribossomos nas células de *E. coli* decorrente do uso de um vetor pET ou esta lipase apresentar alguma toxicidade para este microrganismo.

Os cristais da enzima triose fosfato isomerase de *Naegleria gruberi* forneceram conjuntos de imagens de difração por Raios X que permitiram a solução da sua estrutura tridimensional em três formas cristalinas diferentes (a resolução diferentes também): uma ($P4_122$) a 2,64 Å, uma (C2) a 1,56 Å e outra (C2) a 1,74 Å (o melhor dos conjuntos de dados coletados, que foi priorizado durante o refinamento). Nenhuma evidência de moléculas do substrato di-hidroxiacetona fosfato e nem do produto gliceraldeído-3-fosfato foi encontrada na densidade eletrônica, entretanto, uma molécula de Tris foi modelada no sítio ativo (apontada como responsável pela distorção observada próxima a esta região). Os experimentos de dicroísmo circular possibilitaram identificar a temperatura na qual a NgTIM sofre desnaturação térmica e que este processo é irreversível.

7 Referências Bibliográficas

- ABRIGOR, R. D.; UADIA, P. O.; FOGLIA, T. A.; HAAS, M. J.; SCOTT, K.; SAVARY, B. J. Partial purification and properties of lipases from germinating seeds of *Jatropha curcas* L. **Journal of the American Oil Chemists` Society**, v. 79, n. 11, p. 1123-1126, 2002.
- AFONINE, P. V.; GROSSE-KUNSTLEVE, R. W.; ADAMS, P. D.; The Phenix refinement framework. **CCP4 Newsletter on Protein Crystallography**, v. 42, 2005.
- AFONINE, P. A.; MORIARTY, N. W.; MUSTYAKIMOV, M.; SOBOLEV, O. V.; TERWILLIGER, T. C.; TURK, D.; URZHUMTSEVF, A.; ADAMS, P. D. FEM: Feature-enhanced map. **Acta Crystallographica Section D**, v. D71, p. 646-666, 2015.
- BALASHEV, K.; JENSEN, T. R.; BJORNHOLM, T. Novel methods for studying lipids and lipases and their mutual interaction at interfaces: Part I. Atomic force microscopy. **Biochimie**, v 83. n. 5, p. 387-397, 2001.
- BINO, J.; SALI, A. Comparative protein structure modeling by iterative alignment, model building and model assessment. **Protein Science**, v. 31, p. 3982-3992, 2003.
- BOND, C. S.; WOLFGANG, A.; SCHUTTELKOPF.; ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. **Acta Crystallographica Section D**. v. 65, p. 510-512, 2009.
- BLOW, D. **Outline of Crystallography for Biologists**. 3 ed. Oxford University Press, 2005.
- BOISTELLE, R.; ASTIER, J. P. Crystallization mechanisms in solution. **J. Cryst. Growth**, v.90, p.14-30, 1988.
- BRADFORD, M. M. A rapid sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Anal. Biochem.**, v. 72, p. 248-254, 1976.
- CHAKRABARTI, P.; BHATTACHARYYA, R. Geometry of nonbonded interactions involving planar groups in proteins. **Progress in Biophysics and Molecular Biology**, v. 95, p. 83-137, 2007.
- CDC: Centers for Disease Control and Prevention. Disponível em: <<http://www.cdc.gov/parasites/naegleria/general.html>>. Acesso em: 4 ago. 2015.
- CHEN, V. B.; BRYAN, W. A.; HEADD, D. A.; KEEDY, M. A.; IMMORMINO, M. R.; KAPRAL, G. J.; MURRAY, J. S.; RICHARDSON, D. C. Molprobity: all atoms structure validation for macromolecular crystallography. **Acta Crystallographica Section D**, v. D66, p. 12-21, 2010.
- CHENNA, R.; SUGAWARA, H.; KOIKE, T.; LOPEZ, R.; GIBSON, T. J.; HIGGINS, D. G.; THOMPSON, J. D. Multiple sequence alignment with the clustal series of programs. **Nucleic Acid Research**, v. 31, p. 3497-3500, 2003.

CONTE, L. L.; BRENNER, E. S.; HUBBARD, P. J. T.; CHOTHIA, C.; MURZIN, G. A.; Scop database in 2002: refinements accommodate structural genomics. **Nucleic Acid Research**, v. 30, p. 264-267, 2002.

DE AZEVEDO, W. F. et al. Crystal structure of human purine nucleosidephosphorylase at 2.3 Å resolution. **Biochem. Biophys.**v.303(3), p.545-552, 2003.

DE LA MARE, S.; COULSON, A. F. W.; KNOWLES, J. R.; PRIDDLE, J. D.; OFFORD, R. E. Active-site labelling of triose phosphate isomerase. The reaction of bromohydroxyacetone phosphate with a unique glutamic acid residue and the migration of the label to tyrosine. **Biochemistry Journal**, v. 129, n. 2, p. 321-331.

EMSLEY, P.; COWTAN, K. Coot: Model-building Tools for Molecular Graphics. **Acta Crystallographica Section D**, p. 2126-2132, 2004.

EVANS, P. R.; GARIB, N. M. How good are my data and what is the resolution? **Acta Crystallographica Section D**, v. D69, p. 1204-1214, 2013.

FILHO, O. A. S.; ALENCASTRO, R. B. Modelagem de proteínas por homologia. **Química Nova**. V. 26, n. 2, p. 253-259, 2003.

FRITZ-LAYLIN, L. K.; PROCHNIK, S. E.; GINGER, M. L.; DACKS, J. B.; CARPENTER, M.L.; FIELD, M. C.; KUO, A.; PAREDEZ, A.; CHAPMAN, J.; PHAM, J.; SHU, S.; NEUPANE, R.; CIPRIANO, M.; MANCUSO, J.; TU, H.; SALAMOV, A.; LINDQUIST, E.; SHAPIRO, H.; LUCAS, S.; GRIGORIEV, I. V.; CANDE, W. Z.; FULTON, C.; ROKHSAR, D. S.; DAWSON, S. C. The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. **Cell**, v. 140, p. 631-642, 2011.

FULTON, C. *Naegleria*: A research partner for cell and developmental biology. **J. Eukaryot. Microbio.**, v. 40, p. 520-532, 1993.

FULTON, C. Amebo-flagellates as research partners: The laboratory biology of *Naegleria* and *Tetramitus*. **Methods Cell Physiol.**, v. 4, p. 341-346.

GASTEIGER, E.; HOOGLAND, C.; GATTIKER, A.; DUVAUD, S.; WIKKINS, M. R.; APPEL, R. D.; BAIROCH, A. Protein Identification and Analysis Tools on the ExPasy Server. In: WALKER, J. M. (Ed.). **The Proteomics Protocols Handbook**. New York: Humana Press, 2005. p. 571-607.

GILHAM, D.; LEHNER, R. Techniques to measure lipase and esterase activity in vitro. **Methods**. v. 36, n. 2, p. 139-147, 2005.

GO, M. K.; KOUDELKA, A.; AMYES, T. L.; RICHARD, J. P. The role of Lys-12 in catalysis by triosephosphateisomerase: a two-part substrate approach. **Biochemistry**, v. 49, n. 25, p. 5377-5389, 2010.

GU, H.; JIANG, Y.; ZHOU, L.; GAO, J. Reactive extraction and in-situ self-catalyzed methanolysis of germinated oilseed for biodiesel production. **Energy Environ. Sci.** v.

4, p. 1337-1344, 2010.

HASAN, F.; SHAH, A. A.; HAMEED, A. Methods for detection and characterization of lipases: A comprehensive review. **Biotechnology Advances**, v. 27, n. 6, p. 782-798, 2009.

JAEGER, K. E.; RANSAC, S.; DIJKSTRA, B. W.; COLSON, C.; VANHEUVEL, M.; MISSET, O. Bacterial lipase. **FEMS Microbiol. Rev.** v. 15, p.29-63, 1994.

JAEGER, K. E.; DIJKSTRA, B. W.; REETZ, M. T. Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. **Annual Review of Microbiology**, v.53, p.315-351, 1999.

KABSCH, W. XDS: Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. **Journal of Applied Crystallography**, v. 26, p. 795-800, 1994.

KABSCH, W.; SANDER, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. **Biopolymers**, v. 22, p. 2577–2597, 1983.

KLEYWEGT, G. J.; JONES, T. A. xdi MAPMAN and xdiDATAMAN – programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. **Acta Crystallographica Section D**, v. 52, p. 826-828, 1996.

KLEYWEGT, G. J. Making the most of your search model. **CCP4/ESF-EACBM Newsletter on Protein Crystallography**, v. 32, p. 32-36, 1996.

KNOWLES, J. R. Enzyme catalysis: not different, just better. **Nature**, v. 350, p. 121-124, 1991.

KUMAR, A; SHARMA, S. An evaluation of multipurpose oil seed crop for industrial uses (*Jatropha curcas* L.): A review. **Industrial Crops and Products**, v. 28, p. 1-10, 2008.

LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head bacteriophage T4. **Nature**, v. 227, p. 680-685, 1970.

LASKOWSKI, R.A.; MACARTHUR, M. W.; MOSS, D. S.; THORNTON, J. M. Procheck: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, p. 283-291, 1993.

LASKOWSKI, R. A. PDBsum: summaries and analyses of PDB structures. **Nucleic Acids Research**, v. 29, p. 221-222, 2001.

LEITE, A.; NASS, L. L.; VALOIS, A. C. C.; MELO, I. S. Expressão de proteínas heterólogas em plantas. In Recursos genéticos e Melhoramento. **Plantas**, v. 1. MC, Ed (Rondonópolis: Fundação MT), p. 1057-1083. 2001.

LOLIS, E.; PETSKO, G. A. Transition-state analogues in protein crystallography:

probes of the structural source of enzyme catalysis. **Annual Review of Biochem.**, v. 59, p. 597-630, 1990.

MALABANAN, M. M.; AMYES, T. L.; RICHARD, J. P. A role for flexible loops in enzyme catalysis. **Current Opinon in Strucural Biology**, v. 20, p. 702-710, 2010.

MATTHEWS, B. W. Solvent content of protein crystals. **Journal of Molecular Biology**, v. 33, n. 2, p. 491-497, 1968.

McCOY, A. J.; Grosse-Kunstleve, R. W.; ADAMS, P. D.; WINN, M. D.; STORONI, L. C.; READ, R. J. Phaser crystallographic software. **Journal of Applied Crystallography**, v. 40, p. 658-674, 2007.

National Center for Biotechnology Information (NCBI), disponível em: www.ncbi.nlm.nih.gov.

NELSON, D. L.; COX, M. M. Lehninger princípios de bioquímica. 3 ed. São Paulo: Artmed, 2002.

NOTREDAME, C.; HIGGINS, D. G.; HERINGA, J. T-coffee: a novel method for fast and accurate multiple sequence alignment. **Journal of Molecular Biology**, v. 302, n. 1, p. 205-217, 2000.

OPPERDOES, F. R.; DE JONCKHEERE, J. F.; TIELENS, A. G. M. *Naegleria gruberi* metabolism. **International Journal of Parasitology**, v. 41, p. 915-924, 2011.

OROSZ, F.; OLAH, J.; OVADI, J. Triosephosphateisomerase deficiency: new insights into an enigmatic disease. **Biochimica et Biophysica Acta: Molecular basis of Disease**, v. 1792, n. 12, p. 1168-1174, 2009.

POWELL, K. A.; RAMER, S. W.; DEL CARDAYRÉ, S. B.; STEMMER, W. P. C.; TOBIN, M. B.; LONGCHAMP, P. F.; HUISMAN, G. W. Directed evolution and biocatalysis. **Angewandte Chemie International Edition**, v. 40, p. 3948-3959, 2001.

SHARMA, S.; DHAMIJA, H. K.; PARASHAR, B. *Jatropha curcas*: a review. **Asian Pharma Press**, v. 2, n. 3, p. 107-111, 2012.

PARK, Y.; SHEETLIN, S.; MA, N.; MADDEN, T.; SPOUGE, J. New finite-size correction for local alignment score distributions. **BMC Research Notes**, v. 5, n. 1, p. 1-6, 2012.

PAINTER, J.; MERRIT, E. A. TLSMD web server for the generation of multi-group TLS models. **Journal of Applied Crystallography**, v. 39, p. 109-111, 2008.

Protein Data Bank (PDB), disponível em: www.rcsb.org.

RHODES, G. **Crystallography made crystal clear: A guide for users of macromolecular models**. 2 ed. San Diego: Academic Press, 2000.

ROSE, A. I.; O'CONNELL, E. L. Inactivation and labeling of triose phosphate

isomerase and enolase by glycidol phosphate. **J. Biol. Chem.**, v. 244, p. 6548-6550, 1969.

RUPP, B. **Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology**. 1. ed. New York: Gerland Science, 2010.

SALI, A; SHEN, M. Statistical potential for assessment and prediction of protein structure. **Protein Sci.** v. 15, p. 2507-2524, 2006.

SALI, A.; BLUNDELL, T. L. Comparative protein modelling by satisfaction of spatial restraints. **J. Mol. Biol.**, v. 234, p. 779-815, 1993.

SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. **Molecular Cloning: a laboratory manual**. 2. ed. New York: Cold Spring Harbor Laboratory Press, 1989.

SHARMA, R.; CHISTI, Y.; BANERJEE, U. C. Production, purification, characterization, and applications of lipases. **Biotechnology Advances**, v. 19, p. 627-662, 2001.

KELLY, S. M.; JESS, T. J.; PRICE, N. C. How to study proteins by circular dichroism. **Biochimica et Biophysica Acta**, v. 1751, p. 119-139, 2005.

SILVA, F. H. Módulo Biologia molecular. I Escola Brasileira de Inteligência Artificial e Bioinformática InBio, São Carlos. Centro de biotecnologia molecular estrutural. 2001.

SLABINSKI L, JAROSZEWSKI L, RYCHLEWSKI L, WILSON I.A., LESLEY S.A., GODZIK A. XtalPred: a web server for prediction of protein crystallizability. **Bioinformatics**, 2007 v. 23, p. 3403-3405.

SOUSA, J. S.; CAVALCANTI-OLIVEIRA, E. d'Avila.; ARANDA, D. A. G.; FREIRE, D. M. G. Application of lipase from physic nut (*Jatropha curcas* L.) to a new hybrid (enzyme/chemical) hydroesterification process for biodiesel production. **Journal of Molecular Catalysis B: Enzymatic**, v. 65, p. 133-137, 2010.

SREERAMA, N. e WOODY, R. W.; Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set. **Analytical Biochemistry**, v. 287, n. 2, p. 252-260, 2000.

STAUBMAN, R.; NCUBE, I.; GUBITZ, G. M.; STEINER, W.; READ, J. S. Esterase and lipase activity in *Jatropha curcas* L. seeds. **Journal of Biotechnology**, v. 75, n.2-3, p. 117-126, 1999.

STEIN, N. Chainsaw: a program for mutating pdb files used as templates in molecular replacement. **J. Appl. Cryst.**, v. 41, p. 641-643, 2008.

SCHWARZENBACHER, R. et al. The importance of alignment accuracy for molecular replacement. **Acta Crystallographica**, v. D60, p. 1229-1236, 2004.

STRYER, L.; TYMOCZKO, J. L.; BERG, J. M. **Bioquímica**, 5 ed. Editora: Guanabara,

2004.

TORTORA, G. J.; FUNKE, B. R.; CASE, C. L. **Microbiologia**. 10 ed. Porto Alegre: Artmed, 2012.

TWYMAN, R. M.; STOGER, E.; SCHILLBERG, S.; CHRISTOU, P.; FISCHER, R. Molecular farming in plants: host systems and expression technology. **Trends in Biotechnology**, v. 21, p. 570-578, 2003.

WIEDEMANN, C.; BELLEDT, P.; GORLACH, M. Capito: a web server-based analysis and plotting tool for circular dichroism data. **Bioinformatics**, v. 29, n. 14, p. 1750-1757, 2013.

WIERENGA, R. K. KAPETANIOU, E. G.; VENKATESAN, R. Triosephosphate isomerase: a highly evolved biocatalyst. **Cellular and Molecular Life Sciences**, v. 67, n. 23, p.3961-3982, 2010.

WINN, M. D.; BALLARD, C. C.; COWTAN, K. D.; DODSON, E. J.; EMSLEY, P.; EVANS, P. R.; KEEGAN, R. M.; KRISINEL, E. B.; LESLIE, A. G. W.; McCOY, A.; McNICHOLAS, S. J.; GARIB, N. M.; PANNU, N. S.; POTTERTON, E. A.; POWELL, H. R.; READ, R. J.; VAGIN, A.; WILSON, K. S. Overview of the CCP4 suite and current developments. **Acta Crystallographica Section D**, v. 67, p. 235-242, 2011.

VISVESVARA, G. S.; MOURA, H.; SCHUSTER, F. L. Pathogenic and opportunistic free-living amoebae: *Acanthamoeba* spp. *Balamutia mandrillares*, *Naegleria fowleri* and *Sappiniadiploidea*. **FEMS Immunol. Med. Microbiol.**, v. 50, p. 1-26, 2007.

ZAIA, D. A. M.; ZAIA, C. T. B. V.; LICHTIG, J. Determinação de proteínas totais via espectrofotometria: vantagens e desvantagens dos métodos existentes. **Quim. Nova**, v. 21, n. 6, p. 787-793, 1998.