

# UNIVERSIDADE FEDERAL DO RIO DE JANEIRO CENTRO DE CIÊNCIAS MATEMÁTICAS E DA NATUREZA INSTITUTO DE QUÍMICA



## **BEATRIZ ROSA PENNA**

PROSPECÇÃO DE NOVOS ALVOS BIOTECNOLÓGICOS: CARACTERIZAÇÃO BIOQUÍMICA E ESTRUTURAL DA FOSFATASE HIPOTÉTICA DSM-14977 DE *OCEANITHERMUS PROFUNDUS* 



## PROSPECÇÃO DE NOVOS ALVOS BIOTECNOLÓGICOS: CARACTERIZAÇÃO BIOQUÍMICA E ESTRUTURAL DA FOSFATASE HIPOTÉTICA DSM-14977 DE *OCEANITHERMUS PROFUNDUS*

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Bioquímica, Instituto de Química, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Ciências

Orientadora: Cristiane Dinis Ano Bom (IQ/UFRJ)

Rio de Janeiro

## Penna, Beatriz Rosa

Prospecção de novos alvos biotecnológicos: Caracterização bioquímica e estrutural da fosfatase hipotética DSM-14977 de *Oceanithermus profundus* / Beatriz Rosa Penna - 2018.

Dissertação (Mestrado em Ciências) – Universidade Federal do Rio de Janeiro, Instituto de Química, Programa de Pós-Graduação em Bioquímica, 2018.

Orientador: Cristiane Dinis Ano Bom

1. Estrutura de proteínas. 2. Superfamília HAD. 3. biocatalisadores. 4. Enzimas. 5. Fluorescência. 6. Fosfatases. I. Ano Bom, Cristiane Dinis (Orient.). II. Universidade Federal do Rio de Janeiro. Instituto de Química. Programa de Pós-Graduação em Bioquímica. III. Título.

#### **Beatriz Rosa Penna**

## PROSPECÇÃO DE NOVOS ALVOS BIOTECNOLÓGICOS: CARACTERIZAÇÃO BIOQUÍMICA E ESTRUTURAL DA FOSFATASE HIPOTÉTICA DSM-14977 DE OCEANITHERMUS PROFUNDUS

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Bioquímica, Instituto de Química, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Ciências

| ovada em |   |
|----------|---|
|          |   |
|          |   |
|          |   |
|          |   |
|          | Prof <sup>a</sup> adjunta Cristiane Dinis Ano Bom (DBq/IQ/UFRJ) |
|          |   |
|          |   |
|          | Prof associado Rodrigo Volcan Almeida (DBq/IQ/UFRJ)             |
|          |   |
|          |   |
|          |   |
|          | Prof <sup>a</sup> adjunta Juliana Reis Cortines (PPGM/ UFRJ)    |
|          |   |
|          |   |
|          | Prof <sup>a</sup> Evelin Andrade Manoel (FF/UFRJ)               |
|          |   |
|          |   |
|          |   |
|          | Prof adjunto Fábio César Sousa Nogueira (DBq/IQ/UFRJ)           |

Dedico esta dissertação à querida Márcia de Oliveira Dias, *in memoriam*, pela sua curiosidade e devoção a ciência,

#### **AGRADECIMENTOS**

Agradeço a Deus por iluminar meu caminho até aqui, por colocar pessoas maravilhosas na minha vida e por sempre me trazer fé, positividade e força nos momentos mais difíceis. Aos meus guias espirituais, pela proteção e inspiração.

Agradeço a minha mãe, Rose, e ao meu pai, Wellington, por sempre ter acreditado no meu potencial. Só vocês sabem o tanto de trabalho que eu dei, mas vocês souberam contornar isso da melhor forma possível. Obrigada por todos os ensinamentos, os conselhos e os castigos que me fizeram refletir e amadurecer. E apesar de todas as dificuldades no nosso caminho não faltou amor, carinho, atenção, dedicação e alegria. A nossa perseverança me fez chegar até aqui. Agradeço por todo o apoio e incentivo nos meus estudos e nos momentos difíceis. Obrigada por esse amor incondicional. Vocês são meus exemplos, meu suporte, amo vocês!!!!

Agradeço aos meus irmãos por existirem na minha vida. Ao meu irmão, Felipe, obrigada por sempre estar ao meu lado quando criança, foi com você que aprendi a jogar bola, a brincar de pique, também foi com você que tomei meus primeiros esporros e tive meus primeiros hematomas de briga rs. Obrigada por todas essas boas lembranças, minha infância não teria tido graça sem você. Que o tempo não deixe a gente perder esse espírito de criança e que ele nos aproxime mais! E ao meu irmão mais novo, Cristiano, obrigada por fazer parte da minha vida, você chegou para multiplicar as alegrias e o amor na nossa família. Amo muito vocês!

Agradeço aos meus parentes que indiretamente contribuíram para este momento, principalmente, a tia Ciça e tia Iza, pelos conselhos, incentivos e carinho.

Agradeço a minha orientadora Cris por todo o carinho e pela amizade dentro e fora do laboratório. Obrigada por ser uma professora maravilhosa, que me conquistou na primeira aula e que me fez acreditar em que eu estava no lugar e caminho certo. Obrigada por todo o conhecimento proporcionado ao longo desses tantos anos, pela dedicação como orientadora mesmo com o Matheus e a Clarinha, por estar sempre disponível a me ajudar e por ser esse exemplo de profissional que eu admiro muito. Não poderia deixar de agradecer por todas as broncas, os puxões de orelha, a paciência, os incentivos e os conselhos, sem dúvida foram indispensáveis para o meu crescimento e amadurecimento científico e pessoal. E mesmo com a pretensão de voar um pouco fora do ninho no doutorado, você vai continuar sendo minha mãe científica, te amo!

Agradeço à querida professora Dani por ser essa pessoa meiga e atenciosa com a qual tenho compartilhado bons momentos e estudos espirituais que tem me ajudado muito na compreensão da vida e das perdas também. Obrigada por todo o conhecimento e auxílio na bancada. Você contribuiu extremamente para este momento!

Agradeço à professora Lucia Paiva e ao professor Anderson por toda a ajuda e conselhos prestados.

Agradeço a todos os queridos amigos do laboratório: Biancuda, Bruna, Camila, Carol Matos, Danielzinho, Didi, João, Luisinho, Lulis, e, aos quase internos, Ariana, Maysa e Natália! Vocês tornaram o laboratório a minha segunda casa. Em especial, agradeço a Carolixa, Guiguico, Leo e Rafa por todo auxílio, conselho e amizade prestada durante o mestrado. Todos vocês são muito importantes pra mim. Muito obrigada pela ajuda na bancada, pelos lanchinhos, conversas, bagunças, festinhas, noites das poderosas (apesar de estarmos velhos e cansados), enfim, pela amizade. Amo vocês, LaBEP!

Agradeço aos laboratórios do departamento de Bioquímica, principalmente, ao LaMMP e ao Laboratório de Bioinformática por sempre abrir, gentilmente, as portas para nós do LaBEP e nos receber com grande prestatividade.

Agradeço ao doutor Francisco Gomes Dinis pela preciosa colaboração com os dados de modelagem molecular.

Agradeço as minhas amigas de infância Dutrinha, Gau e Raiza, vocês são minhas irmãs de alma. Agradeço também a minha amiga Ju que não é amiga de infância, mas é muito especial também. Obrigada por sempre compreenderem (às vezes não rs) o "hoje não vai dar, tenho que estudar". Esses últimos meses foram bastante difíceis e seriam muito pior sem a amizade, o carinho, os conselhos, a companhia insubstituível, o ombro amigo, a alegria, os abraços e os sorrisos de vocês. O meu muito obrigado! Amo muito vocês!

Agradeço a todos os meus amigos da faculdade Ana, Deza, Camila, Daniel, Eric, Gabi, Gigi, Janão, Juju, Lolo, Luna, Piri, Pottinho, Raquel do wave, Kelzinha, Sinfo, Thabatinha e Yuri. Uma das melhores coisas que levei da faculdade foi a amizade de vocês, nos tornamos uma família, a família Biofx. Ahh Xerém... temos muuuitas histórias pra contar! Cada um de vocês tem um lugar especial no meu coração, obrigada por fazerem parte da minha caminhada científica.

Agradeço a Tia Olímpia, a minha mãe da creche, que sempre acreditou na minha capacidade. Muito obrigada por cuidar, me mimar e me amar como uma mãe. Não tenho palavras para agradecer tudo o que fez por mim.

Agradeço à agência de fomento CAPES pelo auxílio financeiro que me ajudou e incentivou a desenvolver esse projeto.

Agradeço aos integrantes da banca prof. Rodrigo Volcan, prof<sup>a</sup>. Juliana Cortines, prof<sup>a</sup>. Evelin e o prof. Fábio Nogueira por terem se disponibilizado a ler e avaliar o meu trabalho

E, finalmente, agradeço às pessoas que, direta ou indiretamente, fizeram parte da minha formação. Aos que foram e são importantes na minha vida, sendo a inspiração e o exemplo para que eu me tornasse a pessoa que sou hoje, meu sincero obrigado!

.

#### **RESUMO**

PENNA, Beatriz Rosa. **PROSPECÇÃO DE NOVOS ALVOS BIOTECNOLÓGICOS: CARACTERIZAÇÃO BIOQUÍMICA E ESTRUTURAL DA FOSFATASE HIPOTÉTICA DSM-14977 DE** *OCEANITHERMUS PROFUNDUS*. Rio de Janeiro, 2018. Dissertação (Mestrado em Bioquímica). Instituto de Química, Universidade Federal do Rio de Janeiro, 2018.

O rápido progresso do sequenciamento e deposição de genomas em bancos de dados provocou um grande desafio a comunidade científica, a identificação da função biológica de uma variedade de proteínas hipotéticas. Buscando impulsionar o conhecimento e o desenvolvimento de novos biocatalisadores foi realizada a prospecção da fosfatase hipotética, DSM-14977, de Oceanithermus profundus pertencente à superfamília das dehalogenases haloácidas (HAD). O objetivo deste trabalho consiste na caracterização bioquímica e estrutural da fosfatase putativa DSM-14977. O gene que codifica a DSM-14977 foi clonado e otimizado para expressão em E. coli pela empresa GenScript. A subclonagem foi realizada em três vetores de expressão: pETM30-MBP (pETM30-DSM) e pET32a (pET32a-DSM) e o pET22b (pET22b-DSM). Deste modo, a expressão solúvel destas construções foi alcançada utilizando a cepa BL21(DE3) a 37°C, 1 mM de IPTG em D.O<sub>600</sub> 0,8. Para as construções pETM30-DSM e pET32a-DSM a purificação foi realizada através da cromatografia de afinidade a níquel e para a pET22b-DSM através da cromatografia hidrofóbica seguida da afinidade à heparina. Os espectros de dicroísmo circular evidenciaram que a estrutura da proteína é rica em uma mistura de folhas-β e α-hélices, como já esperado para fosfatases pertencentes a superfamília HAD. Os estudos de estabilidade através de fluorescência indicaram que a DSM-14977 é mais estável na presença de ureia do que em cloreto de guanidina, além de ser termoestável até 85°C. Os estudos bioquímicos usando o p-nitrofenol-fosfato mostraram que a atividade fosfatásica da DSM-14977 é maior em pH neutro e em altas temperaturas. A triagem de substratos fosforilados demonstrou maior atividade para o substrato natural frutose 1,6bisfosfato e, o artificial, polifosfato-75. O modelo estrutural obtido através do I-TASSER foi semelhante ao enovelamento das fosfatases da superfamília HAD. Este modelo mostrou melhor predição funcional para uma fosfatase envolvida no metabolismo de açúcares, a CbbY de Rhodobacter sphaeroides, corroborando os dados bioquímicos obtidos. Esses eventos aliados à alta defosforilação para frutose 1,6-bifosfato podem evidenciar uma importante função da fosfatase DSM-14977 no metabolismo de açúcares da Oceanithermus profundus, sugerindo possíveis aplicações biotecnológicas para esta proteína.

**Palavras-Chave:** Biocatalisadores, superfamília das dehalogenases haloácidas (HAD), organismos extremófilos, estudos estruturais, estudos bioquímicos

#### **ABSTRACT**

PENNA, Beatriz Rosa. **PROSPECTION OF NEW BIOTECHNOLOGICAL TARGETS: BIOCHEMICAL AND STRUCTURAL CHARACTERIZATION OF HYPOTHETICAL ENZYME DSM-14977 FROM** *OCEANITHERMUS PROFUNDUS*.
Rio de Janeiro, 2018. Dissertação (Mestrado em Bioquímica). Instituto de Química,
Universidade Federal do Rio de Janeiro, 2016.

The fast progress of sequencing and deposition of genomes in databases has challenged the scientific community, identifying the biological function of a hypothetical variety proteins. In order to promote the knowledge and development of new biocatalysts, the hypothetical phosphatase DSM-14977 from Oceanithermus profundus belonging to the haloacid dehalogenase (HAD) superfamily was prospected. The goal of this work is biochemical and structural characterization of the putative phosphatase DSM-14977. The gene encoding DSM-14977 was cloned and optimized for expression in E. coli by the company GenScript. Subcloning was performed on three expression vectors: pETM30-MBP (pETM30-DSM), pET32a (pET32a-DSM) and pET22b (pET22b-DSM). Thus, soluble expression of these constructs was achieved using BL21(DE3) strain at 37 ° C, 1 mM IPTG at 0.600 O8. For the pETM30-DSM and pET32a-DSM purification was performed by nickel affinity chromatography and pET22b-DSM by hydrophobic chromatography followed by heparin affinity. Circular dichroism spectra have shown that the protein structure is rich mixture of  $\beta$ -sheet and  $\alpha$ -helix, as already expected for phosphatases belonging to HAD superfamily. Fluorescence stability studies indicated that DSM-14977 is more stable in the presence of urea than guanidine chloride, in addition to being thermostable up to 85 ° C. Biochemical studies using p-nitrophenol-phosphate showed that the phosphatase activity of DSM-14977 is higher at neutral pH and at higher temperatures. The phosphorylated substrates showed higher activity for the natural substrate fructose 1,6-bisphosphate and the artificial substrate, polyphosphate-75. The structural model obtained through I-TASSER was similar to phosphatases folding of HAD superfamily. This model showed better functional prediction for a phosphatase involved in sugars metabolism, CbbY of Rhodobacter sphaeroides, corroborating the obtained biochemical data. These events, coupled with the high dephosphorylation for fructose 1,6-bisphosphate, may demonstrate an important function of DSM-14977 phosphatase in sugars metabolism from Oceanithermus profundus, suggesting possible biotechnological applications for this protein.

**Keywords:** biocatalysts, Haloacid Dehalogenase (HAD) superfamily, extremophile organisms, structural studies, biochemical studies.

## LISTA DE ILUSTRAÇÕES

| Figura 1: Árvore filogenética dos organismos extremófilos e suas características  |
|---|
| resistentes que aparecem em pelo menos uma espécie de cada gênero   |
| Figura 2: Microscopia eletrônica de varredura da bactéria extremófila <i>Oceanithermus</i> profundus $506^{\mathrm{T}}$                                     |
| Figura 3: Árvore filogenética destacando a posição da bactéria <i>O. profundus</i> em relação aos outros tipos de cepas dentro da família <i>Thermaceae</i> |
| Figura 4: Reação geral catalisada por uma fosfatase   |
| Figura 5: Alinhamento de diferentes representantes da superfamília HAD com seuss motivos I-IVcatalíticos.   |
| Figura 6: Domínio Rossmannoid.  |
| Figura 7: Domínio núcleo catalítico observado em membros da Superfamília HAD 52   |
| Figura 8: Estruturas tridimensionais de fosfatases pertencentes à superfamília HAD mostrando a variedade estrutural presente nos diferentes domínios tampa  |
| Figura 9: Mecanismo catalítico das fosfatases da superfamília HAD   |
| Figura 10: Mapa do vetor de estocagem pUC57   |
| Figura 11: Mapa dos vetores de expressão: pETM30-MBP, pET32a e pET22b61   |
| Figura 12: Alinhamento da fosfatase putativa DSM-14977 com outras representantes da superfamília HAD  |
| Figura 13: Predição das regiões desordenadas e do conteúdo de estrutura secundária da fosfatase putativa DSM-1497780  |
| Figura 14: Esquema ilustrativo das construções da fosfatase putativa DSM-14977 utilizadas neste estudo.   |
| Figura 15: Gel de agarose 1% da amplificação do gene sintético de DSM-14977 presente no vetor de estocagem pUC57  |
| Figura 16: Gel de agarose 1% da digestão dos vetores de expressão pETM30-MBP, pET32a e pET22b.  |
| Figura 17: Digestão diagnóstica dos DNAs recombinantes a partir de colônias isoladas em meio de cultura seletivo  |
| Figura 18: Sequenciamento do DNA recombinante   |

| diferentes condições   |
|--|
| Figura 20: SDS-PAGE 12% dos testes de expressão da construção TRX-DSM-14977 em diferentes condições.                             |
| Figura 21: SDS-PAGE 15% dos testes de expressão da construção pelB-DSM-14977 (DSM-14977) em diferentes condições.                |
| Figura 22: SDS-PAGE 12% da lise celular das construções MBP-DSM-14977 e TRX-DSM-14977 após a expressão                           |
| Figura 23: SDS-PAGE 15% da extração celular da construção pelB-DSM-14977 após a expressão da proteína                            |
| Figura 24: Cromatograma da purificação e SDS-PAGE 12 % da purificação da proteína MBP-DSM-1497792                                |
| Figura 25: Cromatograma da purificação e SDS-PAGE 12 % da purificação da proteína TRX-DSM-1497793                                |
| Figura 26: Cromatograma e SDS-PAGE 15 % do primeiro passo de purificação da construção DSM-14977 (pelB-DSM-14977)                |
| Figura 27: Cromatograma e SDS-PAGE 15 % do segundo passo de purificação da construção DSM-14977 (pelB-DSM-14977)                 |
| Figura 28: Espectro de dicroísmo circular da fosfatase putativa DSM-1497797  |
| Figura 29: Estabilidade da DSM-14977 na presença de diferentes concentrações de ureia através de fluorescência.                  |
| Figura 30: Estabilidade da DSM-14977 na presença de diferentes concentrações de cloridrato de guanidina através de fluorescência |
| Figura 31: Termoestabilidade da DSM-14977 através de fluorescência   |
| Figura 32: Atividade fosfatásica da proteína DSM-14977 em diferentes condições 101   |
| Figura 33: Atividade específica da DSM-14977 na presença de diferentes substratos fosforilados a 40°C                            |
| Figura 34: Modelagem Molecular da DSM-14977 através da predição estrutural 104   |
| Figura 35: Modelagem Molecular da DSM-14977 através da predição funcional 105  |
| Figura 36: Estrutura da heparina 108   |
| Figura 37: Modelo da DSM-14977 mostrando em detalhes a cadeia lateral dos aminoácidos aromáticos                                 |

Figura 38: Mapa de superfície eletrostática do modelo estrutural da DSM-14977 ......113

## LISTA DE TABELAS

| Tabela 1- Aplicações industriais de biocatalisadores enzimáticos (ROBINSON, GURUNG et al, 2013)  | ,  |
|--|----|
| Tabela 2 : Microrganismos extremófilos e seus ambientes (adaptado de Dalmas                      |    |
| Tabela 3: Principais reagentes utilizados no desenvolvimento deste estudo                        | 62 |
| Tabela 4: Otimização dos códons para expressão em <i>E.coli</i> obtida pela empresa<br>GenScript |    |
| Tabela 5: Sequência primária da fosfatase putativa DSM-14977                                     | 78 |

#### LISTA DE ABREVIATURAS E SIGLAS

Abs - Absorvância

ATP - Adenosina trifosfafo

ADP - Adenosina difosfato

**BSA** – Albumina de soro bovino

CD - Dicroísmo circular

CTP - Citidina trifosfato

**DO** – Densidade ótica

FMN - Flavina mononucleotídeo

FT - Não retido (flow through)

**GTP** – Guanosina trifosfato

HAD- Superfamília das Dehalogenases haloácidas

**IPTG** - Isopropil-tio-β-D-galactosídeo

**I-TASSER** - Iterative Threading ASSEmbly Refinement

Lav – Lavagem

LB - Meio de cultura Luria-Bertani

LD - Tampão de carregamento LD - DNA Gel Loading Dye (6x)

**MBP** – Proteína ligadora de maltose

MUP - Metil-umbeliferil-fosfato

PCR – Reação em cadeia da polimerase

PMSF - Fluoreto de fenilmetilsulfonil

**PPi** – Pirofosfato inorgânico

**p/v** – Peso/Volume

SDS - Dodecil sulfato de sódio

SDS-PAGE - Eletroforese desnaturante em gel de poliacrilamida

**TAE** – Tampão Tris-acetato-EDTA

TCA – Ácido tricloroacético

**TEV** - Tobacco Etch Virus

**UTP** – Uridina trifosfato

v/v - Volume/Volume

## **SUMÁRIO**

| 1. INTRODUÇÃO   | 33   |
|---|------|
| 1.1 ENZIMAS E SUA IMPORTÂNCIA BIOTECNOLÓGICA                                  | 33   |
| 1.1.1 Hidrolases e seu papel na indústria                                     | 35   |
| 1.1.2 Enzimas de origem microbiana  | 36   |
| 1.2 ORGANISMOS EXTREMÓFILOS E EXTREMOZIMAS                                    | 37   |
| 1.2.1 Oceanithermus profundus   | 41   |
| 1.3 PROSPECÇÃO DE NOVAS ENZIMAS   | 43   |
| 1.4 FOSFATASES  | 45   |
| 1.5 A SUPERFAMÍLIA DAS DEHALOGENASES HALOÁCIDAS (HAD)                         | 47   |
| 1.5.1 Características Estruturais da superfamília HAD                         | 50   |
| 1.5.2 Mecanismo catalítico  | 53   |
| 1.6 ESTUDOS ESTRUTURAIS DE PROTEÍNAS E CONSTRUÇÃO DE NOVO<br>BIOCATALISADORES | S 56 |
| 2. OBJETIVOS  | 59   |
| 2.2 OBJETIVOS ESPECÍFICOS   | 59   |
| 3. METODOLOGIA  | 60   |
| 3.1 MATERIAL  | 60   |
| 3.1.1 Cepas e plasmídeos  | 60   |
| 3.1.2 Reagentes   | 62   |
| 3.2 MÉTODOS   | 63   |
| 3.2.1 Prospecção e análises computacionais da fosfatase putativa DSM-149      |      |
| Oceanithermus profundus   | 63   |

| ••••• | 3.2.2 Clonagem da fosfatase putativa DSM-14977 em vetor de estocagem               | 64         |
|-------|--|------------|
|       | 3.2.3 Subclonagem da fosfatase putativa DSM-14977 em diferentes vetores d          |            |
| expr  |  | 65         |
|       | 3.2.4 Eletroforese   | 67         |
|       | 3.2.4.1 Eletroforese em gel de agarose   | 67         |
|       | 3.2.5 Expressão da fosfatase putativa DSM-14977                                    | 68         |
|       | 3.2.6 Extração da fosfatase putativa DSM-14977 da célula de <i>E.coli</i>          | 69         |
| ••••• | 3.2.6.1 Lise celular   | 69         |
|       | 3.2.8 Purificação da fosfatase putativa DSM-14977                                  | 71         |
|       | 3.2.8.1 Purificação por afinidade a íons metálicos (IMAC)                          | 71         |
| ••••• | 3.2.8.2 Purificação por cromatografia hidrofóbica                                  | 71         |
| ••••• | 3.2.8.3 Purificação por cromatografia de afinidade à heparina                      | 72         |
| ••••• | 3.2.9 Dosagem de proteínas totais  | 72         |
| ••••• | 3.2.10 Dicroísmo circular  | 73         |
| ••••• | 3.2.11 Espectroscopia de Fluorescência intrínseca do triptofano                    | 73         |
| ••••• | 3.2.12 Dosagem da atividade enzimática para fosfatases                             | 74         |
| ••••• | 3.2.13 Triagem de substratos específicos para fosfatases                           | 75         |
|       | 3.2.14 Modelagem computacional   | 76         |
| 4.    | RESULTADOS   | 78         |
| 4.1 P | ROSPECÇÃO DA FOSFATASE PUTATIVA DSM-14977 DE <i>OCEANITHER</i><br><i>PROFUNDUS</i> | RMUS<br>78 |
| 4.2 S | UBCLONAGEM DA FOSFATASE PUTATIVA DSM-14977 DE<br>OCEANITHERMUS PROFUNDUS           | 80         |
| 4.3 E | XPRESSÃO DA ENZIMA PUTATIVA DSM-14977 NAS DIFERENTES<br>CONSTRUÇÕES                | 84         |
| 4.4 E | XTRAÇÃO CELULAR  | 90         |
| 4.5 P | URIFICAÇÃO DA CONSTRUÇÃO MBP-DSM-14977, TRX-DSM-14977 e D<br>14977                 | OSM-<br>91 |
| 4.7 E | STUDOS FUNCIONAIS DA FOSFATASE PUTATIVA DSM-14977<br>100                           |            |
|       | 4.7.1 Dosagem da atividade enzimática<br>100                                       |            |

| <b>4.8</b> I | 4.8 MODELAGEM MOLECULAR DA FOSFATASE DSM-14977<br>102 |     |  |
|--------------|---|-----|--|
| 5.           | DISCUSSÃO   | 106 |  |
| 6.           | CONCLUSÃO   | 116 |  |
| 7.           | REFERENCIAS   | 118 |  |

## 1. INTRODUÇÃO

## 1.1 ENZIMAS E SUA IMPORTÂNCIA BIOTECNOLÓGICA

As enzimas são catalisadores biológicos capazes de acelerar as reações bioquímicas, nos organismos vivos, permitindo sua sobrevivência. O conhecimento estrutural e funcional das diferentes enzimas existentes e a importância do papel biológico destas proteínas especializadas, fazem destas, um importante alvo de estudo na área biotecnológica. As enzimas podem ser extraídas de células de fontes diversas, e então, usadas para catalisar uma ampla gama de processos comercialmente importantes, como, na produção de agentes edulcorantes, na modificação de antibióticos, em produtos de limpeza, na produção de biodiesel, em dispositivos analíticos e ensaios com aplicações clínicas, forenses e ambientais (ROBINSON, 2015).

Há milhares de ano, a fermentação tem sido de grande relevância prática e econômica resultando na produção de cervejas, vinhos e pão (BUCHHOLZ e COLLINS, 2013). O primeiro documento que descreve a técnica de fabricação de cerveja data de 6.000 anos atrás (BUD, 1993). As produções eram realizadas através do uso de microrganismos de crescimento espontâneo ou de preparações de origem animal e vegetal. Entretanto, as ideias iniciais sobre as, até então, "desconhecidas" enzimas e processos biocatalíticos foram introduzidas durante os séculos XVII, XVIII, até meados do XIV. Neste breve histórico, podemos ressaltar Leeuwenhoek, em 1680 que observou microrganismos através do seu microscópio. Spallanzani, em 1783, observou a ação de proteases e, em 1793, Lavoisier e Gay-Lussac (1810) elaboraram correlações quantitativas para a fermentação alcoólica, sem dar explicações para o processo em si. Entretanto, em 1830, através da soma de outros estudos, como aqueles realizados por Schwann (1837) e Cagniard-Latour (1838) houve um acúmulo de evidências que apontaram para a natureza biológica da fermentação (SARMIENTO et al, 2015; ROBINSON, 2015; BUCHHOLZ e COLLINS, 2013).

Grandes avanços na extração, caracterização e exploração comercial de muitas enzimas só foram alcançadas no final do século XIX e início do século XX. O termo enzima só foi adotado em 1877 pelo fisiologista alemão KÜHNE (1877) que descreveu a capacidade do "fermento" para produzir álcool a partir de açúcares. E, somente, 40 anos depois, a primeira preparação enzimática para aplicação comercial foi desenvolvida por OTTO ROHMIN (1914) quando isolou a tripsina do pâncreas de animais e adicionou-a à lavagem

com detergentes para a degradação de proteínas. Na década de 1920, a estrutura de algumas enzimas foram determinadas por Cristalografia de raios-x revelando que estas eram proteínas. A partir da década de 1960, a biocatálise enzimática tornou-se uma opção industrial viável, com a produção em massa de proteases microbianas para uso de sabão em pó. Até a década de 1980, acreditava-se que todas as enzimas eram proteínas, entretanto, nesse período, descobriu-se que algumas moléculas de ácido ribonucleico (RNA) também eram capazes de exercer atividades catalíticas com papel importante na expressão gênica. Na mesma década, também foi desenvolvida a tecnologia para gerar anticorpos que possuem propriedades catalíticas em potencial, tanto como novos catalisadores industriais, quanto na área terapêutica (SARMIENTO et al, 2015; ROBINSON, 2015; BUCHHOLZ e COLLINS, 2013).

Durante a última metade do século XX, o desenvolvimento de processos de fermentação voltados para a produção de enzimas, utilizando cepas selecionadas, tornou possível a obtenção de maiores quantidades de enzimas.

Neste contexto, em 2016, o mercado global de enzimas industriais alcançou US\$ 4,7 bilhões e, em 2020, possui previsão de atingir cerca de US\$ 5,4 bilhões com uma taxa de crescimento anual composta (CAGR) de 6,03% durante os anos de 2018 a 2023. Os Estados Unidos ocupa uma grande participação na utilização das enzimas como biocatalisadores devido à disponibilidade de grandes bases de fabricação e políticas governamentais favoráveis. No entanto, a base de fabricação de quase todos os setores industriais, como o setor farmacêutico e biotecnológico, entre outros, está mudando gradualmente para a região da Ásia-Pacífico, tornando-o mais lucrativo (RESEARCH AND MARKETS, 2015). Em 2017, a América do Norte e Ásia-Pacífico constituíram os maiores mercados, com quase 62% da participação global. As três principais empresas e fornecedores são a Novozymes, DowDuPont e DSM que juntas representam quase 74% do mercado global de enzimas. Outros importantes fornecedores do mercado são a Cargill Incorporated, Dyadic International, Inc., Maps Enzymes Ltd. e Advanced Enzymes. (MORDOR INTELLIGENCE, 2017; BUSINESS WIRE, 2018).

É importante pontuar que este mercado está cada vez mais dinâmico e continuará a crescer como resultado dos avanços no setor de biotecnologia, que necessita progressivamente de processos de fabricação econômico e tecnologias mais ecológicas. Descobertas recentes de novas enzimas e desenvolvimentos em engenharia genética e purificação de proteínas possibilitaram a customização de enzimas adaptadas às condições do bioprocessos desejado, o que vem expandindo a utilização destes biocatalisadores nas indústrias. De fato, a catálise industrial evolui em função das enzimas que se tornaram ferramentas importantes em diversos

mercados industriais, como: alimentos e bebidas, rações para animais, detergentes, têxtil e de enzimas técnicas, incluindo aquelas utilizadas para a produção de biocombustíveis, couro, celulose e papel. Além disto, o uso específico de enzimas também vem aumentando devido à sua utilização em kits de diagnósticos, produtos farmacêuticos e pesquisas e desenvolvimento (SARMIENTO et al, 2015; ROBINSON, 2015).

### 1.1.1 Hidrolases e seu papel na indústria

As hidrolases são classificadas como enzimas que catalisam a quebra de ligações químicas na presença de água. Em geral, estas enzimas possuem propriedades atrativas como biocatalisadores industriais, dentre elas: independência de co-fatores, aceitação de substratos não naturais, alta estereoseletividade e capacidade de realizar reações reversas sob condições livres de água (BORNSCHEUER E KAZLAUSKAS, 2005).

Dentre as diversas enzimas, as hidrolases correspondem a 77,9% de todas as enzimas comercializadas mundialmente em 2014 (RESEARCH AND MARKETS, 2015), isto ocorre devido às diversas aplicações que tais biocatalisadores possuem (Tabela 1) ) (ROBINSON, 2015; GURUNG et al, 2013). Em grande parte destas aplicações industriais, a utilização de hidrolases resulta em um consumo reduzido de matérias-primas, energia, produtos químicos e água, como também, de produção de resíduos e emissões de gases de efeito estufa (SKALS et al, 2008; DE MARIA et al, 2007).

Tabela 1- Aplicações industriais de biocatalisadores enzimáticos (ROBINSON, 2015; GURUNG et al, 2013)

|                  | 2013)           |                   |               |                             |  |
|------------------|-----------------|-------------------|---------------|-----------------------------|--|
| Setor industrial | Enzima          | Reação            | Organismo     | Aplicação                   |  |
|                  |                 |                   |               |                             |  |
|                  | Glicose         | Conversão de      | Espécies de   | Produção de xarope de alto  |  |
| Alimentício      | icomoroco       | glicose a frutose | Streptomyces  | teor de frutose             |  |
|                  | isomerase       | gneose a muiose   | Streptomyces  | teor de frutose             |  |
|                  |                 | hidrólise do      |               |                             |  |
|                  | Collagenase;    | colágeno,         | _             | úlceras na pele, tingimento |  |
| Cosméticos       | peroxidase e    | oxidação de       | C.perfringens | de cabelo                   |  |
|                  | oxidade         | corante           |               | de cabelo                   |  |
|                  | <b>D</b> .      | Corante           |               |                             |  |
| Detergentes e    | Proteases       | Digestão de       | Espécies de   | D                           |  |
| sabão em pó      | alcalinas       | proteínas         | Bacillus      | Remoção de resíduos         |  |
| •                |                 | ^                 |               |                             |  |
|                  |                 | Remoção de 1-     |               |                             |  |
| Farmacêutico     | l-asparaginase; | asparagina        |               | Quimioterapia (leucemia);   |  |
|                  | D! .!!!         | essencial para    |               | D d                         |  |
|                  | Penicilina      | tumor;            | E. coli       | Produção de penicilinas     |  |
|                  | acilase         | Clivagem de       |               | semi-sintéticas             |  |
|                  |                 | cadeia lateral de |               |                             |  |
|                  |                 | penicilina        |               |                             |  |

| Laboratorial | Glicose oxidase | Oxidação de<br>glicose | Aspergillus<br>niger | Detecção de glicose no sangue                   |
|--------------|-----------------|------------------------|----------------------|---|
|              | Proteases       | Digestão de            | Kluyveromyc          | 0 1 ~ 1 1 .                                     |
| Laticínios   | ácidas; β-      | proteínas;             | es lactis;           | Coagulação do leite na<br>Fabricação de queijo; |
| Laticinios   | galactosidase   | Hidrólise de           | espécies de          | hidrólise de lactose em                         |
|              |                 | lactose                | Aspergillus          | leite ou soro de leite                          |

### 1.1.2 Enzimas de origem microbiana

Até a década de 1970, a maior parte da aplicação comercial de enzimas envolvia fontes animais e vegetais. Nesta época, a utilização de enzimas era observada, principalmente, na indústria de processamento de alimentos, onde as enzimas de origem animal e vegetal eram preferidas, por serem consideradas livres de problemas de toxicidade e contaminação que eram associados às enzimas de origem microbiana. No entanto, à medida que a demanda cresceu e a tecnologia de fermentação se desenvolveu, o custo competitivo das enzimas microbianas foi reconhecido, e estas tornaram-se amplamente utilizadas. As enzimas microbianas possuem diversas vantagens quando comparadas com às enzimas de fontes vegetais e animais, dentre elas, vantagens econômicas, técnicas e éticas (ROBINSON, 2015).

Além disto, as enzimas microbianas são utilizadas devido à sua ampla diversidade e propriedades únicas, como, por exemplo, a consistência, reprodutibilidade, grande variedade catalítica, possibilidade de altos rendimentos, fácil manipulação genética, fornecimento regular devido à ausência de flutuações sazonais e crescimento rápido de microorganismos em meios de cultivo baratos (GURUNG et al, 2013). As enzimas microbianas também são mais estáveis do que às suas enzimas correspondentes em vegetais e animais, além de sua produção ser mais conveniente e mais segura.

Neste contexto, as cepas bacterianas, geralmente, são mais usadas, pois oferecem atividades superiores, tendem a ter pH ótimo neutro ou alcalino e, muitas vezes, são termoestáveis. A manipulação genética e ambiental para aumentar o rendimento celular, a atividade enzimática, e a otimização da produção da enzima, pode ser empregada facilmente usando células microbianas, devido ao curto tempo de geração e os requisitos nutricionais relativamente simples e procedimentos fáceis de seleção para a característica desejada (ROBINSON, 2015; HASAN et al, 2006).

Entretanto, apesar da crescente utilização de enzimas ao longo da história, poucos microrganismos foram testados como fontes enzimáticas (HASAN et al, 2003). Portanto,

avanços são necessários para aumentar o desempenho de enzimas como biocatalisadores e estender suas áreas de aplicação. O emprego destes biocatalisadores, em processos industriais, ainda é limitado devido a sua baixa estabilidade sob condições adversas e baixa atividade ou especificidade à substratos específicos ou não-naturais. Algumas abordagens vêm sendo desenvolvidas para minimizar estes problemas, principalmente: "engenheirar" enzimas já conhecidas através do desenho racional ou mutagênese aleatória e prospectar novas enzimas ou microorganismos oriundos de ambientes específicos ou extremos (extremófilos) (SARMIENTO et al, 2015; LI et al., 2012; BÖTTCHER E BORNSCHEUER, 2010), no intuito de aumentar a gama de biocatalisadores utilizáveis e específicos para determinados bioprocessos.

## 1.2 ORGANISMOS EXTREMÓFILOS E EXTREMOZIMAS

Os organismos extremófilos são capazes de sobreviver e proliferar em ambientes com parâmetros físicos (temperatura, pressão, radiação) e geoquímicos (salinidade, pH, potencial redox) considerados extremos pelos padrões humanos. A grande maioria destes microrganismos são procariotos e, portanto, pertencem aos domínios Archaea e Bactéria, como apresentado na árvore filogenética ressaltando os diferentes gêneros e suas características extremofílicas (Figura 1) (DALMASO et al , 2015).

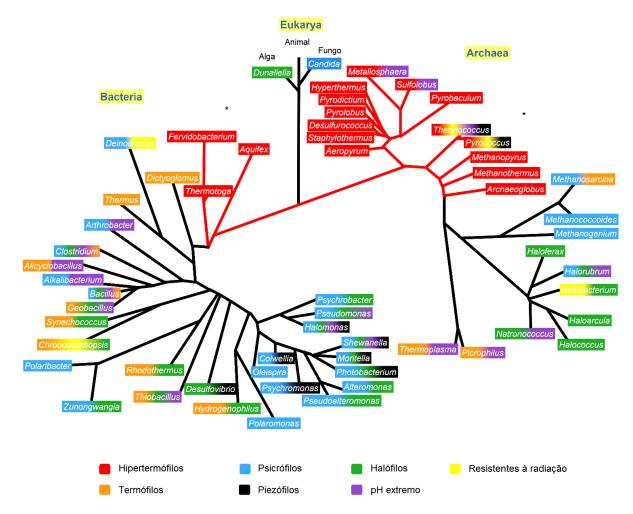


Figura 1: Árvore filogenética dos organismos extremófilos e suas características resistentes que aparecem em pelo menos uma espécie de cada gênero.

As ramificações nos reino Bacteria, Archaea e Eucaria são ilustradas, assim como, a classificação dos organismos, identificada com o código de cor, em hipertermófilos (vermelho), termófilos (laranja), psicrófilos (azul), piezófilos (preto), halófilos (verde), resistentes a pH extremos (magenta) e radiação (amarelo) (DALMASO et al, 2015). O (\*) indica o ramo filogenético de acordo com Lang et al., 2013.

Os microrganismos extremófilos são classificados de acordo com o seu crescimento em ambientes extremos (Tabela 2). Alguns destes microrganismos são conhecidos como poliextremófilos que podem sobreviver em mais de uma dessas condições extremas. Diferentes características estruturais e metabólicas são adquiridas por esses organismos, para que estes possam sobreviver nesses ambientes hostis (DUMORNÉ et al, 2017; SARMIENTO et al, 2015; DALMASO et al, 2015; LÓPEZ-LÓPEZ et al, 2014).

Tabela 2: Microrganismos extremófilos e seus ambientes (adaptado de Dalmaso, 2015).

| Classificação                             | Ambiente   |  |  |
|---|--|--|--|
| Termófilo                                 | Cresce entre 60 °C e 80 °C   |  |  |
| Hipertermófilo                            | Crescimento ótimo > 80 °C  |  |  |
| Halófilo                                  | Requer pelo menos 1M de sal para crescimento                                     |  |  |
| Acidófilo pH ótimo para o crescimento < 3 |  |  |  |
| Alcalófilo                                | pH ideal para o crescimento > 10   |  |  |
| Euripsicrófilo<br>(psicrotolerante)       | Cresce > 25 °C e também < 15 °C  |  |  |
| Estenopsicrófilo<br>(Psicrófilo)          | Cresce entre 10 °C e 20 °C   |  |  |
| Piezófilo                                 | Cresce sob alta pressão > 400 atm (40 MPa)                                       |  |  |
| Endolítico                                | Cresce dentro das rochas   |  |  |
| Hipolítico                                | Cresce em rochas e em desertos frios   |  |  |
| Radioresistente                           | Tolerância a altas doses de radiação   |  |  |
| Toxitolerante                             | Tolera altas concentrações de agentes tóxicos (por exemplo, solventes orgânicos) |  |  |
| Metalotolerante                           | Tolerância a altos níveis de metais pesados                                      |  |  |
| Xerófilo                                  | Cresce em baixa disponibilidade de água, resistente à dessecação                 |  |  |
| Oligotrófico                              | Capaz de crescer em ambientes de nutrientes escassos                             |  |  |

Neste contexto, estes organismos evoluíram e se adaptaram aos seus ambientes de origem, desenvolvendo mecanismos únicos para manter seus componentes celulares estáveis e ativos. No entanto, estes mecanismos são bastante complexos. Dentre eles, podemos citar alguns exemplos: para lidar com altas concentrações de sal, algumas espécies de extremófilos produzem quantidades aumentadas de solutos compatíveis dentro das células ou usam bombas de íons para manter o equilíbrio osmótico; na presença de baixo pH, alguns organismos usam bombas de prótons para manter o pH adequado dentro da célula; e na presença de altas ou baixas temperaturas, os organismos podem modificar a composição de sua membrana

citoplasmática como forma de proteção (DUMORNÉ et al, 2017; SARMIENTO et al, 2015; DALMASO et al., 2015; ELLEUCHE et al, 2014).

As enzimas de extremófilos, conhecidas como extremozimas, evoluíram junto com seus organismos de origem sendo altamente estáveis e funcionais nestes ambientes (JAENICKE E BÖHM, 1998). Essas adaptações correspondem a mudanças importantes na seqüência de aminoácidos, que são traduzidas em variações na estrutura, flexibilidade, carga e/ou hidrofobicidade das extremozimas. Essas mudanças não seguem um padrão ou uma tendência específica. As proteínas extremófilas apresentam variabilidade substancial em adaptações para condições físicas ou químicas extremas semelhantes (REED et al., 2013). Para aumentar sua estabilidade à temperatura, por exemplo, alguns princípios podem ser identificados, como, superfícies externas altamente carregadas, dobras rígidas mantidas por múltiplas redes de pares de íons, núcleo hidrofóbico compacto e estrutura mais densamente compactada, em parte causada pelo aumento do conteúdo de aminoácidos ácidos e básicos. Além dessas características estruturais, os sistemas de reparo de proteínas e de DNA são altamente eficazes e as membranas citoplasmáticas possuem características específicas contribuindo para permitir que estes microrganismos sobrevivam nestas condições extremas (SARMIENTO et al, 2015; ELLEUCHE et al, 2014).

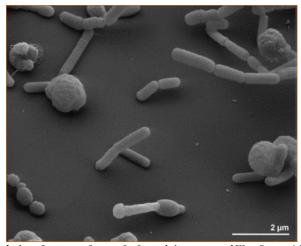
Devido as diversas adaptações e à capacidade de suportar condições extremas, as extremozimas têm sido amplamente investigadas para as aplicações industriais. As termozimas, enzimas de termófilos, estão entre as mais estudadas nas últimas décadas (DUMORNÉ et al, 2017). A possibilidade de ampla utilização destas enzimas decorre da capacidade destas serem estáveis à proteólise, presença de agentes desnaturantes, solventes orgânicos e alta salinidade. O uso dessas enzimas inclui a possibilidade de reduzir o risco de contaminação, mantendo uma baixa adesividade e maior solubilidade dos substratos. Como descrito acima, as termozimas possuem adaptações diferentes como a capacidade de manter sua configuração e função a partir de propriedades físicas e de interações eletrostáticas para manter a atividade em altas temperaturas, como também, a capacidade de aumentar a quantidade de depósitos hidrofóbicos, formando ligação dissulfeto entre dois íons com cargas opostas (DUMORNÉ et al, 2017; SARMIENTO et al, 2015; DALMASO et al , 2015; ELLEUCHE et al, 2014).

Desta forma, as enzimas de organismos extremófilos têm sido utilizadas com frequência pelas indústrias por oferecerem recursos adicionais para o desenvolvimento de bioprocessos, se tornando, assim, uma alternativa para aplicação industrial. As enzimas homólogas de organismos mesófilos ainda são as mais utilizadas até hoje, entretanto, seu uso

é limitado devido a sua baixa estabilidade em condições extremas de temperatura, pressão e pH. Em muitos casos, a solução química tradicional e os catalisadores continuam a ser a única opção viável em condições tão difíceis. Desta forma, a identificação de novas extremozimas com propriedades importantes para as condições de processamento industrial são a chave para o futuro da biocatálise possibilitando a geração de avanços no desenvolvimento de novos biocatalisadores de importância biotecnológica (DUMORNÉ et al, 2017; SARMIENTO et al, 2015; DALMASO et al , 2015; LÓPEZ-LÓPEZ et al, 2014; BÖTTCHER E BORNSCHEUER, 2010).

## 1.2.1 Oceanithermus profundus

A bactéria *Oceanithermus profundus* (Figura 2) é um microrganismo extremófilo, gram-negativa, não esporula e foi isolada após a sua descoberta em fendas hidrotermais no mar do Pacífico Leste a, aproximadamente, 2.600 metros de profundidade. A sequência genômica completa da bactéria *Oceanithermus profundus* 506<sup>T</sup> foi a primeira de um membro do gênero *Oceanithermus* e a quarta sequência da família *Thermaceae*.



**Figura 2: Microscopia eletrônica de varredura da bactéria extremófila** *Oceanithermus profundus* **506**<sup>T</sup> Esta é uma bactéria gram-negativa e foi isolada de fendas hidrotermais no mar do Pacífico Leste a aproximadamente 2.600 m de profundidade (adaptado de Klenk *et al.*, 2011).

Este microrganismo pertence a família *Thermaceae* (Figura 3) possuindo temperatura ótima de crescimento em torno de 60°C e pH ótimo em torno de 7,5 (KLENK et al., 2011; MIROSHNICHENKO et al, 2003). Fenotípicamente, o novo organismo é. em muitos aspectos, próximo a outros representantes da família *Thermaceae*, no entanto, este possui uma série de características que, provavelmente, estão conectadas com seu habitat natural, onde o

fluido vulcânico rapidamente se mistura com a água oceânica. Como outros membros dessa família, a bactéria *O. profundus* é apta a usar uma ampla gama de carboidratos (frutose, maltose, sacarose, trealose, galactose, rhamnose, xilose e outros), alguns substratos proteinaceos, ácidos orgânicos e álcoois (metanol, etanol e manitol). Pode crescer anaerobicamente às custas da redução de nitrato para nitrito, porém, é a primeira representante da família *Thermaceae* capaz de crescer litoheterotróficamente utilizando o hidrogênio molecular e também é a primeira microaerófila da família (PATI et al, 2011; FEDOSOV et al, 2008).

A família *Thermaceae* pertence ao grupo dos extremófilos termofílicos. O primeiro representante desta família, *Thermus aquaticus*, foi isolado a partir de uma fonte de água quente no parque nacional de Yellowstone (BROCK E FREEZE, 1969), mas representantes deste grupo têm sido encontrados em muitos *habitats* de térmicas naturais e artificiais (WILLIAMS E DA COSTA, 1992). O gênero *Thermus* é representado por oito espécies que crescem em elevadas temperaturas (65 a 75°C), enquanto, quatro outras espécies são classificadas como gênero *Meiothermus* que crescem em temperaturas ótimas menores (50-65°C) (NOBRE et al, 1996). A família *Thermaceae* está entre os grupos de procariotos termofílicos mais bem estudados que habitam, principalmente, fendas hidrotermais continentais de vários locais geográficos (FEDOSOV et al, 2008).

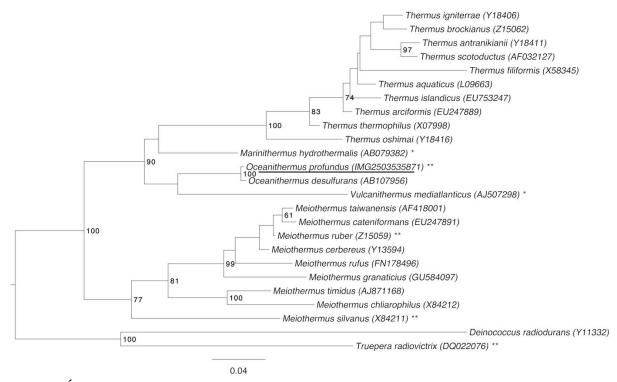


Figura 3: Árvore filogenética destacando a posição da bactéria O. profundus em relação aos outros tipos de cepas dentro da família Thermaceae.

A árvore foi construída a partir de 1.420 caracteres alinhados da sequência do gene 16S rRNA sob o critério da máxima verossimilhança. Os ramos são escalados em termos do número esperado de substituições por sítio. Os números à direita das bifurcações são valores de suporte de 1.000 repetições maiores que 60%. Linhagens contendo cepas obtidas em projetos de sequenciamento de genomas que estão registrados no GOLD(*Genomes OnLine Database*), mas permanecem inéditos são marcados com um asterisco e os genomas publicados, com dois asteriscos (adaptado de Klenk *et al.*, 2011).

A competição pelo espaço e nutrientes no ecossistema marinho constitui uma força seletiva que leva à evolução e a geração de sistemas enzimáticos múltiplos para se adaptarem aos diferentes ambientes. Muitos extremófilos marinhos são capazes de superar tais condições adversas, tornando-se uma fonte de enzimas com características especiais. Os que habitam ecossistemas de fendas hidrotermais, por exemplo, estão expostos à condições extremas, tais como, ambiente hipersalinos, altas pressões e altas temperaturas. Estes organismos e, especialmente, suas enzimas estão adaptadas a sobreviver em tais condições, sendo potencialmente, bem adaptadas para muitos processos industriais, principalmente, na biocatálise (DALMASO et al., 2015).

Esta variação nos *habitats* marinhos levou ao desenvolvimento de novas hidrolases, com especificidades e propriedades diferentes, incluindo a tolerância a algumas condições extremas utilizadas nos processos industriais (SAMUEL et al, 2012; FULZELE et al, 2011). Estudos metagenômicos revelaram que os procariotos extremófilos dos habitats marinhos são uma fonte importante de novos genes e, consequentemente, de novos bioprodutos, incluindo enzimas e outros metabólitos ativos. Entretanto, boa parte dos genes depositados nos bancos de dados não foram expressos e caracterizados, sendo anotados como genes putativos ou hipotéticos. Portanto, é importante estudar e entender esses microrganismos marinhos para poder utilizar todo o seu potencial bioquímico, ecológico, evolutivo e industrial (DALMASO et al , 2015). Assim, as extremozimas presentes nestes microrganismos podem ser empregadas *in vitro*, sendo utilizada diretamente em sua forma pura ou em extratos, ou *in vivo*, utilizada indiretamente via engenheiramento de organismos. Para isso, é necessário a realização de uma abordagem racional que conduza o foco de estudo para potenciais alvos, através da utilização de ferramentas de bioinformática, para a prospecção de novas enzimas, via homologia de sequência e de estudos de caracterização bioquímica e estrutural destas enzimas.

## 1.3 PROSPECÇÃO DE NOVAS ENZIMAS

Atualmente existem milhares de sequências gênicas depositadas nos bancos de dados disponíveis, levando a um número infinito de possibilidades, para a busca de novas enzimas (LI et al., 2012). Com o desenvolvimento e a redução dos custos dos métodos de

sequenciamento, a tendência é que esse número aumente cada vez mais. Além disso, os bancos de dados têm ampliado os seus depósitos com sequências de organelas, plasmídeos e metagenomas (KANEHISA et al., 2015). Porém, apesar do grande potencial para descobertas de novas enzimas, a maioria, permanece como "proteínas hipotéticas", sem função conhecida, pois apenas uma pequena parcela foi experimentalmente caracterizada (KUZNETSOVA et al., 2015; GALPERIN e KOONIN, 2010).

Além disso, várias ferramentas de bioinformática vêm sendo desenvolvidas devido ao grande desafio de processar toda a informação adquirida. Estas ferramentas têm sido aprimoradas gerando plataformas e servidores que buscam identificar genes de proteínas alvo aplicando uma série de algoritmos que integram bioinformática, biologia estrutural e computação com enzimologia, genética e metabolômica. Frequentemente, a estratégia aplicada é baseada em homologia de sequência através de um banco de dados, entretanto, ainda existem muitas falhas devido à falta de informação experimental. Em um novo genoma, estima-se que as técnicas automatizadas falhem em cerca de 40% na identificação da função dos genes, gerando anotações erradas, muito gerais ou hipotéticas (KUZNETSOVA et al., 2015). Para melhorar a identificação da função biológica dos genes é necessário que a caracterização experimental gere dados complementares aos estudos computacionais.

A anotação de uma proteína pela abordagem genômica, se correta, pode fornecer informações gerais quanto à bioquímica e função biológica. Muitas ferramentas computacionais de prospecção são capazes de realizar a predição de domínios catalíticos, famílias, estrutura e função biológica sobre a proteína de interesse. No caso das enzimas, ainda é possível verificar o tipo de reação catalisada, entretanto, para informações mais específicas, como o substrato preferencial e a constante cinética é necessário realizar a caracterização bioquímica. Para os estudos de extremófilos, a utilização da genômica é muito importante, visto que, a grande maioria não é cultivável ou apresenta dificuldades de crescimento devido à condições específicas de cultivos e difícil reprodução em laboratórios, ou por apresentar baixas taxas de crescimento e rendimento, inviabilizando a produção e identificação da atividade catalítica da enzima. Deste modo, a identificação de genes através da abordagem genômica e de metagenoma de extremófilos é uma maneira eficiente de buscar por enzimas de interesse biotecnológico (KUZNETSOVA et al., 2015; LORENZ e ECK, 2005).

Neste contexto, como o cultivo de organismos extremófilos é difícil, o mais utilizado é a expressão heteróloga da proteína prospectada. Isto é possível através do uso de sistemas recombinantes nos quais o gene que codifica a enzima desejada é superexpresso em

hospedeiro mesófilo, como a *Escherichia coli* (HANS et al, 2003; STEWART, 2001;OKUTA et al., 1998). A escolha do hospedeiro varia de acordo com a proteína de interesse e a sua aplicação biotecnológica, entretanto, a expressão heteróloga de proteínas recombinantes permite sua produção em grandes quantidades para uso comercial (ZERBS et al, 2009). Os sistemas de expressão são otimizados e sintetizam cerca de 10% ou mais da proteína recombinante quando comparados ao total de proteínas de uma célula, sendo assim, uma das principais maneiras de se obter grandes concentrações da proteína de interesse (OGAWA E SHIMIZU, 2002; SCHMIDAL. et al, 2001; HANNIG e MAKRIDES, 1998).

#### 1.4 FOSFATASES

As fosfatases pertencem a ampla família de hidrolases, composta por enzimas de interesse biotecnológico, dentre elas, as lipases, esterases, peptidases, entre outras. As fosfatases são enzimas que catalisam a hidrólise de uma variedade de fosfomonoéster e, assim, degradam compostos orgânicos complexos de fósforo liberando fosfato (Figura 4).

Figura 4: Reação geral catalisada por uma fosfatase.

As fosfatases catalisam a hidrólise de diferentes fosfomonoésteres, removendo o fosfato do substrato. A água atua na reação e se divide com o grupo -OH ligado ao íon fosfato, e o H + protona o grupo hidroxil do outro produto. O resultado da reação é a catálise do fosfomonoéster e a produção do íon fosfato e uma molécula com um grupo hidroxil livre (LI et al, 2013).

Estas enzimas foram convencionalmente classificadas de acordo com o seu pH ótimo, em fosfatases alcalinas (EC 3.1.3.1) e ácidas (EC 3.1.3.2), o que acaba definindo sua utilização biotecnológica. Esta classificação foi revista e dividiu-se as fosfatases em cinco famílias distintas: fosfatases alcalinas, fosfatases ácidas roxas, fosfatases ácidas de baixa massa molecular, fosfatases ácidas de alta massa molecular e proteínas fosfatases.

As fosfatases são amplamente distribuídas na natureza e, geralmente, apresentam uma vasta especificidade à substratos e, portanto, são capazes de hidrolisar uma grande variedade de proteínas, açúcares, poliálcoois e nucleotídeos (GUIMARAES et al., 2007). As subunidades ou domínios catalíticos desta enzima são muito diversos possuindo diferentes mecanismos catalíticos, devido à sua evolução em famílias separadas que são estrutural e

mecanicamente distintas (LI et al, 2013). Estas enzimas também estão envolvidas em processos de translocação e regulação no metabolismo de açúcares (FAMUREWA E OLUTIOLA, 1994). As fosfatases são empregadas, principalmente, na indústria de alimentos, fármacos e cosméticos, além de serem importantes em ensaios bioquímicos. Desta forma, as fosfatases são um importante alvo biotecnológico para prospecção de novos biocatalisadores enzimáticos.

Na indústria de alimentos, a fitase, um tipo de fosfatase ácida é utilizada na hidrólise do ácido fítico, molécula encontrada em legumes, sementes oleaginosas e grãos que compõem a dieta de animais. A adição de fitase à ração, nas dietas de animais monogástricos, libera fosfato inorgânico e torna possível que porcos, aves e peixes absorvam esses nutrientes, aumentando a digestibilidade e os levando a uma maior eficiência na utilização destes alimentos. Além disso, as fitases podem degradar componentes nos alimentos que não seriam aproveitados pelo organismo (SPOHNER et al., 2015). Um estudo com a fitase de um termófilo, S. thermphile, mostrou que seu uso em ingredientes alimentares, como, farinha de trigo, bolo de óleo de sésamo e leite de soja levou à redução do ácido fítico com liberação do fosfato inorgânico, aumentando o conteúdo de açúcares redutores e de proteína solúvel quando comparada ao pão controle, apresentando, atividade ótima a 60°C. A adição de αamilase e fitase à massa melhorou a qualidade e as propriedades do pão em comparação com o pão de controle preparado com enzimas comerciais (SINGH e SATYANARAYANA, 2006a, b, 2008a, b). Além da aplicação das fitases como aditivos alimentares e para alimentação animal, elas também podem ser usadas em produtos farmacêuticos, alterações do solo, em síntese de peroxidase e no crescimento de plantas (SINGH E SATYANARAYANA, 2013).

Ainda sobre a indústria alimentícia, as fosfatases podem ser utilizadas como intensificadores de sabor, através da utilização de nucleosídeos 5'-monofosfatos (NMPs) e seus derivados. Esses compostos são produzidos através da atividade fosfotransferásica de algumas fosfatases ácidas, sendo também, importantes na indústria farmacêutica, onde são usados como componentes farmacológicos ativos (MÉDICI et al., 2014). Outra aplicação de fosfatases é na indústria de laticínios, como um marcador de pasteurização do leite de vaca que pode ser testado via mudança de cor do substrato ρ-nitrofenilfosfato em solução tamponada (*Aschaffenburg Mullen Test*). A desnaturação das fosfatases presentes no leite pasteurizado ocorre devido as altas temperaturas do processo, e desta forma, não hidrolisa o

ρ-nitrofenilfosfato, enquanto o leite cru normalmente produz a cor amarela (ASCHAFFENBURG e MULLEN, 1949).

As fosfatases alcalinas e ácidas tornaram-se uma ferramenta muito útil em biologia molecular, já que o DNA possui grupos fosfato na extremidade 5'. Assim, a remoção destes fosfatos evita a religação do DNA (a extremidade 5' se liga à extremidade 3'), mantendo assim, as extremidades da molécula de DNA lineares até o próximo passo do processo. Além disso, a remoção dos grupos fosfato permite a radiomarcação (substituição por grupos de fosfato radioativos) para acompanhar a presença do DNA marcado. Um uso importante da fosfatase alcalina é como marcação para imunoensaios enzimáticos. As fosfatases alcalinas são utilizadas em ensaios imunoabsorventes ligados a enzimas (ELISA), sistemas de sondagem não-isotópica, *blotting* e sequenciamento (SURESH E ARUNAVA DAS, 2014; SINGH E SATYANARAYANA, 2013). As fosfatases também são encontradas em kits de diagnósticos laboratoriais como a pirofosfatase de *T. thermophilus HB8* que pode ser citada como exemplo de aplicação comercial destas enzimas (THAM et al., 2010). As fosfatases ácidas também podem ser utilizadas na detecção e biorremediação de solos e águas que contenham pesticidas (SRIVASTAVA e ANAND, 2014).

## 1.5 A SUPERFAMÍLIA DAS DEHALOGENASES HALOÁCIDAS (HAD)

As fosfatases açúcar-fosfato possuem uma grande diversidade em estruturas proteicas, mecanismos catalíticos e especificidade aos substratos. No entanto, muitas delas estão relacionadas à evolução das enzimas conhecidas como dehalogenases haloácidas (HAD). A superfamília HAD, recebe este nome em homenagem ao primeiro membro desta superfamília, a enzima 2-dehalogenase haloácida de *Pseudomonas* sp. YL, que foi caracterizado estruturalmente (PFEIFFER et al, 2014). Esta superfamília inclui enzimas que catalisam reações de transferência de grupos carbono ou fosfato em uma gama de substratos, composta por mais de 500.000 membros (InterPro IPR023214; PARK et al, 2015) presentes em todos os reinos (BURROUGHS et al., 2006). Apesar da enzima que originou seu nome, cerca de 79% das proteínas desta superfamília, que foram classificadas e estudadas bioquímicamente são compostas por fosfatases e 20% por ATPases. Além disto, as HAD incluem, ainda, as fosfonatases, fosfoesterases e fosfomutases (KUZNETSOVA et al., 2015). Estas enzimas estão envolvidas na transferência do grupamento fosfato e necessitam de um metal bivalente como co-fator, geralmente, o Mg<sup>2+</sup> para catálise. Isto não ocorre com as dehalogenases

haloácidas, nas quais a água atua como nucleófilo atacando a carbonila do aspartato presente no sítio ativo durante a reação de hidrólise parcial (BURROUGHS et al., 2006).

A superfamília HAD representa uma das maiores superfamílias de enzimas, composta por 516.224 sequências depositadas em bancos de dados (InterPro IPR023214) e por 33 grandes famílias. É previsto que a maior parte dos genomas contenha proteínas desta superfamília, incluindo 28 genes em *E. coli*, 39 genes em *S. cerevisiae* e 296 genes em humanos (KUZNETSOVA et al., 2015).

As fosfatases HAD são responsáveis pela maioria das reações metabólicas de hidrólise de fosfomonoéster existentes em todos os reinos da vida, podendo estar envolvida no metabolismo primário e secundário, regulação de *pools* metabólicos, captação de nutrientes e serviço de limpeza celular (KUZNETSOVA et al., 2015). Em humanos, já é descrito que essas enzimas estão envolvidas em várias doenças, como, o câncer, desordens cardiovasculares, distúrbios metabólicos e neurais. Algumas dessas doenças estão relacionadas à mutações nas HADs que regulam o domínio C-terminal da RNA polimerase II. Esse domínio, chamado de CTD, é composto por diversos resíduos de aminoácidos, que ao serem fosforilados ou defosforilados, regulam a expressão gênica através do processamento do mRNA. Outras mutações nas HAD podem levar a doenças relacionadas a alterações nas vias metabólicas, responsáveis pela biossíntese de colesterol, ácido siálico, serina e vitamina B6. Outro exemplo são as proteínas Eya, que são fosfatases envolvidas na formação de muitos tecidos e órgãos. Mutações nestas proteínas podem levar a uma variedade de distúrbios congênitos como: síndromes de displasia branquio-oto-renal (Eya1), cardiomiopatia dilatada-1J e surdez autossômica dominante-10 (Eya4) (SEIFRIED et al, 2013).

A grande diversidade dessas enzimas dificulta muito a identificação de suas funções biológicas específicas, seja por ferramentas de bioinformática, como também, por ensaios enzimáticos devido ao grande número de possíveis substratos a serem experimentados (TERAWAKA et al, 2016; SEIFRIED et al, 2013; BURROUGHS et al, 2006).

Embora a identidade sequencial entre as fosfatases HAD seja muito baixa (< 15%), os membros desta superfamília podem ser identificados através do alinhamento de sequências de aminoácidos baseados na presença de quatro motivos conservados (TERAWAKA et al, 2016; SEIFRIED et al, 2013).

No N-terminal é encontrado o motivo I que contém o resíduo nucleofílico, ácido aspártico (D) e a sequência consenso hhhDxDx(T/V)(L/V)h (onde, h representa um resíduo hidrofóbico e x indica qualquer aminoácido) . Neste motivo, o grupo carboxilato do ácido aspártico nucleofílico e o esqueleto carbonil do segundo Asp (Asp+2) coordenam o

Mg<sup>2+</sup>essencial no sítio ativo para a catálise enzimática . O motivo II formado pela seqüência hhhhhh(S/T) contém um resíduo conservado de serina (Ser) ou treonina (Thr) que ajuda a orientar o substrato para o ataque nucleofílico, formando uma ligação de hidrogênio com o seu grupo fosforil de transferência. O motivo III é menos conservado quando comparado aos outros motivos, contendo um resíduo de lisina (K), que é espaçado por 18-30 resíduos do motivo IV. A função do resíduo de lisina, presente no motivo III, é estabilizar a carga negativa do intermediário da reação junto com os resíduos serina/treonina presentes no motivo II (Figura 4). O motivo IV exibe, tipicamente, a sequência (G/S)(D/S)x3-4(D/E)hhhh, mas uma assinatura DD ao invés de Dx3-4D também pode ser observada (Figura 5A). Assim como os resíduos de ácido aspártico presentes no motivo I, os resíduos ácidos Asp ou Glu do motivo IV também estão envolvidos na coordenação do Mg<sup>2+</sup>. Os motivos I-IV estão dispostos espacialmente em torno de uma única cavidade de ligação na extremidade C-terminal da fita central que forma o sítio ativo das fosfatases HAD (Figura 5B e C) (KUZNETSOVA et al., 2015; SEIFRIED et al, 2013).

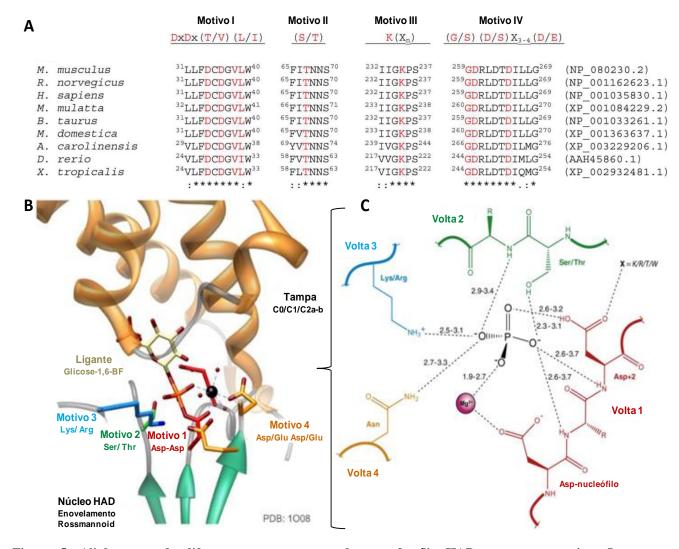


Figura 5: Alinhamento de diferentes representantes da superfamília HAD com seuss motivos I-IVcatalíticos.

(A) Alinhamento de 9 representantes da superfamília HAD onde é possível observar os 4 motivos de sequência conservados destacados pelo asterisco. As letras em vermelho representam os resíduos conservados envolvidos na atividade catalítica.(B) Estrutura cristalográfica de β-fosfoglicomutase composta pelo domínio núcleo catalítico (verde) e o domínio tampa C1 (dourado) fornecendo elementos para a ligação do substrato. Os resíduos dos motivos conservados pertencente a esta superfamília estão ressaltados.(C) Esquema composto pelos resíduos que compõem os motivos I-IV na superfamília HAD envolvidos na formação do sítio catalítico responsável por catalisar a transferência do grupo fosforil mostrando a contribuição de cada aminoácido presente em cada motivo para ligação ao grupo fosforil, ao íon  $Mg^{2+}$  e às moléculas de água que também são importantes na catálise. A volta 4 fornece resíduos ácidos (Asp/Glu) capaz de ajudar a coordenar o metal. Os comprimentos das ligações de hidrogênio em Å são retratadas como linhas tracejadas (respectivamente, SEIFRIED et al., 2014;ALLEN e DUNAWAY-MARIANO, 2009; SEIFRIED et al., 2013).

#### 1.5.1 Características Estruturais da superfamília HAD

Os estudos estruturais da superfamília das fosfatases HAD revelaram que todos os membros pertencentes à esta superfamília compartilham o mesmo arranjo estrutural do domínio núcleo catalítico, caracterizado por um sanduíche  $\alpha/\beta$  formado por três camadas compostas por unidades  $\alpha/\beta$  repetidas formando tipicamente um enovelamento de Rossmann

modificado. A folha-β central é paralela, composta por pelo menos cinco fitas-β na ordem "54123", referidas como S1-S5, orientando, assim, quatro *loops* que contêm os resíduos do núcleo catalítico envolvidos no posicionamento do substrato, do cofator e dos grupos catalíticos. O típico enovelamento Rossmannoid da superfamília HAD possui três domínios estruturais adicionais que permitem que a enzima adote estados conformacionais distintos e que contribuem para a especificidade ao substrato, a saber, os domínios estruturais *squiggle*, *flap* e tampa (Figura 6).

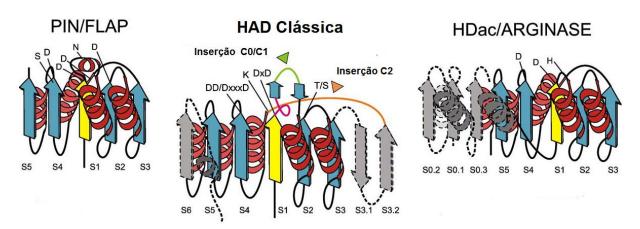


Figura 6: Domínio Rossmannoid.

Diagrama da topologia de domínios representativos das principais divisões do enovelamento de Rossmann com resíduos ácidos catalíticos. Os fios são mostrados como setas com a ponta na extremidade C-terminal e são rotulados de S1 a S6 no clássico HAD. O fio inicial contendo o resíduo catalítico D é representado em amarelo; outras vertentes principais conservadas em todos os membros do domínio estão em azul; elementos não conservados que podem estar ausentes do estado ancestral de um domínio estão em cinza. O ponto de inserção C1 da HAD é representado como uma linha verde onde se encontra o *flap* e o ponto de inserção da tampa C2 é representado com uma linha laranja. As linhas quebradas indicam elementos de estruturas secundárias não presentes em todos os membros que compartilham o domínio. O loop rosa no domínio HAD representa o *squiggle* conservado. Os resíduos conservados em todos os membros de um domínio particular, incluindo o resíduo catalítico inicial D, são mostrados.

O motivo estrutural, *squiggle*, é observado imediatamente após a fita β1, é um motivo formado por 6 resíduos de aminoácidos compreendendo, praticamente, uma volta de hélice completa. O segundo motivo, o *flap*, está localizado logo após ao *squiggle* sendo composto por uma volta em grampo-β formada por duas fitas que se projetam a partir do domínio núcleo catalítico. O *squiggle* helicoidal pode alternar entre conformações mais rígidas ou flexíveis gerando um movimento no *flap*. Uma vez que o grampo-β do *flap* está localizado próximo ao sítio ativo, este movimento, pode cobrir parcialmente a cavidade catalítica (Figura 7). As mudanças conformacionais exercidas pelos motivos *squiggle* e *flap* parecem constituir a maquinaria mínima necessária para exclusão e acesso de solventes ao sítio ativo, mostrando

que estes domínios exercem papéis essenciais na catálise desta superfamília (KUZNETSOVA et al., 2015; SEIFRIED et al, 2013; BURROUGHS et al., 2006).

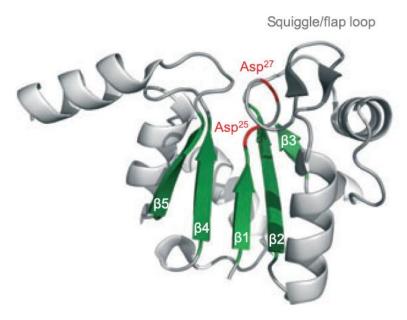


Figura 7: Domínio núcleo catalítico observado em membros da Superfamília HAD. O enovelamento de Rossmann do núcleo catalítico da fosfatase piridoxal fosfato humana Pdxp (PDB: 2P69, modificado). A estrutura é reduzida aos resíduos do núcleo catalítico no qual é observado o enovelamento formado pela repetição de unidades  $\alpha/\beta$ . A folha-β central consiste de cinco fitas β dispostas sequencialmente na ordem '54123' (verde). Os dois ácidos aspárticos do motivo I presentes na superfamília HAD são retratados em vermelho, seguidos dos domínios *squiggle* e flap no primeiro *loop* (cinza escuro) (SEIFRIED et al, 2013).

Os domínios em tampa são as inserções móveis adicionais mais relevantes observadas na superfamília HAD e formam uma maior proteção para a cavidade catalítica quando comparada aos simples elementos do *flap*. Além disso, estes domínios são determinantes para a seletividade ao substrato, além de estarem envolvidos, muitas vezes, na oligomerização de fosfatases. Apesar da sua grande diversidade, as tampas podem ser classificadas em três grandes categorias (C0, C1 e C2). Os representantes menores e estruturalmente mais simples, conhecidos como C0, possuem apenas pequenas inserções, em quaisquer um dos dois pontos de inserção da tampa, podendo consistir em um *loop* ou fita- $\beta$ . As tampas C1 são definidas como inserções, na maioria das vezes,  $\alpha$ -helicoidais de complexidades variáveis que ocorrem no meio do grampo- $\beta$  do domínio *flap* e dobram-se em uma unidade estrutural distinta do núcleo catalítico; são suficientemente grandes para selar completamente o sítio ativo da enzima no estado fechado. As tampas C2 são inserções altamente diversificadas, geralmente, compostas de domínios  $\alpha$ + $\beta$ , com uma folha  $\beta$  central no núcleo de pelo menos três fitas, às quais outros elementos de estrutura secundária simples podem ser adicionados. As estruturas do tipo C0 e C1 são inseridas entre as duas fitas  $\beta$  do *flap*, enquanto as tampas C2 são

incorporadas no ligante, imediatamente, após a folha-β3 do domínio principal. A maioria dos representantes da superfamília HAD tem uma tampa C1 ou C2, embora, em poucos casos, as proteínas possam apresentar simultaneamente tampas C1 e C2 (Figura 8) (KUZNETSOVA et al., 2015; SEIFRIED et al., 2013; ALLEN e DUNAWAY-MARIANO, 2009; BURROUGHS et al., 2006).

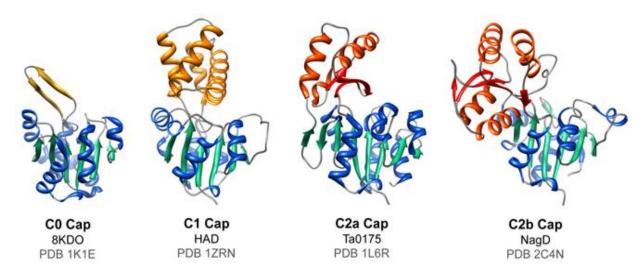


Figura 8: Estruturas tridimensionais de fosfatases pertencentes à superfamília HAD mostrando a variedade estrutural presente nos diferentes domínios em tampa.

Estrutura cristalográfica de quatro representantes da superfamília HAD. O domínio núcleo catalítico (azul/verde) e os domínios em tampa C0 (amarelo), C1 (dourado) e C2 (vermelho/laranja). Estas estruturas demonstram a diversidade encontrada nesta superfamília de enzimas (SEIFRIED et al, 2013).

Em comparação aos representantes da superfamília HAD com tampa C0, os representantes com tampa C1 e C2 são bastante diferenciados, pelo fato, de suas inserções se enovelarem em domínios consideráveis capazes de encapsular o sítio de transferência do grupo fosforil e o grupo liberado do substrato. Este avanço traz um novo nível de sofisticação às enzimas HAD, pois o domínio em tampa aumenta a superfície de contato para a seleção do substrato sendo um meio eficaz para dessolvatar o sítio catalítico. Além disso, esta separação do domínio núcleo catalítico, onde ocorre a transferência do grupo fosforil, e do domínio tampa, que confere especificidade ao substrato, é, possivelmente, a base estrutural para a adaptação destas enzimas a vasta gama de ésteres de fosfato possíveis mantendo a eficiência do catalisador (SEIFRIED et al, 2013; ALLEN e DUNAWAY-MARIANO, 2009).

#### 1.5.2 Mecanismo catalítico

Como mencionado anteriormente, a sequência do motivo I, DxD, está presente no fim da fita β1. O grupo carboxilato do primeiro ácido aspártico e a carbonila do esqueleto carbônico do segundo ácido aspártico coordenam o co-fator Mg<sup>2+</sup>.

O mecanismo catalítico característico das fosfatases HAD é realizado em dois passos. No primeiro passo, o aspartato do motivo I inicia um ataque nucleofílico no grupo fosforil do substrato, resultando na formação de um intermediário fosfoaspartil e o deslocamento do grupo de saída do substrato. No passo subsequente, uma molécula de água exerce um ataque nucleofílico no intermediário fosfoaspartil, liberando fosfato livre e regenerando o aspartato catalítico (Figura 9). As fosfatases HAD contêm um segundo resíduo de ácido aspártico que funciona como um ácido/base geral para protonar o grupo de saída na primeira reação parcial e desprotonar o nucleófilo da água na segunda reação parcial. A vantagem química de um aspartato catalítico é a sua versatilidade: constitui um bom nucleófilo e um bom grupo de saída, e pode operar em pH baixo e alto. Todas as fosfoaspartil transferases utilizam Mg<sup>2+</sup> como um cofator obrigatório. O Mg<sup>2+</sup> ajuda no posicionamento correto do grupo fosforil do substrato em relação ao nucleófilo ácido aspártico e estabiliza eletrostaticamente a aproximação necessária do nucleófilo aniônico ao substrato fosfomonoéster dianiônico. Além disso, o Mg<sup>2+</sup> fornece a neutralização de carga do estado de transição. Juntos, os resíduos catalíticos e o cofator estabilizam o estado de transição de ambas as reações parciais (POSSIK et al, 2017, KUZNETSOVA et al, 2015; PANDYA et al, 2014; SEIFRIED et al, 2013; ALLEN e DUNAWAY-MARIANO, 2009).

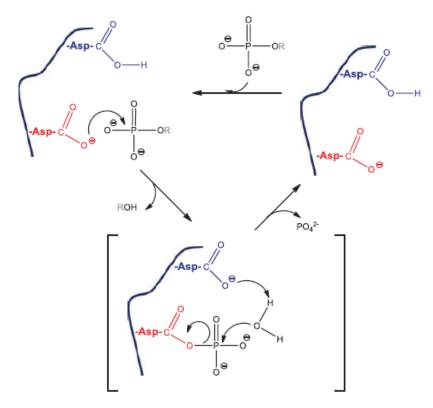


Figura 9: Mecanismo catalítico das fosfatases da superfamília HAD. Essas enzimas utilizam os resíduos de aspartato localizados no sítio ativo para a realização de dois passos da catálise. No primeiro passo ocorre o ataque nucleofílico do aspartato ao substrato, resultando na liberação do grupo de saída e formação do intermediário fosfoaspartil-enzima, e no segundo passo, a regeneração da enzima, via catálise do intermediário, com a liberação do fosfato livre (SEIFRIED et al, 2013).

Em membros da superfamília HAD que são fosfatases e fosfomutases, o segundo resíduo ácido atua como um ácido-base geral. Nas ATPases, a ocorrência de uma treonina nesta posição permite uma taxa reduzida de hidrólise de fosfato-aspartil, a qual permite o atraso necessário para a consequente mudança conformacional. Nas fosfanatases, existe uma alanina no lugar do segundo ácido aspártico, consistente com o papel exclusivo desempenhado pelo intermediário enamina como um ácido-base geral catalisador de hidrólise de fosfato-aspartil por estas proteínas (KUZNETSOVA et al, 2015; SEIFRIED et al, 2013; BURROUGHS et al., 2006).

Em relação a especificidade de substratos das enzimas pertencentes a superfamília HAD pouco é conhecido. Estudos mostraram que resíduos de aminoácidos dos domínios tampa podem ser importantes na escolha do substrato. Além disso, a sobreposição de substratos entre os membros desta superfamília, também dificulta a identificação dos mecanismos de especificidade. Incessantes estudos para a descoberta de novas funções e papéis no metabolismo secundário são desafios a serem superados para o conhecimento mais profundo a cerca dessas enzimas. No entanto, a reação catalisada e a especificidade ao substrato são difíceis de prever necessitando de sua determinação de forma empírica

(KUZNETSOVA et al, 2006). À medida que novos dados estruturais e mecanicistas são descobertos, os princípios unificadores para a atribuição funcional de membros de função desconhecida serão refinados. Ainda são necessárias informações adicionais sobre o "ajuste" estrutural necessário entre a tampa e o núcleo catalítico e as variações na estabilidade termodinâmica das diferentes montagens de tampas para explorar plenamente esse conhecimento na compreensão da evolução enzimática e nos esforços no campo da engenharia proteica (POSSIK et al, 2017, KUZNETSOVA et al, 2015; PANDYA et al, 2014; SEIFRIED et al, 2013; ALLEN e DUNAWAY-MARIANO, 2009).

O estudo das fosfatases da superfamília HAD apresenta importância devido a participação dessas enzimas em doenças como o câncer, doenças cardiovasculares e desordens metabólicas e neurológicas, as funções fisiológicas de muitas destas enzimas ainda são muito pouco conhecidas (SEIFRIED, 2013). Estudos bioquímicos, estruturais e funcionais combinados com as ferramentas de bioinformática são necessários para entender as bases moleculares do reconhecimento e catálise dessas enzimas. O conhecimento profundo destas enzimas poderá expandir a sua aplicabilidade biotecnológica em diversas áreas de estudo.

## 1.6 ESTUDOS ESTRUTURAIS DE PROTEÍNAS E CONSTRUÇÃO DE NOVOS BIOCATALISADORES

O progresso na área de tecnologia do DNA recombinante, genômica e proteômica têm contribuído bastante para o desenvolvimento de novos biocatalisadores e processos biocatalíticos. A engenharia de proteínas com potencial biotecnológico, a fim de desenvolver enzimas mais especializadas e com novas propriedades, cresceu bastante nas últimas décadas. Desta forma, as enzimas podem ser manipuladas, de acordo com suas informações estruturais, aumentando sua estabilidade e versatilidade frente ao conjunto de substratos que podem atuar e, ao mesmo tempo, possuir elevada especificidade para as reações que catalisam (ROBINSON, 2015; GUPTA et al, 2004).

Portanto, a caracterização e, principalmente, a determinação estrutural e dinâmica de proteínas são de grande interesse, possibilitando o desenho racional de novos biocatalisadores.

O enovelamento de proteínas é, na maioria dos casos, o passo crucial para a transformação da informação gênica em sua função biológica específica. Os diversos tipos de enovelamento e as pequenas diferenças nos aminoácidos que compõem uma proteína são essenciais para características como atividade catalítica, especificidade de substrato, oligomerização e estabilidade em diferentes condições reacionais. A relação entre a sequência

de aminoácidos e as propriedades de estrutura e dinâmica da conformação nativa de proteínas representam alguns dos problemas centrais da bioquímica e biotecnologia (JONAS, 2002). A completa compreensão dos mecanismos pelos quais o enovelamento ocorre, portanto, pode representar um grande avanço para o estudo de proteínas (DOBSON, 2012; LINDORFF-LARSEN et al., 2005).

O mecanismo do enovelamento da cadeia polipeptídica em uma estrutura tridimensional específica é estudado em vários trabalhos. Os estados nativos de proteínas quase sempre correspondem às estruturas que são mais termodinamicamente favoráveis em um determinado conjunto de condições (como temperatura, pH, força iônica, dentre outras). A estrutura primária contém a informação que determina a conformação nativa da proteína, sendo um processo espontâneo e reversível. Uma importante consequência dessa reversibilidade é a possibilidade de se estudar as reações de enovelamento em diversas condições experimentais ou fisiológicas (LINDORFF-LARSEN et al., 2005). Além disso, mutações pontuais nos aminoácidos próximos ao sítio ativo já se mostraram importantes para a modificação do substrato de maior afinidade ou para a catálise de novas reações enzimáticas. Como já foram utilizados em enzimas da superfamília HAD, onde frequentemente o domínio tampa é identificado como um dos principais responsáveis pela especificidade de substrato. Esses estudos mostraram que a mutação de resíduos específicos ou a substituição integral do domínio são capazes de gerar alterações na atividade enzimática (SEIFRIED et al., 2014). Dessa forma, as enzimas da superfamília HAD possuem um grande potencial para o desenho racional, já que foram adaptadas evolutivamente a vários substratos e funções biológicas em diversos organismos diferentes.

Neste contexto, o estudo estrutural é de grande importância para a construção de novos biocatalisadores. O desenvolvimento de técnicas espectroscópicas, como a cristalografia de raios-x e a ressonância magnética nuclear (RMN) possibilitou a determinação estrutural de proteínas a nível atômico. A cristalografia de raios-X é um dos métodos mais significativos para a determinação de estruturas tridimensionais de proteínas, sendo responsável pela maioria das estruturas proteicas resolvidas e depositadas (ERICSSON et al, 2006, OWENS, 2016). Já a ressonância magnética nuclear (RMN) é o único método que permite a determinação de estruturas tridimensionais e dinâmica de proteínas em solução. Esta técnica ainda permite o estudo detalhado em nível atômico levando ao entendimento de suas funções biológicas e propriedades relativas aos sítios de ligação e interação de biomoléculas (POULSEN, 2002).

Além disso, outras técnicas espectroscópicas como o dicroísmo circular (CD) e a fluorescência são utilizadas para realizar a caracterização estrutural de proteínas. Apesar de não oferecer tanta informação estrutural também são utilizadas para monitorar e entender fenômenos globais de proteínas. Estas técnicas permitem adquirir dados sobre o conteúdo estrutural secundário e o comportamento terciário de proteínas, respectivamente, possibilitando o monitoramento de importantes alterações estruturais, como o desenovelamento e a agregação proteica (CORRÊA et al, 2009; CRONEY et al, 2001).

Desta forma, a prospecção de novas enzimas com potencial biotecnológico é uma estratégia promissora e cada vez mais utilizada para aumentar o conhecimento e aplicação biotecnológica de novas proteínas. Com a intenção de encontrar novos biocatalisadores em potencial foi realizada uma busca por novas hidrolases utilizando ferramentas de bioinformática. Em meio às sequências obtidas, a fosfatase putativa DSM-14977 de *Oceanithermus profundus* foi selecionada para os estudos de caracterização estrutural e bioquímica devido à sua baixa massa molecular e nenhuma informação funcional, sendo descrita na literatura com uma possível fosfatase hipotética. O desenvolvimento da caracterização estrutural e bioquímica pode gerar o conhecimento necessário a cerca da proteína alvo para sua potencial aplicação como um novo biocatalisador, mais estável e ativo, com características peculiares para determinados bioprocessos.

#### 2. OBJETIVOS

## 2.1 OBJETIVO GERAL

O principal objetivo deste trabalho consiste na caracterização estrutural e bioquímica da enzima putativa DSM-14977 de *Oceanithermus profundus* a fim de promover avanços no conhecimento da família das dealogenases haloácidas e sua possível utilização como biocatalisador.

## 2.2 OBJETIVOS ESPECÍFICOS

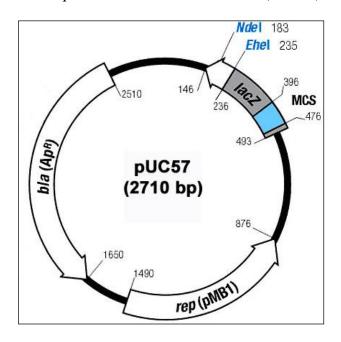
- Prospectar uma nova enzima através de ferramentas de bioinformática;
- Realizar a subclonagem do gene que codifica a enzima putativa DSM-14977 de *Oceanithermus profundus*, inserindo-o nos plasmídeos de expressão pETM30-MBP, pET32a e pET22b;
- Expressar e purificar a proteína, presente em cada uma das construções;
- Monitorar o conteúdo estrutural secundário da proteína DSM-14977 através de dicroísmo circular;
- Estudar as características bioquímicas da enzima DSM-14977 e realizar a triagem de substratos fosfatásicos através de ensaios enzimáticos;
- Estudar as características estruturais e a estabilidade da enzima DSM-14977 na presença de agentes desnaturantes físicos e químicos através de fluorescência intrínseca do triptofano;
- Propor uma função biológica para a enzima DSM-14977 através dos estudos de caracterização bioquímica e modelagem molecular

#### 3. METODOLOGIA

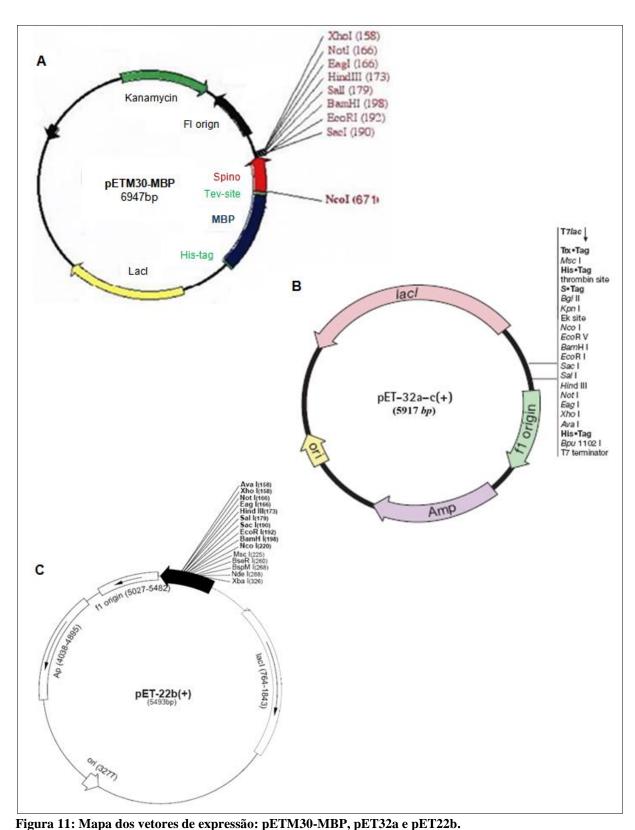
#### 3.1 MATERIAL

## 3.1.1 Cepas e plasmídeos

O gene que codifica a proteína DSM-14977 foi obtido através da empresa GenScript que realizou a síntese, a otimização para códons raros específicos para expressão em *Escherichia coli* e a clonagem no vetor de estocagem pUC57 (Figura 10). Os vetores de expressão pETM30-MBP, pET32a e pET22b foram utilizados na etapa de subclonagem, que foi realizada no Laboratório de Bioquímica Estrutural de Proteínas (LaBEP) utilizando *primers* sintetizados pela empresa *Integrated DNA Technologies* (Figura 11). As enzimas de restrição NcoI e XhoI foram obtidas da empresa *New England Biolab*. O estoque dos plasmídeos, contendo as três construções de interesse, foram transformados na cepa de *E. coli* DH5α, e para a expressão destas construções, foram utilizadas as cepas BL21(DE3), BL21(DE3)pLysS, C43pLysS e *Rosetta gami*. Todas as cepas foram adquiridas do banco de cultura do Laboratório de Bioquímica Estrutural de Proteínas (LaBEP) da UFRJ.



**Figura 10:** Mapa do vetor de estocagem pUC57 O gene que codifica a fosfatase putativa DSM-14977 foi clonado no vetor de estocagem pUC57 pela empresa Genscript (Retirado de GenScript).



A subclonagem do gene que codifica a fosfatase putativa DSM-14977 foi realizada em três diferentes vetores de expressão: pETM30-MBP para a fusão da DSM-14977 à proteína ligadora de maltose (MBP) (A), pET32a para a fusão da DSM-14977 à proteína solúvel Trx (B) e o pET22b para o endereçamento da DSM-14977 para o periplasma (C) (Retirado da GeneOne).

## 3.1.2 Reagentes

Tabela 3: Principais reagentes utilizados no desenvolvimento deste estudo

| Reagentes  | Fornecedores             |
|--|--------------------------|
| Ácido acético                                    | Vetec                    |
| Ácido clorídrico                                 | Sigma                    |
| Ácido fítico                                     | Sigma                    |
| Ácido fosfórico                                  | Vetec                    |
| Ácido tricloroacético                            | Vetec                    |
| Acrilamida                                       | BioRad                   |
| ADP  | Sigma                    |
| ATP  | Sigma                    |
| Bisacrilamida                                    | BioRad                   |
| Brometo de etídeo                                | Vetec                    |
| Canamicina                                       | Sigma                    |
| Cloreto de sódio                                 | Sigma                    |
| Cloreto de magnésio                              | Sigma                    |
| Cloridrato de Guanidina                          | Sigma                    |
| Comassie blue G250                               | Sigma                    |
| Comassie blue R250                               | Sigma                    |
| CTP  | Sigma                    |
| Dodecil sulfato de sódio                         | Sigma                    |
| Fosfato de sódio dibásico                        | Sigma                    |
| Fosfato de sódio monobásico                      | Sigma                    |
| Fosfoserina                                      | Sigma                    |
| Fosfotirosina                                    | Sigma                    |
| Fosfotreonina                                    | Sigma                    |
| FMN  | Sigma                    |
| Frutose 6-fosfato                                | Sigma                    |
| Frutose 1, 6-bisfosfato                          | Sigma                    |
| β-Glicero-fosfato                                | Sigma                    |
| Glicina  | Sigma                    |
| Glicose-1-fosfato                                | Sigma                    |
| Glicose-6-fosfato                                | Sigma                    |
| GTP  | Sigma                    |
| Hidróxido de sódio                               | Merck                    |
| Imidazol   | Sigma                    |
| IPTG   | Sigma                    |
| GeneJet PCR purification Kit-K0702               | Thermo Fisher Scientific |
| GeneJet Plasmid Miniprep Kit                     | Thermo Fisher Scientific |
| Meio Luria-Bertani e Luria-Bertani ágar          | Sigma                    |
| MUP  | Sigma                    |
| Metanol  | Vetec                    |
| α-Naftil-fosfato                                 | Sigma                    |
| ρ-Nitrofenil-fosfato                             | Sigma                    |
| Padrão de massa molecular - Pierce <sup>TM</sup> | Thermo Scientific        |

| Unstained Protein MW Marker - #26610                                 |                          |
|--|--------------------------|
| Padrão de DNA - GeneRuler 1 kb DNA Ladder<br>#SM0313                 | Thermo Scientific        |
| Pierce <sup>TM</sup> 660nm <i>Protein Assay</i>                      | Thermo Fisher Scientific |
| Piridoxal-5-fosfato  | Sigma                    |
| PMSF   | Sigma                    |
| Polifosfato-3  | Sigma                    |
| Polifosfato-25   | Sigma                    |
| Polifosfato-75   | Sigma                    |
| Pirofosfato  | Sigma                    |
| Reagente Bradford  | Sigma                    |
| <i>Pfu</i> DNA polimerase  | QuatroG                  |
| Resina Níquel-Sepharose  | GE                       |
| Resina SP650 M   | TOYOPEARL                |
| T4 DNA ligase  | Takara Bio               |
| Tampão de carregamento LD - <i>DNA Gel Loading Dye</i> (6x) - #R0611 | Thermo Scientific        |
| Tris Base  | Sigma                    |
| Triton X-100   | Sigma                    |
| Ureia  | Sigma                    |
| UTP  | Sigma                    |

Os demais reagentes utilizados e não citados na Tabela 3 foram de grau analítico.

## 3.2 MÉTODOS

## 3.2.1 Prospecção e análises computacionais da fosfatase putativa DSM-14977 de *Oceanithermus profundus*

A prospecção da fosfatase hipotética DSM-14977 foi realizada utilizando o banco de dados de genoma, *GenBank Sequence Database*, através de ferramentas de bioinformática. A busca ocorreu utilizando os parâmetros padrão do programa e foi restrita a hidrolases de microrganismos extremófilos, onde foi realizada uma triagem, a fim de selecionar uma proteína alvo. Além disso, a pesquisa buscou enzimas de baixo peso molecular e com estrutura e função desconhecidas. Essas escolhas foram baseadas na possibilidade de determinar e estudar a estrutura da enzima escolhida por Ressonância Magnética Nuclear (RMN).

As análises computacionais foram realizadas através da Ferramenta de Pesquisa de Alinhamento Local Básico (BLAST) que foi utilizada para procurar regiões conservadas presentes nas fosfatases da superfamília das dealogenases haloácidas.

Após a escolha da proteína DSM14977 de *Oceanithermus profundus*, outros programas e ferramentas foram empregados para prever características físico-químicas e estruturais da enzima selecionada, como IUPred (Precessão de Proteínas Intrinsecamente Não Estruturadas), ExPASy (*Expert Protein Analysis System*) e sua ferramenta ProtParam, e Jpredr4 (*Protein Secondary Structure Prediction Server*). Para usar essas ferramentas e programas de predição foi necessário colocar a sequência primária da proteína de interesse como sequência de entrada.

## 3.2.2 Clonagem da fosfatase putativa DSM-14977 em vetor de estocagem

A clonagem do gene que codifica a proteína DSM-14977 foi realizada pela empresa GenScript em vetor de estocagem pUC57 (Figura 10). O gene codificante da DSM-14977 foi sintetizado e otimizado para expressão em *Escherichia coli* (Tabela 3). Para a estocagem do plasmídeo em *E. coli*, foi realizada a transformação da construção pUC57-DSM-14977 na cepa de estocagem JM109.

Tabela 4: Otimização dos códons para expressão em *E.coli* obtida pela empresa GenScript.

| Sequência                        |  |  |
|----------------------------------|--|--|
| Gene                             | AGACTGCTGTGGCCGAAAGCAAGAGTTAGAAGAGTTACAGAACTGACACCGGAATGGCTGGC   |  |
| Gene com<br>códons<br>otimizados | CGTCTGCTGTGGCCGAAAGCTCGTGTGCGTCGTGTTACGGAACTGACCCCGGAATGGCTGGC   |  |
| Aminoácidos                      | RLLWPKARVRRVTELTPEWLAARGLRGLVVDLDNTLVPYGVRPPAEGELAAWNRALAEAGIPVFIVSN<br>AKPGRTRAWARSLGLEGWGLAGKPFPWSLRRAVRRMGLAPREVAVAGDQLFTDVLGANLIGAYSVLVE<br>PIEPKKGLPHTRWVRALERRILERVPGPQNGGAQRSESGE |  |

O DNA recombinante pUC57-DSM-14977 liofilizado, enviado pela GenScript, foi centrifugado a 6000 x g por 1 minuto. Posteriormente, foi ressuspendido em 20 μL de água deionizada purificada em um sistema Milli-Q e autoclavada resultando em uma concentração final de 200 ng/μL. Apenas 1 μL do plasmídeo foi utilizado para transformação das células de estocagem JM109 competentes e a propagação dos plasmídeos. O restante foi guardado no

congelador a -20°C. A transformação bacteriana foi realizada através de choque térmico acrescentando 1 μL de pUC57-DSM-14977 nas células competentes e deixando-as em banho de gelo por 30 minutos, seguidos de 2 minutos em banho a 42°C e 10 minutos novamente em gelo. Adicionamos 800 μL de LB sem antibiótico e incubamos a 37°C por 1 hora. Em seguida, as células foram centrifugadas por 5 minutos e depois foram ressuspendidas em 50 μL de sobrenadante e plaqueadas em meio LB ágar contendo ampicilina (200 μg/mL). As placas foram incubadas em estufa a 37°C por 16 horas. Algumas colônias isoladas crescidas nas placas foram transferidas para 6 mL de meio LB com antibiótico ficando incubada por 16 horas a 37°C e 200 RPM. Para a armazenagem, as células foram ressuspendidas em glicerol (concentração final de 20%), aliquotadas em 200 μL e armazenadas a -80 °C.

## 3.2.3 Subclonagem da fosfatase putativa DSM-14977 em diferentes vetores de expressão

A subclonagem foi realizada em três vetores de expressão: (*i*) o pETM30-MBP que contém uma cauda com seis histidinas (His<sub>6</sub>) no N-terminal, seguida pela proteína ligadora de maltose (MBP), conhecida por aumentar a solubilidade da proteína de interesse (COSTA et al (2014), e um sítio de clivagem para a protease TEV (*tobacco etch virus*) (Figura 11A); (*ii*) o vetor pET32a que codifica a proteína tioredoxina (TRX) no N-terminal, também conhecida por aumentar a solubilidade da proteína de interesse (LAVALLIE, 2000), seguida pela cauda de His<sub>6</sub> e pelo sítio de clivagem para enteroquinase (EK) (Figura 11B); (*iii*) e o vetor pET22b que contém o peptídeo sinal pelB responsável pelo endereçamento da proteína de interesse para o periplasma da célula (YOON, 2010) (Figura 11C). Os plasmídeos do sistema pET de expressão possuem o promotor T7 que utiliza a RNA polimerase do fago T7 para transcrição do gene de interesse que é induzida por IPTG e contém um marcador de seleção que confere resistência a canamicina para o pETM30-MBP e a ampicilina para o pET22b e pET32a. A estratégia de subclonagem foi realizada no laboratório de bioquímica estrutural de proteínas, através da amplificação, digestão do inserto e dos vetores, defosforilação dos vetores, reação de ligação e diagnóstico por digestão.

Primeiramente foi realizada a amplificação do gene sintético presente na construção pUC57-DSM-14977 através da reação em cadeia da polimerase (PCR) utilizando os oligonucleotídeos *forward* 5'-ATATTT<u>CCATG</u>GCAGACGACACCGGC-3' e *reverse* 5'-ATATGG<u>CTCGAG</u>TTCGCCAGATTCAGAGCGTTG-3' (as sequências sublinhadas indicam o sítio específico para as enzimas de restrição *NcoI* e *XhoI*, respectivamente). Para isto, uma mistura foi preparada com 1 μL de dNTPmix [10 mM dATP, 10 mM dCTP, 10 mM

dGTP, 10 mM dTTP (pH 7,5)], 5 μL de tampão *Pfu* 10x [10 mM Tris-Cl, 10 Mm KCl, 2 mM MgSO<sub>4</sub>, 10 mM (NH4) SO<sub>4</sub>, 0,1% Triton X-100 (pH 8,8)], 37 μL de água Milli-Q autoclavada, 2,5 μL de oligonucleotídeo 5' a 10 μM, 2,5 μL de oligonucleotídeo 3' a 10 μΜ, 1 μL de DNA molde 20 ng/μL (pUC57-DSM-14977) e 1 μL de *Pfu* polimerase, em um volume final de 50 μL. A mistura foi submetida a 35 ciclos de desnaturação, anelamento e amplificação. A desnaturação da fita dupla de DNA ocorreu a 95°C por 1 minuto, o anelamento dos oligonucleotídeos a 57°C por 1 minuto e a amplificação do inserto a 72°C por 1 minuto. Para a confirmação da amplificação do inserto, o produto de PCR foi analisado por eletroforese em gel de agarose 1%. Após a verificação do resultado, o produto de PCR foi purificado com o kit GeneJet<sup>TM</sup> PCR *Purification* de acordo com as especificações do fabricante.

A digestão do inserto (produto de PCR) e de cada um dos três vetores foi realizada com as enzimas de restrição *NcoI* e *XhoI* a 37 °C por 1 hora. Esta etapa, permite a linearização do plasmídeo e a formação de regiões complementares nas extremidades do vetor e do inserto. A mistura da reação foi preparada contendo 50 μL do inserto amplificado ou de um dos vetores, 6 μL de tampão 4 10 x, 1,4 μL de água Milli-Q autoclavada, 0,6 μL de BSA 100x, 1 μL da enzima *NcoI* e 1 μL da *XhoI*. Para verificar a eficiência das enzimas de restrição também foram realizados controles contendo 16,8 μL de vetor (para cada vetor), 2 μL de tampão 4 10x, 0,2 μL de BSA 100x e 1 μL da enzima *NcoI* ou *XhoI*. Para a confirmação da digestão foi preparado um gel de agarose 1 %. Verificada a digestão, as enzimas de restrição foram inativadas por aquecimento a 65°C por 25 minutos.

Em seguida, cada um dos vetores foi defosforilado a 37°C por 1 hora com a adição de 6 μL tampão *Antartic Phosphatase* 10x e 1 μL de *Antartic Phosphatase*. Após a reação, o inserto e os vetores digeridos foram purificados com o kit GeneJetTM PCR *Purification*.

Para a reação de ligação foi feita uma dosagem da concentração do inserto e dos vetores purificados com o Qubit e o kit dsDNA BR *assay* seguindo o protocolo padrão indicado pelo fabricante. Para a reação de ligação foram utilizados os vetores e inserto nas proporções de 1:3 e 1:5, 2 μL de tampão T4 DNA ligase 10x, 1 μL de T4 DNA ligase TAKARA e o restante de água Milli-Q autoclavada, para um volume final de 20 μL. A reação foi incubada a 18°C por 16 h.

Após este período, foi realizada uma transformação em DH5α competente, utilizando 10 μL do produto de ligação obtido. As placas foram incubadas por 16 horas a 37°C em LB ágar (30 g/L) seletivo contendo o antibiótico apropriado. Para a construção pETMBP-DSM-14977 foi utilizada canamicina 100 μg/mL e para as construções, pET22b-DSM-14977 e pET

32a-DSM-14977 foi utilizada a ampicilina 200 μg/mL. Cada colônia crescida foi transferida para 6 mL de meio LB líquido (20 g/L) com antibiótico apropriado, ficando incubada por 16 horas a 37°C e 200 RPM. Posteriormente, os plasmídeos foram extraídos da célula utilizando o kit de extração GeneJET *Plasmid Miniprep*.

Para verificar a inserção do gene de interesse nos vetores de expressão foi realizado um diagnóstico por digestão plasmidial. Esse método seguiu o protocolo de digestão como descrito anteriormente, porém foi adaptado para o volume final de reação de 6 μL. A análise da eficácia da subclonagem foi verificada através de um gel de agarose 1%. As amostras avaliadas como "bem sucedidas" na subclonagem foram enviadas e sequenciadas no laboratório SONDA na Unidade de Genômica do Departamento de Biofísica da Universidade Federal do Rio de Janeiro. As amostras foram enviadas, com aproximadamente, 56 ng de DNA molde para pETMBP-DSM-14977, 58 ng para pET32a-DSM-14977 e 48 ng para pET22b-DSM-14977, e 3,2 pmol de oligonucleotídeos, sendo estes, o promotor T7 ou terminador T7 e o restante de água Milli-Q autoclavada, para um volume final de 7 μL.

#### 3.2.4 Eletroforese

#### 3.2.4.1 Eletroforese em gel de agarose

A agarose à 1% utilizada para os géis de agarose foi solubilizada, após ser esquentada, em tampão TAE (40 mM Tris, 20 mM ácido acético, 1 mM EDTA). Quando a solução de agarose amornou adicionamos 0,01 % de brometo de etídeo para o preparo do gel. Em seguida, o gel de agarose foi posicionado na cuba *Mini-Sub Cell GT Cell* da Bio-Rad e seu volume foi preenchido com tampão TAE. As amostras foram preparadas com a mistura de 5 μL de amostra e 1 μL tampão de carregamento Loading Dye (6x). O padrão de DNA usado foi o *GeneRuler 1 kb DNA Ladder* em uma quantidade de 3 μL. A eletroforese foi realizada a uma voltagem constante de 100 V por, aproximadamente, 30 minutos.

A detecção do DNA no gel foi analisada através de luz UV emitida por um transiluminador sob o gel.

## 3.2.4.2 Eletroforese em gel de poliacrilamida em condição desnaturante (SDS-PAGE)

As amostras apresentadas no trabalho, durante o processo de obtenção da fosfatase putativa DSM-14977, foram analisadas qualitativamente por SDS-PAGE utilizando gel de

poliacrilamida na concentração de 12% ou 15% (p/v), a intensidade de corrente constante foi de 30 mA por gel, em temperatura ambiente, como descrito por Laemmli (1970). Para tal, foi utilizado o sistema *Mini-Protean* da Bio-Rad, tampão de corrida (50 mM Tris-HCl, pH 8.3, 150 mM Glicina e 0,1 % SDS) e padrão de baixo peso molecular (Pierce *Unstained Protein MW Marker*).

A detecção de proteínas no gel foi analisada utilizando o corante *Coomassie brilliant blue*. Os géis foram corados utilizando a solução de *Coomassie brilliant blue* 0,1% em metanol/ácido acético/água (45/10/45, v/v/v) e foram descorados utilizando solução de metanol 40% (v/v) e ácido acético 10% (v/v).

## 3.2.5 Expressão da fosfatase putativa DSM-14977

Após a etapa de subclonagem das construções pETMBP-DSM-14977, pET22b-DSM-14977 e pET32a-DSM-14977 foi realizada a transformação destes plasmídeos em cepas de *E. coli* para a expressão da proteína DSM-14977, seguindo o mesmo protocolo descrito no item 3.2.2 para a transformação em cepa JM109. Culturas de glicerol à 20% foram feitas com cada uma das colônias transformadas.

Os testes de expressão foram realizados visando obter maior expressão proteica da fosfatase putativa DSM-14977. Os parâmetros avaliados e testados foram diferentes temperaturas, tempos de indução e cepas de *E.coli*.

A expressão das três construções foi realizada utilizando as cepas BL21(DE3), BL21(DE3)PLysS, C43pLys-S e rosetta gami. Um pré-inóculo foi realizado contendo 10 μL de cultura de glicerol de cada transformante, 5 mL de LB, 5 μL de canamicina [100 μg/mL] para pETMBP-DSM-14977 e de ampicilina [200 μg/mL] para pET22b-DSM-14977 e pET32a-DSM-14977. Para as cepas rosetta gami e as que contém o plasmídeo pLys-S foi adicionado, ao meio de cultura, o antibiótico cloranfenicol [100 μg/mL]. Este pré-inóculo foi crescido em incubadora rotatória a 37°C a 200 rpm. por aproximadamente, 16 h. Em seguida, 1 mL do pré-inóculo foi repicado em 25 mL de LB, na presença dos antibióticos específicos. e crescido em uma incubadora rotatória a 37°C a 200 rpm até atingir a D.O<sub>600nm</sub> 0.8. Durante o crescimento, alíquotas de 1 mL foram retiradas para acompanhar a densidade óptica (D.O.), em espectrofotômetro, utilizando o comprimento de onda de 600 nm. Ao atingir a DO<sub>600nm</sub> 0,8, a expressão da proteína de interesse foi induzida, com a adição de 1 mM de IPTG. Alíquotas de 1 mL foram retiradas, a cada hora, durante 4 horas e após 16 horas de indução. Todas as alíquotas foram centrifugadas, por 5 minutos, e a massa de células ressuspendida em

30 μL de tampão de amostra [Tris-HCl 12,5 mM, β-mercaptoetanol 25 mM, SDS 0,5%, Azul de bromofenol 0,25%, Glicerol 2,5% (pH 6,8)] para o monitoramento da expressão da DSM-14977 por eletroforese em gel de poliacrilamida (SDS-PAGE 12 ou 15 %).

Após a definição da melhor condição de expressão da DSM-14977, a mesma foi escalonada para um volume maior, para a produção de grandes quantidades da proteína alvo de estudo. Sendo assim, o crescimento e expressão da enzima foi realizada em 1 L de cultura na presença dos antibióticos específicos. Após o melhor tempo de expressão da DSM-14977, o volume total da cultura foi divido em quatro partes (250 mL cada), que foram centrifugadas a 4°C, 10.000 rpm por 20 minutos. O sobrenadante foi descartado e o *pellet* estocado a -20°C.

## 3.2.6 Extração da fosfatase putativa DSM-14977 da célula de *E.coli*

#### 3.2.6.1 Lise celular

Para o passo subsequente de purificação, é necessário a lise celular para a exposição da proteína de interesse na fração solúvel ou insolúvel. Assim, as células que expressaram a proteína DSM-14977 foram lisadas utilizando sonicador de ponta.

Para isto, uma parte do pellet (250 mL) foi descongelada e ressuspendida em 12,5 mL de tampão de lise (50 mM Tris-HCl pH 7,2, 250 mM NaCl, 5 mM imidazol e 250 μL PMSF). A lise celular foi, então, realizada através de sonicação (20 ciclos de 20 segundos *on* e 20 segundos *off* a 225 W) em banho de gelo. O lisado celular foi centrifugado a 8.500 x g por 20 minutos a 4°C. Uma alíquota de 50 μL foi retirada para a análise das proteínas solúveis, presentes no sobrenadante, e insolúveis, presentes no pellet, por SDS-PAGE.

## 3.2.7.1 Choque osmótico

Para as células que foram transformadas com o plasmídeo pET22b-DSM-14977, foram realizados cinco diferentes protocolos de choque osmótico, visando o extravazamento do conteúdo periplasmático. Todos os protocolos resultando em uma mudança específica na permeabilidade da superfície estrutural da bactéria.

No primeiro protocolo testado, foi seguido o modelo de choque osmótico, descrito pelo manual do sistema pET da Novagen, retirado de AUSUBEL e outros (1989). Uma parte do pellet (250 mL) foi ressuspendida em 80 mL de tampão 30 mM de Tris-HCl, pH 8,0 a 20% de sacarose. Em seguida, foi adicionado 1mM de EDTA, pH 8,0 e incubado por 10 minutos em temperatura ambiente sob leve agitação magnética. Posteriormente, a solução contendo a

proteína de interesse foi centrifugada por 10 minutos, a 10.000 x g a 4°C. O *pellet* foi ressuspendido em 100 mL de MgCl<sub>2</sub> 5 mM gelado, e incubado por 10 minutos, em banho de gelo, sob leve agitação magnética. Uma nova centrifugação foi realizada, e o sobrenadante contendo a fração periplasmática, é guardada.

O segundo protocolo testado foi modificado de MATSUSHITA e outros (1981), neste protocolo, foi utilizado 1g de célula úmida que foi ressuspendida em 20 mL de tampão Tris-HCl 100 mM, pH 8,0, 1 mM EDTA e sacarose 20%. Em seguida, 4,8 μg de lisozima foi adicionado à solução. A suspensão celular foi mantida em agitação moderada por 30 minutos a 30°C. A amostra foi centrifugada a 10.000 x g por 10 minutos a 4°C, e o sobrenadante foi reservado.

O terceiro protocolo foi modificado de KOSHLAND & BOTSTEIN (1980). Uma parte do pellet (250 mL) foi ressuspendida em 37,5 mL de solução tampão Tris-HCl 10 mM, pH 7,5 e sacarose 20% (p/v) gelada. Posteriormente, foi adicionado 1,25 mL de EDTA 0,5 M, pH 8,0 e a solução foi incubada por 10 minutos em gelo. As células foram centrifugadas por 5 minutos a 4°C, e o pellet, ressuspendido rapida e vigorosamente em 250 mL de água milliQ, e incubadas por mais 10 minutos, em banho de gelo. Essa suspensão foi centrifugada novamente por 5 minutos, e o sobrenadante guardado a -20 °C.

O quarto protocolo de choque osmótico foi realizado em microescala, como descrito em BORJALILOO e outros (2003). Neste protocolo utilizou-se uma parte do pellet (250 mL) que foi ressuspendida em 2,5 mL de tampão Tris-HCl 200 mM, pH 8,0, 0,5 mM de sacarose e 0,5 mM de EDTA. A suspensão foi mantida sob vigorosa agitação, em banho de gelo, por 10 minutos. Em seguida, foi adicionada 3,8 mL de água milliQ gelada e a solução foi incubada, por 30 minutos, em banho de gelo. Ao final da incubação, a amostra foi centrifugada a 16.000 x g por 20 minutos a 4°C e o sobrenadante guardado a -20 °C.

O quinto protocolo utilizado foi de acordo com WARD e outros (1996). Neste método, uma parte do pellet (250 mL) foi ressuspendida em 50 mL de tampão Tris-HCl 200 mM, pH 7,5, 20% de sacarose, 1 mM de EDTA e lisozima (50, 100 ou 500 μg/mL). A solução foi incubada sob agitação, em temperatura ambiente por 15 minutos. Em seguida, foi adicionado 50 mL de água milliQ gelada, e a solução foi incubada sob agitação, em banho de gelo, por 15 minutos. Posteriormente, a amostra foi centrifugada a 13.000 x g a 4°C por 20 minutos e o sobrenadante foi guardado a -20 °C .

Ao final de todos os protocolos testados, foi retirada uma alíquota de 50 μL para verificar a eficácia do choque osmótico e para analisar as proteínas solúveis, presentes no sobrenadante, e insolúveis, presentes no pellet, através de SDS-PAGE 12 ou 15%.

## 3.2.8 Purificação da fosfatase putativa DSM-14977

O sistema automatizado de cromatografia liquida *Äkta-prime plus* (GE) foi utilizado em todas as purificações realizadas. As colunas cromatográficas foram acopladas ao leitor de UV presente no *Äkta* que permite acompanhar a eluição da proteína da mistura através do monitoramento da absorvância a 280 nm. Além disso, todas as purificações foram monitoradas e analisadas utilizando SDS-PAGE 12 ou 15%. Para isto, as amostras foram preparadas com 20 µL de amostra mais 5 µL de tampão de amostra 4x.

## 3.2.8.1 Purificação por afinidade a íons metálicos (IMAC)

As purificações dos produtos de expressão his<sub>6</sub>-MBP-DSM-14977 e TRX- his<sub>6</sub>-DSM-14977 foram realizadas através da cromatografia de afinidade a íons metálicos utilizando uma coluna de afinidade à níquel, *HisTrap High Performance* de 5 mL (Ni-NTA), devido a cauda de histidina fusionada à porção N-terminal da fosfatase putativa DSM-14977.

O protocolo foi realizado através dos seguintes passos: (*i*) lavagem da coluna com 25 mL de água Milli-Q a 3 mL/min, (*ii*) equilíbrio da coluna com 25 mL de tampão A (50 mM Tris-HCl, pH 7,2, 150 mM de NaCl e 5 mM imidazol) a 3 mL/min, (*iii*) injeção de 25 mL do sobrenadante da lise a 0,5 mL/min, (*iv*) lavagem com 45 mL de tampão A a 2 mL/min para a retirada de ligantes inespecíficos, (*v*) eluição das proteínas através de um gradiente linear crescente de imidazol de 0 a 500 mM. O gradiente foi realizado com o fluxo de 3 mL/min até 25% do tampão B (tampão Tris-HCl 50 mM, pH 7,2, 150 mM NaCl e 2 M imidazol) e teve um volume final de 180 mL, e (*vi*) em seguida, a coluna foi lavada com 100 % do tampão B.

## 3.2.8.2 Purificação por cromatografia hidrofóbica

A purificação da construção pelB-DSM-14977 foi realizada em um primeiro passo através da cromatografia hidrofóbica. A coluna *Hiprep butyl FF 16/10* de 20 mL é constituída por um grupo butil caracteristicamente hidrofóbico. O protocolo utilizado foi realizado através dos seguintes passos: (*i*) lavagem da coluna com 100 mL de água Milli-Q a um fluxo de 5 mL/min, (*ii*) equilíbrio da coluna com 100 mL de tampão A (50 mM Tris-HCl, pH 8,0, 2 M NaCl) a um fluxo de 5 mL/min, (*iii*) injeção de 12,5 mL de amostra a 1 mL/min, (*iv*) lavagem com 150 mL de tampão A a 5 mL/min e (*v*) eluição das proteínas com um gradiente

linear decrescente de 2 a 0 M de NaCl com um volume final de 200 mL. O gradiente decrescente foi realizado com o fluxo de 5 mL/min até 100% do tampão B (tampão Tris-HCl 50 mM, pH 8,0). Posteriormente, foi feita uma lavagem com 150 mL de água milliQ.

## 3.2.8.3 Purificação por cromatografia de afinidade à heparina

A purificação da construção pelB-DSM-14977 foi realizada através da cromatografia de afinidade à heparina. A coluna *HiTrap heparin High Performance* de 5 mL é constituída de heparina, um polissacarídeo polianiônico sulfatado pertencente à familía dos glicosaminoglicanos, e conhecida por ser ligar a proteínas com afinidade a açúcar. O protocolo usado foi realizado através dos seguintes passos: (*i*) lavagem da coluna com 50 mL de água Milli-Q a um fluxo de 5 mL/min, (*iii*) equilíbrio da coluna com 50 mL de tampão A (10 mM Tris-HCl, pH 8,0) a 5 mL/min, (*iii*) injeção de 2 mL de amostra a 1 mL/min, (*iv*) lavagem com 45 mL de tampão A a 5 mL/min e, (*v*) eluição das proteínas com um gradiente em degrau crescente de 0 a 2 M de NaCl. O gradiente crescente foi realizado com fluxo de 5 mL/min até 29 % do tampão B (tampão Tris-HCl 10 mM, pH 8,0 e 2 M NaCl) e, depois, até 75% do tampão B, com um volume final de 50 mL.

## 3.2.9 Dosagem de proteínas totais

A concentração de proteínas totais em solução foi determinada através do método de BRADFORD (1976), utilizando o reagente de Bradford (BioRad) e a albumina sérica bovina (BSA) como padrão. A dosagem foi preparada utilizando uma curva padrão com a BSA nas concentrações de 0,1 mg/mL a 1 mg/mL. Para os ensaios, foi usada uma placa de ELISA com 96 poços onde foram adicionados 10 μL de amostra e 190 μL do reagente de Bradford. A reação foi agitada e colocada no escuro por 15 minutos. A leitura da placa foi realizada em um espectrofotômetro de microplacas SpectraMax M2 (Molecular devices) a 595 nm. Todas as análises foram realizadas em triplicata e a determinação da concentração de proteínas na amostra foi baseada nos valores obtidos na curva padrão de BSA utilizando o coeficiente angular da reta.

#### 3.2.10 Dicroísmo circular

Para verificar características gerais na estrutura secundária da fosfatase putativa DSM-14977 foi empregada a técnica de espectropolarimetria de discroísmo circular. O espectropolarímetro Jasco-715 (Jasco Corporation, Tokyo, Japan) foi utilizado para realizar as medidas de dicroísmo circular. Os espectros foram coletados usando cubetas de quartzo com caminho ótico de 2 mm na região do UV distante (190 a 260 nm). Para a realização dos espectros foi utilizada a concentração de 10 μM da proteína DSM-14977 em tampão Tris-HCl 10 mM, pH 8,0 em um volume final de 400 μL. Cada espectro foi obtido como a média de 3 varreduras em velocidade de 100 nm.min<sup>-1</sup>. As linhas de base (cubeta contendo apenas tampão) foram subtraídas dos espectros. No início de todos os experimentos, o branco apenas do espectropolarímetro foi feito para observar a linha de base do mesmo e a tensão da lâmpada.

## 3.2.11 Espectroscopia de Fluorescência intrínseca do triptofano

Os experimentos de espectroscopia de fluorescência intrínseca do triptofano foram realizados para verificar possíveis mudanças na estrutura terciária da fosfatase putativa DSM-14977 na presença de diferentes agentes químicos e físicos, utilizando o triptofano como sonda intrínseca, geralmente, presente na maioria das proteínas conhecidas.

Os espectros de fluorescência da proteína DSM-14977 foram obtidos em um Fluorímetro Varian CaryEclipse utilizando uma lâmpada de Xenônio de 300 W, fendas de 1 mm para excitação e emissão e cubeta de quartzo com volume final de 500 µL. O comprimento de onda de excitação do triptofano foi fixado em 280 nm, sendo a emissão de fluorescência coletada entre 300 a 420 nm.

Os espectros de emissão de fluorescência da DSM-14977 foram quantificados através do centro de massa espectral (v), como mostrado na equação 1:

$$v = \frac{\sum viFi}{\sum Fi}$$

Onde, Fi é a emissão de fluorescência em um comprimento de onda e vi é o somatório que abrange todos os valores de F avaliados.

Portanto, o centro de massa leva em consideração a intensidade e a área de cada pico. O aumento dos valores de centro de massa, em nm, relacionam o aumento do conteúdo desenovelado das proteínas.

Os experimentos para as análises estruturais da fosfatase putativa DSM-14977 foram realizados com concentração de 1 µM em tampão Tris-HCl 10 mM, pH 8,0, em um volume final de 500 µL. Para estas análises utilizamos agentes desnaturantes químicos, como a ureia e cloridrato de guanidina, e físicos como temperatura. Para tal, foram usadas diferentes concentrações de ureia (0, 1, 2, 3, 4, 5, 6, 7 e 8 M), cloridrato de guanidina (0,5, 1,0, 1,5, 2,0, 2,5, 3,0, 3,5, 4,0, 5,0 e 6,0 M) e temperatura (30, 40, 50, 60, 70, 80 e 90°C). Para os ensaios utilizando os agentes desnaturantes químicos as amostras foram incubadas em temperatura ambiente, por um período de 16h, enquanto que, para a variação de temperatura foram incubadas por 5 minutos.

#### 3.2.12 Dosagem da atividade enzimática para fosfatases

Para a determinação da atividade enzimática da fosfatase putativa DSM-14977 foi realizada a técnica de espectrofotometria, utilizando o substrato geral para fosfatases, ρ-nitrofenil-fosfato (pNPP), no intuito de identificar, se a DSM-14977, possui atividade de fosfohidrolase.

Os ensaios enzimáticos foram realizados através do monitoramento da formação de ρ-nitrofenol (cromóforo) utilizando espectrofotômetro SpectraMax M2 (Molecular devices) a 405 nm. As reações de hidrólise do ρ-nitrofenil-fosfato, catalisadas pela fosfatase hipotética DSM-14977, foram realizadas em triplicata em placas de 96 poços.

O volume reacional final foi de 150  $\mu$ L, destes, foi adicionado 100  $\mu$ L de tampão Tris-Cl 10 mM, pH 6,0, 7,0 ou 8,0, dependendo do ensaio a ser realizado, 300 mM NaCl, 17  $\mu$ L de MgCl<sub>2</sub> a 42,5 mM [5 mM final], 18  $\mu$ L de substrato a 20 mM [2 mM final] e 15  $\mu$ L de DSM-14977 purificada (7,5  $\mu$ g de proteína).

A reação foi incubada a 20°C, 40°C ou 60°C durante 30 minutos e, ao fim do tempo estipulado para cada reação, foram adicionados 20 μL de NaOH 5 M para levar à conversão de todo ρ-nitrofenol em ρ-nitrofenolato, que possui comprimento de onda máximo de absorção de 405 nm.

Nestes casos, a atividade enzimática foi calculada pela fórmula da Equação 2:

Atividade espec {fica (\u03c4moles/min.mg) = 
$$\frac{(MAbs \cdot Vr)}{\frac{(cm \cdot \varepsilon)}{([ptn] \cdot Vp \cdot t)}}$$

Onde, M Abs significa Média da absorvância (Abs–Abs<sub>branco</sub>) lida a 405 nm, Vr é o Volume da reação utilizada (0.17. $10^{-3}$  L), cm é o caminho óptico da cubeta (0,5 cm),  $\epsilon$  é o coeficiente de extinção molar do p-nitrofenolato tabelado (1,75 .  $10^4$  M $^{-1}$  cm $^{-1}$ ), [ptn] é a concentração de proteína utilizada no experimento (0,494. $10^{-3}$  mg/ $\mu$ L), Vp é o volume de proteína utilizado no ensaio que foi 15  $\mu$ L e t é o tempo de reação, que foi realizada em 30 min.

## 3.2.13 Triagem de substratos específicos para fosfatases

Para realizar a identificação de potenciais substratos para a fosfatase hipotética DSM-14977 foi medida atividade fosfatásica, através de espectrofotometria, e como alternativa, através do monitoramento da liberação de fosfato livre utilizando o verde malaquita que possui sensibilidade, de aproximadamente 0,5 μM (KOICHI ITAYA E MICHIO UI, 1966). Estes ensaios foram realizados em um espectrofotômetro SpectraMax M2 (Molecular devices), em placas de Elisa de 96 poços, e em triplicata.

Nestes experimentos foram testados, tanto substratos fosforilados naturais, quanto sintéticos. Os substratos utilizados foram os açúcares—fosforilados (10 mM de  $\alpha$ -D-glicose 1-fosfato, 10 mM de D-glicose 6-fosfato, 10 mM de D-frutose 6-fosfato, 1 mM de D-frutose 1,6-bisfosfato), nucleotídeos (0,1 mM de ADP, 0,1 mM de ATP, 10 mM de CTP, 0,1 mM de GTP e 0,1 mM de UTP), cofator fosforilado (1 mM de FMN), aminoácidos fosforilados (10 mM de fosfotreonina, 10 mM de fosfoserina e 10 mM de fosfotirosina), polifosfatos (0,1 mM de polifosfato-3, 0,01 mM de polifosfato-25, 0,001 mM de polifosfato-75) e outros (1 mM de piridoxal-5-fosfato, 0,1 mM de ácido fítico, 10 mM de pirofosfato de sódio, 10 mM de  $\alpha$ -naftil- fosfato e 10 mM de  $\beta$ -glicerol-fosfato).

O volume final de reação foi de 250  $\mu$ L, composto por 36,67  $\mu$ L de tampão Tris-HCl 10 mM, pH 8,0, 8,33 mM de MgCl<sub>2</sub>, 3,34  $\mu$ L de substrato e 10  $\mu$ L de DSM-14977 purificada (0,64  $\mu$ g de proteína). As reações foram incubadas a 40°C durante 30 min.

Depois da reação, foi adicionado 200 μL de solução de verde malaquita [4,2 % de molibdato de amônio em 4 M de HCl e 0,045 % de verde de malaquita na proporção 1:4 (v/v); a solução é incubada por 20 minutos, e posteriormente, filtrada para sua utilização] e a

mistura é incubada, por 15 minutos, no escuro e em temperatura ambiente. Os produtos da reação foram quantificados através da medida da absorvância a 650 nm.

A taxa de hidrólise inespecífica foi monitorada, em paralelo, usando reação controle sem a presença da enzima (o branco). Uma curva padrão de fosfato inorgânico foi realizada a partir de 40 μM de ácido fosfórico que foi feita entre 0,0025 nmoles e 2 nmoles, onde o fator K foi retirado.

A atividade enzimática foi calculada pela Equação 3:

$$Atividade\ específica\ (\mu moles/min.\ mg) = \frac{(MAbs)}{K\times[ptn]\times Vp\times t}$$

Em que, MAbs é a média da absorvância (Abs–Abs<sub>branco</sub>) lida a 650 nm, K é o fator obtido a partir da curva padrão de fosfato inorgânico que foi de 0.0083 (mg/ml/Abs), [ptn] é a concentração de proteína utilizada no experimento que foi de aproximadamente  $0.064 \times 10^{-3}$  mg/ $\mu$ L, Vp: volume de proteína utilizado no ensaio (10  $\mu$ L) e t: tempo de reação que foi de 30 min.

#### 3.2.14 Modelagem computacional

Para obter informações sobre a fosfatase putativa DSM-14977 foi realizado um modelo baseado na homologia de sequência em colaboração com o Dr. Francisco Gomes da Fundação Oswaldo Cruz. Para tal, foi utilizado o programa I-TASSER (*Iterative Threading ASSEmbly Refinement*) que é uma plataforma *on-line* capaz de implementar algoritmos baseados em I-TASSER para a predição da estrutura e função da proteína de interesse. A partir da sequência primária da DSM-14977, o programa I-TASSER faz uma busca utilizando os 10 melhores programas estruturais, e em cada um desses programas, 10 algoritmos diferentes,FFAS-3D, HHsearch, MUSTER, pGenTHREADER, PPAS, PRC, PROSPECT2, SP3, e SPARKS-X, são usados para selecionar seqüências análogas a proteína de interesse. Assim, entre as 100 seqüências selecionadas, o I-TASSER faz um *rancking* das 10 melhores.

Para a predição da função biológica da proteína de interesse, o servidor I-TASSER combina os modelos 3D previstos para a proteína de interesse à 3 bibliotecas independentes, que consistem em proteínas de número de classificação enzimática (EC), vocabulário de ontologia genética (GO) e sítios de ligação ao ligante conhecidas. Os resultados finais das predições de função são deduzidos, através do consenso das principais combinações estruturais, com os *scores* de função calculados com base no *score* de confiança dos modelos

estruturais do I-TASSER, a similaridade estrutural entre o modelo e os moldes são avaliados pelo *score* TM e a identidade de seqüência nas regiões estruturalmente alinhadas.

#### 4. **RESULTADOS**

## 4.1 PROSPECÇÃO DA FOSFATASE PUTATIVA DSM-14977 DE *OCEANITHERMUS PROFUNDUS*

Para a prospecção de uma nova enzima, através de ferramentas de bioinformática, iniciamos uma pesquisa no banco de dados de genomas, utilizando, como molde, uma hidrolase conhecida como um ótimo biocatalisador. A busca por uma nova enzima, seguiu os seguintes pré-requisitos: ser uma possível enzima de organismo extremófilo, pequena (abaixo de 25 kDa), sem estudos estruturais e funcionais. Desta forma, a proteína DSM-14977 (Tabela 5) foi selecionada (NCBI Reference Sequence: YP\_004056761.1) por preencher todos os requisitos determinados. A DSM-14977 pertence ao organismo extremófilo *Oceanithermus profundus* da cepa DSM-14977, é anotada como uma fosfatase hipotética, de 19,5 kDa, pertencente à superfamília das dealogenases haloácidas (HAD), não existindo, na literatura, quaisquer estudos estruturais ou funcionais sobre a mesma.

Tabela 5: Sequência primária da fosfatase putativa DSM-14977.

| Fosfatase putativa | Sequência primária  |
|--------------------|---|
| DSM-14977          | MRLLWPKARVRRVTELTPEWLAARGLRGLVVDLDNTLVPYGVRP<br>PAEGELAAWNRALAEAGIPVFIVSNAKPGRTRAWARSLGLEGWG<br>LAGKPFPWSLRRAVRRMGLAPREVAVAGDQLFTDVLGANLIGAYS<br>VLVEPIEPKKGLPHTRWVRALERRILERVPGPQNGGAQRSESGE |

O alinhamento da seqüência primária da DSM-14977, e de quatro diferentes fosfatases da superfamília HAD, realizadas com o auxílio do programa Clustal Omega, foi capaz de identificar motivos conservados desta superfamília (Figura 12). Para isso, foi utilizada uma fosfatase previamente caracterizada, a glicerol-1-fosfato fosfohidrolase 2 (GPP2) de *Saccharomyces cerevisiae* (NORBECK et al, 1996), e fosfatases não caracterizadas, dentre elas, a de *Marinithermus hidrothermalis* (NCBI: WP\_013704883.1), *Thermus aquaticus* (NCBI: WP\_053767389.1) e *Bacillus anthracis* (GenBank: AIM08256.1) com 60%, 50% e 43% de identidade, respectivamente.

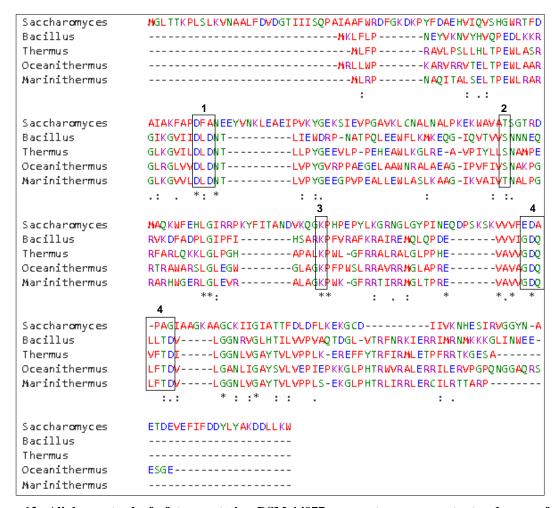


Figura 12: Alinhamento da fosfatase putativa DSM-14977 com outras representantes da superfamília HAD.

O alinhamento da sequência primária da proteína DSM-14977 foi realizado com outras fosfatases da superfamília HAD, dentre elas, a fosfatase caracterizada, glicerol-1-fosfato fosfohidrolase 2 (GPP2) de Saccharomyces cerevisiae, e as fosfatases não caracterizadas de Marinithermus hidrothermalis (NCBI: WP\_013704883.1), Thermus aquaticus (NCBI: WP\_053767389.1) e Bacillus anthracis (GenBank: AIM08256.1). Os motivos característicos e conservados presentes no sítio ativo da superfamília HAD podem ser observados (caixas 1, 2, 3 e 4). A cor vermelho representa resíduos pequenos e hidrofóbicos (AVFPMILW), a azul representa aminoácidos ácidos (DE), a magenta indica resíduos básicos (RK), o verde representa os resíduos hidroxil, sulfidril, amina e G (STYHCNGQ) e cinza representa outros amino/iminoácidos incomuns.

O portal ExPASy (*Expert Protein Analysis System*) foi utilizado para prever características físico-químicas, presentes na fosfatase putativa DSM-14977, através da utilização de sua sequência primária. Com a ferramenta ProtParam do ExPASy foi observado que a enzima de interesse é composta por 177 aminoácidos, possui massa molecular de 19,46 kDa, pI teórico de 11,11 e índice médio de hidropaticidade de -0,200.

Outro programa utilizado foi o IUPred que faz uma análise da sequência de aminoácidos, verificando possíveis interações entre as cadeias laterais dos resíduos e, assim, é capaz de realizar a predição de regiões desenoveladas ou enoveladas da proteína de interesse. O gráfico desta análise evidencia que a maior parte da proteína possui valores de tendência à

desordem próximos de zero, indicando que a proteína é altamente estruturada, e que, somente, entre os aminoácidos 161 ao 177, possui valores acima de zero, o que pode ser esperado, por se tratar da região C-terminal (Figura 13A). Esta análise sugere que a enzima putativa DSM-14977 é enovelada, visto que há possíveis interações entre os aminoácidos e, conseqüentemente, estrutura tridimensional definida.

Outro programa utilizado foi o Jpred4, que contribuiu para uma análise preliminar do conteúdo de estrutura secundária da enzima putativa DSM-14977. Tal ferramenta, inferiu a presença de 7  $\alpha$ -hélices e 5 fitas- $\beta$ , o que se assemelha ao padrão  $\alpha/\beta$  hidrolase característico do enovelamento de Rossmann presente na superfamília HAD (Figura 13B).

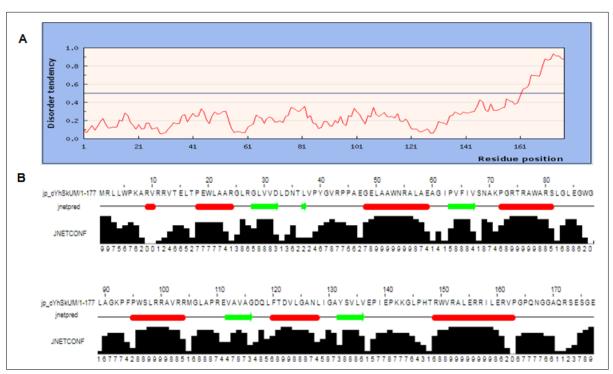


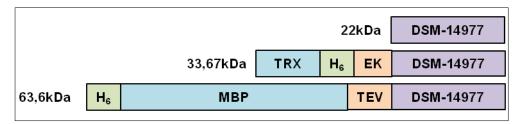
Figura 13: Predição das regiões desordenadas e do conteúdo de estrutura secundária da fosfatase putativa DSM-14977.

Os programas utilizados fazem predições sobre a estrutura da DSM-14977 através da sua estrutura primária. (A) A predição para regiões desordenadas da proteína DSM-14977 pelo programa IUPred mostrou valores baixos de tendência a desordem, indicando que, grande parte da proteína, encontra-se enovelada. (B) A predição do conteúdo de estrutura secundária da fosfatase putativa DSM-14977 analisada pelo programa Jpred4 revelou uma composição de 7  $\alpha$ -hélices e 5 folhas- $\beta$  alternadas ao longo da estrutura, em uma disposição semelhante a encontrada nas  $\alpha/\beta$  hidrolases com enovelamento do tipo Rossmann presentes na superfamília HAD.

## 4.2 SUBCLONAGEM DA FOSFATASE PUTATIVA DSM-14977 DE OCEANITHERMUS PROFUNDUS

A subclonagem do gene que codifica a fosfatase putativa DSM-14977 foi realizada em três vetores de expressão, gerando três construções diferentes (Figura 14). O vetor pETM30-MBP contém uma cauda de seis histidinas (His<sub>6)</sub> no N-terminal, seguida pela proteína ligadora

de maltose (MBP) e um sítio de clivagem para a protease TEV (*tobacco etch virus*) fusionadas à DSM-14977, aqui chamada de pETMBP-DSM-14977. O vetor pET32a, codifica a proteína tioredoxina (TRX) no N-terminal, uma cauda de His<sub>6</sub>, seguida pelo sítio de clivagem para enterokinase (EK), chamada neste estudo de pET32a-DSM-14977. E por fim, o pET22b que fusiona o pelB a proteína de interesse chamada de pET22b-DSM-14977. A MBP e a TRX são proteínas conhecidas por aumentar a solubilidade da proteína de interesse durante o processo de expressão, lise celular e purificação. Já o peptídeo sinal, pelB, endereça a proteína para o compartimento periplasmático da bactéria.



**Figura 14:** Esquema ilustrativo das construções da fosfatase putativa DSM-14977 utilizadas neste estudo. Para a subclonagem foram realizadas diferentes construções da DSM-14977 através do uso de três vetores de expressão: pET22b-DSM-14977, pET32a-DSM-14977 e pETMBP-DSM-14977. O vetor pET22b fusiona o peptídeo sinal pelB, responsável pelo endereçamento da proteína de interesse para o periplasma bacteriano. Este constructo possui massa, de aproximadamente, 22 kDa. O vetor pET32a fusiona a tioredoxina (TRX) à proteína de interesse, além de uma cauda de histidinas (H<sub>6</sub>) e o sítio de clivagem para enterokinase (EK). Este constructo possui massa, de aproximadamente, 34 kDa. O vetor pETM30-MBP fusiona a proteína de interesse, além da cauda de hisdina, a proteína ligadora de maltose (MBP) e o sítio de clivagem para TEV (TEV), formando um constructo com massa, de aproximadamente, 64 kDa.

A primeira etapa da subclonagem da fosfatase putativa DSM-14977 foi realizada através da amplificação do gene sintético codificante presente no vetor intermediário pUC57 (sintetizado pela GenScript) utilizando *primers* específicos através da reação em cadeia de polimerase (PCR) (Figura 15).

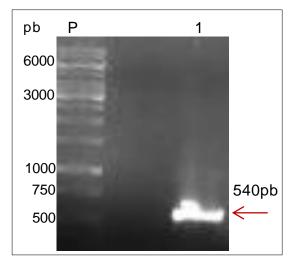
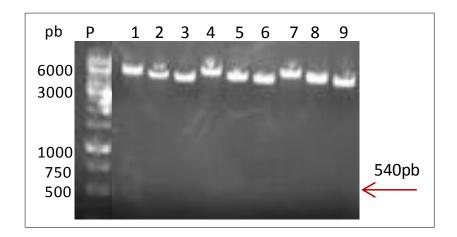


Figura 15: Gel de agarose 1% da amplificação do gene sintético de DSM-14977 presente no vetor de estocagem pUC57.

O gel de agarose mostra o gene sintético da DSM-14977 amplificado através da reação em cadeia da polimerase com massa esperada de aproximadamente 540 pb. P: Padrão de peso molecular, 1: Gene sintético da fosfatase putativa DSM-14977 amplificado (seta).

Em seguida, foi realizada a purificação do produto amplificado e, posteriormente, a digestão de cada um dos três vetores e do inserto purificado utilizando as enzimas de restrição NcoI e XhoI (Figura 16). O resultado da digestão dos vetores pETM30-MBP, pET32a e pET22b (Figura 16 - poços 1, 2 e 3, respectivamente), com ambas as enzimas, mostra uma banda maior, correspondente ao plasmídeo linearizado de tamanho 6947 pb, 5900 pb e 5493 pb, respectivamente, e uma banda menor, de 540 pb, correspondente ao fragmento de DNA que é excisado do plasmídeo, quando este, sofre digestão dupla por estas enzimas. Para o controle do experimento foi realizada a digestão de cada um dos vetores com apenas uma das enzimas, onde é possível observar a linearização do plasmídeo, entretanto, não há excisão do fragmento (Figura 16 - poços 4-9).



**Figura 16:** Gel de agarose 1% da digestão dos vetores de expressão pETM30-MBP, pET32a e pET22b. A digestão foi realizada através do uso das enzimas de restrição NcoI e XhoI com o inserto purificado e com cada um dos três vetores de expressão pETM30-MBP que possui 6947 pb (1), pET32a que possui 5900 pb (2) e pET22b que possui 5493 pb (3), os vetores na mesma ordem apenas com NcoI (4, 5 e 6) e os vetores na mesma ordem apenas com XhoI (7, 8 e 9) P: Padrão de peso molecular.

Após a confirmação da digestão foi realizada a purificação dos produtos digeridos, e, em seguida, a reação de ligação entre o inserto e cada um dos vetores. Os produtos obtidos após a reação foram transformados em cepas de *E. coli* DH5α e plaqueadas em LB ágar e o antibiótico específico para cada uma das construções. As colônias que cresceram isoladas na placa foram propagadas em meio LB líquido seletivo e os plasmídeos foram extraídos por Miniprep. Ao todo foram isoladas sete colônias diferentes de cada uma das construções para a análise da subclonagem. Para a realização de uma PCR diagnóstica foi retirada uma alíquota

de cada um dos plasmídeos extraídos após subclonagem, e estes, foram digeridos com as endonucleases de restrição, NcoI e XhoI, para verificar a inserção do gene que codifica a DSM-14977. Embora sete colônias de cada construção tenham sido testadas, somente cinco das construções pET22b-DSM-14977 e pET32a-DSM-14977, e apenas duas da pETMBP-DSM-14977 mostraram resultado positivo da inserção do gene (Figura 17).

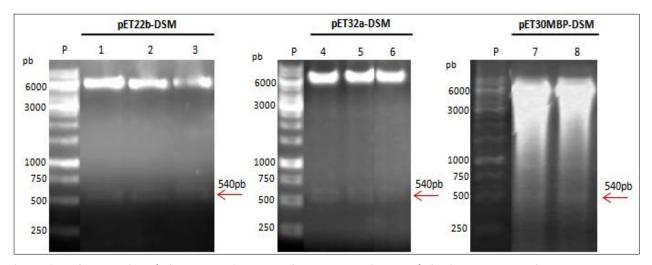


Figura 17: Digestão diagnóstica dos DNAs recombinantes a partir de colônias isoladas em meio de cultura seletivo.

Um diagnóstico por PCR foi realizado para analisar a inserção do gene em cada plasmídeo, ao todo sete colônias isoladas de cada construção foram testadas. Pode-se observar, que após a digestão com as enzimas NcoI e XhoI, ocorre o aparecimento de uma banda do tamanho esperado para o gene (seta). P: Padrão de peso molecular, 1, 2, 3, 4, 5, 6, 7 e 8 : colônias positivas para a inserção do gene no plasmídeo.

Por fim, o DNA recombinate positivo foi enviado ao laboratório SONDA-UFRJ para sequenciamento, confirmando a subclonagem da enzima putativa DSM-14977 em cada uma das construções pET22b-DSM-14977 (DSM-14977), pET32a-DSM-14977 (TRX-DSM-14977) (Figura 18) e pETMBP-DSM-14977 (MBP-DSM-14977) (dado não mostrado).

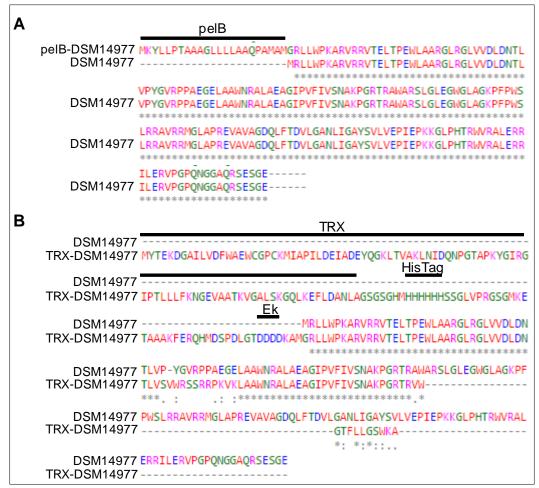


Figura 18: Sequenciamento do DNA recombinante.

O sequenciamento do DNA recombinante foi realizado a partir das colônias positivas das construções pET22b-DSM-14977 (A) e pET32a-DSM-14977 (B). Um alinhamento foi realizado entre a sequência primária da DSM-14977 nativa e cada uma das construções sequenciadas. É possível observar a sequência que corresponde o peptídeo sinal pelB e a exata sequência de aminoácidos da DSM-14977 na construção pET22b-DSM-14977. Já na pET32a-DSM-14977 é mostrada a sequência da TRX, a cauda de hisdina e a sequência de aminoácidos da DSM-14977.

# 4.3 EXPRESSÃO DA ENZIMA PUTATIVA DSM-14977 NAS DIFERENTES CONSTRUÇÕES

Os ensaios de expressão da fosfatase putativa DSM-14977 foram realizados para identificar as melhores condições de expressão da enzima, nas construções, MBP-DSM-14977, TRX-DSM-14977 e pelB-DSM-14977. Para isso, alguns parâmetros foram variados como a temperatura, o tempo de indução e a cepa de *E.coli* transformada. Todos os ensaios foram realizados com a indução da expressão da proteína com 1 mM de IPTG em DO<sub>600nm</sub> 0.8. As alíquotas foram retiradas ao longo do tempo para a análise da expressão através de SDS-PAGE.

### 4.3.1 Expressão da contrução MBP-DSM-14977

Os ensaios de expressão com a MBP-DSM-14977 foram avaliados quanto ao enriquecimento de uma banda na massa molecular esperada, aproximadamente 64 kDa, ao longo do tempo de expressão, nas temperaturas testadas, utilizando diferentes cepas de *E. coli* (Figura 19). O melhor resultado obtido foi observado para a cepa de *E. coli*, BL21(DE3), onde foi possível observar a superexpressão da enzima em todas as temperaturas testadas (37°C, 27°C e 17°C) (Figura 19 A-C). No entanto, a melhor expressão foi observada a 37°C após indução por 4 horas da proteína alvo fusionada à MBP (Figura 19 A). Para a cepa de *E. coli* C43-pLysS não foi observado um aumento da expressão significativo da proteína de interesse em nenhuma das duas temperaturas testadas (37°C e 27°C) (Figura 19 D-E). Cabe ressaltar que não houve crescimento bacteriano para esta cepa a 17°C (Figura 19 F). A expressão da DSM-14977 também não foi observada para as cepas de *E. coli* BL21(DE3)pLysS e rosetta gami, em nehuma das temperaturas testadas.

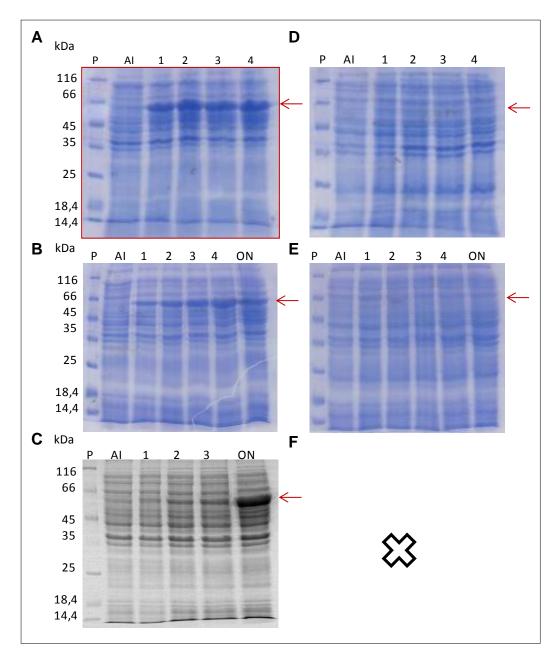


Figura 19: SDS-PAGE 12% dos testes de expressão da construção MBP-DSM-14977 em diferentes condições.

A expressão desta construção foi testada nas cepas de *E.coli* BL21(DE3) nas temperaturas de 37°C (A), 27°C (B) e 17°C (C) e C-43pLysS as 37°C (D), 27°C (E) e 17°C (F) utilizando as diferentes e diversos tempos de indução. O melhor resultado para a expressão da MBP-DSM-14977 foi obtido em BL21 (DE3) a 37°C, 1 mM de IPTG em D.O<sub>600</sub> 0.8 (borda vermelha). P: Padrão de peso molecular; alíquotas retiradas: AI: antes da indução da expressão, 1: 1 h após indução, 2: 2 h após indução, 3: 3 h após indução, 4: 4 h após indução, ON: 16 h após indução. A seta evidencia a expressão da proteína de interesse ao longo do tempo, com a massa molecular esperada. A borda vermelha demonstra o melhor resultado encontrado.

#### 4.3.2 Expressão da contrução TRX-DSM-14977

Para os ensaios com a construção TRX-DSM-14977 foi possível observar o enriquecimento da banda ao longo do tempo de indução, em aproximadamente, 34 kDa, referente a massa molecular esperada, para a DSM-14977 nas diferentes temperaturas testadas. Para a cepa de *E. coli* C43-pLysS não foi observado um aumento significativo da

expressão da proteína de interesse em nenhuma das temperaturas testadas (37 °C e 27 °C e 17 °C) (Figura 20 D-F). A expressão da DSM-14977 também não foi observada para as cepas de *E. coli* BL21(DE3)pLysS e rosetta gami. A cepa BL21(DE3), indicou a superexpressão da proteína alvo fusionada à TRX a 37, 27 e 17 °C (Figura 20 A-C, respectivamente). Entretanto, a maior superexpressão da TRX-DSM-14977 foi observada para a cepa BL21(DE3), evidenciando um melhor resultado a 37°C após a indução por 4 horas (Figura 20 A). As cepas de *E. coli* BL21(DE3)pLysS e rosetta gami não apresentaram crescimento bacteriano para a construção TRX-DSM-14977.

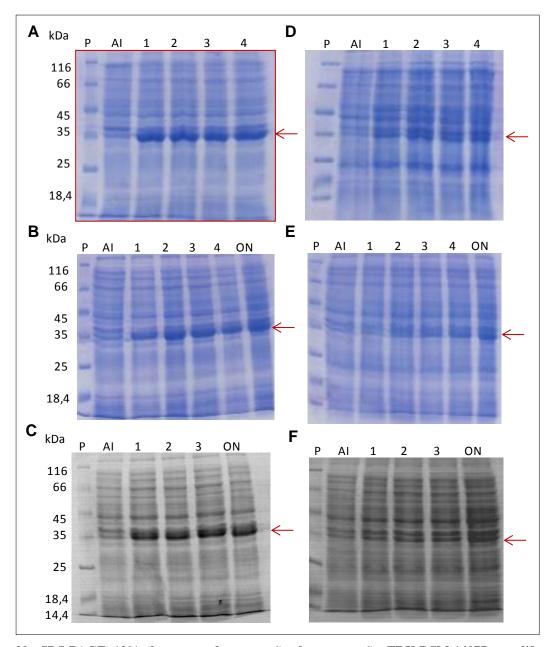


Figura 20: SDS-PAGE 12% dos testes de expressão da construção TRX-DSM-14977 em diferentes condições.

A expressão desta construção foi testada em diferentes cepas de *E.coli* (colunas) BL21(DE3) e C-43pLysS utilizando as diferentes temperaturas de 37°C (A), 27°C (B) e 17°C (C) e C-43pLysS as 37°C (D), 27°C (E) e 17°C (F). O melhor resultado para a expressão da TRX-DSM-14977 foi obtido em BL21 (DE3) a 37°C, 1 mM de IPTG em D.O<sub>600</sub> 0.8 (borda vermelha). P: Padrão de peso molecular; alíquotas retiradas: AI: antes da indução da expressão, 1: 1 h após indução, 2: 2 h após indução, 3: 3 h após indução, 4: 4 h após indução, ON: 16 h após indução. A seta evidencia a expressão da proteína de interesse ao longo do tempo, com a massa molecular esperada. A borda vermelha demonstra o melhor resultado encontrado.

#### 4.3.3 Expressão da contrução pelB-DSM-14977 (DSM-14977)

Para os ensaios com a construção DSM-14977 foi possível observar o enriquecimento da banda de interesse, com massa molecular esperada, de aproximadamente 20 kDa, ao longo do tempo de indução, nas diferentes temperaturas e cepas de *E. coli* testadas (Figura 21). As

cepas C43-pLysS (Figura 21 D-F), rosetta gami (Figura 21 G-I) e, sobretudo, BL21(DE3) (Figura 21 A-C), mostraram uma boa expressão da DSM-14977. A melhor condição de expressão observada para esta construção foi em BL21(DE3) a 37°C após 4 horas de indução (Figura 21 A). A cepa BL21(DE3)pLysS não apresentou crescimento bacteriano para a proteína de interesse (Figura 21).

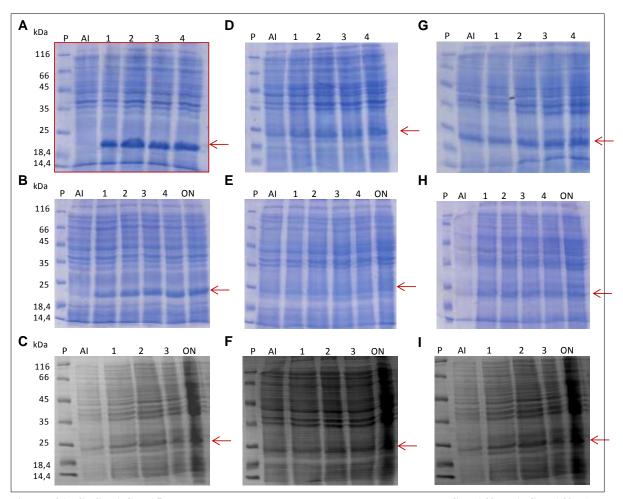


Figura 21: SDS-PAGE 15% dos testes de expressão da construção pelB-DSM-14977 (DSM-14977) em diferentes condições.

A expressão desta construção foi testada em diferentes cepas de *E.coli* (colunas) BL21(DE3), C-43pLysS e BL21(DE3)pLysS respectivamente, utilizando as diferentes temperaturas de 37°C (A,D G), 27°C (B, E e H) 17°C de (C, F, I), e diversos tempos de indução. O melhor resultado para a expressão da DSM-14977 foi obtido em BL21 (DE3) a 37°C, 1 mM de IPTG em D.O<sub>600</sub> 0.8 (borda vermelha). P: Padrão de peso molecular; alíquotas retiradas: AI: antes da indução da expressão, 1: 1 h após indução, 2: 2 h após indução, 3: 3 h após indução, 4: 4 h após indução, ON: 16 h após indução. A seta evidencia a expressão da proteína de interesse ao longo do tempo, com a massa molecular esperada. A borda vermelha demonstra o melhor resultado encontrado.

Portanto, a melhor expressão para todas as três construções foi alcançada utilizando a cepa BL21(DE3) a 37°C por 4 horas após a indução com 1 mM de IPTG em DO<sub>600</sub> 0.8. É importante ressaltar que as amostras em cada gel estão normalizadas por volume em cada poço, permitindo análise quantitativa dos mesmos.

# 4.4 EXTRAÇÃO CELULAR

Após a análise das melhores condições de expressão das construções MBP-DSM-14977 e TRX-DSM-14977 foi necessário identificar, se após a lise celular, a proteína estava na fração solúvel ou insolúvel. Para isto, as células crescidas nas melhores condições foram lisadas por sonicação, e suas frações solúvel e insolúvel, separadas por centrifugação (Figura 22). Uma porção satisfatória da proteína de interesse foi encontrada na fração solúvel.

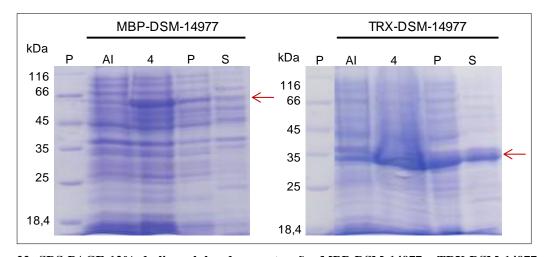


Figura 22: SDS-PAGE 12% da lise celular das construções MBP-DSM-14977 e TRX-DSM-14977 após a expressão.

A expressão da MBP-DSM-14977 e TRX- DSM-14977 foi obtida em BL21 (DE3) a  $37^{\circ}$ C, 1 mM de IPTG em D.O<sub>600</sub> 0.8. A lise celular foi realizada através do protocolo de sonicação onde é possível observar a porção solúvel e insolúvel do lisado, sendo possível observar a proteína de interesse presente na fração solúvel, ou seja, no sobrenadante. P: Padrão de peso molecular; AI: antes de induzir, 4: 4 h após induzir, P: Pellet (insolúvel) e S: Sobrenadante (solúvel). As setas vermelhas indicam a banda no gel referente a massa molecular esperada para a DSM-14977 em cada construção.

A princípio, para verificar a obtenção da proteína de interesse, presente na construção pelB-DSM-14977 (DSM-14977) foram realizados quatro protocolos diferentes de choque osmótico, a fim de extrair, somente, o conteúdo periplasmático (Figura 23A). Isto é necessário devido à presença do peptídeo sinal pelB que possibilita a exportação da DSM-14977 para este compartimento celular. Dentre os protocolos utilizados, o que se mostrou mais eficiente, foi o segundo, o qual emprega a lisozima em um dos passos de extração, assim, a proteína DSM-14977 foi obtida na fração solúvel (Figura 23A). Para investigar se a presença da lisozima era necessária para a extração solúvel realizamos outro ensaio através de um quinto protocolo de choque osmótico utilizando 50, 100 ou 500 μg/mL de lisozima (Figura 23 B). Com 50 μg/mL de lisozima já foi possível observar a presença da DSM-14977 na fração solúvel. Isto indica que a proteína de interesse provavelmente não foi endereçada corretamente para o compartimento programado visto que precisaria da lisozima, que é uma

hidrolase de polissacarídeos da parede celular de bactérias, para ser extraída de dentro da célula. Portanto, com a necessidade da lise celular, utilizamos o mesmo protocolo de lise por sonicação das outras construções para retirar a DSM-14977 que foi observada majoritariamente na fração solúvel (Figura 23 C).

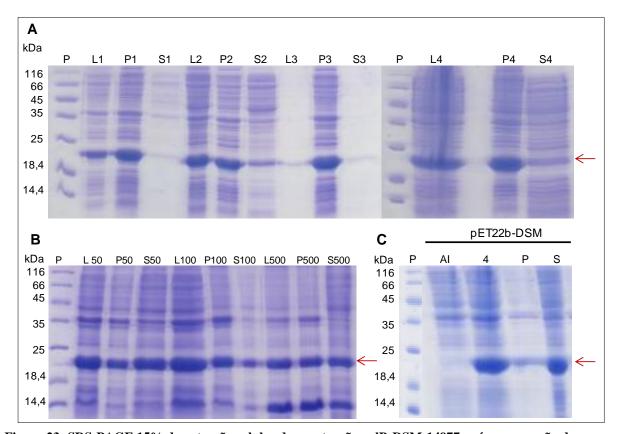


Figura 23: SDS-PAGE 15% da extração celular da construção pelB-DSM-14977 após a expressão da proteína.

A expressão da pelB-DSM-14977 foi obtida em BL21 (DE3) a 37°C, 1 mM de IPTG em D.O<sub>600</sub> 0.8. Para a extração celular da DSM-14977 foram utilizados quatro diferentes protocolos, sendo, o segundo deles, o mais eficiente, consistindo na utilização de lisozima (A). Outro protocolo de choque osmótico empregado fez uso de lisozima nas seguintes concentrações: 50, 100 ou 500 μg/mL (B). Com indícios da necessidade da lise celular, o protocolo de lise por sonicação foi utilizado onde a DSM-14977 foi observada, predominantemente, na fração solúvel (C). Isto sugere que a proteína de interesse não foi endereçada corretamente para o periplasma. P: Padrão de peso molecular; L1, L2, L3 e L4: lisado celular do protocolo 1, 2, 3 e 4, respectivamente; P1, P2, P3 e P4: pellet celular (insolúvel) do protocolo 1, 2, 3 e 4; S1, S2, S3 e S4: sobrenadante (solúvel) do protocolo 1, 2, 3 e 4, respectivamente; L50, 100 ou 500, P50, 100 ou 500 e S50, 100 ou 500: lisado, pellet e sobrenadante com 50, 100 ou 500 μg/mL de lisozima, respectivamente; AI: antes de induzir, 4: 4 h após induzir, P: Pellet, S: Sobrenadante.

Com a constatação da superexpressão solúvel das três construções foi possível dar prosseguimento aos ensaios de purificação de cada uma delas.

## 4.5 PURIFICAÇÃO DA CONSTRUÇÃO MBP-DSM-14977, TRX-DSM-14977 e DSM-14977

Os ensaios de purificação das proteínas solúveis MBP-DSM-14977, TRX- DSM-14977 e DSM-14977 foram realizados a partir da expressão das proteínas em 1L de LB utilizando a cepa BL21(DE3) a 37°C, induzida com 1mM de IPTG em DO<sub>600</sub> 0.8, e induzida por 4 horas. Após a lise celular por sonicação, as amostras presentes na fração solúvel, foram aplicadas nas diferentes cromatografias.

A purificação das construções MBP-DSM-14977 e TRX-DSM-14977 foi realizada através da cromatografia de afinidade a níquel, utilizando a coluna His-Trap (GE), em sistema Äkta-prime, com eluição realizada através de gradiente crescente de imidazol (Figuras 24 e 25, respectivamente).

No cromatograma da purificação da construção MBP-DSM-14977 pode ser observado um primeiro pico que corresponde às proteínas contaminantes de *E.coli* (fração entre 194 - 272 mM de imidazol). O segundo pico cromatográfico é coincidente com a fração 280 - 420 mM de Imidazol, na qual, a construção MBP-DSM-14977 foi eluída e se apresentou bastante enriquecida, como é possível observar no gel SDS-PAGE 12% (Figura 24).

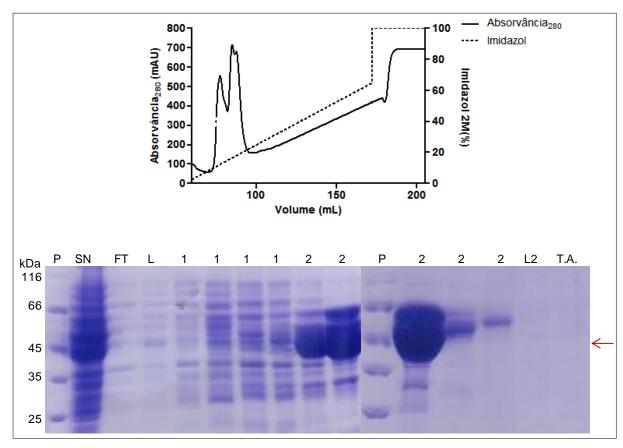


Figura 24: Cromatograma da purificação e SDS-PAGE 12 % da purificação da proteína MBP-DSM-14977.

A purificação foi feita utilizando uma coluna His-trap em sistema Äkta-prime em tampão A (50 mM Tris-HCl, pH 7,2, 150 mM de NaCl, 5mM imidazol). A injeção foi realizada com 25 mL do sobrenadante da lise. A lavagem com tampão A para a retirada de ligantes inespecíficos e a eluição das proteínas foi feita com um

gradiente linear crescente de imidazol de 0 a 1 M. Em seguida, a coluna foi lavada com 100 % do tampão B (Tampão Tris-HCl 50 mM, pH 7,2, 150 mM NaCl e 2 M de imidazol). P: Padrão de peso molecular; SN: sobrenadante injetado, FT: *Flow-through*; L: Lavagem; L2: lavagem com 2 M de imidazol, TA: controle com tampão de amostra. 1: 1º pico cromatográfico: 194-272 mM de imidazol e 2: 2º pico cromatográfico: 280-420 mM imidazol. A seta vermelha indica a banda de proteína na massa molecular esperada.

No cromatograma da purificação da construção TRX-DSM-14977 foi observado um primeiro pico cromatográfico, correspondente às proteínas contaminantes de *E.coli*, observado na fração entre 84 e 118 mM de imidazol. O segundo pico cromatográfico corresponde à fração 214-356 mM de imidazol, na qual a construção TRX-DSM-14977 foi eluída e se apresentou bastante enriquecida e homogênea, como pode ser observado no gel SDS-PAGE 12% (Figura 25).

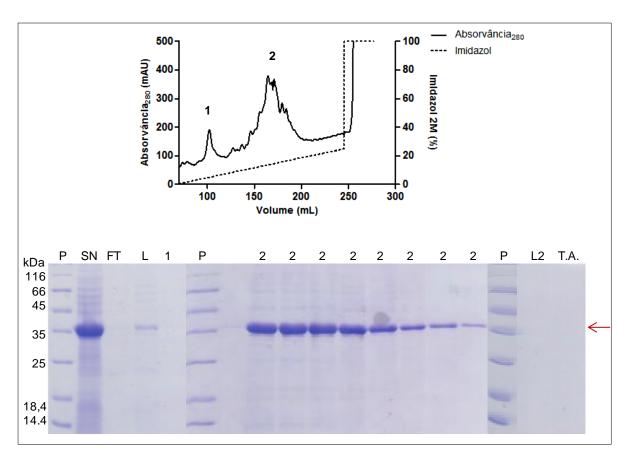


Figura 25: Cromatograma da purificação e SDS-PAGE 12 % da purificação da proteína TRX-DSM-14977

A purificação foi feita utilizando uma coluna His-trap em sistema Äkta-prime em tampão A (50 mM Tris-HCl, pH 7,2, 150 mM de NaCl e 5mM imidazol). A injeção foi realizada com 25 mL do sobrenadante da lise. A lavagem com tampão A para a retirada de ligantes inespecíficos e a eluição das proteínas foi feita com um gradiente linear crescente de imidazol de 0 a 0,5 M. Em seguida, a coluna foi lavada com 100 % do tampão B (Tampão Tris-HCl 50 mM, pH 7,2, 150 mM NaCl e 2 M de imidazol). P: Padrão de peso molecular; SN: sobrenadante injetado, FT: *Flow-through*; L: Lavagem; L2: lavagem com 2 M de imidazol, TA: controle com tampão de amostra. 1: 1º pico cromatográfico: 84-118 mM de imidazol e 2: 2º pico cromatográfico: 214-356 mM imidazol. A seta vermelha indica a banda de proteína na massa molecular esperada.

Para os ensaios de purificação da construção pelB-DSM-14977, as propriedades bioquímicas da proteína alvo foram exploradas e diferentes métodos de fracionamento foram aplicados, como: a cromatografia de troca catiônica (*HiPrep Sepharose FastFlow 16/10* e *SP 650M*) e a cromatografia hidrofóbica (*HiTrap Butyl FF* e *HiTrap phenyl FF*) utilizando diferentes resinas e diferentes protocolos (dados não mostrado).

Após diversos protocolos testados de purificação, esta foi alcançada através da utilização de dois passos: a cromatografia hidrofóbica (coluna *Hiprep butyl FF 16/10*) seguida da cromatografia de afinidade a heparina (coluna *Hi-Trap heparin HP*) em sistema Äktaprime com eluição realizada através de gradiente decrescente e crescente de sal, respectivamente (Figura 26 e 27).

No primeiro passo de purificação da DSM-14977 um primeiro relevo cromatográfico condiz com as proteínas contaminantes de *E.coli*, que foi observado na fração entre 1,6 M-220 mM de sal, durante o gradiente decrescente de sal. A DSM-14977 foi eluída em dois picos, um correspondente à fração 160 mM - 0 M de sal e, em um outro, na lavagem da coluna com água, o que demonstra uma forte ligação hidrofóbica como observado no gel SDS-PAGE 15% (Figura 26).

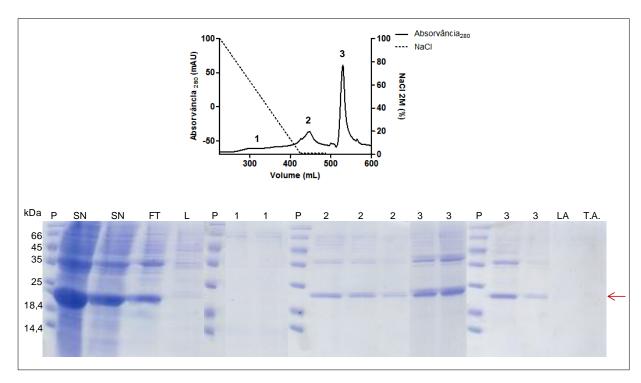


Figura 26: Cromatograma e SDS-PAGE 15 % do primeiro passo de purificação da construção DSM-14977 (pelB-DSM-14977).

A purificação foi feita utilizando coluna Hiprep butyl FF 16/10 em sistema Äkta-prime em tampão A (50 mM Tris-HCl, pH 8,0, 2 M de NaCl). A injeção foi realizada com 12,5 mL do sobrenadante da lise. A lavagem foi feita com tampão A para a retirada de ligantes inespecíficos e a eluição das proteínas realizada com um gradiente linear decrescente de sal de 2 a 0 M. Em seguida, a coluna foi lavada com 100% do tampão B (tampão Tris-HCl

50 mM, pH 8,0). P: Padrão de peso molecular; SN: sobrenadante injetado, FT: Flow-through; L: Lavagem; LA: lavagem com água, T.A.: controle com tampão de amostra. 1: 1º pico cromatográfico (1.6 M - 220 mM de sal), 2: 2º pico cromatográfico (160 mM - 0 M de sal) e 3: 3º pico cromatográfico (lavagem com água). A seta mostra a banda relativa a massa molecular esperada para a DSM-14977.

A cromatografia de afinidade a heparina foi realizada como um segundo passo de purificação da DSM-14977. No cromatograma é possível observar um "ombro" na fração entre 520 mM - 600 mM de sal, correspondente as proteínas contaminantes de *E.coli*. A DSM-14977 foi eluída em dois picos, um equivalente à fração 620 - 920 mM de sal e, em um outro, na fração 1 – 1.5 M de sal, mostrando que a DSM-14977 se liga com alta afinidade ao componente da coluna como pode ser observado no gel SDS-PAGE 15% (Figura 27).

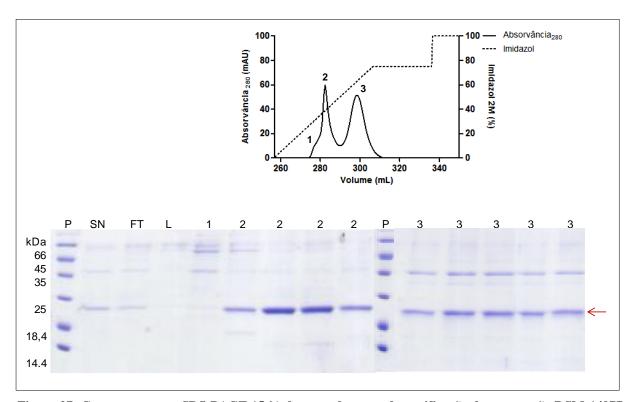


Figura 27: Cromatograma e SDS-PAGE 15 % do segundo passo de purificação da construção DSM-14977 (pelB-DSM-14977).

A purificação foi feita utilizando coluna Hitrap-heparin HP em sistema Äkta-prime em tampão A (10 mM Tris-HCl, pH 8,0). A injeção foi realizada com 70 mL da amostra obtida na cromatografia hidrofóbica. A lavagem foi realizada com tampão A para a retirada de ligantes inespecíficos e a eluição das proteínas feita com um gradiente crescente de sal de 0 a 2 M. Em seguida, a coluna foi lavada com 100% do tampão B (10 mM Tris-HCl, pH 8,0 e 2 M de NaCl). P: Padrão de peso molecular; SN: sobrenadante injetado, FT: Flow-through; L: Lavagem. 1: 1º pico cromatográfico (520 - 600 mM de sal), 2: 2º pico cromatográfico (620 - 920 mM de sal) e 3: 3º pico cromatográfico (1 - 1.5 M de sal). A seta mostra a banda relativa a massa molecular esperada para a DSM-14977.

As frações do pico 2 foram utilizadas paras os ensaios de caracterização bioquímica e estrutural da DSM-14977, visto que a proteína de interesse se encontra pura e na forma livre,

sem a necessidade de passos de clivagem e de uma nova purificação, como seria o caso para as construções MBP-DSM-14977 e TRX- DSM-14977.

Desta forma, todos os ensaios obtidos a partir daqui foram realizados com a proteína DSM-14977 oriunda da construção pelB-DSM-14977.

#### 4.6 ESTUDOS ESTRUTURAIS DA FOSFATASE PUTATIVA DSM-14977

Os estudos estruturais de proteínas são de grande importância, pois o enovelamento correto está intimamente ligado à função biológica desempenhada pela enzima. Portanto, sua caracterização estrutural por técnicas espectroscópicas, como, o dicroísmo circular (CD) e fluorescência permitem o monitoramento do conteúdo estrutural secundário e terciário de proteínas, respectivamente, permitindo avaliar a estabilidade da enzima frente à diferentes condições.

#### 4.6.1 Dicroísmo circular

O conteúdo da estrutura secundária da fosfatase putativa DSM-14977 foi analisado por dicroísmo circular (CD). O espectro de CD da DSM-14977 evidenciou uma mistura de  $\alpha$ -hélice e folha- $\beta$ , com um pico positivo próximo a 195 nm e um vale em 208 e 222 nm, condizente com o esperado para fosfatases pertencentes à superfamília das dealogenases haloácidas, mostrando que a proteína encontra-se estruturada após a purificação. Os ensaios foram realizados com a DSM-14977 à 10  $\mu$ M em tampão Tris 10 mM em pH 8,0, com o intuito de identificar se a proteína DSM-14977 estava (Figura 28).

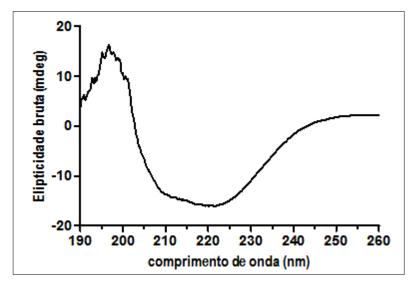


Figura 28: Espectro de dicroísmo circular da fosfatase putativa DSM-14977.
Os ensaios foram realizados com tampão 10 mM Tris-HCl pH 8. O espectro obtido apresenta o perfil de estrutura secundária da DSM-14977 característico de α-hélice com dois vales em 222 e 208 nm e um pico próximo a 195 nm, característico de folha-β.

### 4.6.2 Espectroscopia de fluorescência

Para estudar o comportamento estrutural terciário da fosfatase putativa DSM-14977 foram realizados ensaios de fluorescência intrínseca do triptofano frente a agentes desnaturantes físicos (temperatura) e químicos (ureia e cloridrato de guanidina). Estes resíduos são sondas intrínsecas, geralmente, presentes em centros hidrofóbicos proteicos, que podem ser monitorados através da excitação a 280 nm de comprimento de onda. O espectro de emissão de fluorescência do triptofano muda conforme a polaridade do meio. Quando em ambiente apolar ele emite em comprimentos de onda menores, deslocados para o azul. Já em ambiente polar, o triptofano emite em comprimento de onda maiores, deslocados para o vermelho. Desta forma, este aminoácido pode ser utilizado como sonda intrínseca para os estudos de enovelamento proteico. Os espectros de fluorescência podem ser avaliados quanto ao centro de massa espectral, que leva em consideração a intensidade e a área do pico estudado. A partir das intensidades máximas de emissão de fluorescência, os cálculos de centro de massa foram realizados em função da concentração do agente desnaturante utilizado.

As amostras de DSM-14977, na concentração de 1µM, utilizadas para os ensaios de estabilidade foram incubadas, por 16 h, em tampão Tris-HCl 10 mM, pH 8,0, em diferentes concentrações de agentes desnaturantes.

O comportamento estrutural da DSM-14977 na presença de 0 a 8 M de ureia evidenciou que a proteína perde gradativamente sua estrutura, mostrando centro de massa deslocado para comprimentos de onda maiores com o aumento da concentrações de ureia. Cabe ressaltar, que o espectro de emissão de fluorescência já se inicia um pouco deslocado para comprimentos de onda maiores, em torno de 346 nm, isto indica, que parte dos sete triptofanos existentes na estrutura primária da proteína, podem estar expostos ao solvente levando a este deslocamento do centro de massa espectral. Entretanto, com o aumento da concentração de ureia, é possível observar maior exposição destes triptofanos, indicando uma tendência ao desenovelamento em função do aumento da concentração de ureia, principalmente, após 3 M deste agente desnaturante (Figura 29).

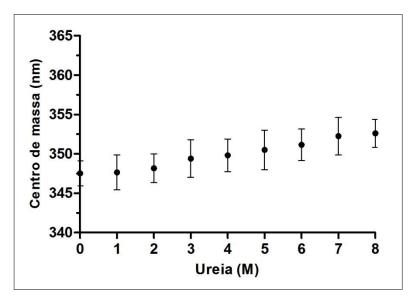


Figura 29: Estabilidade da DSM-14977 na presença de diferentes concentrações de ureia através de fluorescência.

A estabilidade estrutural da proteína DSM-14977 foi acompanhada através do monitoramento de mudanças no centro de massa espectral. O aumento da concentração de ureia, especificamente após 3 M, leva a um deslocamento do centro de massa para valores maiores de comprimento de onda, indicando que há uma tendência à perda estrutural. Cabe ressaltar, que em 0 M de ureia a proteína já possui centro de massa deslocado para o vermelho, indicando a presença de triptofanos expostos ao solvente. A análise foi realizada com 1 μM de proteína em tampão Tris-HCl 10 mM, pH 8,0. Todos os experimentos foram realizados em triplicata.

Os estudos de estabilidade da DSM-14977 também foram realizados na presença de cloridrato guanidina (0 a 6 M) através de fluorescência (Figura 30). A avaliação do centro de massa espectral da proteína indica a perda completa de sua estrutura a partir de 2,0 M deste agente desnaturante, mostrando uma forte alteração estrutural, logo nas primeiras concentrações de desnaturante testadas.

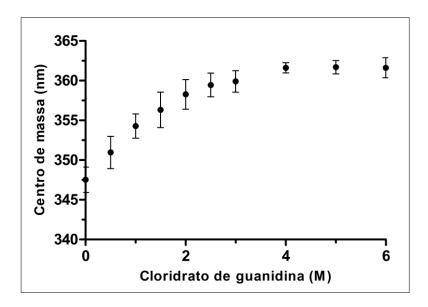
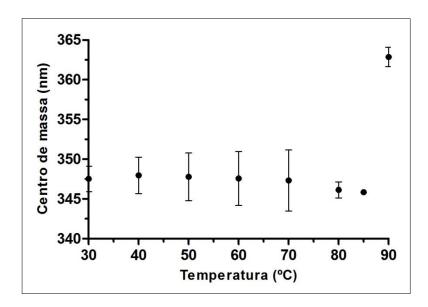


Figura 30: Estabilidade da DSM-14977 na presença de diferentes concentrações de cloridrato de guanidina através de fluorescência.

A estabilidade estrutural da proteína DSM-14977 foi acompanhada através do monitoramento de mudanças no centro de massa espectral. O aumento da concentração de cloridrato de guanidina, especificamente após 0,5 M, leva a um grande deslocamento do centro de massa, para comprimentos de onda maiores, indicando que há uma grande perda estrutural. Cabe ressaltar, que em 0 M de cloridrato de guanidina a proteína já possui centro de massa deslocado para o vermelho, indicando a presença de triptofanos expostos ao solvente. A análise foi realizada na presença de 1  $\mu$ M de proteína em tampão Tris-HCl 10 mM, pH 8,0. .Todos os experimentos foram realizados em triplicata.

Para avaliar a termoestabilidade da DSM-14977 foi realizado ensaios de fluorescência intrínseca do triptofano em diferentes temperaturas: 30, 40, 50, 60, 70, 80, 85 e 90°C. Neste ensaio, o monitoramento do centro de massa espectral mostrou que a proteína é altamente termoestável até 85°C. No entanto, a proteína perde totalmente sua estrutura quando a temperatura atinge 90°C, indicando que desestruturação, após 85°C, é muito drástica (Figura 31).



**Figura 31: Termoestabilidade da DSM-14977 através de fluorescência.**A termoestabilidade da proteína DSM-14977 foi acompanhada através do monitoramento de mudanças em seu

A termoestabilidade da proteina DSM-14977 foi acompanhada atraves do monitoramento de mudanças em seu centro de massa espectral em diferentes temperaturas. A proteína possui alta termoestabilidade mostrando-se estável até 85°C. Entretanto, a proteína perde completamente sua estrutura quando a temperatura atinge 90°C. A análise foi realizada na presença de 1 µM de proteína em tampão Tris-HCl 10 mM, pH 8,0. Todos os experimentos foram realizados em triplicata.

Os dados referentes à estabilidade estrutural da DSM-14977 na presença de agentes desnaturantes indicaram que a proteína é bem mais estável em ureia do que em cloridrato de guanidina. Esta proteína possui uma forte tendência à formação de oligômeros em solução. Provavelmente, estes oligômeros são rapidamente desfeitos na presença de guanidina, devido ao seu efeito drástico em proteínas, sendo este efeito minimizado na presença de ureia.

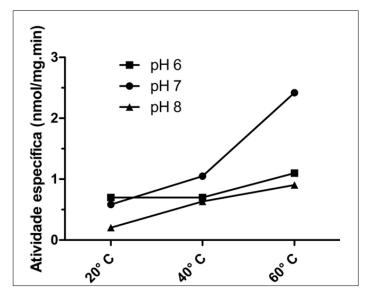
A DSM-14977 mostrou-se altamente termoestável condizente com o organismo termófilo ao qual esta proteína pertence. Cabe ressaltar, que os resultados obtidos também estão de acordo com a possível oligomerização da enzima, geralmente observada para organismos termófilos, o que poderia explicar a drástica perda estrutural após 85°C.

#### 4.7 estudos funcionais da fosfatase putativa dsm-14977

#### 4.7.1 Dosagem da atividade enzimática

As enzimas da superfamília HAD que já foram caracterizadas, possuem dependência à metal, sendo este, necessário para sua atividade enzimática (Proudfoot et al, 2004; Godinho et al, 2011, Kuznetsova et al, 2004). Desta forma, como a fosfatase putativa DSM-14977 foi prospectada, possui função desconhecida, e como não temos evidências da melhor condição de atividade, foi necessário fazermos alguns testes para sua caracterização enzimática.

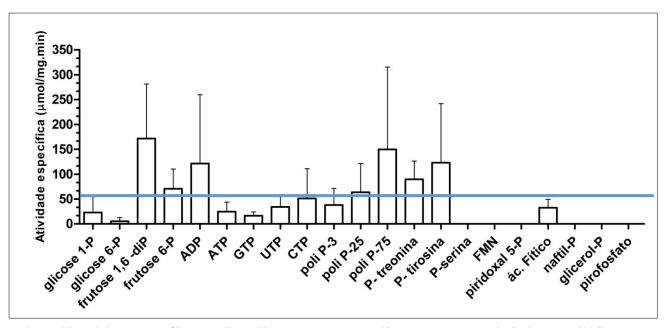
Inicialmente, utilizamos o ρ-nitrofenil-fosfato, substrato artificial comumente utilizado na caracterização de fosfatases, para testar a atividade fosfatásica da DSM-14977. Para isto, foram testadas três temperaturas (20, 40 e 60°C) e pHs (6, 7 e 8) diferentes na presença de 5 mM de MgCl<sub>2</sub> (Figura 32).



**Figura 32: Atividade fosfatásica da proteína DSM-14977 em diferentes condições.** A atividade fosfatásica da fosfatase putativa DSM-14977 foi verificada utilizando o p-nitrofenil-fosfato que é um substrato geral catalisado por fosfatases utilizando temperaturas (20, 40 e 60°C) e pHs (6, 7 e 8) diferentes. A concentração da proteína e substrato para este ensaio foi de 0,5 mg/mL e 2,5 mM,, respectivamente.

Observamos que houve uma pequena atividade em todas as condições testadas, entretanto, em pH 7,0 a 60°C, a atividade dobrou quando comparada com as demais. Com este resultado, foi possível realizar novos ensaios enzimáticos utilizando diferentes substratos comuns às fosfatases.

Com o intuito de identificar o substrato ótimo para esta fosfatase, realizamos a triagem de diferentes substratos através do monitoramento da produção de fosfato livre usando verde malaquita. Assim, a atividade da fosfatase do DSM-14977 purificada foi testada e avaliada em condições padrão como descrito na seção materiais e métodos utilizando substratos fisiológicos, como: ribo- e nucleosídeos tri- e difosfatos, fosfatos de açúcar e outros metabólitos fosforilados, bem como, polifosfatos (Figura 33).



**Figura 33: Atividade específica da DSM-14977 na presença de diferentes substratos fosforilados a 40°C.** Neste ensaio, foram utilizados substratos naturais e artificiais para caracterizações de fosfatases através do monitoramento da produção de fosfato livre medido pelo método verde de malaquita. Os ensaios foram realizados com aproximadamente 0,06 mg/mL de DSM-14977 em tampão Tris-HCl 10 mM, pH 8,0 na presença de 5 mM de MgCl<sub>2</sub> por 30 minutos. A concentração de cada substrato foi realizada conforme o item 3.2.11. Acima da linha azul podemos observar os substratos com resultados mais significativos e os resultados abaixo da linha azul foram desconsiderados.

A DSM-14977 mostrou o mesmo padrão de catálise para os substratos testados nas replicatas, evidenciando uma maior atividade fosfohidrolásica para os substratos: frutose 1,6 difosfato, ADP, poli P-75 e fosfotirosina, e uma menor atividade para os substratos: frutose 6P, CTP, polifosfato-3, polifosfato-25 e fosfotreonina. Embora o desvio padrão tenha ficado grande para alguns substratos, aqueles supracitados foram os que ficaram acima da média. A DSM-14977 não mostrou atividade detectável para fosfoserina, FMN, piridoxal 5-fosfato, naftil-fosfato, glicerol-fosfato e pirofosfato.

#### 4.8 MODELAGEM MOLECULAR DA FOSFATASE DSM-14977

Para explorar os dados de caracterização bioquímica e estrutural da fosfatase DSM-14977, utilizamos a ferramenta BLAST da plataforma NCBI, para buscar proteínas com sequências similares, no intuito de propor um modelo estrutural por homologia para a DSM-14977. No entanto, não houve resultados com algo grau de homologia de sequência primária com as estruturas depositas no PDB (*Protein data bank*), desta forma, a utilização de programas computacionais simples para a modelagem automática não serviria ao nosso propósito.

Sendo assim, utilizamos o programa I-TASSER, que a partir da sequência de aminoácidos da DSM-14977, é capaz de gerar automaticamente, predições de modelos estruturais de alta qualidade e função biológica para a proteína de interesse. Este programa utiliza 10 diferentes algoritmos de comparação de sequências (*profile-profile*): FFAS-3D, HHsearch, MUSTER, pGenTHREADER, PPAS, PRC, PROSPECT2, SP3, e SPARKS-X.

Estes algoritmos representam a sequência de interesse e as do banco de dados por um perfil, também chamados de *position-secific scoring sequences* (PSSMs). Os perfis são calculados através do múltiplo alinhamento de sequências relacionadas que foram coletadas. O *profile* é uma matriz contendo, para cada posição da sequência de estudo, um *score* de similaridade para os 20 aminoácidos. Estes *scores* são calculados de acordo com as frequências de cada resíduo de aminoácido em cada posição em múltiplos alinhamentos de sequência. Como os *profiles* contêm muito mais informações do que um simples alinhamento, os métodos *profile-profile* são muito mais sensíveis do que métodos de comparação sequência-sequência, como por exemplo, o BLAST, ou profile-sequência, como o PSI-BLAST. Assim, cada um dos 10 algoritmos coleta 10 alinhamentos gerando os 100 melhores alinhamentos gerais. O I-TASSER então ranqueia os 10/100 melhores alinhamentos e utiliza o melhor *score* para a modelagem molecular automática. Desta forma, as modelagens são realizadas de forma mais específica e direcionada para proteínas que possuem baixa homologia de sequência com àquelas disponíveis nos bancos de dados, como é o caso da DSM-14977.

O melhor análogo estrutural observado para a DSM-14977 foi a fosfatase de pirofosfato inorgânico, fosfolisina e fosfohistidina humana, LHPP (PDB 2x4d), e com este modelo, a predição estrutural da DSM-14977 foi realizada pelo próprio I-TASSER. O modelo da DSM-14977 evidenciou um padrão de enovelamento próximo ao característico para esta classe de enzimas, com uma folha-β central e fitas-β ligadas por α-hélices. Além disso, é possível observar a posição dos seguintes resíduos catalíticos, a saber, o ácido aspártico D32 que é o nucleófico, o ácido aspártico D34 (motivo I), a serina S68 conservada (motivo II), a lisina K92 (motivo III) e os resíduos de glicina G116, ácido aspártico D117 e D122 (motivo IV) (Figura 34).

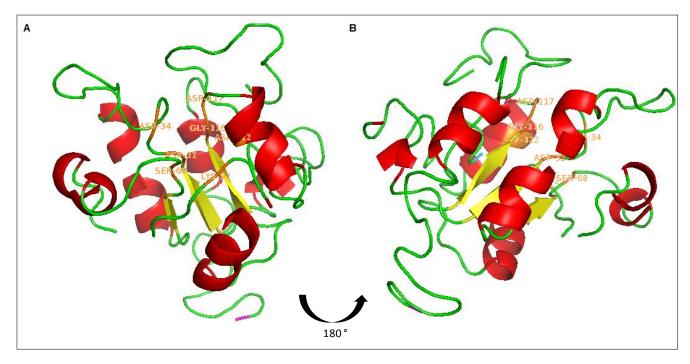


Figura 34: Modelagem Molecular da DSM-14977 através da predição estrutural.

O modelo da DSM-14977 pode ser observado evidenciando o padrão de enovelamento semelhante ao observado para a superfamília HAD, com uma folha-β central e fitas-β (amarelo) paralelas ligadas por α-hélices (vermelho) e a presença de *loops* (verde). Em (A) visão do N- para o C-terminal e em (B) observamos a rotação do modelo em 180°. Além disso pode ser observado os resíduos catalíticos que compõem o sítio ativo (laranja), o N-terminal (ciano) e o C-terminal (magenta). O melhor análogo estrutural obtido para a DSM-14977 foi a proteína LHPP (PDB 2X4D), uma fosfatase de pirofosfato inorgânico, fosfolisina e fosfohistidina humana.

Nossos resultados indicaram a proteína CbbY (PDB 4UAS), uma fosfatase de *Rhodobacter sphaeroides*, como análogo funcional. A proteína CbbY descrita como uma fosfatase ligadora da xilulose-1,5-bifosfato (XuBP) é altamente seletiva e realiza a conversão do XuBP em xilulose-5-fosfato (BRACHER et al., 2015). O XuBP é um grande inibidor da ribulose-1,5-bifosfato carboxilase/oxigenasse, a proteína Rubisco, que hidrolisa a conversão do dióxido de carbono atmosférico em compostos orgânicos em organismos fotossintéticos. Neste trabalho, os pesquisadores realizaram uma mutação no ácido aspártico (Asp 10) conservado para Asparagina (Asn 10) para estudar a especificidade da CbbY pelo XuBP (PDB: 4UAT). Este mutante revelou que a CbbY coopera com a atividade da Rubisco evitando um acúmulo prejudicial do XuBP que inibe o funcionamento desta proteína.

Desta forma, o modelo funcional da fosfatase DSM-14977 mostrou perfeita sobreposição dos aminoácidos presentes no sítio catalítico (Figura 35), ou seja, onde a xilulose-1,5-bifosfato da fosfatase CbbY se liga. Estes resultados sugerem que a DSM-14977 possui características estruturais de uma fosfatase ligadora de açúcar. Cabe ressaltar, que estes dados corroboram com aqueles obtidos para a caracterização bioquímica da DSM-14977, que mostrou que a enzima tem a capacidade de defosforilar alguns substratos, dentre eles, a

frutose 1,6-bisfosfato, podendo estar envolvida no metabolismo de açúcar da bactéria extremófila *Oceanithermus profundus*, sendo esta, provavelmente, a função biológica da fosfatase DSM-14977.

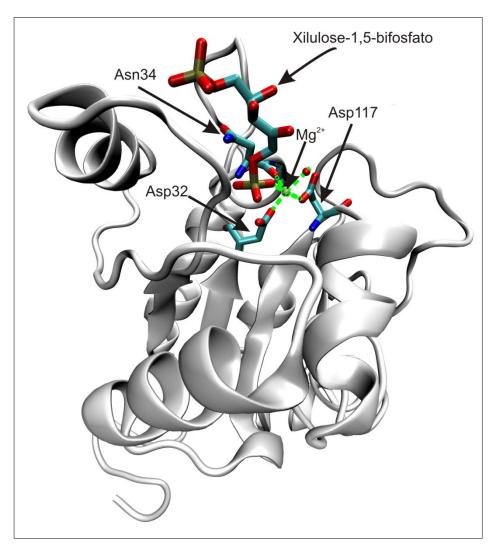


Figura 35: Modelagem Molecular da DSM-14977 através da predição funcional.

O modelo da DSM-14977 pode ser observado, em cinza, mostrando padrão de enovelamento semelhante ao observado para a superfamília HAD, com uma folha-β central e fitas-β ligadas por α-hélices O melhor análogo funcional obtido para a DSM-14977 foi a proteína CbbY (PDB 4UAS) de *Rhodobacter sphaeroides*, uma fosfatase ligadora de açúcares, especificamente a xilulose 1,5-bifosfato, envolvida com o metabolismo fotossintético desta bactéria e com o controle da Rubisco. Aqui apresentamos o modelo da DSM-14977 e a sobreposição da xilulose 1,5-bifosfato a partir das coordenadas do PDB (4UAT), mostrando a sobreposição perfeita das cadeias laterais do sítio ativo e a coordenação do átomo de Magnésio, cofator para a atividade enzimática desta proteína.

# 5. DISCUSSÃO

As enzimas são cada vez mais utilizadas como biocatalisadores em diversos setores industriais devido ao seu potencial em várias aplicações biotecnológicas gerando grande interesse no desenvolvimento de novas pesquisas nesta área. Os biocatalisadores são moléculas biodegradáveis, mais estáveis, além de gerar menos emissões e subprodutos comparados com os processos químicos convencionais. Em virtude disso, resultam em um baixo impacto ambiental e são mais ecológicos, sendo uma solução para muitos desafios industriais (SARMIENTO, 2015; SANDERSON, 2011; LI, 2012; GAVRILESCU E CHISTI, 2005).

Contudo, para a maior utilização destes biocatalisadores é necessário torná-los mais eficientes e vantajosos em relação ao custo e produção, visto que, a substituição dos catalisadores químicos pelos biocatalisadores ainda é restrita devido ao seu uso bem estabelecido no mercado. Neste contexto, o estudo e o desenvolvimento de novos biocatalisadores, especialmente, aqueles obtidos de organismos extremófilos, são essenciais visando enzimas com maior estabilidade, reprodutibilidade, seletividade e melhor desempenho em condições adversas para uso em diversos processos industriais (DUMORNÉ, 2017; BÖTTCHER E BORNSCHEUER, 2010).

A bioprospecção é uma estratégia utilizada para o conhecimento e desenvolvimento de novas enzimas contribuindo para o fornecimento de biocatalisadores. Através desta estratégia diversos biocatalisadores já foram produzidos e são utilizados em vários setores industrias como a hidantoinase de *Agrobacterium* e carbonil redutase de *Sporobolomyces salmonicolor* (indústria farmacêutica), tirosina fenol-liase de *E. herbicola* e nucleosideo fosfotransferase de *Morganella morganii* (indústria alimentícia), nitrila hidratases (indústria química) (OGAWA E SHIMIZU, 2002; SCHMID, 2001). Neste contexto, a enzima DSM-14977 de *Oceanithermus profunus* foi prospectada para o desenvolvimento deste estudo. Apesar de não existir nenhum dado experimental relativo à esta proteína, ela está anotada como uma fosfatase putativa, possuindo os motivos catalíticos conservados para a superfamília das dealogenases haloácidas (HAD), composta principalmente por fosfatases. As fosfatases possuem aplicações em diferentes ramos industriais, como na indústria de alimentos, fármacos e cosméticos, além de serem importantes em ensaios bioquímicos, na detecção e biorremediação de solos e agricultura, tornando-se, assim, um excelente alvo biotecnológico para o estudo de novos biocatalisadores. Neste trabalho, o objetivo principal foi realizar a

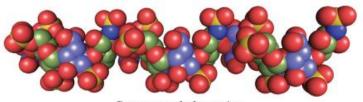
prospecção e a caracterização bioquímica e estrutural da fosfatase putativa DSM14977 para o conhecimento de um nova enzima de interesse biotecnológico.

Para isto, o gene sintético codificante da DSM-14977 foi subclonado nos plasmídeos de expressão pET22b, pETM30-MBP e pET32a. Assim, como descrito por COSTA et al (2014) e LAVALLIE (2000) a fusão da proteína de interesse a MBP e TRX, respectivamente, permitiu alcançar elevados níveis de expressão solúvel, bem como, com a inserção do pelB (SOCKOLOSKYA E SZOKAA, 2013; YOON, 2010). Os três DNA recombinantes foram expressos em *E. coli* e a melhor condição de expressão paras as três construções foi em BL21 (DE3) a 37°C com 1 mM de IPTG em D.O<sub>600nm</sub> 0,8.

Após a expressão, as construções MBP-DSM-14977 e TRX-DSM-14977 foram recuperadas na fração solúvel através de lise celular por sonicação. Já para a construção pelB-DSM-14977 a extração da proteína do espaço periplásmico foi realizada com a utilização de protocolos de choque osmótico. No entanto, a quantidade de proteína de interesse extraída se mostrou muito baixa quando comparada com a adição da lisozima, indicando a ineficiência do endereçamento da mesma para o periplasma. A lisozima é uma enzima que catalisa a hidrólise da ligação glicosídica que une os resíduos de açúcares aminados presentes no peptidioglicano da parede celular de bactérias, levando a permeabilidade da parede bacteriana e o extravasamento das proteínas do periplasma (LOPES, 2004; RATHORE, 2003). Diferentes fatores podem interferir na passagem da proteína para o compartimento periplasmático, o mais provável é que a sequência sinalizadora utilizada, pelB, não tenha sido capaz de realizar tal proposta. Muitas vezes, isto ocorre, devido a impedimentos pertinentes à própria constituição da proteína, como, a composição intrínseca de aminoácidos ou o próprio enovelamento (BYUNG, 2013; KAJAVA, 2000). Portanto, como alternativa, a lise celular por sonicação foi utilizada de forma eficiente para a obtenção desta construção na fração solúvel.

As três construções MBP-DSM-14977, TRX-DSM-14977 e DSM-14977 (pET22b) foram purificadas de acordo com as características intrínsecas de cada construção. Para MBP-DSM-14977 e TRX-DSM-14977 foi utilizada a cromatografia de afinidade a níquel devido ao tag de histidina presente nestas contruções. Para a construção DSM-14977 foram feitos inúmeros protocolos e diferentes tipos de cromatografia utilizando as propriedades bioquímicas da proteína para isto, como, seu alto pI (11,11), alta hidrofobicidade e possível ligação a açúcar. O melhor protocolo de purificação foi realizado em duas etapas: uma cromatografia hidrofóbica seguida de uma cromatografia de afinidade a heparina. A heparina é um glicosaminoglicano altamente sulfatado, ou seja, altamente aniônico, de origem natural,

usado como anticoagulante atuando na cascata da coagulação sanguínea (Figura 36) (UDEH CI et al, 2013). Além disso, a heparina é capaz de inibir fosfatases como a fosfatase de polinucleotídeo quinase T4 (T4 PNKP), descrito por (ZHAOA, 2017), que é uma enzima bifuncional com atividades 5'-quinase e 3'-fosfatase possuindo papel importante na resposta celular de danos ao DNA e no metabolismo do ácido nucleico. Desta forma, a grande afinidade da DSM-14977 à heparina pode estar associada a sua atuação como ligante específico, como, um papel inibitório ou ainda como um ligante inespecífico devido ao caráter positivo da proteína, evidenciado pelo seu alto pI.



Segmento de heparina

Figura 36: Estrutura da heparina

A heparina é um glicosaminoglicano altamente sulfatado e é usada na indústria farmacêutica como anticoagulante. Ela é uma molécula extremanete negativa devido ao seu conteúdo de grupos sulfatados e carboxilatos. A estrutura representa em vermelho o oxigênio, em amarelo o enxofre, em azul o carbono sulfatado do ácido idurônico e em verde carbono sulfatado da glicosamina que compõem este glicosaminoglicano.

Os estudos bioquímicos e estruturais foram realizados com a DSM-14977 oriunda da construção pelB, visto que, a mesma encontra-se na forma ivre sem a necessidade de um passo de clivagem e de uma nova purificação, como no caso das construções MBP-DSM-14977 e TRX-DSM-14977.

Os estudos estruturais da enzima DSM-14977 por dicroísmo circular, fluorescência e modelagem molecular foram realizados e geraram informações importantes quanto ao conteúdo de estrutura secundária presente na proteína, como também, quanto sua estabilidade frentes a agentes desnaturantes químicos e físicos.

O espectro de dicroísmo circular da DSM-14977 evidenciou um conteúdo de estrutura secundária rico em  $\alpha$ -hélices e folhas- $\beta$ , como já esperado, por ser tratar de uma fosfatase putativa pertencente a superfamília HAD que possui enovelamento do tipo Rossmann composto por um sanduíche de  $\alpha/\beta$ . Os estudos de modelagem molecular da DSM-14977 realizados pelo programa computacional I-TASSER evidenciou um padrão de enovelamento próximo ao característico para a superfamília HAD, com uma folha- $\beta$  central e fitas- $\beta$  ligadas por  $\alpha$ -hélices. Estes dados identificaram como melhor modelo funcional, a fosfatase de

pirofosfato inorgânico, fosfolisina e fosfohistidina humana, LHPP, e como melhor análogo funcional proteína CbbY (PDB 4UAS) que é uma fosfatase de *Rhodobacter sphaeroides* ligadora de açúcar envolvida.

Os estudos estruturais da DSM-14977 por espectroscopia de fluorescência indicaram um centro de massa deslocado para comprimentos de onda maiores, em torno de 346-348 nm, diferente do padrão encontrado para a maioria das proteínas enoveladas, observado em torno de 330 e 340 nm. Entretanto, isto pode ser explicado pela exposição de alguns resíduos de aminoácidos aromáticos, que são as sondas intrínsecas utilizadas para estes estudos, constituintes da proteína, a saber, sete triptofanos, uma tirosina e três fenilalaninas (EFTINK, 1995) (Figura 37). Cabe ressaltar que 3 triptofanos e 1 tirosina e 1 fenilalanina estão voltados para o interior desta proteína permitindo os estudos de estabilidade da proteína por fluorescência. A disposição dos aminoácidos aromáticos desta proteína pode apontar para um possível mecanismo de oligomerização, sugerindo a existência de um possível oligômero funcional, o que também é encontrado para outras fosfatases pertencentes a superfamília HAD (KUZNETSOVA, 2015). Cabe ressaltar que nossos estudos apontam para a oligomerização da enzima DSM-14977 em solução (dados não mostrados)

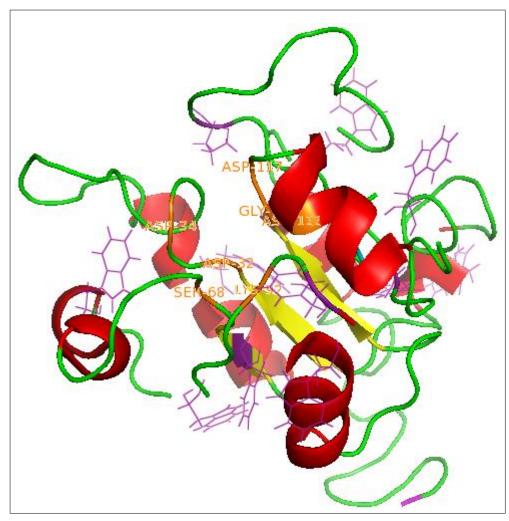


Figura 37: Modelo da DSM-14977 mostrando em detalhes a cadeia lateral dos aminoácidos aromáticos No modelo da DSM-14977 pode ser observado padrão de enovelamento semelhante ao encontrado para a superfamília HAD, com uma folha- $\beta$  central e fitas- $\beta$  (amarelo) ligadas por  $\alpha$ -hélices (vermelho), a presença de *loops* (verde), o sítio ativo (laranja). O N-terminal (ciano) e o C-terminal (magenta) podem ser observados. As cadeias laterais dos resíduos aromáticos podem ser evidenciadas no modelo em roxo, dentre eles, 7 triptofanos, 1 tirosina e 3 fenilalaninas. Na figura podemos observar a exposição de alguns destes resíduos para o solvente, podendo explicar o centro de massa deslocado da proteína.

Os estudos de estabilidade estrutural da DSM-14977 foram realizados por fluorescência na presença de agentes desnaturantes físicos e químicos. A investigação desses resultados mostrou que a DSM-14977 é mais estável na presença de uréia do que em cloreto de guanidina. Além disto, a enzima mostrou-se termoestável até 85°C confirmando se tratar de uma enzima extremófila do tipo termofílica. Após esta temperatura a enzima perde rápida e completamente sua estrutura a 90°C. Uma vez desnaturada, a enzima não mostrou evidências de retomar sua estrutura nativa a 30°C (dado não mostrado), o que provavelmente, também não ocorre com sua função biológica.

A anotação funcional de enzimas representa um grande desafio, principalmente, para aquelas pertencentes a famílias cujos membros catalisam reações semelhantes ou iguais, mas que evoluíram para diferentes preferências de substrato e funções biológicas específicas. A

gama de substratos fosforilados para a superfamília HAD é geralmente muito ampla, e inclui vários compostos fosfatados com ou sem açúcar (KUZNETSOVA, 2015; PFEIFFER, 2014).

Com a finalidade de confirmar a inserção da enzima putativa DSM-14977 na superfamília de fosfatases HAD, estudos bioquímicos foram realizados com a DSM-14977 purificada a partir da pET22b-DSM recombinante. Uma atividade geral de fosfatase foi identificada utilizando pNPP cuja melhor condição foi observada a 60°C em pH 7,0. Esta maior atividade em pH neutro também é observada na literatura para os membros da superfamília HAD, principalmente, devido ao mecanismo catalítico que utiliza o ácido aspártico como nucleófilo (BURROUGHS, 2006). Uma triagem com diferentes substratos fosforilados foi realizada para investigar a atividade fosfatásica da DSM-14977 e sua especificidade. Para isso foram utilizados substratos fisiológicos, como ribo- e nucleosídeos tri-e difosfatos, fosfatos de açúcar e outros metabólitos fosforilados, bem como polifosfatos. O monitoramento da atividade foi realizado através da quantificação da produção de fosfato livre usando verde malaquita, onde o fosfato livre forma um complexo com molibdato e o complexo fosfomolibdato é reduzido pelo meio ácido produzindo o cromóforo que pode ser medido a 650 nm. As maiores atividades catalíticas observadas foram para a fosfotirosina, ADP, e, principalmente, para polifosfato-75 e frutose 1,6-bifosfato.

De uma forma geral, a literatura específica para a superfamília HAD aponta para a dificuldade em identificar um substrato específico para estas fosfatases, visto que eles possuem a capacidade de hidrolisar um vasto espectro de compostos fosforilados, tais como: açúcares, coenzimas, ácidos orgânicos e nucleotídeos (PARK, 2015; SEIFRIED et al, 2013; PFEIFFER, 2014; BURROUGHS et al, 2006). KUZNETSOVA (2015) relata em seu trabalho sobre a multifuncionalidade que estas enzimas podem ter, na qual algumas são extremamente específicas, catalisando a hidrólise de apenas um substrato, enquanto outras são "promíscuas". Estas enzimas atuam, principalmente, como fosfatases com diversas possibilidades para substratos, mas com perfis diferentes entre si. Funcionalmente, esta "promiscuidade" atribui relevância fisiológica importante podendo propiciar benefícios para a célula através de diferentes mecanismos, como, a eliminação de nutrientes, revisão, balanceamento de pools metabólicos e remoção de antimetabólitos (POSSIK et al, 2017; PANDYA et al., 2014; LAHIRI et al, 2006; LU et al, 2005 e ZHANG et al, 2005). A especificidade das fosfatases da superfamília HAD é ditada por estruturas flexíveis do loop nos domínios central catalítico e tampa (NGUYEN et al, 2010). A flexibilidade conformacional das proteínas dificulta ainda mais os esforços para inferir especificidade ao substrato a partir de informações da sequência primária ou da estrutura tridimensional. O domínio tampa é um importante alvo para mutação

de resíduos que interagem com uma parte do substrato sem afetar a catálise de transferência do grupo fosforil realizada no domínio núcleo catalítico podendo gerar um biocatalisador mais eficiente e específico (DUNAWAY-MARIANO, 2010).

No mapa de superfície eletrostática da DSM-14977 pode ser observado uma distribuição de cargas formando porções bastante negativas/neutras ou positivas na proteína (Figura 38). Na região do sítio catalítico é formada uma cavidade de ligação ao substrato com caráter neutro e levemente negativo constituindo uma importante informação para compreender a especificidade e o mecanismo catalítico da DSM-14977. Entretanto, o que chamou mais atenção foi a observação de uma região extremamente positiva na DSM-14977 o que pode explicar sua forte interação com a heparina que é extremamente negativa. É importante ressaltar, que a modelagem evidenciou do lado oposto ao da região positiva, uma região rica em resíduos neutros/hidrofóbicos sugestivo de uma região de oligomerização. Muitas enzimas de organimos extremófilos tem a tendência a oligomerizar como forma de proteção em ambientes extremos, como, por exemplo, em altas temperaturas.

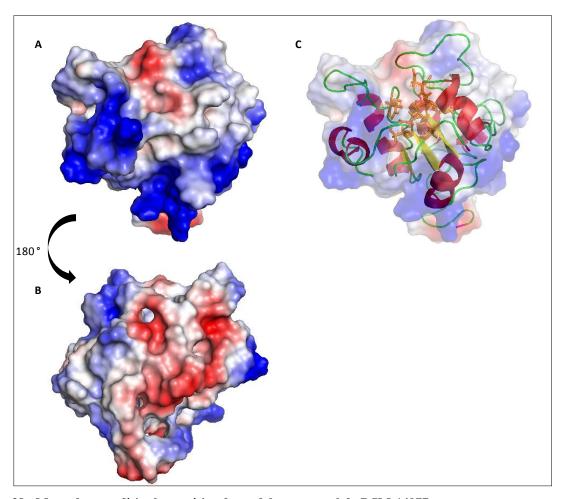


Figura 38: Mapa de superfície eletrostática do modelo estrutural da DSM-14977 Neste modelo é observado o padrão de enovelamento semelhante ao encontrado para a superfamília HAD, com uma folha- $\beta$  central e fitas- $\beta$  (amarelo) ligadas por  $\alpha$ -hélices (vermelho), a presença de *loops* (verde), o sítio ativo (laranja); o N-terminal (ciano) e o C-terminal (magenta). O mapa de superfícies eletrostática ilustra a distribuição de carga da proteína, mostrando regiões mais negativas (vermelho), mais positivas (azul) e neutras (branco).

Como mencionado anteriormente, o modelo da DSM-14977 mostrou enovelamento característico das fosfatases da superfamília HAD, além da perfeita organização dos resíduos de aminoácidos que formam o sítio ativo característico desta superfamília. Tal modelo, tanto estrutural quanto funcional, teve predição para enzimas envolvidas no metabolismo de açúcares, o que corrobora os nossos resultados experimentais obtidos. O análogo funcional para o modelo da DSM-14977 foi a fosfatase CbbY (PDB: 4UAS), altamente seletiva conhecida como fosfatase de xilulose-1,5-bifosfato (XuBP) (BRACHER, 2015). O XuBP é um potente inibidor da Rubisco que catalisa a conversão do dióxido de carbono atmosférico em compostos orgânicos em organismos fotossintéticos. O grupo produziu um mutante (PDB: 4UAT) que mostrou que a CbbY coopera com a atividade da Rubisco evitando um acúmulo prejudicial do XuBP. Além disso, eles observaram que a CbbY foi capaz de liberar um equivalente de fosfato livre a partir da frutose 1,6-bifosfato. Esses eventos aliados à alta

defosforilação pela DSM-14977 para a frutose 1,6-bifosfato podem evidenciar seu possível papel no metabolismo de açúcares.

Desta forma, os estudos bioquímicos e estruturais mostraram que a DSM-14977 é uma fosfatase pertencente a superfamília das dealogenases haloácidas indicando um possível papel no metabolismo de açúcar na bactéria *Oceanithermus profundus*.

Além disso, a alta catálise para o polifosfato-75 pode indicar sua possível aplicação biotecnológica na manutenção da estabilidade de alimentos industriais. SILVA (2008) descreveu a utilização do polifosfato melhorando o rendimento do produto final avaliado, uma lingüiça de frango, levando ao aumento da suculência e auxiliando na sua conservação. Os polifosfatos também são utilizados no preparo de misturas de nutrição parenteral, diminuindo a formação de precipitados que podem causar obstruções pulmonares, levando a consequências clínicas severas (MARKS E CRILL, 2004; ALLWOOD E KEARNEY, 1998). Além disto, os polifosfatos também são utilizados na indústria agrícola como fontes de fósforo para o desenvolvimento das plantas.

Outra aplicação interessante seria voltada para a indústria farmacêutica, visto que o polifosfato está envolvido no desenvolvimento do tecido ósseo, inflamação e outros processos no organismo humano (TRILISENKO et al, 2015). Recentemente, a descoberta de que o polifosfato é armazenado em grânulos densos de plaquetas e segregado durante a ativação plaquetária gerou um grande de interesse no papel desta molécula antiga na biologia humana (TRAVERS et al, 2015). Além disso, foi descoberto que o polifosfato atua como um chaperona para evitar a agregação de proteínas danificadas oxidativamente e também está intimamente envolvido na regulação das vias gerais de resposta ao estresse oxidativo em diferentes bactérias (GREY et al, 2015). Desta forma, é de grande interesse o estudo de enzimas envolvidas no metabolismo do polifosfato.

O contínuo interesse e o progresso no estudo de enzimas proporcionam maior sucesso em áreas de biocatálise industrial, principalmente, na área de biotransformações. Existem diversos fatores que influenciam esse crescente interesse, dentre eles, a promiscuidade enzimática, através de métodos computacionais robustos combinados com evolução direta e tecnologias de triagem para o melhoramento das propriedades enzimáticas. A ideia é atender melhor às demandas do processo utilizando catalisadores multifuncionais, ou seja, selecionando enzimas que catalisam qualquer reação química desejada. Investigações futuras usarão combinações de enzimas projetadas combinadas com a química, para gerar mais produtos químicos e materiais de recursos mais baratos (materiais renováveis), o que, consequentemente, contribuirá para o estabelecimento de uma economia baseada em biologia

gerando um menor impacto ambiental (MARTINEZ, 2017; ADRIO E DEMAIN, 2014; PFEIFFER et al, 2014; PANDYA et al, 2014).

#### 6. CONCLUSÃO

- A enzima putativa DSM-14977 foi prospectada e o gene sintético codificante da DSM-14977 foi subclonado nos plasmídeos de expressão pETM30-MBP, pET32a e pET22b. Assim, a fusão da proteína de interesse a MBP e TRX, respectivamente, permitiu alcançar elevados níveis de expressão solúvel, bem como, com a inserção do pelB.
- A melhor condição de superexpressão encontrada para as construções MBP-DSM-14977, TRX -DSM-14977 e pelB-DSM-14977 foi 37°C com 1 mM de IPTG em DO<sub>600</sub> 0,8.
- As construções MBP-DSM-14977 e TRX- DSM-14977 foram obtidas com alto grau de pureza após cromatografia de afinidade a níquel.
- A construção pelB-DSM-14977 foi obtida em alto grau de homogeneidade após dois passos de fracionamento: cromatografia hidrofóbica seguida de cromatografia de afinidade à heparina.
- O espectro de dicroísmo circular da fosfatase putativa DSM-14977 apresentou estrutura secundária com um conteúdo rico em  $\alpha$ -hélices e folhas- $\beta$  característica de fosfohidrolases da superfamília HAD,.
- O comportamento estrutural da fosfatase putativa DSM-14977 mostrou que há uma tendência ao desenovelamento em função do aumento da concentração de ureia, principalmente, após 3 M deste agente desnaturante
- O comportamento estrutural da DSM-14977 na presença de cloridrato de guanidina mostrou que rapidamente a proteína se desenovela, mesmo em baixas concentrações de guanidina.
- Ensaios enzimáticos utilizando o pNPP como substrato mostrou maior atividade para o meio reacional neutro, pH 7 a 60°C.
- Ensaios enzimáticos utilizando substratos fosforilados, comumente utilizados em caracterização bioquímica de fosfatases, demonstrou uma atividade de fosfohidrolase pela DSM-14977 relativamente alta em relação à fosfotirosina e ADP e possuiu uma atividade

moderada em relação à fosfotreonina, CTP, frutose 6-fosfato e polifosfato-25. A enzima apresentou maior atividade em direção ao polifosfato-75 e frutose 1,6-bifosfato

- A predição estrutural selecionou como melhor análogo estrutural a fosfotase de pirofosfato inorgânico, fosfolisina e fosfohistidina humana. Este modelo foi utilizado para a modelagem da estrutura da DSM-14977. O modelo mostrou o enovelamento de Rossmann, característico para esta classe de enzimas, com uma folha-β central e fitas-β ligadas por α-hélices.
- A predição funcional selecionou o modelo da fosfatase ligadora de açúcar CbbY de *Rhodobacter sphaeroides*. O modelo apartir do mutante de CbbY mostrou uma perfeita sobreposição do substrato no sítio catalítico da DSM-14977. Esses eventos aliados à alta defosforilação pela DSM-14977 para a frutose 1,6-bifosfato podem evidenciar uma importante função da fosfatase DSM-14977 no metabolismo de açúcares da bactéria *Oceanithermus profundus*.

# 7. REFERÊNCIAS

ALLEN, K. N.; DUNAWAY-MARIANO, D. Markers of Fitness in a Successful Enzyme Superfamily. **Current Opinion in Structural Biology**, 19:658–665, 2009.

Aschaffenburg R, Mullen JEC J Dairy Res 16:58–67, 1949.

BYUNG HOON JO, IM GYU KIM, JEONG HYUN SEO, DONG GYUN KANG, HYUNG JOON CHAA. Engineered Escherichia coli with Periplasmic Carbonic Anhydrase as a Biocatalyst for CO2 Sequestration. **Applied and Environmental Microbiology** p. 6697–6705, 2013.

BÖTTCHER D AND BORNSCHEUER U T Protein engineering of microbial enzymes **Current Opinion in Microbiology** 13:274–282, 2010.

BORJALILOO, S.; ZOMORODIPOUR, A.; YAKHCHALI, B.; SHOJAI, S. Comparison of T7- and lac-based systems for the periplasmic expression of human granulocyte macrophage colony stimulating factor in Escherichia coli. *Iranian Journal of Biotechnology*, v. 1 (2), p. 101-108, 2003.

BORNSCHEUER UT1, KAZLAUSKAS RJ Catalytic promiscuity in biocatalysis: using old enzymes to form new bonds and follow new pathways. **Angew Chem Int Ed Engl.** 43(45):6032-40, 2005.

BRACHER A, A SHARMA, A STARLING-WINDHOF, F. U HARTL AND M HAYER-HART Degradation of potent Rubisco inhibitor by selective sugar phosphatase **NATURE PLANTS** VOL 1, 2015

BRADFORD, M.M.; A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-dye Binding; **Analytical Biochemistry**; v.72.p. 248-554, 1976.

BROCK, T. D. & FREEZE, H. *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. **J Bacteriol** 98,289–297, 1969.

BUCHOLZ K, AND COLLINS The roots - a short history of industrial microbiology and biotechnology **Appl Microbiol Biotechnol** 97:3747–3762, 2013

BUD R The uses of life, a history of biotechnology, Cambridge University Press, Cambridge, 1993

BURROUGHS, A. M. et al. Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. **J Mol Biol**, v. 361, n. September 2015, p. 1003–1034, 2006.

BUSINESS WIRE, The "Global Industrial Enzymes Market 2018-2022" report has been added to ResearchAndMarkets.com's, 2018.

CORRÊA, D. H. A; RAMOS, C. H. I. The use of circular dichroism spectroscopy to study protein folding, form and function. **African Journal of Biochemistry Research Vol.3**, v. 3,

- n. 5, p. 164–173, 2009.
- COSTA S, ALMEIDA A, CASTRO A, DOMINGUES L Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: the novel Fh8 system **Front Microbiol**. 5: 63, 2014.
- CRONEY, J. C.; JAMESON, D. M.; LEARMONTH, R. P. Fluorescence spectroscopy in biochemistry: Teaching basic principles with visual demonstrations. **Biochemistry and Molecular Biology Education**, v. 29, n. 2, p. 60–65, 2001.
- DALMASO, G.; FERREIRA, D.; VERMELHO, A. Marine Extremophiles: A Source of Hydrolases for Biotechnological Applications. **Marine Drugs**, v. 13, n. 4, p. 1925–1965, 2015.
- DE MARIA, L., J. VIND, K.M. OXENBOLL, A. SVENDSEN, AND S. PATKAR. Phospholipases and their industrial applications. **Appl. Microbiol. Biotechnol**.74(2): p. 290-300, 2007
- DOBSON, C. M. Protein Folding and Misfolding. v. 426, n. December, 2012. FLEMMING M. POULSEN. A brief introduction to NMR spectroscopy of proteins. p. 1–33, 2002.
- DUMORNÉ K, CÓRDOVA D C, ASTORGA-ELÓ M, AND RENGANATHAN P Extremozymes: A Potential Source for Industrial **Applications Microbiol. Biotechnol.**, 27(4), 649–659, 2017
- EFTINK M. R, GHIRON C. A.; Exposure of tryptophanyl residues in proteins. Quantitative determination by fluorescence quenching studies. **Biochemistry** 15(3): 672–680, 1976 FISKE, C. H.; SUBBAROW, Y. The Colorimetric Determination of Phosphorus. **The Journal of biological chemistry**, v. LXVI, p. 1689–1699, 1925.
- ELLEUCHE, S. et al. Extremozymes—biocatalysts with unique properties from extremophilic microorganisms. **Current Opinion in Biotechnology**, v. 29, p. 116–123, 2014.
- ERICSSON U B, HALLBERG B M, DETITTA G T, DEKKER N, NORDLUND P,Thermofluor-based high-throughput stability optimization of proteins for structural studies, **Anal. Biochem**. 357 289–298, 2006
- FAMUREWA O, et al. Acid phosphatase synthesis in Aspergillus flavus. **Folia Microbiol**, 1994.
- FEDOSOV D.V., D.A. PODKOPAEVA, M.L. MIROSHNICHENKO, E.A. BONCH-OSMOLOVSKAYA, A.V. LEBEDINSKY, M.YU. GRABOVICH, Metabolism of the Thermophilic Bacterium *Oceanithermus profundus*, **Mikrobiologiya** Vol. 77, No. 2, pp. 188–195, 2008.
- FULZELE, R.; DESA, E.; YADAV, A.; SHOUCHE, Y.; BHADEKAR, R. Characterization of novel extracellular protease produced by marine bacterial isolate from the Indian Ocean. **Braz. J. Microbiol.**, 42, 1364–1373, 2011.

GALPERIN, MICHAEL Y.; KOONIN, E. V. From complete genome sequence to "complete" understanding? **Trends Biotechnology**, v. 28, n. 8, p. 398–406, 2010.

GAVRILESCU, M.; CHISTI, Y. Biotechnology—a sustainable alternative for chemical industry. **Biotechnology Advances**, v. 23, n. 7-8, p. 471–499, 2005.

GUIMARAES LHS, TERENZI HF, JORGE JA, LEONE FA, POLIZELI MLTM **Biotechnol Appl Biochem** 40:201–207, 2007

GUPTA, R.; GUPTA, N.; RATHI, P. Bacterial lipases: an overview of production, purification and biochemical properties. **Applied Microbiology and Biotechnology**, v. 64, n. 6, p. 763–781, 2004.

GURUNG, N., RAY, S., BOSE, S., AND RAI, V. A broader view :microbial enzymes and their relevance in industries, medicine, and beyond. **Biomed Res. Int.** 329121, 2013.

HANS E. SCHOEMAKER *et al.*Dispelling the Myths--Biocatalysis in Industrial Synthesis *Science* 299, 1694, 2003.

HASAN, F.; SHAH, A. A.; HAMEED, A. Methods for detection and characterization of lipases: A comprehensive review. **Biotechnology advances**, v. 27, n. 6, p. 782–98, 2009.

HASAN, F.; SHAH, A. A.; HAMEED, A. Industrial applications of microbial lipases **Enzyme and Microbial Technology** 39 235–251, 2006

HENNE, A. et al. The genome sequence of the extreme thermophile Thermus thermophilus. **Nature Biotechnology**, v. 22, n. 5, p. 547–553, 2004.

InterPro IPR023214 [https://www.ebi.ac.uk/interpro/entry/IPR023214/taxonomy]

JAENICKE R, BÖHM G. The stability of proteins in extreme environments. **Curr Opin Struct Biol.** (6):738-48, 1998.

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.

KAJAVA, A.V., ZOLOV, S.N., KALININ, A.E. AND NESMEYANOVA, M.A. **J. Bacteriol**. 182, 2163–2169, 2000.

KANEHISA, M. et al. KEGG as a reference resource for gene and protein annotation. **Nucleic Acids Research**, p. gkv1070, 2015.

KLENK H.P., PATI1 A., ZHANG X., LAPIDUS A., NOLAN M., LUCAS S., TIJANA GLAVINA DEL RIO, TICE H., CHENG J.F, TAPIA R., HAN C., E LAND M. Complete genome sequence of *Oceanithermus profundus* type strain (506<sup>T</sup>). **Standards in Genomic Sciences** 4:210-220, 2011.

- KOICHI ITAYA E MICHIO UI, A NEW MICROMETHOD FOR THE COLORIMETRIC DETERMINATION OF INORGANIC PHOSPHATE **Clin. Chim. Acta**, 14 (1966) 361-366, 1966
- KOSHLAND, D.; BOTSTEIN, D. Secretion of beta-lactamase requires the carboxy end of the protein. **Cell**, v. 20, p. 331-34, 1997.
- KUZNETSOVA, E. et al. Functional Diversity of Haloacid Dehalogenase Superfamily Phosphatases from Saccharomyces cerevisiae: Biochemical, Structural, and Evolutionary Insights. **The Journal of biological chemistry**, p. 1–45, 2015.
- YOON SH, KIM SK, KIM JF. Secretory production of recombinant proteins in Escherichia coli. **Recent Pat Biotechnol**. 4(1):23-9. 2010
- LAEMMLI, U. K.; Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4, **Nature**, v. 193, p. 680-685, 1970.
- LANG, J.M.; DARLING, A.E.; EISEN, J.A. Phylogeny of bacterial and archaeal genomes using conserved genes: Supertrees and supermatrices. **PLoS ONE** 8, e62510, 2013
- LAVALLIE, E. R., LU, Z. J., DIBLASIO-SMITH, E. A., COLLINS-RACIE, L. A., AND MCCOY, J. M. Thioredoxin as a fusion partner for production of soluble recombinant proteins in Escherichia coli. **Appl. Chimeric Genes Hybrid Proteins** Pt A 326, 322–340, 2000.
- LAHIRI S D, G ZHANG,D DUNAWAY-MARIANO, AND K N. ALLEN Caught in the Act: The Structure of Phosphorylated â-Phosphoglucomutase from Lactococcus lactis Biochemistry, 41, 8351-8359, 2002
- LI, S. et al. Technology prospecting on enzymes: application, marketing and engineering. **Computational and structural biotechnology journal**, v. 2, n. September, p. e201209017, 2012.
- LI, X, M WILMANNS, J THORNTON, M KÖHN Elucidating Human Phosphatase-Substrate Networks **SCIENCESIGNALING** Vol 6 Issue 275 rs10, 2013
- LINDORFF-LARSEN, K. et al. Protein folding and the organization of the protein topology universe. **Trends in biochemical sciences**, v. 30, n. 1, p. 13–19, jan. 2005.
- LÓPEZ-LÓPEZ, O., CERDÁN, M.E., SISO, M.I.G., New extremophilic lipases and esterases from metagenomics, **Current protein & peptide science**, v. 15, n° 5, p.445-455, 2014.
- LORENZ, P.; ECK, J. Metagenomics and industrial applications. **Nature**, v. 3, n. June, p. 510–516, 2005.
- LU, Z.; DUNAWAY-MARIANO, D.; ALLEN, K. N. The catalytic scaffold of the haloalkanoic acid dehalogenase enzyme superfamily acts as a mold for the trigonal bipyramidal transition state. **Proceedings of the National Academy of Sciences of the United States of America**, v. 105, n. 15, p. 5687–92, 2008.

MATSUSHITA K.; SHINAGAWA E.; ADACHI O.; AMEYAMA M. Spheroplast of acetic acid bacteria. *Agricultural and biological chemistry*, v. 45 (6), p. 1515-18, 1981.

MÉDICI, R. et al. A comparative study on phosphotransferase activity of acid phosphatases from Raoultella planticola and Enterobacter aerogenes on nucleosides, sugars, and related compounds. **Applied Microbiology and Biotechnology**, v. 98, n. 7, p. 3013–3022, 2014.

MIROSHNICHENKO M.L., L'HARIDON S., JEANTHON C., ANTIPOV A.N., KOSTRIKINA N.A., TINDALL B.J., SCHUMANN P., SPRING S., STACKEBRANDT E., BONCH-OSMOLOVSKAYA E.A. Oceanithermus profundus gen. nov., sp. nov., a thermophilic, microaerophilic, facultatively chemolithoheterotrophic bacterium from a deep-sea hydrothermal vent. **Int J Syst Evol Microbiol**.53(Pt 3):747-52, 2003.

MORDOR INTELLIGENCE Industrial Enzymes Market - Growth, Share, Trends and Forecasts (2018 - 2023), 2017

NOBRE, M. F., TRU" PER, H. G. & DA COSTA, M. S. Transfer of *Thermus ruber* (Loginova et al. 1984), *Thermus silvanus* (Tenreiro et al. 1995), and *Thermus chliarophilus* (Tenreiro et al. 1995) to *Meiothermus* gen. nov. *Meiothermus ruber* comb. nov., *Meiothermus silvanus* comb. nov., and *Meiothermus chliarophilus* comb. nov., respectively, and emendation of the genus *Thermus*. **Int J Syst Bacteriol** 46, 604–606, 1996.

NORBECK J, A PÅHLMAN, N AKHTAR, ABLOMBERG, AND L ADLER Purification and Characterization of Two Isoenzymes of DL-Glycerol-3-phosphatase from Saccharomyces cerevisiae SIGNAL TRANSDUCTION PATHWAY **THE JOURNAL OF BIOLOGICAL CHEMISTRY** Vol. 271, No. 23, pp. 13875–13881, 1996

OKUTA A, OHNISHI K, HARAYAMA S: PCR isolation of catechol 2,3-dioxygenase gene fragments from environmental samples and their assembly into functional genes. **Gene**, 212:221-228, 1998.

OWENS, R. Methods in integrated structural biology. **Methods**, v. 95, p. 1–2, 2016.

PANDYA, C.; FARELLI, J. D.; ALLEN, K. N. Enzyme Promiscuity: Engine of Evolutionary Innovation \*. **Journal of Biological Chemistry**, v. 289, n. 44, p. 30229 – 30236, 2014.

PARK J, A M. GUGGISBERG, A R. ODOM AND N H. TOLIA Cap-domain closure enables diverse substrate recognition by the C2-type haloacid dehalogenaselike sugar phosphatase Plasmodium falciparum HAD1 **Acta Cryst**. D71, 1824–1834, 2015.

PFEIFFER M, P WILDBERGER, B NIDETZKY Yihx-encoded haloacid dehalogenase-like phosphatase HAD4 fromEscherichia coli is a specific d-glucose 1-phosphate hydrolase usefulfor substrate-selective sugar phosphate transformations **Journal of Molecular Catalysis B: Enzymatic** 110 39–46, 2014

POSSIK E, S.R. M MADIRAJU, M PRENTKI Glycerol-3-phosphate phosphatase/PGP: Role in intermediary metabolism and target for cardiometabolic diseases **Biochimie** 1e11, 2017.

PUSEY, M. L. et al. Life in the fast lane for protein crystallization and X-ray crystallography.

Progress in Biophysics and Molecular Biology, v. 88, n. 3, p. 359–386, jul. 2005.

RATHORE, A.S.; BILBREY, R.E.; STEINMEYER D.E. Optimization of an osmotic shock procedure for isolation of a protein product expressed in E. coli. **Biotechnology Progress**, v 19 (5), p. 1541-1546, 2003.

RESEARCH AND MARKETS, 2015

REED M.S., G. PODESTA, I. FAZEY, N. GEESON, R. HESSEL,, K. HUBACEK, D. LETSON, D. NAINGGOLAN, C. PRELL, M.G. RICKENBACH, C. RITSEMA, G. SCHWILCH, L.C. STRINGER, A.D. THOMAS Combining analytical frameworks to assess livelihood vulnerability to climate change and analyse adaptation options **Ecological Economics** 94 66–77, 2013

ROBINSON, P K Enzymes: principles and biotechnological applications **Essays Biochem**. 59, 1–41, 2015

SAMUEL, P.; RAJA, A.; PRABAKARAN, P. Investigation and application of marine derived microbial enzymes: Status and prospects. Int. **J. Ocean. Mar. Ecol. Syst.**, 1, 1–10, 2012

SANDERSON, K. Enzyme expertise. Nature, v. 471, p. 397–398, 2011.

SARMIENTO F, R PERALTA AND J M BLAMEY 1,2 Cold and hot extremozymes: industrial relevance and current trends **Frontiers in Bioengineering and Biotechnology** Volume 3 Article 1481, 2015

SCHMID, A et al. Industrial biocatalysis today and tomorrow. **Nature**, v. 409, n. 6817, p. 258–268, 2001.

SCHNOES, A. M. et al. Annotation Error in Public Databases: Misannotation of Molecular Function in Enzyme Superfamilies. **PLoS Computational Biology**, v. 5, n. 12, p. e1000605, 2009.

SEIFRIED, A.; SCHULTZ, J; GOHLA, A; Human HAD phosphatases: structure, mechanism, and roles in health and disease. **FEBS Journal**, v. 280, p. 549–571, 2013.

SEIFRIED, A. et al. Evolutionary and structural analyses of mammalian haloacid dehalogenase-type phosphatases AUM and chronophin provide insight into the basis of their different substrate specificities. **The Journal of biological chemistry**, v. 289, n. 6, p. 3416–31, 7 fev. 2014.

SILVA LEONARDO PESSANHA, LOPES MÁRCIA MARTINS, MANO SÉRGIO, MÁRSICO ELIANE TEIXEIRA, CONTE-JÚNIOR CARLOS ADAM, TEODORO ANDERSON JUNGER, GUEDES WAGNER SOUZA Influence of polyphosphate addition in chicken sausage **R. bras. Ci. Vet.**, v. 15, n. 1, p. 50-55, 2008

SINGH B, SATYANARAYANA T Phytase production by thermophilic mold *Sporotrichum thermophile* in solid-state fermentation and its application in dephytinization of sesame oil cake. **Appl Biochem Biotechnol** 133:239–250, 2006a

- SINGH B, SATYANARAYANA T J Appl Microbiol 101:344–352, 2006b
- SINGH B, SATYANARAYANA T Biores Technol 99:2824–2830, 2008a
- SINGH B, SATYANARAYANA T Biores Technol 99:824–830, 2008b
- SINGH B. AND SATYANARAYANA T. Phytases and Phosphatases of Thermophilic Microbes: Production, Characteristics and Multifarious Biotechnological Applications, **Thermophilic Microbes in Environmental and Industrial Biotechnology: Biotechnology of Thermophiles**, DOI 10.1007/978-94-007-5899-5\_25, 2013
- SKALS, P., A. KRABEK, P. NIELSEN, AND H. WENZEL. Environmental assessment of enzyme assisted processing in pulp and paper industry. **Int. J. Life Cycle Assess**. 13(2): p. 124-132., 2008
- SPOHNER, S. C. et al. Expression of enzymes for the usage in food and feed industry with Pichia pastoris. **Journal of biotechnology**, v. 202, p. 118–134, 2015.
- SOCKOLOSKYA, JONATHAN T. AND SZOKAA, FRANCIS C. Periplasmic production via the pET expression system of soluble, bioactive human growth hormone **Protein Expr Purif**. 87(2): 129–135, 2013
- SRIVASTAVA, P. K.; ANAND, A. Immobilization of acid phosphatase from Vigna aconitifolia seeds on chitosan beads and its characterization. **International journal of biological macromolecules**, v. 64, p. 150–4, 2014.
- SURESH N AND ARUNAVA DAS, PURIFICATION AND CHARACTERIZATION OF PHOSPHATASES FROM ASPERGILLUS FUMIGATUS FOR APPLICATIONS IN INDUSTRIAL USES **Rasayan J. Chem,** Vol. 7 No.2 118 -128, 2014
- TERAKAWA A, A NATSUME, A OKADA, S NISHIHATA, J KUSE, K TANAKA, S TAKENAKA, S ISHIKAWA AND K YOSHIDA Bacillus subtilis 5'-nucleotidases with various functions and substrate specificities **BMC Microbiology** 16:249, 2016.
- THAM, S. J. et al. Biochemical Characterization of an Acid Phosphatase from Thermus thermophilus. **Biosci Biotechnol Biochem**, v. 74, n. 4, p. 727–735, 2010.
- UDEH CI, DOUGLAS A, UDEH BL, HATA JS. Heparin-induced thrombocytopenia: a clinical and economic review. **OA Anaesthetics**.1(1):3, 2013
- WILLIAMS, R. A. D. & DA COSTA, M. S.. The genus *Thermus* and related microorganisms. **The Prokaryotes**, 2nd edn, vol. 1, pp. 3746–3751,,1992.
- WOODY, R. W.; Circular dichroism; Methods in Enzymology, v. 246, p. 34-71, 1995.
- ZERBS, S.; FRANK, A. M.; COLLART, F. R. **Bacterial systems for production of heterologous proteins**. 1. ed. [s.l.] Elsevier Inc., 2009. v. 463
- ZHANG, Y. et al. Simultaneous production of 1,3-dihydroxyacetone and xylitol from

glycerol and xylose using a nanoparticle-supported multi-enzyme system with in situ cofactor regeneration. **Bioresource Technology**, v. 102, n. 2, p. 1837–1843, 2011.

ZHAO A H, Q LIUB, M LIUB, Y JINA, B LIA, Label-free fluorescent assay of T4 polynucleotide kinase phosphatase activity based on G-quadruplexe—thioflavin T complex **Talanta** 165 653–658, 2017.