

**MARCO ANTONIO SEIKI KADOWAKI**

**ANÁLISE ESTRUTURAL DAS PROTEÍNAS PHBF, PHAP1 e HFQ  
DE *Herbaspirillum seropedicae***

**Tese apresentada ao curso de Pós-graduação  
em Bioquímica da Universidade Federal do  
Paraná como requerimento parcial para a  
obtenção do grau de Doutor em Ciências-  
Bioquímica.**

**Orientadora: Prof.<sup>a</sup> Maria Berenice R. Steffens**

**CURITIBA**

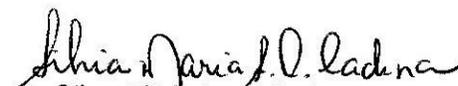
**2012**

## DECLARAÇÃO

Declaro que Marco Antônio Seike Kadowaki, em data de 15 de fevereiro de 2012 defendeu a tese de doutorado intitulada: “*Análise estrutural das proteínas PHBF, PHAPI e HFQ de Herbaspirillum seropedicae*”, desenvolvida sob a orientação do Prof.<sup>a</sup> Dra. Maria Berenice Reynaud Steffens, sendo a Comissão Examinadora constituída pelos professores: Dra. Elaine Benelli, Dr<sup>a</sup>. Fabiane Rego, Dr. Hernán Terenzi e Dr. Ronaldo Alves Pinto Nagem, tendo o mesmo sido aprovado, concluindo assim as exigências do Regulamento do Programa de Pós-Graduação em Ciências - Bioquímica da Universidade Federal do Paraná, referente ao Título de **DOUTOR em Ciências – Bioquímica**.

Esta declaração é válida até 15 de abril de 2012, data limite para a entrega da versão definitiva da tese de doutorado.

Curitiba, 15 de fevereiro de 2012.

  
Silvia M<sup>a</sup> Suter C. Cadena  
Matrícula 1326672  
Vice Coordenadora do Programa de  
Pós-Graduação em Ciências - Bioquímica - UFPR

## **AGRADECIMENTOS**

Agradeço enormemente a Professora Dr<sup>a</sup> Maria Berenice Reynaud Steffens pela orientação e paciência durante o desenvolvimento não só deste trabalho, mas também durante toda a minha formação científica. Lembrarei sempre a confiança depositada e o incondicional incentivo.

Ao Professor Dr. Jorge Iulek pelos preciosos ensinamentos sobre cristalografia confiados a mim com diferencial paciência e principalmente por sua amizade e hospitalidade.

A Professora Dr. Fabiane G. M. Rego pela paciência, atenção e disponibilidade durante a realização dos desafiadores experimentos envolvendo a manipulação de material radioativo.

Ao professor Dr. Marcelo Santos Muller e alunos de pós-graduação Evandro Freire Tirapelle e Cícero Silvano Teixeira, desbravadores do metabolismo de PHB, pelas discussões e por me mostrarem a grande importância do trabalho em equipe.

A aluna de Doutorado Cecília Fabiana da Gama Ferreira pela prontidão e auxílio durante o desenvolvimento dos experimentos de microscopia de força atômica.

Aos demais professores do grupo Fixação de Nitrogênio pelas sugestões e auxílios em todos os momentos.

A todos os colegas do Núcleo de Fixação de Nitrogênio pelas discussões, momentos de descontração e auxílios prestados.

As professoras Dr<sup>a</sup> Elaine Machado Benelli e Dr<sup>a</sup> Leda Satie Chubatsu pelas contribuições e sugestões prestadas ao acompanharem o desenvolvimento desta tese de Doutorado.

A coordenação do programa de Pós-graduação em Bioquímica e aos demais professores do departamento de Bioquímica.

Ao Genopar/SETI-PR, Fundação Araucária, Institutos do Milênio, INCT da Fixação Biológica de Nitrogênio e CNPq pelo apoio financeiro.

## SUMÁRIO

|  |           |
|--|-----------|
| AGRADECIMENTOS .....   | II        |
| LISTA DE QUADROS.....  | IX        |
| LISTA DE TABELAS .....   | X         |
| LISTA DE ABREVIATURAS E SÍMBOLOS .....   | XI        |
| RESUMO.....  | XIII      |
| ABSTRACT.....  | XV        |
| <i>Herbaspirillum seropedicae</i> SmR1 COMO ORGANISMO MODELO .....   | XVII      |
| <b>CAPÍTULO 1: ANÁLISE ESTRUTURAL DAS PROTEÍNAS RELACIONADAS À SÍNTESE DO BIOPOLÍMERO POLIHIDROXIBUTIRATO EM <i>Herbaspirillum seropedicae</i> SmR1.....</b> | <b>1</b>  |
| <b>1.1 INTRODUÇÃO .....</b>  | <b>2</b>  |
| 1.1.1 Biopolímeros .....   | 2         |
| 1.1.2 Polihidroxicanoatos .....  | 4         |
| 1.1.3. Metabolismo do PHB.....   | 9         |
| 1.1.4 Proteínas envolvidas na síntese e degradação de PHB .....  | 10        |
| 1.1.4.1 $\beta$ -cetoacil-CoA tiolase (PhbA).....  | 10        |
| 1.1.4.2 Acetoacetil-CoA redutase (PhbB) .....  | 10        |
| 1.1.4.3 PHB sintase (PhbC).....  | 11        |
| 1.1.4.4 Depolimerase (PhaZ).....   | 11        |
| 1.1.5 Proteínas envolvidas na regulação transcricional e estrutural no metabolismo de PHB.....   | 12        |
| 1.1.5.1 Regulação transcricional .....   | 12        |
| 1.1.5.2 Regulação estrutural dos grânulos de PHB.....  | 16        |
| <b>1.2. OBJETIVOS.....</b>   | <b>21</b> |
| 1.2.1 Objetivos gerais.....  | 21        |
| 1.2.2 Objetivos específicos.....   | 21        |
| <b>1.3. MATERIAL E MÉTODOS .....</b>   | <b>22</b> |
| 1.3.1 Análise utilizando ferramentas de bioinformática .....   | 22        |
| 1.3.1.1 Análise dos genes e proteínas alvo .....   | 22        |
| 1.3.1.2 Análise das estruturas secundárias e geração dos modelos estruturais <i>ab initio</i> .....  | 23        |
| 1.3.2 Materiais gerais e métodos de microbiologia.....   | 23        |
| 1.3.2.1 Bactérias e plasmídeos.....  | 23        |
| 1.3.2.2 Meios de cultura, condições de cultivo e estoque dos microrganismos.....   | 25        |
| 1.3.2.3 Antibióticos .....   | 27        |
| 1.3.2.4 Extração de PHB de <i>H. seropedicae</i> .....   | 27        |
| 1.3.3 Métodos de manipulação de DNA.....   | 28        |
| 1.3.3.1 Extração e purificação de DNA genômico de <i>H. seropedicae</i> .....  | 28        |

|             |  |           |
|-------------|--|-----------|
| 1.3.3.2     | Reação em cadeia da DNA polimerase (PCR).....  | 28        |
| 1.3.3.3     | Clonagem em vetor de expressão .....   | 31        |
| 1.3.3.4     | Clonagem em vetor contendo o gene repórter <i>lacZ</i> .....                                   | 31        |
| 1.3.3.5     | Minipreparação de plasmídeo por lise alcalina .....  | 31        |
| 1.3.3.6     | Sequenciamento de DNA .....  | 32        |
| 1.3.3.7     | Clivagem do DNA com enzimas de restrição.....  | 33        |
| 1.3.3.8     | Ligação DNA inserto ao vetor .....   | 33        |
| 1.3.3.9     | Eletroforese de DNA.....   | 33        |
| 1.3.3.10    | Transformação bacteriana .....   | 33        |
| 1.3.3.11    | Transferência de DNA plasmídial por conjugação bacteriana biparental.....                      | 35        |
| 1.3.4       | Métodos de expressão e manipulação de proteínas .....  | 35        |
| 1.3.4.1     | Determinação das condições de expressão .....  | 35        |
| 1.3.4.2     | Análise de proteínas por eletroforese em gel de poliacrilamida.....                            | 36        |
| 1.3.4.3     | Determinação do grau de pureza das proteínas.....  | 36        |
| 1.3.4.4     | Análise espectrométrica de massa das proteínas expressas .....                                 | 37        |
| 1.3.4.5     | Purificação das proteínas de <i>H. seropedicae</i> .....                                       | 38        |
| 1.3.4.6     | Quantificação da concentração protéica.....  | 38        |
| 1.3.4.7     | Concentração das proteínas por ultrafiltração .....  | 39        |
| 1.3.5       | Métodos <i>in vitro</i> de caracterização de proteínas .....                                   | 39        |
| 1.3.5.1     | Ensaio de ligação das proteínas PhbF e PhaP1 ao PHB .....                                      | 39        |
| 1.3.5.2     | Ensaio de retardamento eletroforético de banda em gel .....                                    | 40        |
| 1.3.5.3     | Ensaio de proteção de digestão pela DNase I.....   | 41        |
| 1.3.6       | Métodos <i>in vivo</i> de caracterização das proteínas alvo .....                              | 41        |
| 1.3.6.1     | Imunodeteção .....   | 41        |
| 1.3.6.2     | Determinação da atividade de $\beta$ -galactosidase .....                                      | 44        |
| 1.3.7       | Métodos de caracterização estrutural das proteínas alvo.....                                   | 45        |
| 1.3.7.1     | Dicroísmo circular .....   | 45        |
| 1.3.7.2     | Determinação do estado de oligomerização das proteínas .....                                   | 45        |
| 1.3.7.3     | Espalhamento de raios-X a baixo ângulo (SAXS) .....  | 46        |
| 1.3.7.4     | Análise da adesão da proteína PhaP1 ao PHB por Microscopia de força atômica (AFM).....         | 48        |
| <b>1.4.</b> | <b>RESULTADOS E DISCUSSÃO .....</b>  | <b>50</b> |
| 1.4.1       | Organização dos genes envolvidos na biossíntese de PHAs em <i>H. seropedicae</i> SmR1.....     | 50        |
| 1.4.2       | Análise <i>in silico</i> das proteínas PhbF e PhaP1 de <i>H. seropedicae</i> SmR1 .....        | 52        |
| 1.4.3       | Caracterização estrutural das proteínas PhbF e PhaP1 de <i>H. seropedicae</i> SmR1.....        | 56        |
| 1.4.3.1     | Amplificação e clonagem dos genes <i>phbF</i> e <i>phaP1</i> em vetores de expressão.....      | 56        |
| 1.4.3.2     | Testes de expressão e confirmação da identidade das proteínas PhaP e PhbF.....                 | 57        |
| 1.4.3.3     | Otimização das condições de expressão das proteínas PhbF e PhaP1.....                          | 59        |
| 1.4.3.4     | Purificação das proteínas fusionadas a cauda de histidinas.....                                | 62        |
| 1.4.3.5     | Determinação do estado de oligomerização das proteínas PhbF e PhaP1 .....                      | 65        |
| 1.4.3.6     | Análise por microscopia de força atômica da interação da proteína PhaP1 com filmes de PHB..... | 69        |
| 1.4.3.7     | Análise por dicroísmo circular .....   | 72        |

|  |  |            |
|--|--|------------|
| 1.4.3.8  | Análise por espalhamento de raios-X a baixo ângulo (SAXS) .....  | 75         |
| 1.4.3.9  | Análise por modelagem estrutural utilizando o método <i>ab initio</i> .....  | 80         |
| 1.4.4  | Caracterização funcional das proteínas PhbF e PhaP1 de <i>H. seropedicae</i> SmR1  | 85         |
| 1.4.4.1  | Ensaio de interação das proteínas PhbF e PhaP1 com o polímero PHB .....  | 85         |
| 1.4.4.2  | Interação da proteína PhbF com as regiões promotoras dos genes relacionados ao metabolismo de PHB de <i>H. seropedicae</i> SmR1..... | 87         |
| 1.4.4.3  | Caracterização do padrão de expressão dos promotores <i>phbF</i> e <i>phaP1</i> de <i>H. seropedicae</i> SmR1 .....                  | 96         |
| 1.4.4.4  | Expressão dos promotores <i>phbF</i> e <i>phaP1</i> em <i>E. coli</i> estirpe ET8000.....  | 96         |
| 1.4.4.5  | Caracterização do padrão de expressão das proteínas PhbF e PhaP1 em <i>H. seropedicae</i> SmR1 .....                                 | 98         |
| <b>1.5</b>   | <b>CONSIDERAÇÕES FINAIS .....</b>  | <b>104</b> |
| <b>1.6</b>   | <b>CONCLUSÕES .....</b>  | <b>108</b> |
| <br>   |  |            |
| <b>CAPÍTULO 2: DETERMINAÇÃO DA ESTRUTURA DA PROTEÍNA HFQ DE <i>Herbaspirillum seropedicae</i> SmR1 POR DIFRAÇÃO DE RAIOS-X .....</b> |  | <b>109</b> |
| <br>   |  |            |
| <b>2.1</b>   | <b>INTRODUÇÃO .....</b>  | <b>110</b> |
| <b>2.2</b>   | <b>OBJETIVO .....</b>  | <b>116</b> |
| 2.2.1  | Objetivo geral .....   | 116        |
| 2.2.2  | Objetivos específicos.....   | 116        |
| <b>2.3</b>   | <b>MATERIAIS E MÉTODOS .....</b>   | <b>117</b> |
| 2.3.1  | Cristalização, coleta de dados de difração de raios-X e análise estrutural da proteína Hfq de <i>H. seropedicae</i> SmR1 .....       | 117        |
| 2.3.2  | Espalhamento de raios-X a baixo ângulo e construção do modelo a baixa resolução .....  | 118        |
| <b>2.4</b>   | <b>RESULTADOS.....</b>   | <b>119</b> |
| 2.4.1  | Processamento das imagens e indexação .....  | 119        |
| 2.4.2  | Resolução da estrutura cristalográfica, refinamento e validação.....   | 121        |
| 2.4.3  | Descrição da estrutura geral da proteína Hfq de <i>H. seropedicae</i> SmR1 .....   | 125        |
| 2.4.4  | Análise estrutural comparativa .....   | 131        |
| 2.4.5  | Estrutura a baixa resolução da proteína Hfq em solução .....   | 134        |
| <b>2.5</b>   | <b>CONCLUSÕES .....</b>  | <b>138</b> |
| <b>3</b>   | <b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>  | <b>139</b> |

## LISTA DE FIGURAS

|   |    |
|---|----|
| FIGURA 1 – Estruturas subcelulares formadas por biopolímeros .....  | 4  |
| FIGURA 2 – Estrutura geral dos polihidroxicanoatos .....  | 5  |
| FIGURA 3 – Proteínas associadas ao grânulo de PHA .....   | 8  |
| FIGURA 4 – Via de síntese do PHB nos organismos que possui o operon <i>phbCAB</i> completo .....  | 10 |
| FIGURA 5 – Modelos de regulação transcricional no metabolismo de PHB .....  | 15 |
| FIGURA 6 – Modelos estruturais da fasina .....  | 19 |
| FIGURA 7 – Rede formada pelas fasinas na superfície dos grânulos de PHA .....   | 20 |
| FIGURA 8 – Genes envolvidos na biossíntese de PHAS em <i>H. seropedicae</i> SmR1 .....  | 51 |
| FIGURA 9 – Análise dos domínios conservados das proteínas PhbF e PhaP1 de <i>H. seropedicae</i> SmR1 .....                                  | 53 |
| FIGURA 10 – Análise de similaridade da proteína PhbF de <i>H. seropedicae</i> SmR1 .....  | 54 |
| FIGURA 11 – Análise de similaridade da proteína PhaP1 de <i>H. seropedicae</i> SmR1 .....   | 55 |
| FIGURA 12 – Eletroforese do extrato bruto visando à confirmação da expressão das proteínas PhbF e PhaP1 de <i>H. seropedicae</i> SmR1 ..... | 58 |
| FIGURA 13 – Perfil eletroforético do teste de expressão da proteína PhbF de <i>H. seropedicae</i> SmR1 .....                                | 60 |
| FIGURA 14 – Perfil eletroforético do teste de expressão da proteína PhaP1 fusionada de <i>H. seropedicae</i> SmR1 .....                     | 61 |
| FIGURA 15 – Purificação da proteína PhbF fusionada a cauda de histidinas .....  | 63 |
| FIGURA 16 – Purificação da proteína PhaP1 fusionada a cauda de histidinas .....   | 64 |
| FIGURA 17 – Pureza e homogeneidade das proteínas HisPhbF e HisPhaP1 após purificação .....  | 65 |
| FIGURA 18 – Determinação do estado de oligomerização de PhbF por gel filtração .....  | 67 |
| FIGURA 19 – Determinação do estado de oligomerização de PhaP1 por gel filtração .....   | 68 |
| FIGURA 20 – Imagens de afm da deposição de PHB sobre silício .....  | 70 |
| FIGURA 21 – Imagens de AFM da deposição da proteína HisPhaP1 sobre filme de PHB extraído de <i>H. seropedicae</i> Smr1 .....                | 71 |
| FIGURA 22 – Predição da estrutura secundária da proteína PhbF .....   | 73 |

|   |     |
|---|-----|
| FIGURA 23 – Predição da estrutura secundária da proteína PhaP1 .....  | 74  |
| FIGURA 24 – Análises de SAXS da proteína PhbF .....   | 76  |
| FIGURA 25 – Análises de SAXS da proteína PhaP1 .....  | 77  |
| FIGURA 26 – Modelos <i>ab initio</i> das proteínas PhbF e PhaP1 construídos a partir dos dados de SAXS .....  | 79  |
| FIGURA 27 – Modelo <i>ab initio</i> do monômero da proteína PhbF .....  | 82  |
| FIGURA 28 – Modelo <i>ab initio</i> do monômero da proteína PhaP1 .....   | 83  |
| FIGURA 29 – Ensaio de interação entre as proteínas PhbF e PhaP1 com PHB .....   | 86  |
| FIGURA 30 – Ensaio de retardamento da migração de banda do referente a interação da proteína PhbF com os promotores <i>phbF</i> e <i>phaP1</i> marcados com [ <sup>32</sup> P]..... | 88  |
| FIGURA 31 – Amplificação dos promotores alvos utilizando oligonucleotídeos marcados com as sondas fluorescentes VIC e 6-FAM.....  | 89  |
| FIGURA 32 – Ensaio de retardamento da migração de banda do complexo DNA promotor – proteína PhbF .....  | 90  |
| FIGURA 33 – Predição do provável motivo de ligação da proteína PhbF às regiões promotoras.....  | 93  |
| FIGURA 34 – Interação da proteína HisPhbF com a região promotora <i>phbF</i> .....  | 94  |
| FIGURA 35 – Região promotora de <i>phbF</i> é protegida por HisPhbF da ação da DNaseI..   | 95  |
| FIGURA 36 – Sequência de nucleotídeos das regiões promotoras dos genes <i>phbF</i> e <i>phaP1</i> de <i>H. seropedicae</i> SmR1 .....   | 97  |
| FIGURA 37 – A proteína PhbF reprime a expressão dos promotores <i>phbF</i> e <i>phaP1</i> de <i>H. seropedicae</i> SmR1 em <i>E. coli</i> ET8000 .....                              | 98  |
| FIGURA 38 – Perfil de produção de polihidroxibutirato (PHB) em <i>H. seropedicae</i> .....  | 99  |
| FIGURA 39 – Padrão transcricional e traducional da expressão de PhbF em <i>H. seropedicae</i> SmR1 .....  | 100 |
| FIGURA 40 – Padrão transcricional e traducional da expressão de PhaP1 em <i>H. seropedicae</i> SmR1 .....   | 103 |
| FIGURA 41 – Modelo de regulação desempenhado pela proteína PhbF em <i>H. seropedicae</i> SmR1 .....   | 107 |
| FIGURA 42 – Mecanismos de ação da proteína Hfq .....  | 111 |
| FIGURA 43 – Estrutura representativa da proteína Hfq de <i>E. coli</i> .....  | 113 |

|   |     |
|---|-----|
| FIGURA 44 – Avaliação do padrão de difração de raios-X da proteína HFQ.....                         | 119 |
| FIGURA 45 – Qualidade do ajuste da estrutura da proteína Hfq na densidade eletrônica                | 124 |
| FIGURA 46 – Qualidade estereoquímica da estrutura cristalográfica de Hfq .....                      | 125 |
| FIGURA 47 – Estrutura monomérica da proteína Hfq de <i>H. seropedicae</i> SmR1 .....                | 126 |
| FIGURA 48 – Estrutura geral da proteína Hfq.....  | 128 |
| FIGURA 49 – Interações intermoleculares na proteína Hfq de <i>H. seropedicae</i> SmR1               | 129 |
| FIGURA 50 – Empacotamento cristalino das proteínas Hfq de <i>H. seropedicae</i> e de <i>E. coli</i> | 132 |
| FIGURA 51 – Análise comparativa de proteínas Hfq de diferentes microrganismos .....                 | 133 |
| FIGURA 52 – Análises de SAXS da proteína Hfq .....  | 135 |
| FIGURA 53 – Estrutura a baixa resolução da proteína Hfq em solução .....                            | 137 |

## LISTA DE QUADROS

|  |     |
|--|-----|
| QUADRO 1 – Propriedades dos biopolímeros com síntese independente de molécula molde..... | 3   |
| QUADRO 2 – Alguns tipos de monômeros que compoem os PHAs.....                            | 7   |
| QUADRO 3 – Lista de estirpes de bactérias utilizadas .....                               | 23  |
| QUADRO 4 – Lista de plasmídeos utilizados .....  | 24  |
| QUADRO 5 – Oligonucleotídeos utilizados para a amplificação dos genes alvo .....         | 29  |
| QUADRO 6 – Oligonucleotídeos utilizados para a amplificação das regiões promotoras.      | 30  |
| QUADRO 7 – Proteínas Hfq com estrutura cristalográfica depositada no PDB .....           | 112 |

## LISTA DE TABELAS

|  |     |
|--|-----|
| TABELA 1 – Características bioquímicas das proteínas PhbF e PhaP1 de <i>H. seropedicae</i> SmR1..... | 56  |
| TABELA 2 – Parâmetros estruturais obtidos com os dados de SAXS das proteínas PhbF e PhaP1 .....      | 80  |
| TABELA 3 – Resumo dos dados estatísticos de processamento de imagens .....                           | 120 |
| TABELA 4 – Grupos espaciais sugeridos para o processamento dos dados .....                           | 120 |
| TABELA 5 – Estatísticas de coleta de dados e refinamento da estrutura cristalográfica                | 123 |
| TABELA 6 – Similaridade estrutural entre os monômeros .....  | 127 |
| TABELA 7 – Interações intermoleculares na proteína Hfq de <i>H. seropedicae</i> SmR1 .....           | 130 |
| TABELA 8 – Parâmetros estruturais obtidos com os dados de SAXS da proteína Hfq...                    | 137 |

## LISTA DE ABREVIATURAS E SÍMBOLOS

|                     |  |
|---------------------|--|
| AFM                 | - <i>Atomic force microscopy</i>               |
| Å                   | - Angstrom                                     |
| BSA                 | - Albumina de soro bovino                      |
| CD                  | - Dicroísmo Circular                           |
| CHCA                | - Ácido $\alpha$ -ciano-4-hidroxicinamímico    |
| CoA                 | - Coenzima A                                   |
| Da                  | - Dalton                                       |
| DMSO                | - Dimetilsulfóxido                             |
| DNA                 | - Ácido desoxirribonucleico                    |
| D.O. <sub>600</sub> | - Densidade óptica a 600 nm                    |
| EDTA                | - Ácido etileno-diamino-tetracético            |
| EMSA                | - <i>Electro mobility shift assay</i>          |
| F                   | - Faraday                                      |
| Hz                  | - Hertz  |
| HTH                 | - <i>Helix-Turn-Helix</i>                      |
| IPTG                | - Isopropil $\beta$ -D-1-tiogalactopiranosideo |
| K                   | - Kelvin                                       |
| K <sub>av</sub>     | - Constante de partição                        |
| m/v                 | - Relação massa e volume                       |
| MW                  | - Massa molecular                              |
| NSD                 | - <i>Normalized Spatial Discrepancy</i>        |
| $\Omega$            | - Ohm  |
| ONPG                | - O-nitrofenil- $\beta$ -D-galactosídeo        |
| ORF                 | - <i>Open reading frame</i>                    |
| PAGE                | - <i>Polyacrilamide gel electrophoresis</i>    |
| pb                  | - pares de base                                |
| PCR                 | - <i>Polymerase chain reaction</i>             |
| PEG                 | - Polietileno glicol                           |
| PDB                 | - <i>Protein Data Bank</i>                     |

|               |  |
|---------------|--|
| PHA           | - Polihidroxialcanoato                 |
| PHB           | - Polihidroxibutirato                  |
| pl            | - Ponto isoelétrico                    |
| ppm           | - Partes por milhão                    |
| PVDF          | - Fluoreto de polivinilideno           |
| RBS           | - <i>Ribosome binding site</i>         |
| Rg            | - raio de giro                         |
| RMSD          | - <i>Root mean square deviation</i>    |
| RNA           | - Ácido ribonucleico                   |
| rpm           | - Rotações por minuto                  |
| RSCC          | - <i>Real Space Correlation Factor</i> |
| RSR           | - <i>Real Space Factor</i>             |
| SAXS          | - <i>Small-angle X-ray scattering</i>  |
| SDS           | - Dodecilsulfato de sódio              |
| $\sigma^{70}$ | - Fator sigma 70 da RNA polimerase     |
| TAE           | - Tampão Tris-acetato-EDTA             |
| TBE           | - Tampão Tris-borato-EDTA              |
| Tris          | - Tris(hidroximetil)-aminometano       |
| tRNA          | - RNA transportador                    |
| UV            | - Ultravioleta                         |
| VC            | - Volumes de coluna                    |
| xg            | - Unidades de Gravidade                |

## RESUMO

*Herbaspirillum seropedicae* SmR1 é uma  $\beta$ -proteobactéria capaz de colonizar de forma endofítica diversas plantas de interesse comercial, sugerindo o seu uso como um potencial biofertilizante. Este trabalho é apresentado em dois blocos que descrevem a caracterização de proteínas de *H. seropedicae* SmR1 envolvidas no metabolismo do polímero polihidroxibutirato (PHB) e na regulação pós-transcricional da expressão gênica, com o objetivo de explorar os mecanismos regulatórios destas proteínas do ponto de vista molecular e estrutural. O capítulo 1 descreve a caracterização das proteínas PhbF e PhaP1 envolvidas no metabolismo de PHB. A capacidade de produzir este poliéster alifático é um importante fator associado à sobrevivência no ambiente em condições de estresse e competitividade com outras bactérias. Este polímero é estocado no interior das células na forma de grânulos cobertos por uma camada de várias classes de proteínas que regulam desde a produção/degradação do polímero, passando por seu tamanho e chegando a transcrição. PhbF está relacionada com a regulação transcricional e PhaP1 ou fasina está envolvida na regulação da relação volume/superfície dos grânulos. Estas proteínas foram expressas de forma heteróloga em *Escherichia coli* e purificadas com sucesso na forma fusionada a uma cauda de histidinas, utilizando o detergente TritonX-100. Experimentos de gel filtração revelaram PhbF como uma proteína tetramérica e PhaP1 como uma proteína pentamérica, ambas em solução. A proteína PhbF foi capaz de ligar-se a onze prováveis regiões promotoras de genes envolvidos com o metabolismo de PHB em *H. seropedicae*, confirmando a sua atividade de ligação ao DNA. A análise *in silico* destas regiões promotoras indicou a sequência consenso 5`-TG[N]TGC[N]<sub>3</sub>GCAA-3` que foi protegida da ação da DNaseI na região promotora de *phbF*. Esta proteína foi também capaz de ligar-se ao polímero PHB extraído de *H. seropedicae* assim como PhaP1. A fasina interagiu fortemente com o PHB e foi capaz de aderir a um filme deste polímero onde foram observadas várias estruturas proteicas individualizadas com dimensão lateral média de  $177,8 \pm 11,7$  Å. Esta dimensão é compatível com a distância máxima de 200 Å calculada para PhaP1 em solução pelo método de SAXS. Esta técnica permitiu também acessar o envelope molecular das proteínas PhbF e PhaP1. Aspectos *in vivo* relacionados à correlação entre o metabolismo de PHB e a transcrição dos genes

*phbF* e *phaP1* foram abordados através de fusões transcricionais ao gene *lacZ*. PhbF foi capaz de reprimir a sua própria expressão e a do gene *phaP1* em *E. coli* sem a interferência do polímero PHB. Em *H. seropedicae*, foi possível observar o acoplamento entre a expressão dos genes *phbF* e *phaP1* ao metabolismo de PHB assim como a mobilidade da proteína PhbF entre as formas ligada ao DNA e ligada ao grânulo de PHB. O capítulo 2 descreve a caracterização da chaperona de RNA Hfq de *H. seropedicae*. Esta é uma proteína homohexamérica, identificada em *E. coli* como um fator do hospedeiro envolvido na replicação do bacteriófago Q $\beta$  e como um importante regulador pós-transcricional de vários tipos de RNA afetando uma série de funções bacterianas. A estrutura da proteína Hfq foi determinada por cristalografia de raios-X e espalhamento de raios-X a baixo ângulo (SAXS). A estrutura cristalográfica revelou uma topologia do tipo Sm conservada, composta por uma  $\alpha$ -hélice N-terminal seguida de cinco fitas  $\beta$ , e uma nova interação intra-subunidades do tipo empilhamento  $\pi$ - $\pi$  entre dois resíduos de histidinas ausente em outras Hfqs com estrutura resolvida. Além disso, o envelope molecular calculado com base nos dados de SAXS concordou com a estrutura cristalográfica, sugerindo que a proteína tem a mesma conformação em solução.

## ABSTRACT

*Herbaspirillum seropedicae* SmR1 is an endophytic  $\beta$ -proteobacterium able to colonize several economically important crops, suggesting its potential use as a biofertilizer. This work is presented in two blocks describing the characterization of proteins related to polyhydroxybutyrate (PHB) metabolism and post-transcriptional regulation in *H. seropedicae* with the aim to explore the protein regulatory mechanisms at molecular and structural points of view. The chapter 1 describes the characterization of PhbF and PhaP1 proteins related to PHB metabolism. The ability to produce this polyester is an important factor associated with survival under stress conditions and in competitive environments. This polymer is stored intracellularly as granules which are covered by several proteins involved in granule synthesis/degradation, size and transcription regulation. PhbF is a transcriptional regulatory protein and PhaP1 or phasin regulates the granule volume/surface ratio. These proteins were expressed in *Escherichia coli* and purified using the detergent Triton-X100. The gel filtration assay described PhbF as a tetramer and PhaP1 as a pentamer in solution. PhbF was able to bind to eleven putative promoters of genes involved in PHA metabolism in *H. seropedicae* confirming its DNA binding property. *In silico* analysis indicated the consensus sequence 5'-TG[N]TGC[N]3GCAA-3' which was protected in DNA footprinting assay using *phbF* promoter region. This protein and PhaP1 were able to bind to PHB polymer extract from *H. seropedicae*. PhaP1 was also able to bind to PHB film as several structures with lateral dimension of  $177.8 \pm 11.7 \text{ \AA}$  compatible with  $D_{\text{max}}$  of  $200 \text{ \AA}$  from SAXS measurements. This approach also described the PhbF and PhaP1 molecular envelopes. *In vivo* aspects concern PHB metabolism and *phbF/phaP1* transcription were approached by *lacZ* transcriptional fusions. PhbF was able to repress *phaP1* expression as well as its own in *E. coli*. This system allowed the study of *phbF* and *phaP1* without PHB metabolism interference. The *H. seropedicae* environment allowed to show that *phbF* and *phaP1* expression are coupled to PHB metabolism and the PhbF mobility between DNA and PHB binding states. The chapter 2 describes the RNA chaperone Hfq from *H. seropedicae*. This protein is a homohexamer identified in *E. coli* as a host factor involved in phage Q $\beta$  replication and also as an important post-transcriptional regulator of several types of RNA,

affecting a plethora of bacterial functions. The Hfq protein structure was determined by X-ray crystallography and small angle X-ray scattering (SAXS). The crystal structure revealed a conserved Sm topology, composed of one N-terminal  $\alpha$ -helix followed by five twisted  $\beta$ -strands, and a novel  $\pi$ - $\pi$  stacking intra-subunit interaction of two histidine residues, absent in other Hfq proteins. Moreover, the calculated envelope based on SAXS data agreed with the Hfq crystal structure, suggesting that the protein has the same folding structure in solution.

## ***Herbaspirillum seropedicae* SmR1 COMO ORGANISMO MODELO**

O gênero *Herbaspirillum* pertencente à subdivisão  $\beta$  das Proteobactérias (YOUNG, 1992) e inclui principalmente bactérias diazotróficas com potencial para colonização endofítica de diversas plantas. *Herbaspirillum seropedicae* é um representante que apresenta elevado potencial como biofertilizante para culturas de interesse comercial como milho (*Zea mays*), arroz (*Oryza sativa*), sorgo (*Sorghum bicolor*) e cana-de-açúcar (*Saccharum officinarum*) (BALDANI *et al.*, 1986; PIMENTEL *et al.*, 1991; BODDEY *et al.*, 1995). Diversos estudos têm mostrado que este potencial concentra-se na capacidade deste microrganismo sintetizar o fitormônio ácido-3-indol acético (AIA) que pode afetar a morfologia e o desenvolvimento da planta colonizada, e também converter o nitrogênio gasoso em amônia que poderia ser disponibilizada para as plantas como fonte de nitrogênio (BALDANI & BALDANI, 2005). Em função disto, o Núcleo de Fixação de Nitrogênio (NFN-UFPR), iniciado em 1984 pelo professor Dr. Fábio de Oliveira Pedrosa, dedica-se ao estudo da bioquímica, fisiologia e genética da fixação de nitrogênio deste organismo (BALDANI & BALDANI, 2005; BENELLI *et al.*, 1997, 2001 e 2002; BONATTO *et al.*, 2005 e 2007; COUTO, 2005; GALVÃO *et al.*, 2003; HUERGO, 2010; KLASSEN *et al.*, 1997 e 1999; MONTEIRO *et al.*, 1999a, 1999b e 2003; NOINDORF *et al.*, 2006, 2011; PEDROSA *et al.*, 1997 e 2001; PERSHUN *et al.*, 2000; REGO *et al.*, 2006; RODRIGUES *et al.*, 2011; RONCATO-MACCARI *et al.*, 2003; SOUZA *et al.*, 1991a, 1991b, 1999 e 2000; SOUZA *et al.*, 2010; STEFFENS *et al.*, 1993; WASSEN *et al.*, 2000). Estudos envolvendo a análise do perfil proteômico de referência e diferencial também foram conduzidos (CHAVES *et al.*, 2007). O mecanismo de interação bactéria-planta envolvendo o gênero *Herbaspirillum* também está sendo investigado (MONTEIRO *et al.*, 2008; BALSANELLI *et al.*, 2010; TADRA-SFEIR *et al.*, 2011; SERRATO *et al.*, 2010; SCHMIDT *et al.*, 2011). Em 2011, a sequência genômica de *H. seropedicae* estirpe SmR1 foi publicada pelo Programa de Sequenciamento do Genoma de *H. seropedicae* ([www.genopar.org](http://www.genopar.org)) (PEDROSA *et al.*, 2011). O genoma está composto por um cromossomo circular de 5.513.887 pares de bases e contém um total de 4.804 genes.

Durante os anos, o estudo do metabolismo nitrogenado em *H. seropedicae* forneceu importantes informações sobre o metabolismo geral deste microrganismo (CHUBATSU *et al.*, 2011) e permitiu que novas áreas de investigação fossem abertas como o estudo do metabolismo do polímero polihidroxibutirato (PHB) e mecanismos de regulação do tipo pós-transcricional. O genoma de *H. seropedicae* SmR1 revelou a sua capacidade em metabolizar uma grande variedade de fontes de carbono dentre eles o PHB (PEDROSA *et al.*, 2011). A produção de PHB em *H. seropedicae* foi relatada para três estirpes diferentes (Z67, Z69 e Z78) (CATALAN *et al.*, 2007). A produção deste polímero foi otimizada a 36% da massa celular, sendo que a sua síntese é iniciada no começo da fase estacionária e atinge um pico após 63 horas de cultivo utilizando glucose como fonte de carbono. Além de PHB foi também constatado a produção do copolímero P(3HB-co-3HV) quando o meio de cultivo foi suplementado com ácido nanóico (CATALAN *et al.*, 2007). A presença dos grânulos formados por este polímero foi evidenciada em *H. seropedicae* por microscopia eletrônica utilizando a técnica de congelamento sob alta pressão (SILVA *et al.*, 2003). Outra evidência da produção deste polímero foi obtida através do proteoma de referência de *H. seropedicae* SmR1 onde foram identificadas as proteínas PhaP, PhbA, PhbB e PhbF (CHAVES *et al.*, 2007). Estas proteínas desempenham papel central na regulação e síntese. O papel regulatório da proteína PhbF foi relatada por KADOWAKI e colaboradores (2011a) e estudos estruturais envolvendo outras proteínas deste metabolismo estão sendo conduzidos. Estes estudos estão sendo realizados através da cooperação entre o Núcleo de Fixação de Nitrogênio e o Centro de Biologia Molecular Estrutural - Grupo de Cristalografia de Proteínas, sediado no Laboratório Nacional de Luz Síncrotron (Campinas – SP), que permitiu aprofundar a investigação da estrutura e função de proteínas de *H. seropedicae*, baseada na clonagem dos genes de interesse seguida da superexpressão e caracterização funcional dos produtos, aliada à determinação estrutural das proteínas por cristalografia de raios-X. O primeiro resultado desta colaboração foi a determinação da estrutura cristalográfica da proteína Hfq *H. seropedicae* (KADOWAKI *et al.*, 2011b). A proteína Hfq, considerada pleiotrópica e com função de chaperona de RNA, tem sido reconhecida recentemente por sua importância na regulação pós-transcricional em várias bactérias.

**CAPÍTULO 1: ANÁLISE ESTRUTURAL DAS  
PROTEÍNAS RELACIONADAS À SÍNTESE DO  
BIOPOLÍMERO POLIHIDROXIBUTIRATO EM  
*Herbaspirillum seropedicae* SmR1**

## 1.1 INTRODUÇÃO

### 1.1.1 Biopolímeros

Os sistemas vivos são capazes de sintetizar uma grande variedade de polímeros que podem responder pela maior parte da massa celular e que são conhecidos como biopolímeros. Segundo PÖTTER e STEINBÜCHEL (2005), os biopolímeros podem pertencer a oito classes diferentes: 1) ácidos nucleicos como o DNA e o RNA; 2) poliamidas como as proteínas e a cianoficina; 3) polissacarídeos como a celulose e o amido; 4) politioésteres como o poli(3-mercaptopropionato); 5) polianidridos como polifosfato; 6) poliisoprenoides como o látex; 7) polifenóis como a lignina e 8) poliésteres orgânicos como os polihidroxicanoatos (PHA). Estas classes diferem na composição química e na regularidade e variabilidade de seus elementos, os monômeros (PARRY & BAKER, 1984), que conferem a identidade a cada classe de polímero. Esta identidade é determinada pela frequência com que cada tipo de monômero compõe a estrutura do polímero o que também confere a complexidade e especificidade de sua função (PARRY & BAKER, 1984).

Por muitos anos os bioquímicos tradicionalmente focaram seus estudos nos biopolímeros lineares e informacionais como proteínas e ácidos nucleicos além dos polissacarídeos destinados a reserva energética, como o amido e o glicogênio. Nas últimas décadas outros biopolímeros produzidos a partir de um processo de polimerização independente de uma molécula molde estão sendo melhor investigados e dentre estes encontram-se os PHAs. Uma comparação entre diferentes polímeros está mostrada no quadro 1, onde é possível observar a variedade de composição monomérica, molécula iniciadora para o processo de polimerização, tamanho do polímero e tamanho do grânulo formado por estes.

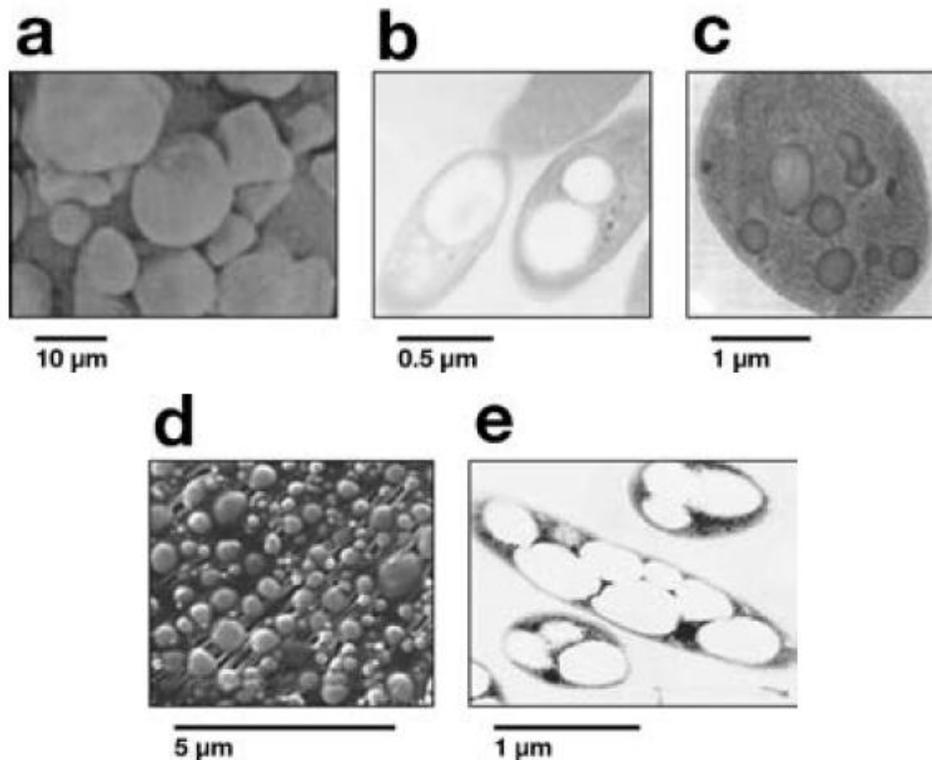
**QUADRO 1 – Propriedades dos biopolímeros com síntese independente de molécula molde**

| <b>Polímero (fonte)</b>                         | <b>Monômero</b> | <b>Molécula Iniciadora</b>      | <b>Tamanho do grânulo</b> | <b>Tamanho do polímero</b>   |
|---|-----------------|---------------------------------|---------------------------|--|
| Amido (Vegetal)                                 | ADP-Glc         | Maltooligossacarídeo ou amilose | Variável (1-100 µm)       | Amilose: (9x10 <sup>4</sup> a 3x 10 <sup>6</sup> Da)<br>Amilopectina: (3x10 <sup>8</sup> Da) |
| Glicogênio (Humano)                             | UDP-Glc         | Glicogenina                     | Variável                  | > 10 <sup>6</sup> Da   |
| Glicogênio ( <i>Agrobacterium tumefaciens</i> ) | ADP-Glc         | ADP-Glc                         | Nenhum grânulo detectável | > 10 <sup>6</sup> Da   |
| Polifosfato (Bactéria)                          | ATP             | Oligômeros de fosfato           | Depende da fonte          | > 7x10 <sup>4</sup> Da   |
| Cianoficina (Cianobactérias)                    | Asp e Arg       | Oligômeros de Asp-Arg           | 0,2-0,5 µm                | 2,5x10 <sup>4</sup> a 10 <sup>5</sup> Da   |
| Látex ( <i>Hevea brasiliensis</i> )             | IPP             | Geranylgeranyl pirofosfato      | ~1,0 µm                   | ~ 10 <sup>6</sup> Da   |
| Polihidroxibutirato (Bactéria)                  | HB-CoA          | HB-CoA                          | ~0,5 µm                   | ~ 10 <sup>6</sup> Da   |

FONTE: Modificado de STUBBE *et al.*, 2005.

Vários biopolímeros citados anteriormente são encontrados na forma de estruturas subcelulares denominadas grânulos ou corpos de inclusão. Estas estruturas compreendem os grânulos de polifosfatos, carboxissomos, grânulos de poliglucose (glicogênio e amido), vacúolos de gás, grânulos de polihidroxibutirato (PHB), corpos lipídicos e grânulo de cianoficina (Figura 1).

**FIGURA 1 – Estruturas subcelulares formadas por biopolímeros**



a) Grânulo de amido em *Zea mays*. b) Grânulo de polifosfato em *Vibrio cholerae*. c) Grânulo de cianoficina em *Aphanocapsa* 6308. d) Grânulo de látex em *Hevea brasiliensis*. e) Grânulo de PHB em *Wautersia eutropha*. A barra indica a escala da Figura. Fonte: STUBBE *et al.*, 2005.

A estocagem destes polímeros na forma de estruturas insolúveis permite que a osmolaridade celular não seja alterada (OBST & STEINBÜCHEL, 2006), o que possibilita o acúmulo dos mesmos em quantidades significativas. Na seção seguinte, será apresentada a classe de biopolímero foco deste estudo que se caracteriza por apresentar uma biossíntese de natureza não ribossomal e dependente de uma molécula iniciadora, o 3-hidroxi butiril-CoA.

### 1.1.2 Polihidroxicanoatos

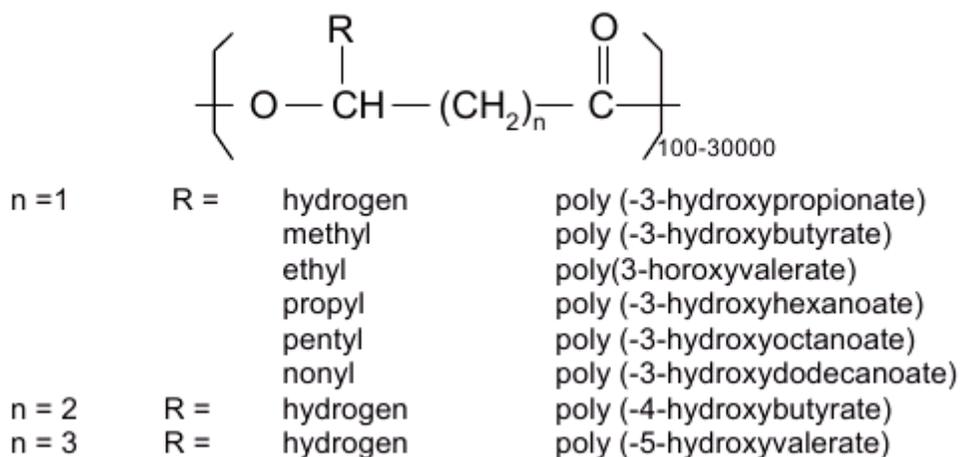
Polihidroxicanoatos (PHAs) são poliésteres lineares, compostos por monômeros do tipo ácido (R)-3-hidroxi graxos, sintetizados em várias bactérias gram-positivas e negativas (MADISON & HUISMAN, 1999) (Figura 2). Nestes polímeros o

grupo carboxil de um monômero forma uma ligação do tipo éster com o grupo hidroxil de outro monômero (MADISON & HUISMAN, 1999).

Estas macromoléculas constituem uma forma de estocagem de carbono e energia e são produzidas em condições limitantes de nutrientes (ANDERSON & DAWES, 1990). Além deste papel, os PHAs podem proteger as células em diferentes condições de estresse como, por exemplo, choque térmico, irradiação por UV, exposição a agentes oxidantes e choque osmótico (KADOURI *et al.*, 2003). A massa molecular destes polímeros varia entre  $2 \times 10^5$  a  $3 \times 10^6$  Daltons dependendo do tipo de microrganismo e das condições de crescimento (BYROM *et al.*, 1994). De modo geral sua produção ocorre durante o final da fase exponencial podendo atingir 90% da massa seca celular.

A estrutura final dos PHAs varia de acordo com o número de carbonos (1-13) do grupo alquil (R) na posição C3 do monômero (MADISON & HUISMAN, 1999) (Figura 2 e Quadro 2). Estes polímeros subdividem-se em duas classes de acordo com o comprimento do grupo alquil de cada monômero: PHAs de cadeia lateral curta [C<sub>3</sub>-C<sub>5</sub>] (ssc-PHA ou PHA<sub>SCL</sub> = *short chain lenght*) ou PHAs de cadeia lateral média [C<sub>6</sub>-C<sub>14</sub>] (msc-PHA ou PHA<sub>MCL</sub> = *medium chain lenght*) (MADISON & HUISMAN, 1999).

**FIGURA 2 – Estrutura geral dos polihidroxicanoatos**



FONTE: OJUMU *et al.*, 2004.

A diversidade na composição monomérica dos PHAs depende da fonte de carbono ofertada, como por exemplo, glucose, alcanos ou alcenos. A glucose depois de ser processada pela via glicolítica gera um *pool* de moléculas de acetil-CoA que são utilizadas, em parte, para a síntese de PHAs de cadeia curta. Os PHAs de cadeia média são formados a partir de ácidos 3-hidroxi graxos resultantes da síntese *de novo* de ácidos graxos ou intermediários da  $\beta$ -oxidação. Este tipo de PHA foi descoberto em 1983 quando WITHOLT e colaboradores cultivaram o microrganismo *Pseudomonas oleovorans* na presença de 50% de octano (MADISON & HUISMAN, 1999). N-alcanos ou n-alcenos são oxidados nas células em seus respectivos ácidos graxos que então podem sofrer outros passos de oxidação na via da  $\beta$ -oxidação. Intermediários metabólicos podem ser convertidos em PHAs em condições limitantes de elementos essenciais como o nitrogênio, caso observado em *Pseudomonas oleovorans*. Quando o meio de cultura deste organismo foi suplementado com dodecano, os PHAs possuíram composição monomérica variável: 2% com seis carbonos, 31% com oito carbonos, 36% com dez carbonos e 31% com doze carbonos (LAGEVEEN *et al.*, 1988). Conclui-se, portanto, que a fonte de carbono disponível e metabolizada reflete-se no tipo de PHA encontrado na célula.

Até o momento, foram reportados mais de 100 tipos diferentes de monômeros que podem compor os PHAs (STEINBÜCHEL & VALENTIN, 1995; KESSLER *et al.*, 2001) e alguns deles estão listados no quadro 2. O comprimento da cadeia lateral assim como o tipo de grupo funcional que a compõe influencia nas propriedades físicas como temperatura de fusão, temperatura de transição para o estado vítrio e cristalinidade (EGGINK *et al.*, 1995). A variabilidade monomérica reflete-se também nas propriedades termoplásticas apresentadas por polímeros com R (H, CH<sub>3</sub>, CH<sub>3</sub>CH<sub>2</sub>) e propriedades elastômeras apresentadas por polímeros com R (C<sub>3</sub>H<sub>7</sub> a C<sub>14</sub>H<sub>29</sub>) (STUBBE *et al.*, 2005). Estas últimas propriedades fazem deste material um bom substituto para os plásticos a base de derivados de petróleo e por este motivo estão recebendo maior atenção da área de bioengenharia.

Dentre os PHAs o poli(3-hidroxi butirato), ou PHB, é historicamente o mais estudado. Este polímero foi descoberto por LEMOIGNE, no Instituto Pasteur (1926), no lisado da bactéria *Bacillus megaterium* e é um dos mais simples PHAs onde o grupo R

corresponde a um metil (CH<sub>3</sub>). Este polímero é altamente estereoespecífico com um carbono assimétrico dotado de configuração do tipo D, característica que contribui para a sua cristalinidade (STEINBÜCHEL, 1991). Possui uma temperatura de fusão de 175 °C, um pouco abaixo de sua temperatura de degradação (185 °C), fato que dificulta o processo de modelagem para fins industriais. Este problema tem sido contornado com a produção do copolímero polihidroxibutirato-hidroxicvalerato (P(3HB-co-3HV)) que resulta na diminuição da temperatura de fusão e consequente diminuição da degradação térmica (OJUMU *et al.*, 2004).

## QUADRO 2 – Alguns tipos de monômeros que compoem os PHAs

| Ácidos 3-hidroxi | Ácidos 3-hidroxi (insaturados) | Ácidos 3-hidroxi (ramificados) | Ácidos 3-hidroxi (*CL substituída) | Ácidos 3-hidroxi (outros) |
|------------------|--------------------------------|--------------------------------|------------------------------------|---------------------------|
| Propiônico       | 2-Butanóico                    | 2-metilbutírico                | Ciclohexilbutírico                 | 4-hidróxibutírico         |
| Butírico         | 4-Pentanóico                   | 2-medilvalérico                | 5-fenilvalérico                    | 4-hidróxivalérico         |
| Valérico         | 4-Hexanóico                    | 2,6-dimetil-5-heptanóico       | 7-fluoroheptanóico                 | 4-hidróxihexanóico        |
| Hexanóico        | 5-Hexanóico                    | 4-metilhexanóico               | 9-fluorononanóico                  | 4-hidróxiheptanóico       |
| Heptanóico       | 6-Heptanóico                   | 5-metilhexanóico               | 6-clorohexanóico                   | 4-hidróxiheptanóico       |
| Octanóico        | 6-Octanóico                    | 6-metiloctanóico               | 8-clorooctanóico                   | 4-hidróxiocatnóico        |
| Nonanóico        | 7-Octanóico                    | 7-metiloctanóico               | 11-bromoundecanóico                | 5-hidróxivalérico         |
| Decanóico        | 8-Nonanóico                    | 6-metilnonanóico               | 7-cianoheptanóico                  | 5-hidróxihexanóico        |
| Undecanóico      | 9-Decanóico                    | 7-metilnonanóico               | 9-cianononanóico                   | Ácido málico              |
| Dodecanóico      | 10-Undecanóico                 | 8-metilnonanóico               | 12-hidróxi-dodecanóico             |                           |
| Tetradecanóico   | 6-Dodecanóico                  | 7-metildecanóico               | Ácido succínico metil éster        |                           |
|                  | 5-Tetradecanóico               | 9-metildecanóico               | Ácido adípico metil éster          |                           |
|                  | 5,8-Tetradecadienóico          |                                | Ácido adípico metil éster          |                           |
|                  | 5,8,11-Tetradecatrienóico      |                                | Ácido subérico metil éster         |                           |
|                  | 4-Hexadecanóico                |                                | Ácido subérico etil éster          |                           |
|                  | 4,7-Hexadecadienóico           |                                | Ácido pimélico propil éster        |                           |
|                  |                                |                                | Ácido sebácico benzil éster        |                           |

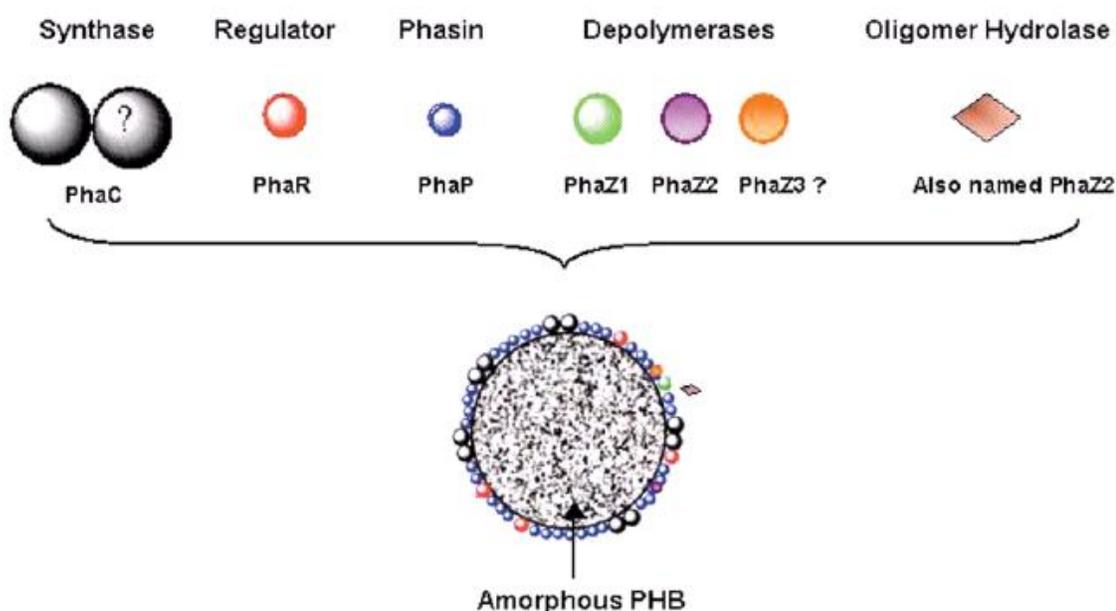
\*CL: cadeia lateral. FONTE: ZINN *et al.*, 2001.

PHAs são acumulados intracelularmente na forma de grânulos, podendo chegar a 90% da massa celular (MADISON & HUISMAN, 1999). Estes grânulos são estruturas subcelulares capazes de responder a variações do estresse fisiológico através da síntese ou degradação do polímero (POTTER & STEINBÜCHEL, 2005). Análises de

inclusões de *B. megaterium* mostraram que elas são formadas por 97,5% de PHA, 2% de proteínas e 0,5% de lipídeos (GRIEBEL *et al.*, 1968). Pode-se concluir então que os grânulos são cobertos por uma camada composta basicamente por lipídeos e proteínas. Os grânulos de PHA assim como as proteínas associadas a estes têm sido estudados em *Pseudomonas oleovorans*, *Paracoccus putida* e *Ralstonia eutropha* e em estirpes recombinantes de *E. coli* (FULLER *et al.*, 1992; HUIJBERTS *et al.*, 1992; POTTER *et al.*, 2004; SCHUBERT *et al.*, 1988).

Segundo STEINBÜCHEL e colaboradores (1995), as proteínas associadas aos grânulos de PHA podem ser classificadas em três classes: (I) PHA sintases, que catalizam a polimerização dos monômeros de hidroxiacil-CoA, (II) PHA depolimerases, responsáveis pela degradação intracelular e mobilização do PHA, (III) fasinas, que formam uma camada sobre os grânulos juntamente com os fosfolipídeos e (IV) outras proteínas, dentre as quais se destaca uma proteína regulatória, específica da síntese do PHA, chamada de PhbF ou PhaR (MAEHARA *et al.*, 2002; YORK *et al.*, 2002) (Figura 3).

**FIGURA 3 – Proteínas associadas ao grânulo de PHA**



FONTE: STUBBE & TIAN, 2003.

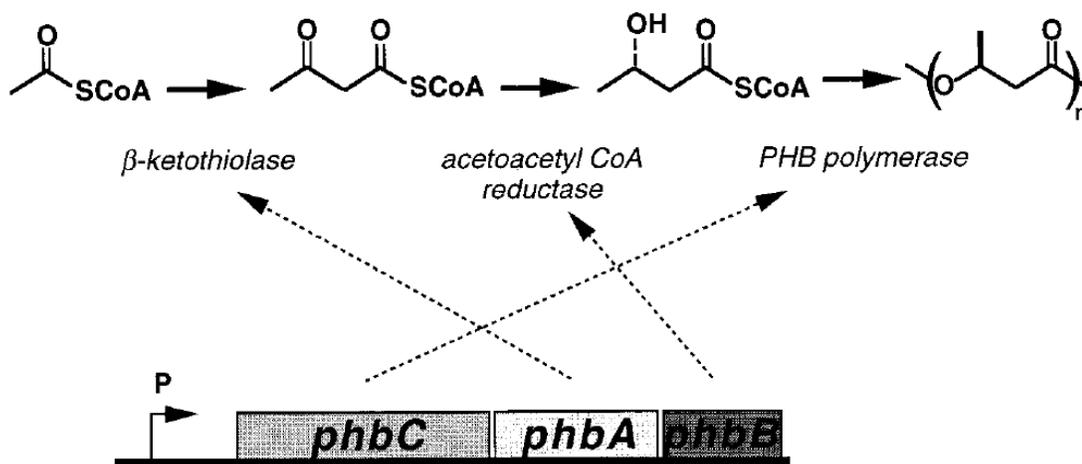
A presença de diversas classes de proteínas associadas ao grânulo de PHA levou JENDROSSEK (2009) a chamar esta complexa estrutura subcelular de carbonossomo.

### 1.1.3. Metabolismo do PHB

O PHB é sintetizado por uma via de três passos catalisados pela ação sucessiva das enzimas  $\beta$ -cetoacil-CoA tiolase (PhbA), acetoacetyl-CoA redutase (PhbB) e PHB polimerase (PhbC) (Figura 4). Já a degradação deste polímero, de forma extracelular ou intracelular, é catalisada pela PHB depolimerase (PhaZ). A organização dos genes envolvidos na biossíntese do PHB possui uma considerável divergência, reflexo da diversidade nas vias de biossíntese. MADISON & HUISMAN (1999), após caracterizarem os genes envolvidos na síntese de PHB em 18 espécies diferentes, relataram seis tipos de *loci* para os genes estudados: 1- operon completo *phbCAB*; 2- *loci phb* interrompido; 3- *loci phb* incompleto; 4- *loci phb* de organismos que codificam uma PHB polimerase com duas subunidades; 5- *loci phbJI* envolvido na síntese de P(3HB-3HH) e 6- *loci pha* para a síntese de PHA de cadeia média. Desta forma, conclui-se que os genes que codificam as proteínas envolvidas no metabolismo do PHB não estão, necessariamente, formando um operon e a organização gênica varia de espécie para espécie.

As principais características das proteínas envolvidas no metabolismo do PHB assim como na regulação transcricional e na regulação do tamanho granular serão detalhadas a seguir.

**FIGURA 4 – Via de síntese do PHB nos organismos que possui o operon *phbCAB* completo**



O PHB é sintetizado em três etapas sucessivas e catalisadas respectivamente pelas enzimas  $\beta$ -cetotiolase (PhbA), acetoacetyl-CoA redutase (PhbB) e PHB polimerase (PhbC). Fonte: MADISON & HUISMAN, 1999.

#### 1.1.4 Proteínas envolvidas na síntese e degradação de PHB

##### 1.1.4.1 $\beta$ -cetoacil-CoA tiolase (PhbA)

A primeira etapa da formação do PHB é catalisada pela  $\beta$ -cetoacil-CoA tiolase (PhbA). Esta enzima pertence à classe de tiolases envolvidas exclusivamente com processos de biossíntese de corpos cetônicos, esteróides, isoprenóides e PHB. O mecanismo de reação resulta na condensação de duas moléculas de acetil-CoA formando um acetoacetyl-CoA (MADISON & HUISMAN, 1999).

##### 1.1.4.2 Acetoacetyl-CoA redutase (PhbB)

PhbB é uma acetoacetyl-CoA redutase que catalisa a segunda etapa da síntese do PHB. Esta desidrogenase converte o acetoacetyl-CoA em 3-hidroxiacetyl-CoA. Esta enzima, juntamente com PhbA, fornece a molécula monomérica (3HB) para o processo de polimerização (MADISON & HUISMAN, 1999). Assim como PhbA, PhbB é encontrada no citoplasma durante a síntese ativa de PHB e, no caso de *Alcaligenes*

*eutrophus*, não foram encontradas em associação com o grânulo deste polímero (HAYWOOD *et al.*, 1988).

#### **1.1.4.3 PHB sintase (PhbC)**

PhbC é uma PHB sintase ou polimerase considerada a enzima chave para a produção do PHB. Esta enzima catalisa a conversão enantioselectiva de (R)-3-hidroxi-butiril-CoA em polihidroxi-butirato com a concomitante liberação de Coenzima A (REHM, 2003). De forma geral as sintases são divididas em duas classes principais de acordo com o seu produto: sintases de cadeia curta e sintases de cadeia média (REHM, 2006). Na literatura é frequente a utilização do termo PhbC para as enzimas que catalizam a síntese do PHB enquanto que o termo PhaC é destinado as sintases de polímeros de cadeia média. Estas proteínas podem ser ainda classificadas de acordo com a especificidade pelo substrato, composição das subunidades e diferença de massa molecular das mesmas. PhbC pode ser encontrada no citoplasma nas formas monomérica e dimérica sendo que a segunda forma é aumentada na presença de (3-hidroxi-butiril)<sub>3</sub>-CoA (REHM, 2003). JIA e colaboradores (2000) sugeriram uma homologia estrutural entre a PHA sintase de *Chromatium vinosum* e lipases bacterianas, e seus estudos permitiram a identificação dos resíduos de histidina, cisteína e aspartato como sendo importantes para a catálise mediada pela sintase. Este estudo destacou a variedade de mecanismos de reações realizadas pelas polimerases e discutiu um modelo diferente do mecanismo da  $\beta$ -cetoacil-ACP sintase, envolvido na síntese de lipídeos e considerado um paradigma.

#### **1.1.4.4 Depolimerase (PhaZ)**

PHA depolimerases (PhaZs) catalisam a hidrólise dos poliésteres de PHA com a liberação de ácidos hidroxialcanóicos mono ou oligoméricos que servirão como fonte de carbono e energia para os microrganismos. A degradação do PHA pode ocorrer de duas formas: intracelularmente por iPHA depolimerases em períodos de falta de nutrientes ou extracelularmente por ePHA depolimerases que utilizam como substrato

PHA liberado por células mortas. A ação de cada depolimerase é específica ao tipo de PHA, sendo a ePHA depolimerase capaz de degradar o PHA parcialmente cristalizado enquanto que a iPHA depolimerase atua sobre o PHA amorfo ou nativo (JENDROSSEK & HANDRICK, 2003). Até o momento, somente a estrutura tridimensional de uma ePHA depolimerase de *Penicillium funiculosum* (HISANO *et al.*, 2006) e de uma iPHA depolimerase de *Paucimonas lemoignei* (PAPAGEORGIU *et al.*, 2008) foram resolvidas por métodos experimentais de difração de raios-X. A PhaZ7 de *P. lemoignei* revelou uma conformação de  $\alpha/\beta$  hidrolase e uma similaridade estrutural com a lipase A de *Bacillus subtilis*. Nesta proteína também foi identificada a tríade catalítica Ser-His-Asp característica das proteínas pertencentes à família das lipases serino-esterases. O sítio catalítico que abriga estes resíduos de aminoácidos encontra-se bloqueado, sugerindo uma mudança conformacional quando há a ligação ao substrato. O contrário foi observado para a ePHA depolimerase de *Penicillium funiculosum*, cujo sítio ativo encontra-se acessível ao substrato além de apresentar diferenças estruturais consideráveis em relação a iPHA depolimerase (PAPAGEORGIU *et al.*, 2008).

### **1.1.5 Proteínas envolvidas na regulação transcricional e estrutural no metabolismo de PHB**

A regulação no metabolismo do PHB pode ser descrita a nível enzimático com o envolvimento de intermediários metabólicos diretamente sobre as enzimas, a nível transcricional com a participação de reguladores da expressão gênica e a nível estrutural com a participação de proteínas que regulam o tamanho e o número de grânulos de PHB intracelulares. Nesta seção serão descritos com detalhes os níveis de regulação transcricional e estrutural assim como as proteínas que participam destes processos.

#### **1.1.5.1 Regulação transcricional**

A diversidade nos mecanismos de regulação transcricional no metabolismo de PHAs envolve sistemas globais e específicos. Dentre os globais pode-se destacar a participação do regulador LuxR de *Vibrio harveyi* relacionado aos mecanismos de

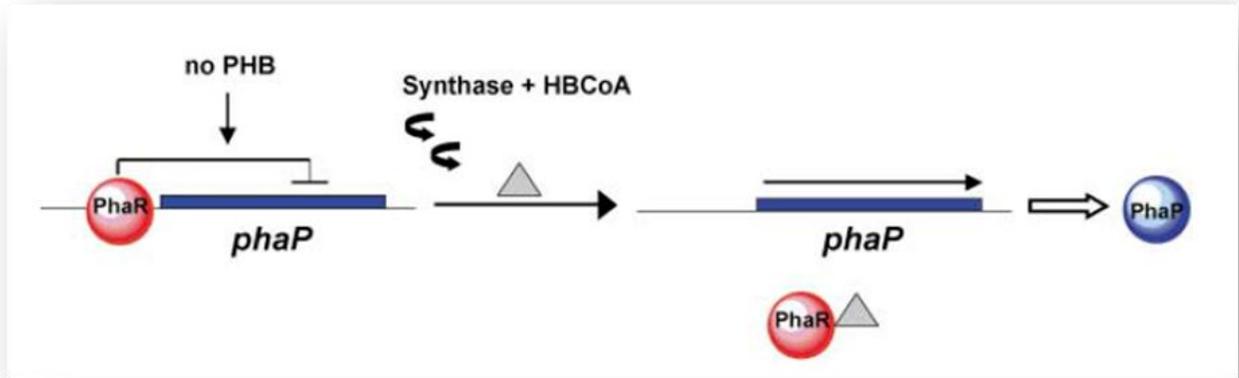
*quorum sensing* (McCOOL & CANNON, 1999) e as proteínas NtrB e NtrC de *Azospirillum brasilense* Sp7 relacionadas ao metabolismo de nitrogênio (SUN *et al.*, 2000). A participação destes reguladores mostra a importância e a relação do metabolismo de PHAs com outras vias metabólicas. Além destes reguladores tem sido descritos outras proteínas que atuam exclusivamente nesta via e que podem desempenhar o papel de ativadores ou repressores transcricionais dependendo do microrganismo (KESSLER & WITHOLT, 2001). As proteínas PhbR de *Azotobacter vinelandii* UW136 (PERALTA-GIL *et al.*, 2002; HERNANDEZ-ELIGIO *et al.*, 2011) e PhaD de *Pseudomonas putida* KT2442 (DE EUGENIO *et al.*, 2010) são relatadas como ativadoras transcricionais. Em *Azotobacter vinelandii* UW136 *phbR* codifica uma proteína da família AraC que teve a função de ativadora observada na estirpe mutante *phbR*, a qual apresenta redução na produção de PHB e na expressão do gene *phbB* (acetoacetil-CoA redutase) (PERALTA-GIL *et al.*, 2002). A proteína PhaD pertencente a família TetR e seu modelo estrutural levou DE EUGENIO e colaboradores (2010) a sugerirem que a ativação transcricional pode ser induzida por um intermediário da via de  $\beta$ -oxidação de ácidos graxos em *Pseudomonas putida* KT2442 (Figura 5B). As proteínas PhaF de *P. oleovorans* GPo1 (PRIETO *et al.*, 1999), PhaR de *R. eutropha* (PÖTTER *et al.*, 2002), *P. denitrificans* (MAEHARA *et al.*, 2002) e *Rhodobacter sphaeroides* FJI (CHOU & YANG, 2010) são relatadas como repressoras transcricionais e tem sido encontradas em associação aos grânulos de PHB. Dois modelos de regulação transcricional têm sido sugeridos para estas proteínas. O modelo de regulação mais difundido conta com a participação de apenas uma proteína (PhbF ou PhaR) envolvida principalmente na regulação da expressão do gene *phaP*, que codifica a proteína fasina e cuja função é descrita na seção seguinte (Figura 5A). Este papel também foi observado via expressão da proteína PhaP nas estirpes de *R. eutropha* e *R. sphaeroides* FJI com o gene *phaR* mutado (YORK *et al.*, 2002; CHOU & YANG, 2010). O mutante *phaR*<sup>-</sup> em *R. eutropha* apresentou diminuição da produção de PHA sugerindo o papel regulatório de PhaR sobre a produção de PHA através da modulação da expressão do gene *phaP*. O outro modelo envolve a participação da proteína PhaF de *P. oleovorans* GPo1 e uma proteína semelhante a fasina (Phal) (Figura 5C). PhaF foi relatada como um regulador negativo da expressão do gene *phaC1* cujo mecanismo de

repressão depende do tipo de fonte de carbono disponível (PRIETO *et al.*, 1999). Estas duas proteínas foram também encontradas em *Pseudomonas putida* KT2442 onde PhaF foi descrita como uma proteína importante para a localização celular dos grânulos assim como na segregação dos mesmos durante a divisão celular (GALÁN *et al.*, 2011).

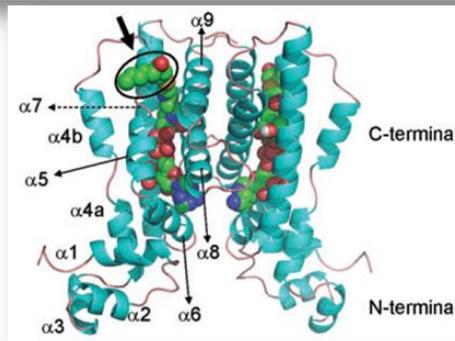
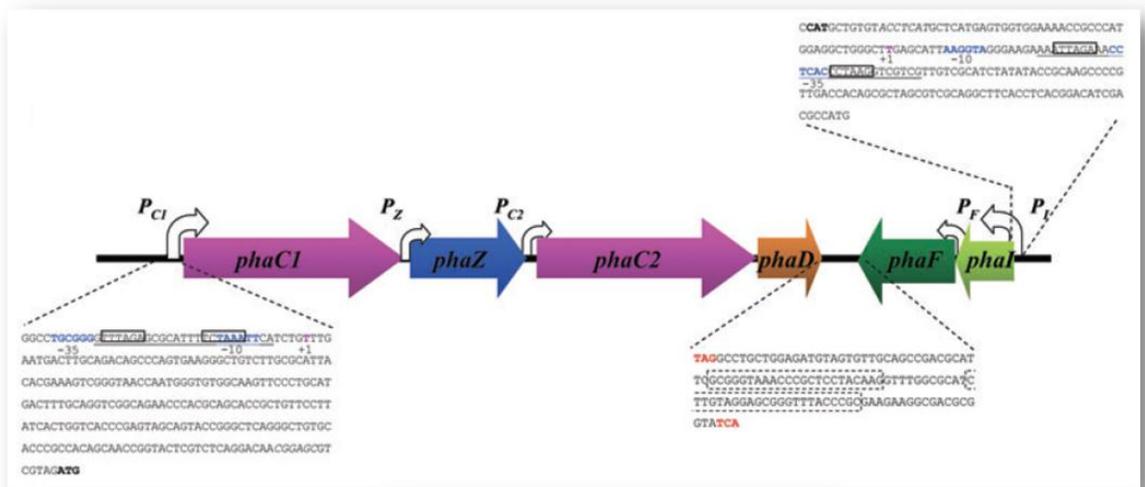
Dentre as proteínas citadas acima PhaR é o regulador mais estudado *in vitro*, quanto a sua atividade de interação com DNA e PHB. PhaR é uma proteína bifuncional, ou seja, liga-se a regiões específicas do DNA e diretamente aos grânulos de PHB. Esta última atividade é considerada não específica, pois PhaR pode ligar-se a todas as formas de PHB (cristalina, amorfa ou oligômeros) (MAEHARA *et al.*, 2002). Em *P. denitrificans* é uma proteína de 22 kDa capaz de ligar-se a uma região a montante dos genes *phaP* e *phaR* (MAEHARA *et al.*, 2002). Esta ligação resulta em um mecanismo de autorepressão de PhaR assim como a repressão da expressão do gene *phaP* como demonstrado *in vivo* (MAEHARA *et al.*, 1999) e *in vitro* (MAEHARA *et al.*, 2001). Um provável motivo de ligação desta proteína ao DNA, composto de sequências de TGC repetitivas e que sobrepõe parcialmente o promotor do gene *phaR*, foi sugerido por estudos de digestão utilizando a técnica de DNaseI *footprinting* em *Paracoccus denitrificans* (MAEHARA *et al.*, 2002). Em *R. sphaeroides* FJI, a análise das regiões promotoras dos genes *phaZ*, *phaC*, *phaP* e *phaR* revelaram a presença da sequência consenso CTGCN<sub>3-4</sub>GCAG (CHOU *et al.*, 2009) e, mais tarde, a importância do espaçamento de 3 a 4 bases entre os módulos CTGC e GCAG através de experimentos de deleção e ligação a proteína PhaR *in vitro* (CHOU *et al.*, 2010). Em *R. eutropha* H16, PhaR liga-se ao promotor -35/-10 do gene *phaP* em dois sítios distintos: em uma região onde está presente o sítio de início de transcrição e em uma outra região acima do módulo -35 do promotor, sugerindo novamente um papel repressor ao bloquear o acesso da RNA polimerase (PÖTTER *et al.*, 2002). Estes dois sítios compreendem uma sequência de 12 pares de base não relacionados à sequência observada em *P. denitrificans* e *R. sphaeroides* FJI, sugerindo que a variação no sítio de ligação observado resulta em variação nos mecanismos de regulação.

FIGURA 5 – Modelos de regulação transcricional no metabolismo de PHB

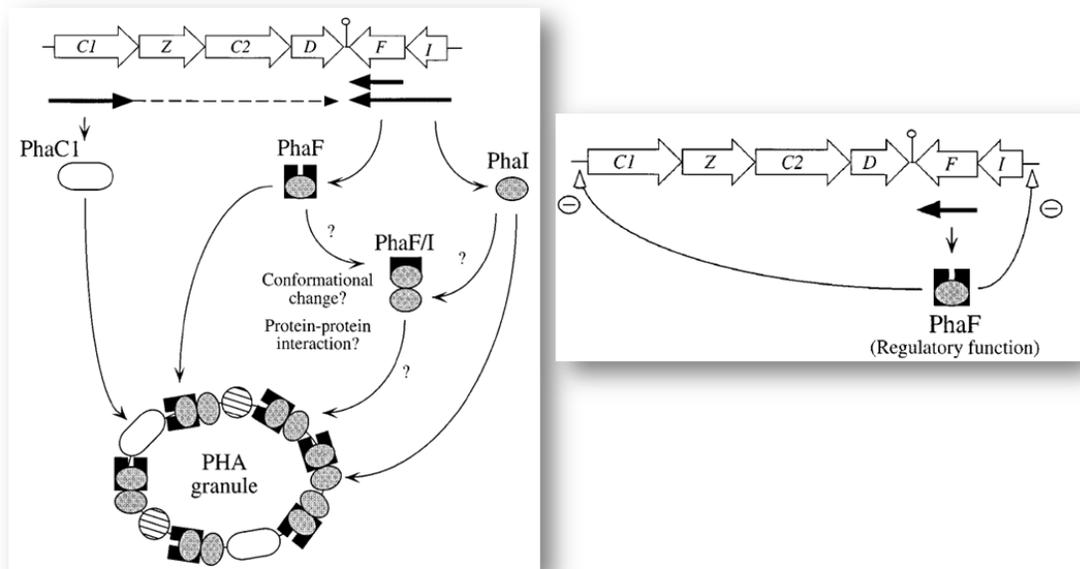
(A)



(B)



(C)



Modelos de regulação transcricional relatados na literatura. (A) Modelo de repressão transcricional descrito em *Ralstonia eutropha* (STUBBE & TIAN, 2003). O esquema descreve a indução da expressão do gene *phaP* na presença de PHA. (B) Modelo de ativação transcricional descrito em *Pseudomonas putida* (DE EUGENIO *et al.*, 2010). As sequências de pares de base destacadas no grupo de genes relacionados ao metabolismo de PHA indicam as regiões promotoras dos genes *phaC1* e *phal* que contém sítos de ligação a proteína PhaD. Na parte inferior da figura encontra-se o modelo estrutural da proteína PhaD obtida por modelagem molecular. (C) Modelo de repressão transcricional envolvendo a participação da fasina Phal em *Pseudomonas oleovorans* (PRIETO *et al.*, 1999). A figura descreve a indução da expressão dos genes *phaC1*, *phal* e *phaF* na presença grânulos de PHA e a repressão transcricional na ausência do mesmo.

### 1.1.5.2 Regulação estrutural dos grânulos de PHB

Os PHAs podem ser encontrados nos estados amorfo ou parcialmente cristalino dentro dos grânulos (JENDROSSEK & HANDRICK, 2003). O baixo grau de cristalinidade encontrado no grânulo nativo é considerado um fator importante para a biodegradabilidade uma vez que torna o polímero mais acessível à ação das depolimerases (TOKIWA *et al.*, 2009). Este estado é mantido pela camada de proteínas que cobrem os grânulos e cuja alteração decorrente do processo de extração pode

levar a um aumento de 50-60% na cristalinidade (PÖTTER & STEINBÜCHEL, 2006). Esta camada também isola o PHA hidrofóbico do contato com o citoplasma (PIEPER-FURST *et al.*, 1995) e impede que outras proteínas hidrofóbicas associem-se de forma inespecífica ao grânulo, o que poderia causar efeitos deletérios à célula (BOHMERT *et al.*, 2002). A proteína sem propriedade enzimática mais abundante associada ao grânulo é denominada fasina. Este termo foi empregado por Steinbüchel e colaboradores (1995) para descrever as proteínas associadas aos grânulos e análogas as oleosinas (proteínas anfífilas que formam uma camada na superfície de inclusões oleosas de células vegetais) (JURASEK & MARCHESSAULT, 2002). Esta proteína caracteriza-se pela baixa massa molecular (11-25 kDa), por possuir um caráter anfífilo, pela alta afinidade por PHA e por contribuir significativamente para a fração total de proteínas celulares (TRAINER *et al.*, 2006). O efeito direto sobre o número e o tamanho dos grânulos foi observado na estirpe mutante *phaP* de *Alcaligenes eutrophus* que resultou na formação de um único grânulo (WIECZOREK *et al.*, 1995) e com a superexpressão da proteína PhaP em *Aeromonas hydrophila* que levou à formação de muitos grânulos com tamanho reduzido (TIAN *et al.*, 2005). O papel desta proteína na síntese de PHA foi demonstrado por Pötter e colaboradores (2005) que detectaram a redução nos níveis deste polímero nas estirpes mutantes nos genes *phaP* em *R. eutropha*. A importância das fasinas na síntese de PHA pode ser explicada através de seu papel na regulação da razão superfície/volume e também pela interação com PHA sintases (YORK *et al.*, 2001). A primeira propriedade tem o papel de prevenir a coalescência entre os grânulos (JURASEK & MARCHESSAULT, 2002) permitindo maior superfície de ação para as enzimas responsáveis pela mobilização da fonte de carbono.

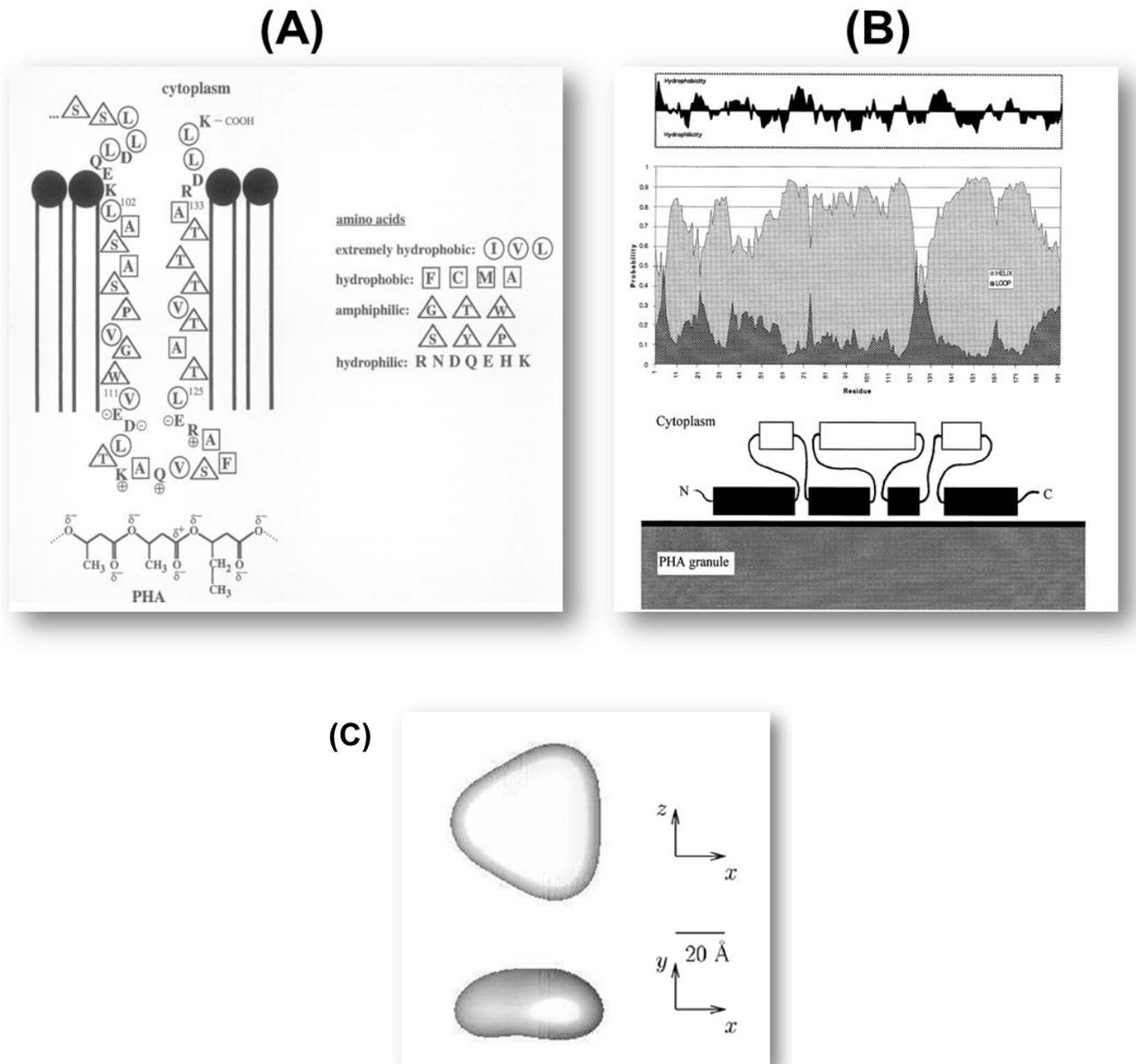
A estrutura cristalográfica destas proteínas é desconhecida até o momento, contudo alguns modelos têm sido propostos com base na sequência de aminoácidos. O primeiro deles foi proposto por Pieper-Fürst (1994) para a fasina GA14 de *Rhodococcus ruber*. Esta proteína é um tetrâmero determinado por técnicas de gel filtração e microscopia eletrônica e cujo monômero possui uma massa molecular de 15,5 kDa. A técnica de microscopia ainda revelou partículas triangulares e quadradas com dimensões entre 7 a 12 nm de diâmetro. Baseados na presença de dois segmentos de

aminoácidos hidrofóbicos e anfifílicos, os mesmos autores sugerem que a proteína encontra-se ancorada a superfície dos grânulos de PHA (Figura 6A).

A presença destes módulos também foi observada na fasina GA24 de *Alcaligenes eutrophus* (WIECZOREK *et al.*, 1995). Esta proteína possui massa molecular de 24 kDa, maior do que a anterior, composto por trechos de aminoácidos hidrofóbicos na porção C-terminal cuja função estaria relacionada a interação da proteína ao polímero. Hanley e colaboradores (1999) reavaliando a estrutura primária da fasina de *R. eutropha* concluíram que o modelo descrito para *A. eutrophus* não se aplicava neste organismo. Este estudo descreveu esta fasina como composta predominantemente por  $\alpha$ -hélices constituídas de regiões hidrofílicas flanqueadas por regiões hidrofóbicas. Estas características levaram os autores a propor outro modelo de interação fasina-grânulo onde os trechos hidrofílicos estariam voltados para o citoplasma e os trechos hidrofóbicos associando-se ao polímero (Figura 6B).

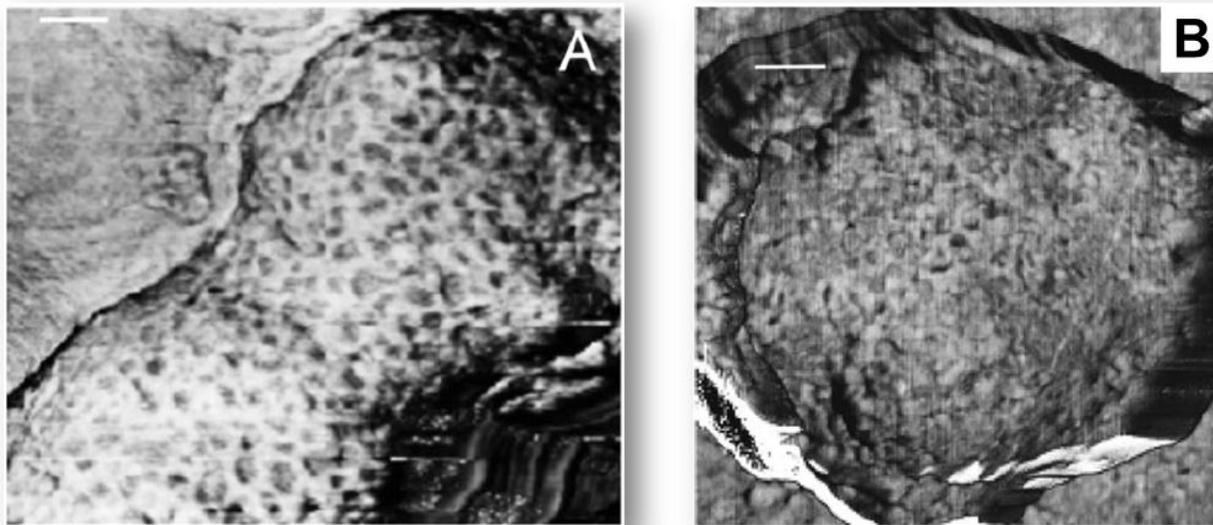
Nos últimos anos informações estruturais das fasinas em solução ou associadas aos grânulos de PHA tem sido reportadas. A informação mais próxima da estrutura tridimensional de uma fasina foi obtida por experimentos de difração de raios-X a baixo ângulo (SAXS) (NEUMANN *et al.*, 2008). Este estudo descreve a fasina de *R. eutropha* como uma proteína planar, triangular e com uma distância máxima de  $86 \pm 3 \text{ \AA}$  em solução (Figura 6C). Quando associada aos grânulos, PhaP forma estruturas em forma de rede (Figura 7) confirmada pela ausência da formação desta estrutura nanométrica na estirpe mutante *phaP* de *Cupriavidus necator* H16 (DENNIS *et al.*, 2008).

FIGURA 6 – Modelos estruturais da fasina



Informações estruturais das fasinas de *R. ruber* e *R. eutropha* (B-C). (A) Modelo de ancoramento da fasina GA14 de *R. ruber* com a camada de fosfolipídeos que cobrem os grânulos de PHA (PIEPER-FÜRST *et al.*, 1994). (B) Predições com base na sequência de aminoácidos da proteína PhnP de *R. eutropha*. Quadro superior mostra o perfil de hidrofobicidade. Quadro intermediário mostra a probabilidade de cada resíduo de aminoácido compor uma estrutura secundária em  $\alpha$ -hélice ou alça. Quadro inferior ilustra de forma esquemática o arranjo das regiões hidrofílicas e hidrofóbicas assim como a interação com a superfície do grânulo de PHA (HANLEY *et al.*, 1999). (C) Envelope molecular proposto para a proteína PhnP com base em experimentos de difração de raio-X a baixo ângulo (NEUMANN *et al.*, 2008).

**FIGURA 7 – Rede formada pelas fasinas na superfície dos grânulos de PHA**



Imagens de fase da superfície de grânulos extraídos de *C. necator* H16 utilizando a técnica de microscopia de força atômica (AFM) (DENNIS *et al.*, 2008). Inclusões extraídas da estirpe selvagem (A) e mutante *phaP* (B).

## 1.2. OBJETIVOS

### 1.2.1 Objetivos gerais

Caracterizar funcional e estruturalmente as proteínas PhbF e PhaP1 envolvidas no metabolismo do biopolímero polihidroxi butirato em *Herbaspirillum seropedicae* SmR1.

### 1.2.2 Objetivos específicos

- Expressar e purificar as proteínas PhbF e PhaP1 de *H. seropedicae*.
- Determinar a atividade *in vitro* de ligação das proteínas ao polímero PHB.
- Determinar a atividade de ligação ao DNA da proteína PhbF.
- Determinar o motivo de ligação ao DNA da proteína PhbF.
- Obter informações estruturais das proteínas PhbF e PhaP1 por SAXS.
- Propor um modelo estrutural para as proteínas PhbF e PhaP1.
- Determinar a função regulatória da proteína PhbF *in vivo*.

## 1.3. MATERIAL E MÉTODOS

### 1.3.1 Análise utilizando ferramentas de bioinformática

#### 1.3.1.1 Análise dos genes e proteínas alvo

A análise da sequência dos genes *phbF* e *phaP1* assim como das regiões a montante e a jusante destes foram realizadas por programas de bioinformática de acesso público na internet. O Artemis (RUTHERFORD *et al.*, 2000) foi utilizado para a identificação das ORFs e o PRODORIC (MÜNCH *et al.*, 2005) foi utilizado para a identificação de possíveis sequências promotoras. Os mapas de restrição e as sequências de aminoácidos traduzidas a partir das sequências de pares de base foram obtidas e editadas com o programa BioEdit (HALL, 2001).

Índices de similaridade assim como de identidade das sequências de aminoácidos foram obtidos pela submissão das sequências ao programa BlastP (<http://ncbi.nlm.nih.gov/BLAST>). O processo de alinhamento e avaliação da conservação de aminoácidos e domínios foi realizado pelo programa ClustalW (THOMPSON *et al.*, 1994). As sequências de comparação para o alinhamento foram obtidas do banco de dados GeneBank DataBase (<http://ncbi.nlm.nih.gov/Genbank>). A árvore de similaridade entre as sequências de aminoácidos foi obtida com o programa MEGA (TAMURA *et al.*, 2011) utilizando o modelo de *Neighbor-joining*. Características bioquímicas úteis durante os experimentos laboratoriais como massa molecular, ponto isoelétrico (pI) e coeficiente de extinção molar foram obtidas pelo programa PROTPARAM disponível no site [www.expasy.org](http://www.expasy.org).

A identificação de motivos de ligação ao DNA nas regiões promotoras foi realizada pelo programa MEME (BAILEY *et al.*, 2006). O programa foi configurado para identificar um motivo contendo de 6 a 50 pb. O motivo conservado foi representado no formato LOGO.

### 1.3.1.2 Análise das estruturas secundárias e geração dos modelos estruturais *ab initio*

A análise *in silico* da estrutura secundária das proteínas alvo foi realizada pelo programa Psipred (JONES, 1999). Os modelos estruturais *ab initio* foram obtidos pela submissão das sequências de aminoácidos das proteínas em estudo ao servidor Robetta (KIM *et al.*, 2004). As estruturas foram manipuladas e as imagens registradas utilizando o programa Pymol (SCHRODINGER, 2010). O potencial eletrostático dos modelos foi calculado utilizando o programa DelPhi (ROCCHIA *et al.*, 2001).

## 1.3.2 Materiais gerais e métodos de microbiologia

### 1.3.2.1 Bactérias e plasmídeos

As estirpes de bactérias e plasmídeos que foram utilizados neste trabalho estão relacionadas nos quadros 3 e 4, respectivamente.

#### QUADRO 3 – Lista de estirpes de bactérias utilizadas

| Estirpe                     | Genótipos e Características Relevantes   | Referência/ Fonte               |
|-----------------------------|--|---------------------------------|
| <b><i>E. coli</i></b>       |  |                                 |
| TOP10                       | F- <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-crBC</i> ) $\Phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74</i> <i>recA1</i> <i>ara</i> $\Delta$ 139 $\Delta$ ( <i>ara-leu</i> )7697 <i>galU galK rpsL</i> (StrR) <i>endA1 nupG</i> | Invitrogen Inc.                 |
| BL21 (DE3)                  | <i>Hsd gal</i> ( $\lambda$ <i>clts 857 ind 1 Sam7 nin5 lacUV5-T7 gene 1</i> )<br>(produtora da RNA polimerase do fago T7).   | New England Biolabs.            |
| ET8000                      | <i>rbs lacZ::IS1 gyrA hutCc k</i> (wild-type)  | MACNEIL <i>et al.</i> , 1982.   |
| S17.1                       | Sm <sup>R</sup> Tra <sup>+</sup>   | SIMON, PRIEFFER & PUHLER, 1983. |
| <b><i>H.seropedicae</i></b> |  |                                 |
| Smr1                        | Estirpe parental, Sm <sup>R</sup> , Nif <sup>+</sup>   | SOUZA <i>et al.</i> , 1995.     |

#### QUADRO 4 – Lista de plasmídeos utilizados

| Plasmídeos           | Características Relevantes  | Referência/ Fonte             |
|----------------------|---|-------------------------------|
| pCR <sup>®</sup> 2.1 | <i>LacZ</i> , Amp <sup>R</sup> , Km <sup>R</sup>  | Invitrogen Inc.               |
| pET-28a(+)           | Promotor T7, <i>lacI</i> , Km <sup>R</sup> , sequência codificadora para a cauda de histidinas.   | Novagen Inc.                  |
| pET-29a(+)           | Promotor T7, <i>lacI</i> , Km <sup>R</sup> .  | Novagen Inc.                  |
| pMP220               | Tc <sup>R</sup> , vetor contendo a fusão transcricional <i>lacZ</i> sem promotor e com o sítio de policlonagem do vetor pPW452 invertido. | SPAINK <i>et al.</i> , 1987   |
| pPW452               | Tc <sup>R</sup> , vetor contendo a fusão transcricional <i>lacZ</i> sem promotor.   | WOODLEY, PAUL                 |
| pKADO3               | pET-28a(+) contendo o gene <i>phbF</i> clonado.   | Este trabalho                 |
| pKADO4               | pET-29a(+) contendo o gene <i>phbF</i> clonado.   | Este trabalho                 |
| pKADO6               | pET-28a(+) contendo o gene <i>phbP1</i> clonado.  | Este trabalho                 |
| pKADO7               | pET-29a(+) contendo o gene <i>phbP1</i> clonado.  | Este trabalho                 |
| pKADO5               | 353 pb da região promotora do gene <i>phbF</i> clonado no vetor pMP220 resultando na fusão transcricional <i>phbF::lacZ</i> .             | Este trabalho                 |
| pKADO8               | 381 pb da região promotora do gene <i>phaP1</i> clonado no vetor pMP220 resultando na fusão transcricional <i>phbP1::lacZ</i> .           | Este trabalho                 |
| pMMS31               | Derivado do vetor pDK6 codificando o gene <i>phbF</i> de <i>H. seropedicae</i> SmR1.  | KADOWAKI <i>et al.</i> , 2011 |
| pKADO9               | 409 pb da região promotora do gene <i>phaP1</i> clonado no vetor pCR2.1   | Este trabalho                 |
| pKADO10              | 300 pb da região promotora do gene <i>phbF</i> clonado no vetor pCR2.1  | Este trabalho                 |
| pKADO11              | 542 pb da região promotora do gene <i>phbC_2999</i> clonado no vetor pCR2.1   | Este trabalho                 |
| pKADO12              | 404 pb da região promotora do gene <i>phaZ_0639</i> clonado no vetor pCR2.1   | Este trabalho                 |
| pKADO13              | 545 pb da região promotora do gene <i>phaC_2625</i> clonado no vetor pCR2.1   | Este trabalho                 |
| pKADO14              | 514 pb da região promotora da ORF <i>Hsero_0262</i> clonado no vetor pCR2.1   | Este trabalho                 |
| pKADO15              | 575 pb da região promotora do gene <i>dksA_2404</i> clonado no vetor pCR2.1   | Este trabalho                 |
| pKADO16              | 543 pb da região promotora do gene <i>fadB_0254</i> clonado no vetor pCR2.1   | Este trabalho                 |
| pKADO17              | 638 pb da região promotora do gene <i>phaB_3315</i> clonado no vetor pCR2.1   | Este trabalho                 |
| pKADO18              | 512 pb da região promotora do gene <i>phaZ_1622</i> clonado no vetor pCR2.1   | Este trabalho                 |

### 1.3.2.2 Meios de cultura, condições de cultivo e estoque dos microrganismos

As estirpes de *Escherichia coli* foram cultivadas em meio LB (SAMBROOK *et al.*, 1989) sob agitação (130 rpm) a 37 °C ou em estufa a 37 °C em meio LA. O meio LB possui a seguinte composição:

|                     |        |
|---------------------|--------|
| Triptona            | 10 g/L |
| Extrato de levedura | 5 g/L  |
| NaCl                | 10 g/L |
| pH                  | 7,0    |

O meio LA é composto de meio LB mais 15 g/L de ágar. Para o cultivo celular de *E.coli* com a finalidade de preparar células eletrocompetentes foi utilizado o meio SOB (Gibco-BRL-Invitrogen) que possui a seguinte composição:

|                     |           |
|---------------------|-----------|
| Triptona            | 20 g/L    |
| Extrato de levedura | 5 g/L     |
| NaCl                | 0,584 g/L |
| KCl                 | 0,186 g/L |
| pH                  | 7,0       |

Após o processo de transformação por eletroporação as células de *E. coli* foram recuperadas em meio SOC (Gibco-BRL-Invitrogen), que possui a seguinte composição:

|                     |           |
|---------------------|-----------|
| Triptona            | 20 g/L    |
| Extrato de levedura | 5 g/L     |
| NaCl                | 0,06 g/L  |
| KCl                 | 0,019 g/L |
| MgCl <sub>2</sub>   | 0,094 g/L |
| MgSO <sub>4</sub>   | 0,12 g/L  |
| Glucose             | 0,36 g/L  |

Os ensaios de  $\beta$ -galactosidase em *E. coli* ET8000 foram realizados em meio mínimo M9 (SAMBROOK *et al.*, 1989) suplementado com tiamina (1 $\mu$ g/mL) sob agitação (120 rpm) a 30 °C. O meio M9 (pH 7,0) possui a seguinte composição:

|   |          |
|---|----------|
| Na <sub>2</sub> HPO <sub>4</sub> .7H <sub>2</sub> O | 12,8 g/L |
| KH <sub>2</sub> PO <sub>4</sub>                     | 3,0 g/L  |
| NaCl  | 0,5 g/L  |
| NH <sub>4</sub> Cl                                  | 1,0 g/L  |
| MgSO <sub>4</sub>                                   | 2 mM     |
| CaCl <sub>2</sub>                                   | 0,1 mM   |
| Glucose   | 0,2%     |

A estirpe de *H. seropedicae* SmR1 foi cultivada a 30 °C, sob agitação a 120 rpm, durante 16 horas em meio NFbHPN (KLASSEN *et al.*, 1997), na presença de estreptomicina (80  $\mu$ g/mL). O meio NFb-malato (pH 6,8) possui a seguinte composição:

|   |                             |
|---|-----------------------------|
| MgSO <sub>4</sub> .7H <sub>2</sub> O                | 2 x 10 <sup>-1</sup> g/L    |
| NaCl  | 1 x 10 <sup>-1</sup> g/L    |
| CaCl <sub>2</sub>                                   | 2 x 10 <sup>-2</sup> g/L    |
| Ácido nitrilo-triacético                            | 5,6 x 10 <sup>-2</sup> g/L  |
| FeSO <sub>4</sub> . 7H <sub>2</sub> O               | 2 x 10 <sup>-2</sup> g/L    |
| Malato de sódio                                     | 5 g/L                       |
| Biotina   | 1 x 10 <sup>-4</sup> g/L    |
| Na <sub>2</sub> MoO <sub>4</sub> .2H <sub>2</sub> O | 2 x 10 <sup>-3</sup> g/L    |
| MnSO <sub>4</sub> .H <sub>2</sub> O                 | 2,35 x 10 <sup>-3</sup> g/L |
| H <sub>3</sub> BO <sub>3</sub>                      | 2,8 x 10 <sup>-3</sup> g/L  |
| CuSO <sub>4</sub> .5H <sub>2</sub> O                | 8 x 10 <sup>-5</sup> g/L    |
| ZnSO <sub>4</sub> .7H <sub>2</sub> O                | 2,4 x 10 <sup>-4</sup> g/L  |

No momento do uso, o meio NFbHPN foi obtido por adição de 50 mL/L de solução estéril de fosfatos (159,5 g/L de  $\text{KH}_2\text{PO}_4$  e 17,8 g/L de  $\text{K}_2\text{HPO}_4$ ) e de 20 mL/L de cloreto de amônio 1 mol/L ao meio NFb-Malato.

As células de *E. coli* foram estocadas a  $-20\text{ }^\circ\text{C}$  em glicerol 50% ou a  $-80\text{ }^\circ\text{C}$  em glicerol 15%.

### 1.3.2.3 Antibióticos

Os antibióticos utilizados nos meios de cultura foram:

| Antibiótico    | Abreviatura | Concentração Final   | Organismo                             |
|----------------|-------------|----------------------|---------------------------------------|
| Ampicilina     | Amp         | 250 $\mu\text{g/mL}$ | <i>E. coli</i>                        |
| Canamicina     | Km          | 50 $\mu\text{g/mL}$  | <i>E. coli</i>                        |
|                |             | 500 $\mu\text{g/mL}$ | <i>H.seropedicae</i>                  |
| Cloranfenicol  | Cm          | 30 $\mu\text{g/mL}$  | <i>E. coli</i>                        |
|                |             | 100 $\mu\text{g/mL}$ | <i>H.seropedicae</i>                  |
| Estreptomicina | Sm          | 80 $\mu\text{g/mL}$  | <i>E. coli</i> e <i>H.seropedicae</i> |
| Tetraciclina   | Tc          | 10 $\mu\text{g/mL}$  | <i>E. coli</i> e <i>H.seropedicae</i> |

As soluções estoque foram preparadas e estocadas segundo SAMBROOK *et al.*, 1989.

### 1.3.2.4 Extração de PHB de *H. seropedicae*

Para a extração de PHB de *H.seropedicae* foi utilizado o protocolo de tratamento com hipoclorito e extração com solvente proposto por WILLIAMSON & WILKINSON (1958) com modificações. *H. seropedicae* foi cultivado em 1000 mL de NFb-malato suplementado com 5 mM de glutamato, condição de baixa disponibilidade de nitrogênio, e solução de fosfatos como descrito no item 1.3.2.2. O cultivo foi realizado durante 16h a  $30\text{ }^\circ\text{C}$ . Em seguida, a cultura foi centrifugada a 5.000 xg por 10 minutos. O pellet celular foi ressuspensionado em uma solução de hipoclorito de sódio 3% (m/v) a uma

proporção final de 1% (m/v). Após ressuspensão, a solução foi incubada por 1h a 30 °C e centrifugada a 4000 xg por 10 minutos para a remoção dos restos celulares. O PHB recuperado na fração sedimentada foi lavado com água e em seguida com acetona. Entre cada lavagem a solução foi incubada a 30 °C por 1h. Após a última lavagem, o PHB foi seco e estocado a temperatura ambiente.

### **1.3.3 Métodos de manipulação de DNA**

#### **1.3.3.1 Extração e purificação de DNA genômico de *H. seropedicae***

*H. seropedicae* foi cultivado em meio NFbHPN e as células coletadas por centrifugação a 12.000 xg por 1 minuto. A massa de células foi ressuspensa, por agitação em “vortex”, em tampão GET (50 mmol/L glicose, 25 mmol/L Tris-HCl pH 8,0; 10 mmol/L EDTA pH 8,0). As células foram lisadas por tratamento com lisozima 100 µg/mL a 37 °C por 20 minutos e por adição de SDS 1% seguido de incubação a temperatura ambiente por cinco minutos. O material foi incubado com Proteínase K 200 µg/mL a 37 °C por 12 horas e a mistura foi centrifugada e o sobrenadante coletado foi extraído sequencialmente uma vez com fenol equilibrado, uma vez com fenol:clorofórmio (1:1) e duas vezes com clorofórmio: álcool isoamílico (24:1). O DNA genômico foi precipitado pela adição de 2 volumes de etanol 96%, lavado com etanol 80%, seco e ressuspensado em água ultra pura estéril.

#### **1.3.3.2 Reação em cadeia da DNA polimerase (PCR)**

O sistema de reação (20 µL) foi composto de: 1-5 ng/mL de DNA genômico de *H. seropedicae*, 0,5 pmol/µL de cada *primer*, 0,2 mmol/L de dNTPs, tampão de reação contendo 20 mmol/L (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> ou 50 mmol/L KCl, 1,5 mmol/L de cloreto de magnésio e 0,3 µL de Taq DNA Polimerase (5 U/µL). A amplificação foi realizada com o seguinte programa de 30 ciclos: 30 segundos a 94 °C, 30 segundos para anelamento e 45 segundos a 72 °C para a extensão. A princípio, a amplificação dos genes e regiões promotoras selecionados foi realizada utilizando uma curva de temperatura de 45 a 65 °C para o anelamento dos oligonucleotídeos iniciadores.

Os oligonucleotídeos iniciadores utilizados para a amplificação dos genes *phbF* e *phaP1* para a expressão e purificação das proteínas estão listados no quadro 5. Para a amplificação das regiões promotoras destinadas aos estudos *in vitro* e *in vivo* foram utilizados os oligonucleotídeos listados no quadro 6.

#### QUADRO 5 – Oligonucleotídeos utilizados para a amplificação dos genes alvo

| Oligonucleotídeo | ORF              | Sequência                                | Sítio de Restrição |
|------------------|------------------|--|--------------------|
| <i>phbF/F</i>    | <i>Hsero2997</i> | 5' GACTGGACTT <u>CATATG</u> ACTACTGC 3'  | <i>NdeI</i>        |
| <i>phbF/R</i>    | <i>Hsero2997</i> | 5' CAACAGGATCCGGCAGAATG 3'               | <i>BamHI</i>       |
| <i>phaP/F</i>    | <i>Hsero1639</i> | 5` GGATGTTTCC <u>CATATG</u> ACGACCTAC 3` | <i>NdeI</i>        |
| <i>phaP/R</i>    | <i>Hsero1639</i> | 5' GGCGCAGGATCCAATTACTTG 3'              | <i>BamHI</i>       |

\* *F*: oligonucleotídeo 5', *R*: oligonucleotídeo 3'. Os sítios de restrição introduzidos estão sublinhados.

**QUADRO 6 – Oligonucleotídeos utilizados para a amplificação das regiões promotoras**

| Oligonucleotídeo    | ORF              | Sequência                            | Sítio de Restrição |
|---------------------|------------------|--------------------------------------|--------------------|
| <i>PphaP1/F</i>     | <i>Hsero1639</i> | 5` <u>AGATCT</u> CACACCACACTCTCG 3`  | <i>Bgl</i> II      |
| <i>PphaP1/R</i>     | <i>Hsero1639</i> | 5` <u>CTGCAG</u> CGGAAAATTGCTC 3`    | <i>Pst</i> I       |
| <i>PphbC/F</i>      | <i>Hsero2999</i> | 5` CTGAGTTTCTGGATCGAAGG 3`           | -----              |
| <i>PphbC/R</i>      | <i>Hsero2999</i> | 5` <u>GGATCC</u> GGGATTGGAATTG 3`    | <i>Bam</i> HI      |
| <i>PphaZ/F</i>      | <i>Hsero0639</i> | 5` CGATGTCTTCCTCGAAGGAG 3`           | -----              |
| <i>PphaZ/R</i>      | <i>Hsero0639</i> | 5` <u>CTGCAG</u> TAGTTCTGGTAGAGC 3`  | <i>Pst</i> I       |
| <i>PphaC/F</i>      | <i>Hsero2625</i> | 5` ATTGTCTTCACCGACCATGC 3`           | -----              |
| <i>PphaC/R</i>      | <i>Hsero2625</i> | 5` <u>CTGCAG</u> TTCTTGCGGTTGATC 3`  | <i>Pst</i> I       |
| <i>PHsero0262/F</i> | <i>Hsero0262</i> | 5` <u>TCTAGAT</u> GAGCAAGCCATTGG 3`  | <i>Xba</i> I       |
| <i>PHsero0262/R</i> | <i>Hsero0262</i> | 5` <u>CTGCAG</u> GTCCGGAATAGTAC 3`   | <i>Pst</i> I       |
| <i>PdksA/F</i>      | <i>Hsero2404</i> | 5` <u>AGATCT</u> ATCATGGGCATGCTC 3`  | <i>Bgl</i> II      |
| <i>PdksA/R</i>      | <i>Hsero2404</i> | 5` <u>TCTAGAC</u> CAGCTGTTGTGTC 3`   | <i>Xba</i> I       |
| <i>PfadB/F</i>      | <i>Hsero0254</i> | 5` <u>TCTAGA</u> ATCGGCGACAAGGTC 3`  | <i>Xba</i> I       |
| <i>PfadB/R</i>      | <i>Hsero0254</i> | 5` <u>CTGCAG</u> ATGAAATTGGACACGC 3` | <i>Pst</i> I       |
| <i>PphaB/F</i>      | <i>Hsero3315</i> | 5` <u>AGATCT</u> CCGTCTCATCATCGAG 3` | <i>Bgl</i> II      |
| <i>PphaB/R</i>      | <i>Hsero3315</i> | 5` <u>TCTAGAC</u> CGCTGATGAAGACG 3`  | <i>Xba</i> I       |
| <i>PphaZ/F</i>      | <i>Hsero1622</i> | 5` GCGACATCAACACCTATCCG 3`           | -----              |
| <i>PphaZ/R</i>      | <i>Hsero1622</i> | 5` <u>CTGCAG</u> TCTGGTTTCAGAAACG 3` | <i>Pst</i> I       |
| <i>PphbF/F</i>      | <i>Hsero2997</i> | 5` <u>GGATCCT</u> TCGGACTTCTCCC 3`   | <i>Bam</i> HI      |
| <i>PphbF/R</i>      | <i>Hsero2997</i> | 5` <u>ATGCAT</u> GCGATTGGGGTATTTC 3` | <i>Nsi</i> I       |
| <i>PphbF_foot/F</i> | <i>Hsero2997</i> | 5` TTCGGACTTCTCCCTCAACG 3`           | -----              |
| <i>PphbF_foot/R</i> | <i>Hsero2997</i> | 5` TTGATCTTGCCCTCTCCTTG 3`           | -----              |
| Fluor-6FAM/F        | -----            | 5`-GTTGTAAAACGACGGCCAGT-3`           | -----              |
| Fluor-VIC/R         | -----            | 5`-TTCACACAGGAAACAGCTATGAC-3`        | -----              |

\* *F*: oligonucleotídeo 5', *R*: oligonucleotídeo 3'. ORF: código do gene anotado no genoma de *H. seropedicae*. Os sítios de restrição introduzidos estão sublinhados.

Todos os fragmentos amplificados foram em seguida clonados no vetor pCR@2.1 e confirmados por sequenciamento.

#### **1.3.3.3 Clonagem em vetor de expressão**

Os genes *phbF* e *phaP1* foram subclonados nos vetores pET-28a(+) e pET-29a(+) nos sítios de restrição *NdeI* e *BamHI*. Estes vetores permitem a superexpressão dos genes alvo utilizando o promotor do fago T7. A expressão a partir do vetor pET-28a(+) permite expressar a proteína alvo fusionada a uma cauda contendo seis resíduos de histidinas consecutivos importantes para o processo de purificação. O vetor pET-29a(+) permitiu expressar a proteína alvo desprovida de fusão traducional. Estes vetores permitem ainda controlar a expressão utilizando como indutor o IPTG.

#### **1.3.3.4 Clonagem em vetor contendo o gene repórter *lacZ***

Os vetores pMP220 e pPW452 foram utilizados para os estudos de expressão dos genes *phbF* e *phaP1* *in vivo*. Estes vetores possuem o gene repórter *lacZ* desprovido de promotor. A clonagem de regiões promotoras de interesse permitem observar a expressão gênica como uma medida indireta da atividade da proteína  $\beta$ -galactosidase descrita na seção 1.3.6.2.

As regiões promotoras dos genes *phbF*, contendo 353 pb incluindo 54 pb da região codificadora, e *phaP1*, contendo 381 pb incluindo 28 pb da região codificadora, foram clonadas no vetor pMP220 resultando nos plasmídeos denominados pKADO5 e pKADO8 (Quadro 4).

#### **1.3.3.5 Minipreparação de plasmídeo por lise alcalina**

A extração dos plasmídeos das células de *E. coli* foi realizada pelo método de lise alcalina (SAMBROOK *et al.*, 1989). Os plasmídeos purificados foram analisados por eletroforese em gel de agarose (SAMBROOK *et al.*, 1989).

### 1.3.3.6 Sequenciamento de DNA

A reação de sequenciamento foi realizada segundo o método de Sanger (SANGER & COULSON, 1977), onde os dideoxynucleotídeos foram utilizados como terminadores de cadeia. O sistema de reação foi constituído de aproximadamente 0,5 µg de DNA de fita dupla, 2 pmol de oligonucleotídeo iniciador, 3 µL de ET *terminador mix* (Sequencing Reagent Premix - DYEnamic™ ET Dye Terminator Cycle Sequencing Kit, GE HealthCare) que contém os cromóforos fluorescentes como terminadores de cadeia e água ultra pura suficiente para 7,5 µL. O sistema foi amplificado em termociclador utilizando a seguinte sequência de temperaturas por 30 ciclos:

| Temperaturas | Tempo       |
|--------------|-------------|
| 95 °C        | 1 minutos   |
| 94 °C        | 30 segundos |
| X            | 90 segundos |
| 60 °C        | 1 minuto    |

\* X: temperatura de anelamento diferente para cada oligonucleotídeo.

A seguir o produto das reações foi transferido para um novo tubo contendo 2 µL de acetato de amônio (7,5 mol/L) e 60 µL de etanol absoluto para a sua purificação. Após precipitação, a amostra de DNA foi centrifugada a 12.000 xg por 20 minutos, lavada uma vez com etanol 80% e secada. O DNA foi ressuspendido em 4 µL de *Formamide Loading Dye* (Applied Biosystems), desnaturado por 2 minutos a 96 °C e aplicado no gel do sequenciador automático de DNA ABI-PRISM 377 (Applied Biosystems).

### **1.3.3.7 Clivagem do DNA com enzimas de restrição**

O sistema de restrição foi composto das enzimas apropriadas, tampão de reação e DNA purificado, sendo em seguida incubado em temperatura e tempo sugeridos pelo fabricante.

### **1.3.3.8 Ligação DNA inserto ao vetor**

A reação de ligação foi composta de: 2 µL de vetor (aproximadamente 50 ng/µL); 10 µL de inserto (aproximadamente 100 ng/µL); 2 µL de tampão para a enzima T4 DNA Ligase (10X) e 0,5 U de T4 DNA Ligase, para um volume final de 20 µL. A reação foi incubada a 4 °C durante um período de 12 horas. O produto de ligação foi introduzido por choque elétrico ou térmico na estirpe *E. coli* TOP10, conforme item 1.3.3.9, e os transformantes foram selecionados pela resistência a antibiótico(s).

### **1.3.3.9 Eletroforese de DNA**

Eletroforeses de DNA foram realizadas em gel de agarose em cuba horizontal como descrito por SAMBROOK *et al.* (1989). O DNA separado em gel foi tratado com brometo de etídeo (0,5 µg/mL) e o perfil eletroforético foi registrado utilizando transluminador de luz ultravioleta com sistema de vídeo-imagem acoplado (UVP).

### **1.3.3.10 Transformação bacteriana**

#### **3.3.10.1 Transformação por eletroporação em *E. coli***

DNA plasmidial foi transferido para células de *E. coli* tornadas eletrocompetentes conforme protocolo do fabricante do eletroporador Cell-Porator Voltage Booster (Gibco-BRL-Invitrogen), com modificações.

Uma cultura saturada de *E. coli* foi adicionada a 100 mL de meio SOB na proporção de 1:100. A cultura foi mantida sob agitação (130 rpm) a 37 °C até atingir

D.O.<sub>600</sub> de aproximadamente 0,5. A cultura foi mantida então em gelo por 30 minutos. Em seguida, as células foram coletadas por centrifugação a 5000 xg por 5 minutos a 4 °C e lavadas duas vezes com água ultrapura estéril gelada e uma vez com glicerol 15% estéril gelado. Em seguida foram ressuspensas em 100 µL de glicerol 15%, separadas em alíquotas de 35 µL e estocadas a -70 °C ou transformadas imediatamente.

Para a transformação por eletroporação, 35 µL de células competentes e 1 µL de solução de plasmídeo (5 ng - 0,5 µg/µL) foram misturados e transferidos para uma cubeta de eletroporação. O aparelho foi ajustado para a aplicação do choque elétrico numa diferença de potencial de 16 a 21 Kv/cm e um campo elétrico de 4 kΩ/330 µF à câmara de eletroporação, com tempo de pulso de 6 a 10 ms. Após a aplicação deste processo, as células de *E. coli* foram inoculadas em 1 mL de meio SOC, para recuperação, por 60 minutos a 37 °C e 130 rpm. Em seguida, foram espalhadas em meio LA (contendo os antibióticos necessários para a seleção dos clones) e incubadas a 37 °C durante um período de 12 horas.

#### 1.3.3.10.2 Transformação por choque térmico em *E. coli*

O preparo das células de *E. coli* competentes ao choque térmico foi realizado segundo CHUNG e colaboradores (1989). Uma cultura saturada de *E. coli* foi reinoculada em 10 mL de meio LB na proporção de 1:100. A cultura foi mantida sob agitação (130 rpm) a 30 °C até atingir D.O.<sub>600</sub> entre 0,3-0,4. A cultura foi centrifugada a 5000 xg por 1 minutos a 4 °C e o pellet celular ressuspensado em 1 mL da solução TSS filtrada e gelada. Esta solução possui a seguinte composição:

|  |                           |
|--|---------------------------|
| PEG 3350 ou 8000                       | 10% (massa/vol)           |
| DMSO                                   | 5% (vol/vol)              |
| MgSO <sub>4</sub> ou MgCl <sub>2</sub> | 20-50 mM Mg <sup>+2</sup> |
| Em LB, pH 6,5                          |                           |

A solução de células foi separada em alíquotas de 100 µL, misturada a 1 µL de solução de plasmídeo (5 ng - 0,5 µg/µL).

A transformação foi realizada incubando-se a mistura por 30 minutos em gelo, 45 segundos a 42 °C em banho-maria e 2 minutos em gelo. Após o choque térmico, as células foram recuperadas em 0,9 mL de meio LB por 60 minutos a 37 °C. Em seguida, foram espalhadas em meio LA (contendo os antibióticos necessários para a seleção dos clones) e incubadas a 37 °C durante um período de 12 horas.

#### **1.3.3.11 Transferência de DNA plasmídial por conjugação bacteriana biparental**

Plasmídeos contendo as fusões transcricionais *phbF::lacZ* e *phaP1::lacZ* foram transformados em *E. coli* S17.1. As estirpes *H. seropedicae* SmR1 (receptora) e *E. coli* (doadora) foram cultivadas por 12 horas em meio NFbHPN e LB, respectivamente, a partir de culturas saturadas. Em seguida, 1 mL da cultura de *H. seropedicae* SmR1 foi centrifugado a 2.000 *xg* por 30 segundos e o pellet celular ressuscitado em 100 µL de meio NFbHPN. Esta solução de células foi misturada com 50 µL da cultura de *E. coli* S17.1 contendo o plasmídeo desejado e depositada em forma de gota em meio LA/NFbHPN na proporção de 1:3, na ausência de antibióticos e incubada a 30 °C por 12h. A massa celular foi raspada da placa e ressuscitada em 1 mL de meio NFbHPN sendo em seguida submetida às diluições de 1:100 e 1:1000. As diluições foram espalhadas em meio NFbHPN contendo os antibióticos adequados e as placas foram incubadas a 30 °C.

#### **1.3.4 Métodos de expressão e manipulação de proteínas**

##### **1.3.4.1 Determinação das condições de expressão**

Os plasmídeos recombinantes obtidos foram transformados na estirpe de *E. coli* BL21 (DE3) destinada à expressão. Três colônias de *E. coli* contendo o plasmídeo de interesse foram inoculadas em 3 mL de meio LB com os antibióticos adequados, e cultivadas a 37 °C por 12 horas a 130 rpm. Cem microlitros deste cultivo foram re-inoculados em 10 mL de meio LB (proporção 1:100) e incubados a 37 °C e 130 rpm, a

fim de se obter um cultivo celular com D.O.<sub>600nm</sub> entre 0,4 e 0,5. A partir desta fase foram variadas as condições de expressão quanto à temperatura (20, 30 e 37 °C), período de expressão (3h ou 12h) e tipo indutor (1 mmol/L de IPTG ou 14 mmol/L de lactose). A lactose é um indutor que pode ser utilizado, no lugar do IPTG, com o objetivo de aumentar a solubilidade de proteínas durante a expressão (MONTEIRO *et al.*, 2000). Após a expressão, as culturas (10 mL) foram centrifugadas a 7.000 xg durante 5 minutos a 4 °C, o meio de cultura foi descartado, e o pellet celular foi ressuscitado em 400 µL de tampão de ressuspensão. Neste tampão foram variadas as concentrações e os tipos de sal, diferentes tipos de tampões e presença de determinados detergentes quando necessário. Uma alíquota de 20 µL da fração bruta foi estocada para posterior análise em gel SDS-PAGE. O restante da fração bruta foi submetido à lise celular no aparelho Ultrasonic Processor XL Heat Systems. A sonicação foi realizada em 8 ciclos de 20 segundos, alternando-se iguais períodos de repouso em gelo. Em seguida as amostras foram centrifugadas a 12.000 xg durante 15 minutos para a separação das frações solúvel e insolúvel. A fração insolúvel foi ressuscitada em tampão de ressuspensão com igual volume ao da fração solúvel.

#### **1.3.4.2 Análise de proteínas por eletroforese em gel de poliacrilamida**

As amostras protéicas foram analisadas através de eletroforese em gel de poliacrilamida 12% (SDS-PAGE) (LAEMMLI, 1970) ou em gel nativo na ausência de SDS. As proteínas foram visualizadas após a coloração por Coomassie Blue (LAEMMLI, 1970), sendo em seguida foto-documentadas no aparelho UVP-Biolmaging Systems.

#### **1.3.4.3 Determinação do grau de pureza das proteínas**

Para avaliar o grau de pureza das proteínas superexpressas e purificadas, uma amostra foi submetida à eletroforese SDS-PAGE. O gel foi corado com Coomassie Blue e submetido à análise densitométrica no aparelho Personal Densitometer SI –

Molecular Dynamics. A imagem do gel foi capturada e armazenada pelo programa Scan Control. A pureza foi determinada por densitometria pelo programa ImageQuant™.

#### **1.3.4.4 Análise espectrométrica de massa das proteínas expressas**

A confirmação da identidade das proteínas pelo método MALDI-TOF/MS foi realizada como descrito por WESTERMEIER & NAVEN (2004) com adaptações.

##### **1.3.4.4.1 Tratamento das amostras por digestão trípica**

A banda correspondente à proteína de interesse foi recortada do gel e o corante Coomassie removido utilizando a solução descolorante (75 mmol/L de bicarbonato de amônio e 40% etanol). O gel foi em seguida desidratado utilizando uma solução de acetonitrila 100%. A proteína aprisionada no gel foi digerida com uma solução de tripsina (10 µg/mL) e incubada por 20 minutos em gelo. O gel foi em seguida coberto com solução tampão contendo 40 mmol/L de bicarbonato de amônio e 10% de acetonitrila e incubado por 12h a 37 °C. 2 µL de amostra foram misturadas a igual volume de solução contendo a matriz CHCA (Ácido  $\alpha$ -ciano-4-hidroxicinamímico) e aplicada sobre uma placa metálica. A análise dos peptídeos foi realizada pelo método MALDI-TOF/MS (Espectrometria de Massas de Ionização a Laser com Tempo de Vôo) no espectrômetro de massa Bruker Daltonics Inc., no modo refletor positivo. A calibração do aparelho foi realizada utilizando-se uma mistura de padrões de peptídeos (Bruker Daltonics Inc.). Após a análise de cada amostra foram gerados espectrogramas que foram visualizados utilizando o programa Flexanalysis (Bruker Daltonics Inc.).

##### **1.3.4.4.2 Identificação das proteínas**

A identificação das proteínas foi realizada através do programa Aldente (<http://www.expasy.org/tools/aldente/>), que utiliza os bancos de dados do Swiss-Prot e do TrEMBL, e do programa Protein Prospector v. 4.0 (<http://prospector.ucsf.edu/>) que utiliza o banco contendo dados dos peptídeos provenientes da digestão trípica *in silico*

de todas as ORFs anotadas do genoma de *H. seropedicae*. Foram utilizados os valores dos picos monoisotópicos referentes à m/z para a submissão nos bancos de dados.

#### **1.3.4.5 Purificação das proteínas de *H. seropedicae***

A purificação das proteínas alvo foi realizada por cromatografia de alto desempenho (HPLC) no cromatógrafo Äkta Basic (GE HealthCare). Para cada purificação as soluções foram cuidadosamente degaseificadas utilizando sistemas de vácuo. As proteínas alvo foram purificadas na forma fusionada a cauda de histidinas.

A fração solúvel do extrato celular foi coletada e submetida à cromatografia de afinidade. A coluna de purificação utilizada foi a Hi-Trap Chelating (GE Health Care), carregada manualmente com solução de Ni (100 mmol/L de acetato de amônio pH 4,0 e 100 mmol/L de cloreto de níquel) e em seguida acoplada ao equipamento. A solução de proteínas foi injetada, num fluxo de 1 mL/min, na coluna Hi-Trap Chelating (1 mL ou 5 mL) equilibrada com 3 VC de tampão A contendo 10 mmol/L de imidazol, tampão e sal apropriados para cada proteína. A proteína foi eluída utilizando um gradiente contínuo de 10 a 1000 mmol/L de imidazol equivalente a 20 VC. As amostras foram coletadas em alíquotas de 1 mL. Amostras de 10 µL das frações obtidas e correspondentes ao pico no cromatograma foram coletadas para análise em gel SDS-PAGE. As frações que apresentaram maior pureza foram reunidas e dialisadas contra 1 L de solução de troca. A diálise foi realizada a temperatura de 4 °C por 12 horas.

#### **1.3.4.6 Quantificação da concentração protéica**

As proteínas foram quantificadas pelo método de BRADFORD (1976) ou utilizando o coeficiente de extinção molar (PACE *et al.*, 1995) no caso da proteína purificada.

Pelo método de Bradford, o sistema foi composto de 170 µL de solução corante (100 mg de Coomassie Blue G-250; 50 mL de etanol 95% e 100 mL ácido fosfórico 85%) e 30 µL de amostra. Esta solução foi homogeneizada e sua absorbância a 595 nm

medida. A curva padrão foi determinada utilizando albumina de soro bovino (BSA) como padrão.

Pelo coeficiente de extinção molar, uma alíquota (1  $\mu$ L) da solução de proteína foi diluída inicialmente 100X em tampão de purificação (50 mmol/L de Tris-HCl pH 7,5, 100 mmol/L de NaCl e 0,05% de TritonX-100) e a sua absorbância a 280 nm foi medida. O valor de absorbância foi aplicado na fórmula ( $C=A/\epsilon$ ), onde “C” corresponde à concentração molar protéica, A corresponde à absorção a 280 nm e  $\epsilon$  corresponde ao valor do coeficiente de extinção molar. O valor de  $\epsilon$  foi obtido pela submissão da sequência de aminoácidos da proteína alvo ao programa ProtParam, disponível no site do EXPASY ([www.expasy.ch/tools/protparam.html](http://www.expasy.ch/tools/protparam.html)). Esta função foi utilizada tanto para a determinação da concentração molar quanto para a determinação da concentração em mg/mL.

#### **1.3.4.7 Concentração das proteínas por ultrafiltração**

As proteínas foram concentradas por ultrafiltração utilizando tubos do tipo Centricon (Millipore) com limite de exclusão de 10 kDa. Amostras entre 2 e 10 mL foram transferidas para os tubos e submetidas a ciclos centrifugações de 7.500  $xg$  por 10 minutos até atingir o volume final desejado.

#### **1.3.5 Métodos *in vitro* de caracterização de proteínas**

##### **1.3.5.1 Ensaio de ligação das proteínas PhbF e PhaP1 ao PHB**

As proteínas alvo purificadas foram testadas quanto a sua atividade de interação com PHB na sua forma cristalina purificada. O ensaio de interação foi baseado nos protocolos propostos por MIYAZAKI e colaboradores (2002) e PÖTTER e colaboradores (2004).

As reações de ligação das proteínas alvo (PhbF e PhaP1) ao PHB foram realizadas em um volume final de 100  $\mu$ L contendo: 25  $\mu$ g de proteína purificada, 1 mg de PHB, 50 mM Tris-HCl pH 7,5. A reação foi incubada a 37 °C por 10 minutos. Em seguida a solução foi centrifugada a 10.000  $xg$  por 1 minuto e a solução foi reservada

para posterior análise. O material sedimentado foi lavado duas vezes com 400 µL de solução contendo 50 mM Tris-HCl pH 7,5. A atividade de ligação ao PHB foi observada pela permanência da associação após estas lavagens. As proteínas associadas ao PHB foram dissociadas por adição de 20 µL de tampão contendo 2% SDS, 10% glicerol, 0,01% azul de bromofenol, 0,0625 M Tris-HCl pH 6,8 e 5% β-mercaptoetanol. As frações reservadas foram analisadas por SDS-PAGE 12%. Como controle foi utilizado uma fração contendo só o PHA que mostrou a ausência de proteínas previamente associadas. Outro controle utilizado foi BSA na mesma concentração das proteínas PhbF e PhaP1.

### 1.3.5.2 Ensaio de retardamento eletroforético de banda em gel

As regiões promotoras selecionadas (1.3.1.1) e relacionadas ao metabolismo de PHAs foram amplificadas utilizando os oligonucleotídeos Fluor-6FAM/F e Fluor-VIC/R (Quadro 6) marcados em suas extremidades 5' com as sondas fluorescentes FAM e VIC. As reações de amplificação utilizaram como molde as regiões promotoras clonadas no vetor pCR2.1. Para alguns ensaios, as regiões promotoras dos genes *phbF* e *phaP1* foram amplificadas utilizando oligonucleotídeos específicos sendo um deles marcado radioisotopicamente utilizando [<sup>32</sup>P]γ-ATP por ação da enzima T4 polinucleotídeo quinase (SAMBROOK *et al.*, 1989). Os ensaios de ligação ao DNA foram realizados em 10 µL contendo 20 nmol/L de DNA marcado, 100 ng de DNA de timo de bezerro e quantidades crescentes de proteína purificada em tampão de ligação (10 mmol/L de Tris-HCl pH 7,5, 80 mmol/L de NaCl, 1 mmol/L de EDTA, 10 mmol/L de β-mercaptoetanol e 5% (m/v) de glicerol). A incubação foi conduzida por 5 minutos a 30 °C. O complexo DNA-proteína foi separado do DNA livre em eletroforese em gel de poli(acrilamida em condição nativa 8% (acrilamida/bisacrilamida 29:1 (m/m)) em 2X TAE). O DNA fluorescente foi observado por excitação com luz UV (254 nm) e o DNA marcado com [<sup>32</sup>P] foi detectado utilizando o sistema PhosphorImager e registrado com o scanner STORM. As constantes de dissociação (K<sub>d</sub>) da proteína PhbF utilizando os promotores *phbF* e *phaP1* marcados com [<sup>32</sup>P] foram estimadas medindo-se a

concentração proteica correspondente a 50% de formação do complexo DNA-proteína com base na análise densitométrica das autoradiografias.

### **1.3.5.3 Ensaio de proteção de digestão pela DNase I**

Um fragmento de 325 pb contendo parte da região promotora do gene *phbF* foi amplificado a partir do DNA genômico extraído de *H. seropedicae* SmR1 utilizando os oligonucleotídeos *PphbF\_foot/F* e *PphbF\_foot/R* marcado previamente com [<sup>32</sup>P] (Quadro 6) (SAMBROOK *et al.*, 1989). O DNA marcado foi purificado utilizando o kit Wizard (Promega) e incubado com PhbF em 50 mmol/L de Tris-Acetato pH 8, 10 mmol/L de KCl, 0,8 mmol/L de acetato de magnésio e a 30 °C por 5 minutos. Para a hidrólise parcial, 1 U de DNaseI (Invitrogen) foi adicionada e a reação foi incubada a 30 °C por 1 minuto. Em seguida a reação foi interrompida por adição de 0,2 volume de 500 mmol/L de EDTA e o sistema aquecido a 80 °C por 5 minutos. O produto da reação foi precipitado na presença de tRNA de levedura (0,5 mg/mL) e etanol 96% a -80 °C por 1 hora. A amostra seca foi ressuspendida em 6 µL de tampão contendo 47% de formamida (v/v), 10 mmol/L de EDTA, 0,05% azul de bromofenol (m/v), 0,05% de xileno cianol (m/v). A mostra foi em seguida desnaturada a 80 °C por 5 minutos e aplicada em gel de poliacrilamida desnaturante de sequenciamento 6% (contendo 8 mol/L de uréia) em TBE 1X. A região promotora do gene *phbF* foi sequenciada utilizando o kit T7 sequencing (GE Healthcare). As autoradiografias foram obtidas utilizando o sistema PhosphorImager e registradas utilizando o scanner Storm. O perfil densitométrico foi analisado utilizando o programa ImageQuant v. 5.2 (Molecular Dynamics).

### **1.3.6 Métodos *in vivo* de caracterização das proteínas alvo**

#### **1.3.6.1 Imunodeteção**

##### **1.3.6.1.2 Preparo das amostras**

Para seguir a expressão das proteínas PhbF e PhaP1 ao longo do crescimento celular assim como a mobilidade das mesmas do citoplasma para os grânulos de PHB,

*H. seropedicae* SmR1 foi cultivado em meio NFbHPN (100 mL) a partir de uma D.O.<sub>600</sub> inicial de 0,01. A partir das 4 horas de crescimento, 10 mL de cultura foram coletados a cada 2 horas e as células separadas por centrifugação a 5.000 xg por 10 minutos. O *pellet* celular foi ressuscitado em 300 µL de tampão contendo 50 mM Tris-HCl pH 8,0 e 100 mM NaCl e 100 µL da amostra foram reservados. Os demais 200 µL foram submetidos à lise celular pelo método de sonicação utilizando 8 ciclos de 20 segundos, alternando-se com o mesmo período de repouso em gelo. Em seguida as amostras foram centrifugadas a 12.000 xg durante 15 minutos para a separação da fração de proteínas solúveis. O *pellet* contendo proteínas insolúveis e proteínas agregadas aos grânulos de PHB foi ressuscitado em 200 µL de tampão de amostra (2% SDS, 10% glicerol, 0,01% azul de bromofenol, 0,06 mol/L Tris-HCl pH 6,8 e 5% β-mercaptoetanol). A amostra equivalente às proteínas totais e solúveis foram tratadas com NaOH a uma concentração final de 0,1 mol/L por 30 minutos e a concentração protéica determinada pelo método de Bradford.

As amostras protéicas foram separadas em gel do tipo SDS-PAGE 12% utilizando 5 µg de proteína por poço.

#### 1.3.6.1.3 Transferência de proteínas e tratamento das membranas

Após a corrida eletroforética, as proteínas foram transferidas para membranas de fluoreto de polivinilideno (PVDF Hybond P – GE Healthcare) utilizando um sistema de transferência do tipo semi-seco.

As membranas de PVDF foram ativadas por 10 segundos em metanol, lavadas por 5 minutos em água e incubadas juntamente com os géis em tampão de transferência (pH 7,6) por 10 minutos. Este tampão possui a seguinte composição:

|         |           |
|---------|-----------|
| Tris    | 1,51 g/L  |
| Glicina | 9,37 g/L  |
| Metanol | 10% (v/v) |

A placa de anodo foi coberta com 4 camadas de papel filtro embebidas em tampão de transferência, seguido da membrana de PVDF, do gel e mais 4 camadas de papel filtro. A placa do catodo foi colocada sobre as folhas de filtro e conectada ao sistema. A transferência foi realizada utilizando uma voltagem máxima de 20 V e um mínimo de 100 mA por gel durante 1 hora. As membranas foram bloqueadas por 12 horas em tampão TBS contendo 5% de leite em pó desnatado. O tampão TBS possui a seguinte composição:

|      |                   |
|------|-------------------|
| Tris | 12,1 g/L (pH 7,6) |
| NaCl | 40 g/L            |

A membrana foi lavada duas vezes durante 10 minutos com tampão TBST (TBS + 0,1% Tween 20) e uma vez com tampão TBS. A incubação com anticorpo primário foi realizado em 10 mL de solução de leite em pó desnatado (5%) utilizando uma diluição de 1:25000. Os anticorpos policlonais para as proteínas PhbF e PhaP1 foram obtidas por imunização de coelhos, foram comprados da empresa Célula B - Desenvolvimento de Anticorpos (UFRGS). Para diminuir o reconhecimento inespecífico de proteínas, extrato bruto de proteínas de *E. coli* BL21 (DE3) a uma concentração final de 100 µg/mL foi também adicionado à solução. O sistema foi incubado a temperatura ambiente e sob agitação durante 1 hora.

Após a incubação com anticorpo primário, as membranas foram lavadas duas vezes durante 10 minutos com tampão TBST (TBS + 0,1% Tween 20) e 1 vez com tampão TBS. As membranas foram em seguida incubadas em 10 mL de solução de leite em pó desnatado (5%) contendo anticorpo secundário (anti IgG de coelho conjugado com a enzima peroxidase – GE Healthcare) a uma diluição de 1:5000 e mantidas sob agitação a temperatura ambiente por 1 hora. Após 5 lavagens com tampão TBST, as membranas foram secadas e submetidas a revelação utilizando o kit ECL (GE Healthcare) e a imagem registrada no sistema de foto documentação UVP.

### 1.3.6.2 Determinação da atividade de $\beta$ -galactosidase

A atividade de  $\beta$ -galactosidase em células de *E. coli* ET8000 e *H. seropedicae* SmR1, cultivadas respectivamente nos meios mínimo M9 e NFbHPN, foi determinada segundo MILLER (1990).

O sistema de reação foi composto de 900  $\mu$ L de tampão Z, 20  $\mu$ L de clorofórmio e 100  $\mu$ L de cultura que foi em seguida misturado e incubado por 5 minutos a 30 °C. O tampão Z completo possui a seguinte composição:

|                                  |            |
|----------------------------------|------------|
| Na <sub>2</sub> HPO <sub>4</sub> | 60 mmol/L  |
| NaH <sub>2</sub> PO <sub>4</sub> | 40 mmol/L  |
| KCl                              | 10 mmol/L  |
| MgSO <sub>4</sub>                | 1 mmol/L   |
| $\beta$ -mercaptoetanol          | 50 mmol/L  |
| SDS                              | 0,1% (m/v) |
| pH                               | 7,0        |

A reação foi iniciada pela adição de 200  $\mu$ L do substrato ONPG (4 mg/mL) e interrompida pela adição de 500  $\mu$ L de Na<sub>2</sub>CO<sub>3</sub> (1 mol/L). As amostras foram submetidas à leitura de absorvância do o-nitrofenol (420 nm), turbidez da mistura (550 nm) e densidade celular (600 nm).

A atividade foi expressa em unidades Miller segundo a fórmula:

$$\text{Unidades Miller} = 1000 \times (\text{DO}_{420\text{nm}} - 1,75 \times \text{DO}_{550\text{nm}}) / (\text{D.O.}_{600\text{nm}} \times t \times v)$$

onde t = tempo de reação (minutos) e v = volume de cultura (mL).

Nos ensaios em *E. coli* as atividades foram realizadas em triplicata de leitura após 8 horas de crescimento. Em *H. seropedicae*, a atividade foi seguida durante o crescimento celular e os ensaios foram realizados em triplicata de leitura e em triplicata biológica.

### **1.3.7 Métodos de caracterização estrutural das proteínas alvo**

#### **1.3.7.1 Dicroísmo circular**

As medidas de dicroísmo circular (CD) foram realizadas no espectropolarímetro Jasco model J-810, localizado no Laboratório Nacional de Luz Síncrotron (Campinas-SP). O espectro UV (200-260 nm) foi coletado utilizando as proteínas alvo a uma concentração de 16,9  $\mu\text{mol/L}$  (PhbF) e 41  $\mu\text{mol/L}$  (PhaP1) no tampão de purificação a 25 °C. As cubetas de quartzo tinham um caminho óptico de 0,1 cm. Os dados foram coletados com uma resolução de 0,5 nm e registrados com uma velocidade de 100 nm/min. Cada espectro resultou de uma média de vinte registros com uma resposta de 1s. Os dados correspondentes ao branco foram coletados usando somente o tampão da proteína. Cada espectro foi corrigido à linha de base e normalizado a  $[\theta]_{\text{mrw}}$  (graus  $\text{cm}^2 \text{dmol}^{-1}$ ) de acordo com ADLER *et al.*, 1973.

#### **1.3.7.2 Determinação do estado de oligomerização das proteínas**

A massa molecular do oligômero das proteínas em estudo foi estimada por gel filtração na coluna Superose 12HR 10/30 (GE HealthCare). Esta coluna tem as seguintes dimensões: 1 cm de diâmetro, 30 cm de altura e 23,52 ml de volume total. A eluição foi realizada no cromatógrafo ÄKTA Basic System (GE HealthCare). A coluna foi equilibrada com 3 VC de tampão apropriado para cada proteína a um fluxo de 0,5 mL/min. Os padrões proteicos utilizados foram: citocromo C (12,4 kDa), anidrase carbônica (29 kDa), albumina de soro bovino (BSA) (66 kDa), álcool desidrogenase (150 kDa) e  $\alpha$ -amilase (200 kDa) (GE Healthcare). Após a eluição dos padrões e da proteína em estudo (1 mg) foram determinados seus volumes de eluição e respectivas constantes de partição. Com estes dados, foi montado um gráfico contendo as informações de volume de partição em função do log da massa molecular de cada proteína padrão. A determinação da equação de reta foi realizada por regressão linear usando o programa Excel. As proteínas em estudo foram eluídas em seus respectivos

tampões e seus volumes de eluição foram registrados. A massa molecular foi estimada pela interpolação da constante de partição no gráfico preparado anteriormente.

### **1.3.7.3 Espalhamento de raios-X a baixo ângulo (SAXS)**

#### **1.3.7.3.1 Introdução à técnica**

O espalhamento de raios-X a baixo ângulo é um método que permite caracterizar a conformação geral de proteínas ordenadas ou desordenadas de diferentes conformações e tamanhos em solução a uma resolução entre 50 e 10 Å (FEIGIN & SVERGUN, 1987; BERNADÓ & SVERGUN, 2011; PUTNAM *et al.*, 2007). Esta técnica apresenta-se como uma complementação dos métodos de alta resolução como cristalografia de raios-X, limitada a produção de cristais, e ressonância magnética nuclear (RMN) limitada ao tamanho da proteína (< 30 kDa). A técnica de SAXS pode fornecer informações importantes sobre massa molecular, raio de giro, volume, conformação geral, organização espacial dos domínios proteicos e mudanças estruturais relacionadas à função (TREWHELLA, 1997; TSUTAKAWA *et al.*, 2007).

O experimento consiste na exposição da solução de proteínas, acomodadas em um capilar ou cubeta, a um feixe colimado de raios-X e cuja intensidade espalhada é registrada em função do ângulo de espalhamento  $2\theta$  por detectores bidimensionais. A mesma medida é realizada apenas para o tampão na qual se encontra a proteína (BERNADÓ & SVERGUN, 2011). A subtração do espalhamento observado para a solução de proteínas em relação ao tampão resulta na informação de espalhamento apenas da macromolécula. Esta diferença deve ser superior a três ordens de magnitude para gerar dados com precisão adequada para os estudos estruturais (PUTNAM *et al.*, 2007). Em solução, as proteínas apresentam um espalhamento radialmente isotrópico e proporcional ao espalhamento de uma única partícula média, calculada em relação a todas as orientações adotadas aleatoriamente. Esta afirmativa é verdadeira para sistemas monodispersos (FEIGIN & SVERGUN, 1987; BERNADÓ & SVERGUN, 2011). A informação contida no padrão de espalhamento é submetida a uma média circular e a intensidade é plotada em função do vetor de espalhamento  $q$  ( $(4\pi \sin\theta)/\lambda$ ) gerando uma

curva de espalhamento. As unidades  $q$  tem uma relação inversa ao comprimento de onda ( $q=2\pi/d$ ) e fornecem informação sobre a resolução dos dados (PUTNAM *et al.*, 2007).

A curva de espalhamento pode fornecer várias informações. A baixa resolução, o plot de Guinier ( $\ln(I)$  vs.  $q^2$ ) mostra uma função linear próxima ao ângulo zero e cuja inclinação é usada para calcular o raio de giro ( $R_g$ ). Sendo o  $R_g$  a raiz quadrada do quadrado das distâncias médias de cada átomo em relação ao centro da partícula. Este parâmetro permite avaliar a forma, o tamanho e a compactação da partícula em estudo (GUINIER & FOURNET, 1955). Esta função pode ainda fornecer uma ferramenta para monitorar a qualidade da amostra em termos de agregação. A maior resolução, a intensidade de espalhamento é dependente do formato da macromolécula. As proteínas estruturadas apresentam uma queda na intensidade em função de  $q^4$ , relação conhecida como lei de Porods. A partir desta relação é possível deduzir o plot de Kratky ( $q^2 \times I(q)$  vs.  $q$ ). Este plot possui padrões específicos que permitem avaliar se a proteína está estruturada, parcialmente desordenada e desnaturada (PUTNAM *et al.*, 2007). A informação global contida na curva de espalhamento pode ser submetida à operação de transformada de Fourier para gerar a função de distribuição de pares  $P(r)$ . Esta função descreve todas as distâncias entre os elétrons em uma proteína e permite a reconstrução tridimensional através de modelos de envelope molecular denominados *ab initio* (SVERGUN *et al.*, 2001).

#### **1.3.7.3.2 Descrição do experimento**

Os dados de SAXS das proteínas PhbF e PhaP1 fusionadas a uma cauda de histidinas foram coletados utilizando a linha de luz SAXS1 ( $\lambda$ , 1,488 Å) no Laboratório Nacional de Luz Síncrotron utilizando um detector 2D (Pilatus-Dectris). O experimento foi realizado a 25 °C com uma distância amostra-detector de 1534,5 mm que permitiu registrar os dados em um intervalo de  $0,009 < q < 0,22 \text{ \AA}^{-1}$ . As concentrações proteicas utilizadas foram 5, 2,5 e 1 mg/mL com um tempo total de exposição de 150 segundos. Os dados foram escalonados em relação às variações na intensidade do feixe e subtraídos da radiação de fundo e do espalhamento do tampão utilizando o programa

FIT2D (HAMMERSLEY, 1997). O raio de giro foi avaliado segundo o método de Guinier e a função de distribuição de distâncias  $p(r)$  foi calculada utilizando o programa GNOM (SVERGUN, 1992). A massa molecular experimental foi determinada pelo programa SAXS MoW (FISCHER *et al.*, 2010). Vinte modelos de envelope molecular de baixa resolução foram gerados utilizando o programa DAMMIF (FRANKE & SVERGUN 2009). Os modelos gerados foram alinhados e um mapa de probabilidades foi calculado pelo programa DAMAVER (VOLKOV & SVERGUN, 2003).

#### **1.3.7.4 Análise da adesão da proteína PhaP1 ao PHB por Microscopia de força atômica (AFM)**

##### **1.3.7.4.1 Introdução à técnica**

AFM ou microscopia de força atômica é uma técnica que permite mapear as superfícies dos materiais quanto à sua topografia, elasticidade e plasticidade. Nesta análise, a superfície é varrida em linhas paralelas por uma sonda que mede as interações locais do tipo van der Waals (BINNING, 1986; SANTOS & CASTANHO, 2004). A sonda de dimensões nanométricas é presa a um braço flexível (*cantilever*) reponsável pela transmissão do sinal. À medida que a sonda percorre a superfície da amostra ocorre a deflexão e ou torção do *cantilever* que é detectada por um feixe de laser apontado na outra extremidade do mesmo. A reflexão do laser é direcionada em um fotodiodo que detecta os movimentos do cantilever e permite reconstruir uma imagem bidimensional ou tridimensional da amostra. Esta deflexão é proporcional à força de interação entre a sonda e a superfície (SANTOS & CASTANHO, 2004).

##### **1.3.7.4.2 Descrição do experimento**

###### **1.3.7.4.2.1 Preparo do material**

O substrato utilizado para a adesão do polímero PHB foi o silício (111) com dimensão média de  $0,5 \text{ cm}^2$  pré-tratado por 10 minutos com solução ácida de

H<sub>2</sub>O:H<sub>2</sub>O<sub>2</sub>:HCl (86:11:3) para a remoção de contaminação orgânica. Foi utilizada uma solução de PHB 1% (m/v) comercial (Sigma) ou extraído de *H. seropedicae* (seção 1.3.2.4) solubilizado em clorofórmio a uma temperatura de 90 °C por 10 minutos. Um volume de 50 µL foi depositado de forma estática sobre o silício utilizando o método de *spin coating* onde o material foi submetido a uma rotação de 4000 rpm por 1 minuto. Este material foi em seguida submetido à análise por microscopia de força atômica.

A deposição da proteína sobre o filme de PHB, previamente analisado por AFM, foi realizada por deposição estática de 50 µL de uma solução da proteína PhaP1 fusionada a uma cauda de histidinas (0,3 µg/mL) em tampão de purificação (50 mmol/L de Tris-HCl pH 7,5, 100 mmol/L de NaCl e 0,05% de TritonX-100) seguida de incubação a temperatura ambiente por 5 minutos. O material foi submetido a uma rotação de 4000 rpm por 1 minuto para a retirada das proteínas não adsorvidas ao polímero e submetida a análise por microscopia de força atômica.

#### **1.3.7.4.2.2 Análise por AFM**

Os dados foram coletados com um microscópio Shimadzu modelo SPM9500J3 situado no Departamento de Física (UFPR) com a colaboração do professor Dr. Paulo César Camargo e a aluna de Doutorado Cecília Fabiana da Gama Ferreira. O equipamento foi operado em modo dinâmico com uma frequência de 1,3 Hz a temperatura e umidade ambientes. Um *cantilever* de silício e raio de curvatura de 10 nm (Nanosensor) foi utilizado nas análises e as imagens tratadas com o programa da Shimadzu com o objetivo de retirar ruídos mecânicos e corrigir a inclinação da cerâmica piezoelétrica. Em cada amostra três regiões distintas foram analisadas e a dimensão das estruturas correspondentes à proteína depositada resultou da média de três estruturas em uma mesma imagem.

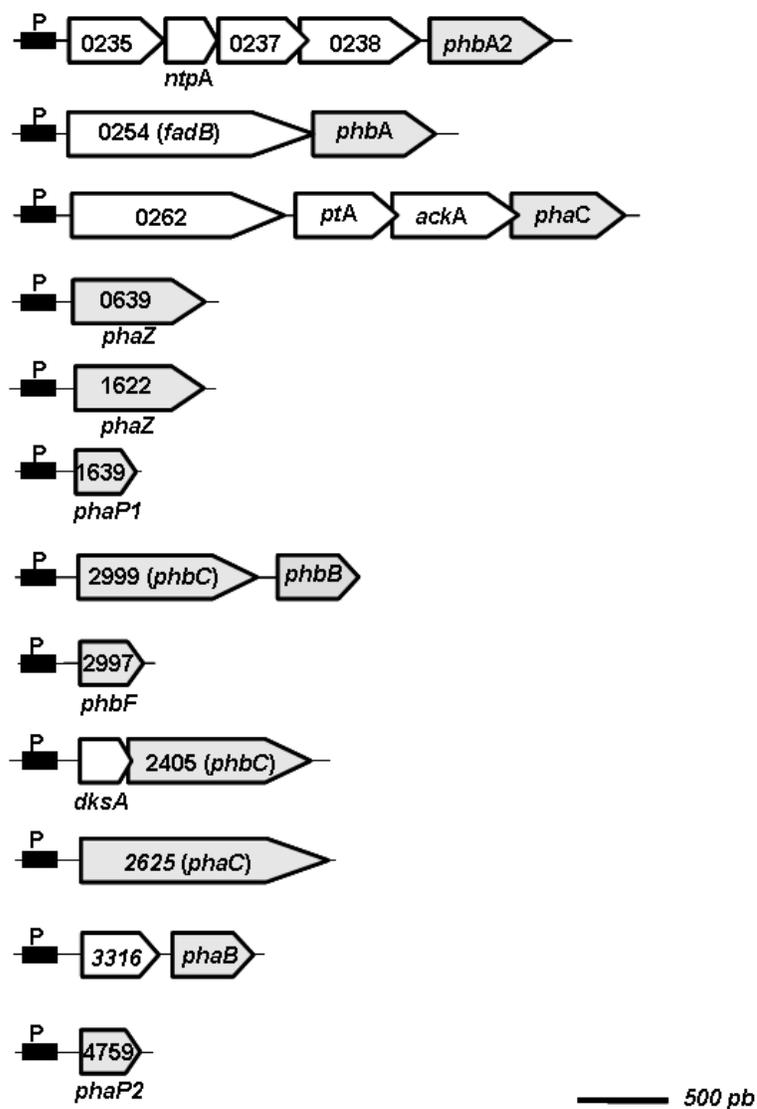
## 1.4. RESULTADOS E DISCUSSÃO

### 1.4.1 Organização dos genes envolvidos na biossíntese de PHAs em *H. seropedicae* SmR1

O genoma de *H. seropedicae* SmR1 (PEDROSA *et al.*, 2011) foi pesquisado para identificar os genes envolvidos no metabolismo de polihidroxialcanoatos. Segundo MADISON & HUISMAN (2002) a organização dos genes envolvidos na biossíntese de PHAs possui uma considerável divergência e esta também foi observada em *H. seropedicae*. Neste organismo os genes encontram-se espalhados pelo genoma como pode ser observado na figura 8. Foi possível localizar um provável operon que reúne os genes *phbC*, *phbB* e *phbF* envolvidos respectivamente na síntese de PHB, na síntese do monômero 3HB e na regulação da expressão gênica. Dentre os treze genes identificados encontram-se: duas polimerases de cadeia curta (*phbC*), duas polimerases de cadeia média (*phaC*), duas  $\beta$ -cetoacil-CoA tiolase (*phbA*), duas acetoacetyl-CoA redutase (*phbB*), duas fasinas (*phaP*), duas depolimerases (*phaZ*) e uma proteína regulatória (*phbF*).

Quatro proteínas foram selecionadas para a expressão e estudos funcionais e estruturais: PhbF, PhaP1, PhaC1 e PhaZ1. Em função da impossibilidade de solubilização dos dois últimos alvos, apenas PhbF e PhaP1 foram estudadas em detalhe. A fasina estudada foi denominada PhaP1 uma vez que existe mais um gene no genoma de *H. seropedicae* SmR1 que codifica para outra fasina. Esta proteína foi escolhida uma vez que é a mais expressa, com base no proteoma de referência deste organismo (CHAVES *et al.*, 2007).

**FIGURA 8 – Genes envolvidos na biossíntese de PHAs em *H. seropedicae* SmR1**



O esquema mostra através das setas em cinza genes relacionados ao metabolismo de PHAs em *H. seropedicae*. As setas indicam os genes e o sentido de transcrição e os retângulos pretos as regiões promotoras estudadas neste trabalho. As ORFs estão indicadas por números de acordo com a anotação do genoma de *H. seropedicae* (número de acesso no GenBank: CP002039). *phbA*: acetil-CoA acetiltransferase; *phbB*: acetoacetil-CoA redutase; *phbC*: poli-beta-hidroxiacanoato sintase; *phaP*: fasina; *phaZ*: poli-beta-hidroxiacanoato depolimerase; *phaB*: 3-cetoacil-CoA redutase; *phaC*: poli(3-hidroxiacanoato) sintetase; *phbF*: proteína regulatória. Os demais genes vizinhos são: *ackA*: acetato quinase; *dksA*: DnaK supressora; *fadB*: 3-hidroxiacil-CoA dehidrogenase oxidoreductase; *ntpA*: dATP pirofosfohidrolase; *pta*: fosfato acetiltransferase.

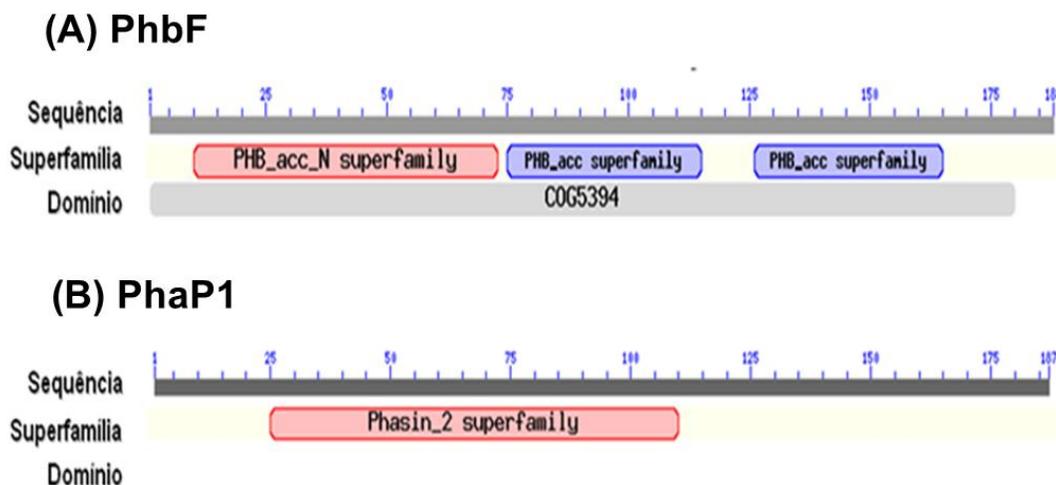
#### 1.4.2 Análise *in silico* das proteínas PhbF e PhaP1 de *H. seropedicae* SmR1

As proteínas selecionadas para estudo possuem baixa identidade com proteínas cujas estruturas tridimensionais já foram resolvidas por métodos experimentais, revelando-se assim importantes alvos para estudos. A proteína PhaP1 apresenta uma identidade de 23% com a DNA metiltransferase Dnmt2 humana e a proteína PhbF uma identidade de 28% com a proteína Dead-box de *Drosophila vasa*. A análise *in silico* foi realizada submetendo as sequências de aminoácidos ao programa BlastP contra o banco de dados de estruturas PDB (*Protein Data Bank*). Os resultados sugerem que as proteínas estudadas possuem características estruturais inéditas.

Simultaneamente aos estudos de identidade estrutural foram analisados os domínios que compõem as proteínas alvo com o objetivo de inferir as possíveis atividades desempenhadas (Figura 9). A proteína regulatória PhbF é composta basicamente por três domínios, sendo dois (PHB\_acc) na região carboxi terminal relatados como domínios importantes para a ligação da proteína ao PHB, enquanto o domínio amino terminal (PHB\_acc\_N) está envolvido na ligação ao DNA. PhaP1 possui um domínio (Phasin\_2) conservado entre as proteínas que cobrem em maior proporção os grânulos de PHB.

PhbF é uma proteína de 188 aminoácidos que possui maior identidade com a proteína PhbF de *Oxalobacteriaceae bacterium* IMCC9480 (84%). Quando comparada as demais proteínas com atividade de regulação transcricional descritas na literatura, encontra-se agrupada com a proteína repressora PhaR de *Ralstonia eutropha* H16 (Figura 10A). A maior similaridade foi verificada com esta proteína de 183 aminoácidos (83% identidade e 90% de similaridade) e a menor para as proteínas PhaR de *Rhodobacter sphaeroides* FJ1 (41% de identidade e 59% similaridade) (Figura 10B) e *Paracoccus denitrificans* (37% de identidade e 56% similaridade). A similaridade entre as proteínas PhbF de *H. seropedicae* e PhaR de *P. denitrificans* é restrita a região N-terminal (Figura 10C) onde provavelmente encontra-se o motivo hélice-volta-hélice relacionado a ligação ao DNA.

**FIGURA 9 – Análise dos domínios conservados das proteínas PhbF e PhaP1 de *H. seropedicae* SmR1**

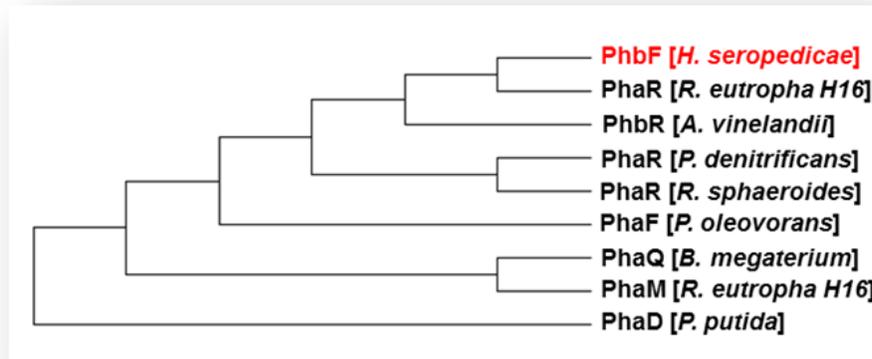


Representação da arquitetura das proteínas selecionadas para os estudos estruturais. A linha sequência representa a escala referente ao tamanho de cada proteína em aminoácidos. Na linha Superfamília está representada na forma de blocos a posição da região conservada representativa de uma família de proteínas específica. A linha domínio representa a região da proteína que possivelmente possui um enovelamento independente. A figura foi obtida com o programa BlastP.

PhaP1 é uma das duas fasinas de *H. seropedicae* com 187 aminoácidos que possui maior identidade com a fasina de *Collimonas fungivorans* Ter331 (63%). Quando comparada às demais fasinas descritas na literatura, PhaP1 encontra-se agrupada com as quatro fasinas de *R. eutropha* H16 (Figura 11). A maior similaridade foi obtida com PhaP de *Cupriavidus necator* N-1 (49% identidade e 65% de similaridade) seguida da fasina 2 de *R. eutropha* H16 (49% identidade e 64% de similaridade).

As proteínas PhbF e PhaP1 possuem considerável hidropaticidade, representada por números negativos (Tabela 1), reflexo da hidrofobicidade destas moléculas associadas com os grânulos de PHB.

**FIGURA 10 – Análise de similaridade da proteína PhbF de *H. seropedicae* SmR1 (A)**



**(B)**



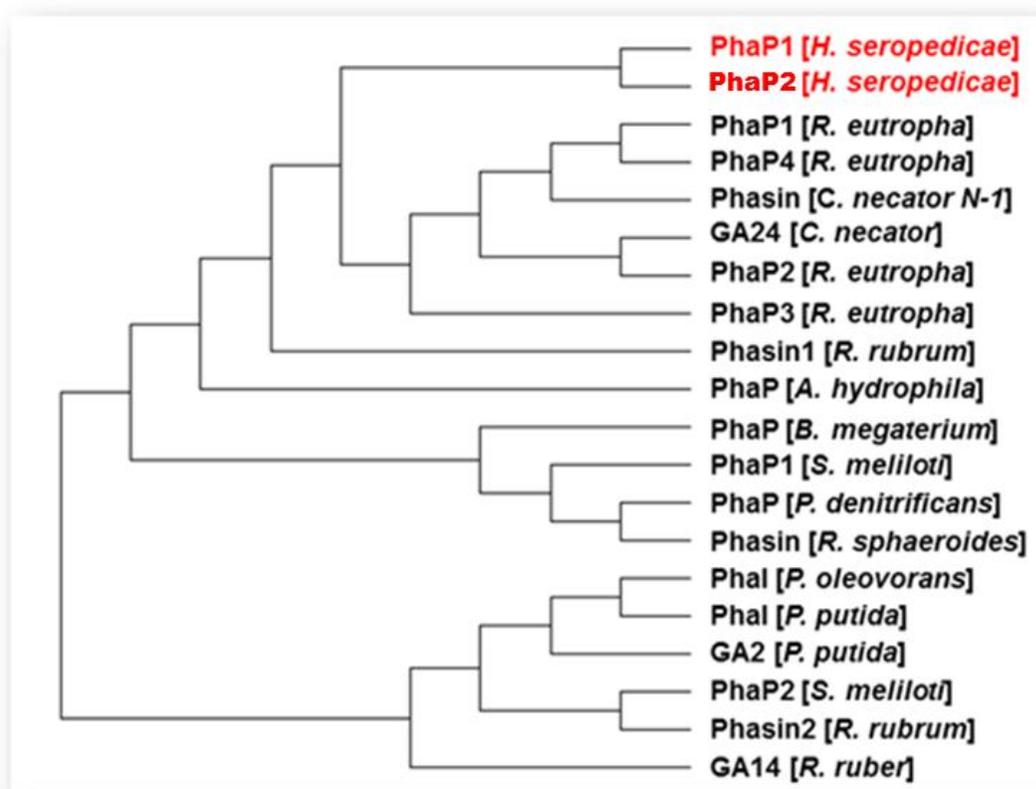
**(C)**



Legenda na próxima página.

(A) Árvore de identidade entre a proteína PhbF de *H. seropedicae* e proteínas regulatórias com função descrita na literatura. PhbF de *H. seropedicae* SmR1 (YP\_003776393.1), PhbR de *A. vinelandii* DJ (YP\_002799531.1), PhaD de *P. putida* (AAV68230.1), PhaF de *P. oleovorans* (CAA09109.1), PhaM de *R. eutropha* H16 (YP\_724661.1), PhaR de *P. denitrificans* (BAA77259.1), PhaR de *R. eutropha* H16 (YP\_725943.1), PhaR de *R. sphaeroides* (AAY51768.1), PhaQ de *B. megaterium* (AAD05257.1). (B) Alinhamento entre a sequência de aminoácidos das proteínas PhbF de *H. seropedicae* e PhaR de *R. eutropha* H16. (C) Alinhamento entre a sequência de aminoácidos das proteínas PhbF de *H. seropedicae* e PhaR de *P. denitrificans*. Os alinhamentos foram obtidos com o programa ClustalW (THOMPSON et al., 1994). (\*) Aminoácidos idênticos; (:) Aminoácidos de alta similaridade; (.) Aminoácidos de baixa similaridade. Os aminoácidos estão coloridos de acordo com a classificação do grupo R.

**FIGURA 11 – Análise de similaridade da proteína PhaP1 de *H. seropedicae* SmR1**



Árvore de identidade entre as proteína PhaP de *H. seropedicae* e proteínas PhaP1 de *R. eutropha* H16 (NP\_942840.1), PhaP2 de *R. eutropha* H16 (CAJ92517.1), PhaP3 de *R. eutropha* H16 (AAR38792.1), PhaP4 de *R. eutropha* H16 (AAR38793.1), GA24 de *C. necator* (2111322A), PhaP de *S. meliloti* 1021 (CAC46064.1), PhaP2 de *S. meliloti* 1021 (NP\_384828.1), PhaP de *B. megaterium* (AAD05256.1), PhaP de *A. hydrophila* (AAT77260.1), phasin1 de *R. rubrum* ATCC 11170 (YP\_428365.1), phasin2 de *R. rubrum* ATCC 11170 (YP\_427901.1), Phal de *P. putida* (AAV68228.1), GA2 de *P. putida* (AD02215.1), Phal de *P. oleovorans* (CAA09108.1), PhaP de *P. denitrificans* (BAA77258.1), phasin de *R. sphaeroides* (AAY51767.10) e phasin de *C. necator* N-1 (AEI82199.1), GA14 de *R. ruber* (CAA47036.1).

**TABELA 1 – Características bioquímicas das proteínas PhbF e PhaP1 de *H. seropedicae* SmR1**

|                                | <b>PhbF</b> | <b>PhaP1</b> |
|--------------------------------|-------------|--------------|
| <b>Massa molecular (Da)</b>    | 21472       | 19799        |
| <b>Número de aminoácidos</b>   | 188         | 187          |
| <b>pl</b>                      | 5,72        | 6,76         |
| <b>Hidropaticidade (GRAVY)</b> | -0,556      | -0,214       |

### **1.4.3 Caracterização estrutural das proteínas PhbF e PhaP1 de *H. seropedicae* SmR1**

#### **1.4.3.1 Amplificação e clonagem dos genes *phbF* e *phaP1* em vetores de expressão**

Com o objetivo de estudar as proteínas PhbF e PhaP1 de *H. seropedicae*, as regiões codificadoras destes genes foram amplificadas por PCR e clonadas em vetores de expressão que foram posteriormente transformados em *E. coli* estirpe BL21(DE3) para os primeiros ensaios de expressão.

Os oligonucleotídeos iniciadores planejados para os experimentos de amplificação estão listados no quadro 5. Nos oligonucleotídeos 5' foram introduzidos sítios de restrição para a enzima *NdeI* e nos oligonucleotídeos 3' sítios para a enzima *BamHI*. O DNA genômico de *H.seropedicae* foi utilizado como molde para as reações de PCR. A reação de PCR para todos os genes foi realizada utilizando um gradiente de temperatura (45-65°C) de anelamento. Como resultado obteve-se os fragmentos *phaP1* (590 pb) e *phbF* (605 pb) que foram em seguida clonados no vetor pCR<sup>®</sup>2.1 (Invitrogen<sup>TM</sup>) e sequenciados. Este vetor caracteriza-se por facilitar a clonagem de produtos de PCR, devido a presença de uma timina em cada ponta 3' que anela-se com facilidade as adeninas 5' adicionadas pela Taq DNA polimerase ao amplificado. Os plasmídeos pCR21::*phaP1* e pCR21::*phbF* foram digeridos com as enzimas *NdeI* e *BamHI* para a liberação de seus fragmentos e para serem, em seguida, clonados nos vetores pET-28a(+) e pET-29a(+), gerando a série pET-28::*(phaP* ou *phbF)* e pET-

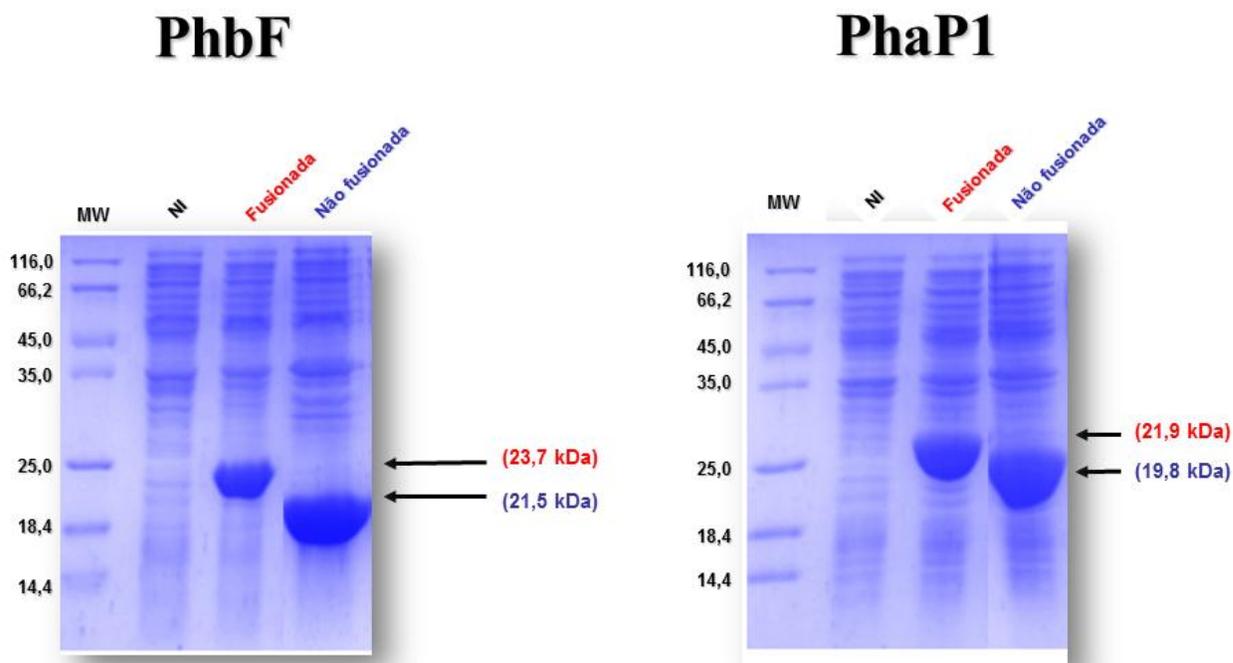
29::(*phaP* ou *phbF*). Estes vetores permitiram a expressão das proteínas nas formas fusionadas a cauda de histidinas e não fusionadas respectivamente.

#### **1.4.3.2 Testes de expressão e confirmação da identidade das proteínas PhaP e PhbF**

Para uma análise inicial da expressão das proteínas PhbF e PhaP1, os plasmídeos foram transformados em *E. coli* estirpe BL21 (DE3) e células transformantes foram inoculadas em meio LB e induzidas com IPTG (1 mmol/L) durante 3 horas. Os perfis de expressão das frações brutas estão representados na figura 12. Nota-se que a superexpressão das proteínas PhbF e PhaP1, em ambas as formas, foi satisfatória para a purificação e os ensaios de caracterização.

As bandas correspondentes às proteínas superexpressas foram retiradas dos géis e analisadas por MALDI-TOF/MS quanto a suas identidades. A relação entre o número de m/z obtidas e aquelas que contribuíram para a identificação foram 64/11 para PhbF e 71/22 para PhaP1. As proteínas PhbF e PhaP1 foram identificadas com coberturas de 70 e 42% respectivamente. As análises foram realizadas assumindo-se um desvio de 209 ppm.

**FIGURA 12 – Eletroforese do extrato bruto visando à confirmação da expressão das proteínas PhbF e PhaP1 de *H. seropedicae* SmR1**



Perfil eletroforético em gel de poliacrilamida (SDS-PAGE) 12% das proteínas expressas na bactéria *E. coli* estirpe BL21(DE3) a 37 °C nas formas nativa ou fusionada a cauda de histidinas, utilizando 1 mM de IPTG como indutor. As células foram ressuspensas em 50 mM Tris-HCl pH 7,5 e 300 mM NaCl e em seguida lisadas por sonicação. O gel foi corado com azul de Coomassie. MW: padrão de massa molecular (kDa). NI: extrato celular bruto não induzido. A direita de cada perfil de expressão estão indicadas a posição das formas fusionada e não fusionada assim como as suas respectivas massas moleculares calculadas.

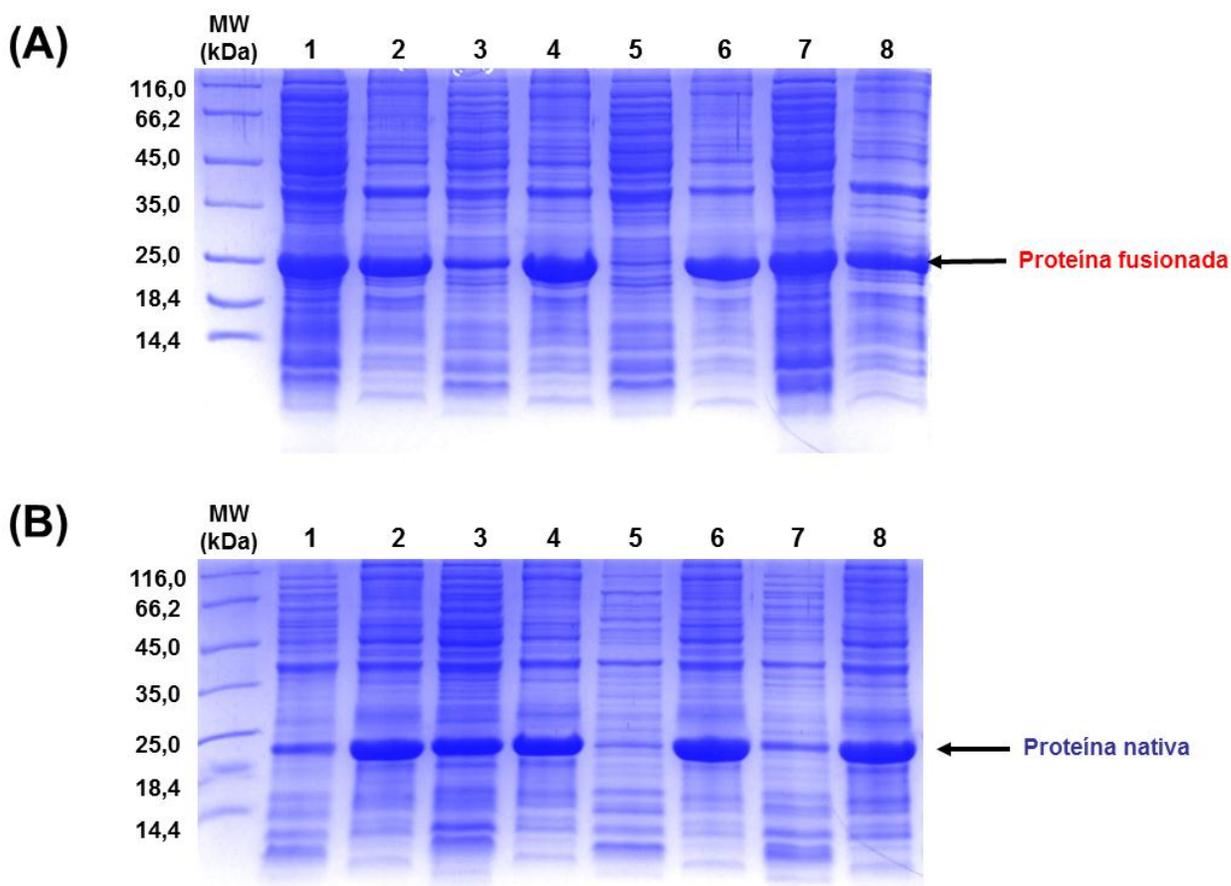
### 1.4.3.3 Otimização das condições de expressão das proteínas PhbF e PhaP1

As proteínas PhbF e PhaP1 apresentaram uma grande insolubilidade quando as células foram lisadas em tampão sem detergente. Esta característica já era esperada em função do perfil de hidrofobicidade sugerido pelas análises baseadas na sequência de aminoácidos e pelo fato destas proteínas encontrarem-se associadas aos grânulos hidrofóbicos de PHB, como já relatado na literatura. A solubilidade de PhbF e PhaP1 foi obtida através da adição de TritonX-100 a uma concentração de 0,05% ao tampão de lise. Este detergente foi escolhido por ser não iônico e interferir menos na estrutura proteica quando comparado com o detergente SDS. A concentração de 0,05% foi otimizada para a solubilização das proteínas e encontra-se abaixo das concentrações recomendadas para minimização de interações hidrofóbicas pouco específicas (0,2%) e dissolução de precipitados proteicos na forma de corpos de inclusão (2%) (GE Health Care).

Em seguida diferentes temperaturas de indução assim como dois tipos de indutores (IPTG e lactose) foram utilizados com o objetivo de aumentar a solubilidade das proteínas. Os resultados obtidos podem ser observados nas figuras 13 e 14. A proteína PhbF fusionada apresentou um considerável aumento na solubilidade com a diminuição da temperatura de indução quando comparado a condição inicial a 37 °C (Figura 13). A proteína PhaP1 fusionada também apresentou um aumento significativo na solubilidade a 30 °C, independente do tipo de indutor (Figura 14). A expressão de PhaP1 por 12 horas a 20 °C levou a um perfil de degradação da mesma, sendo que longos períodos de indução para esta proteína devem ser evitados.

Estes experimentos permitiram determinar as condições de expressão e solubilização das proteínas alvo como pré-requisito para a obtenção de proteína necessária para os ensaios de caracterização *in vitro*.

**FIGURA 13 – Perfil eletroforético do teste de expressão da proteína PhbF de *H. seropedicae* SmR1**



Perfil eletroforético em gel de poliacrilamida (SDS-PAGE) 12% dos extratos de células expressando a proteína PhbF na forma fusionada (A) e não fusionada a cauda de histidinas (B). As células foram ressuspensas em 50 mM Tris-HCl pH 7,5, 300 mM NaCl e 0,05% TritonX-100. O gel foi corado com azul de Coomassie. MW: padrão de massa molecular (kDa).

Linhas 1-4: expressão a 30 °C durante 3h.

Linha 1: fração solúvel após sonicação da cultura induzida com 0,5% lactose.

Linha 2: fração insolúvel após sonicação da cultura induzida com 0,5% lactose.

Linha 3: fração solúvel após sonicação da cultura induzida com 1 mM IPTG.

Linha 4: fração insolúvel após sonicação da cultura induzida com 1 mM IPTG.

Linhas 5-6: expressão a 37 °C durante 3h.

Linha 5: fração solúvel após sonicação da cultura induzida com 0,5% lactose.

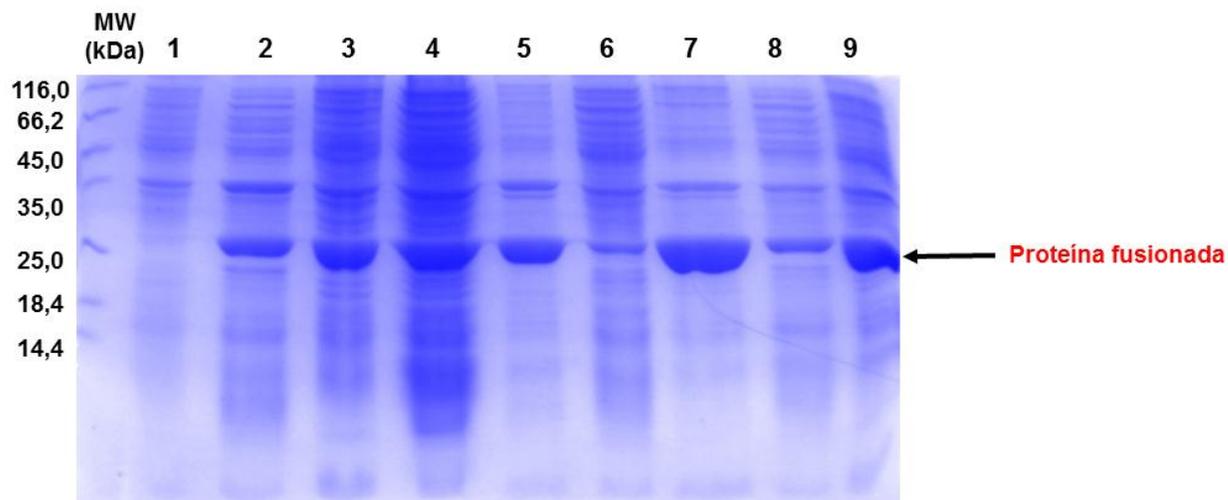
Linha 6: fração insolúvel após sonicação da cultura induzida com 0,5% lactose.

Linhas 7-8: expressão a 20 °C durante 12h.

Linha 7: fração solúvel após sonicação da cultura induzida com 1 mM IPTG.

Linha 8: fração insolúvel após sonicação da cultura induzida com 1 mM IPTG.

**FIGURA 14 – Perfil eletroforético do teste de expressão da proteína PhaP1 fusionada de *H. seropedicae* SmR1**



Perfil eletroforético em gel de poliacrilamida (SDS-PAGE) 12% dos extratos das células expressando a proteína PhaP na forma fusionada a cauda de histidinas. As células foram ressuspensas em 50 mM Tris-HCl pH 7,5, 300 mM NaCl e 0,05% TritonX-100. Neste experimento foram variados temperatura de expressão e tipo de indutor. O gel foi corado com azul de Coomassie. MW: padrão de massa molecular (kDa).

Linha 1: extrato bruto das células antes da indução.

Linhas 2-5: expressão a 30 °C durante 3h.

Linha 2: fração solúvel após sonicação da cultura induzida com 0,5% lactose.

Linha 3: fração insolúvel após sonicação da cultura induzida com 0,5% lactose.

Linha 4: fração solúvel após sonicação da cultura induzida com 1 mM IPTG.

Linha 5: fração insolúvel após sonicação da cultura induzida com 1 mM IPTG.

Linhas 6-9: expressão a 37 °C durante 3h.

Linha 6: fração solúvel após sonicação da cultura induzida com 0,5% lactose.

Linha 7: fração insolúvel após sonicação da cultura induzida com 0,5% lactose.

Linha 8: fração solúvel após sonicação da cultura induzida com 1 mM IPTG.

Linha 9: fração insolúvel após sonicação da cultura induzida com 1 mM IPTG.

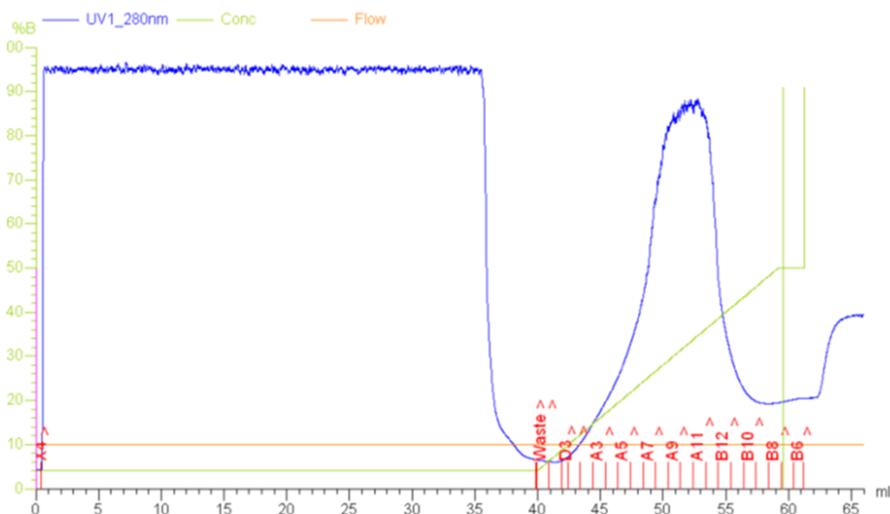
#### 1.4.3.4 Purificação das proteínas fusionadas a cauda de histidinas

A estirpe BL21(DE3) foi transformada com os plasmídeos pET-28::*phbF* e pET-28::*phaP1* e a expressão foi induzida com 0,5 mM de IPTG por 12h a 20 °C para a proteína HisPhbF e por 3h a 30 °C para a proteína HisPhaP1. O *pellet* celular foi ressuspensionado em solução de lise composta de 100 mM NaCl, 50 mM Tris-HCl pH 7,5 e 0,05% Triton X-100 e cuja composição final foi obtida após algumas tentativas de purificação na ausência de detergente, o que resultou na precipitação da mesma após eluição.

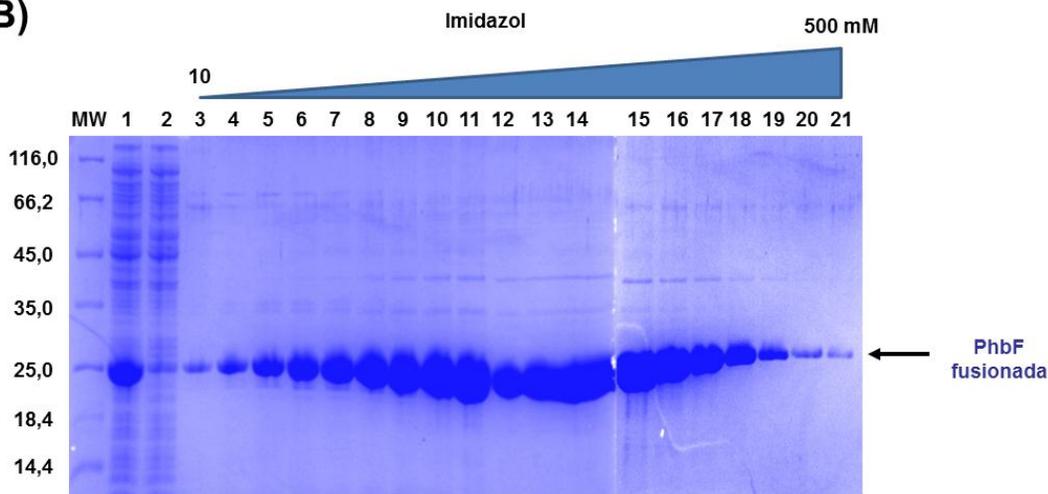
As proteínas PhbF e PhaP1 foram purificadas em apenas um passo cromatográfico utilizando a coluna Hi-Trap Chelating (5 mL – GE HeathCare). As proteínas foram eluídas com aproximadamente 300 mM de imidazol e o perfil de eluição pode ser observado nos cromatogramas das figuras 15A e 16A e nas eletroforeses das figuras 15B e 16B. As amostras com maior grau de pureza e correspondentes ao pico no cromatograma foram coletadas, reunidas e dialisadas contra 1 L de solução de diálise composta de 100 mM NaCl, 50 mM Tris-HCl pH 7,5 e 0,05% Triton X-100 para remoção do imidazol. As proteínas apresentaram-se estruturalmente homogêneas (Figura 17B) e com pureza superior a 93% (Figura 17A).

**FIGURA 15 – Purificação da proteína PhbF fusionada a cauda de histidinas**

**(A)**

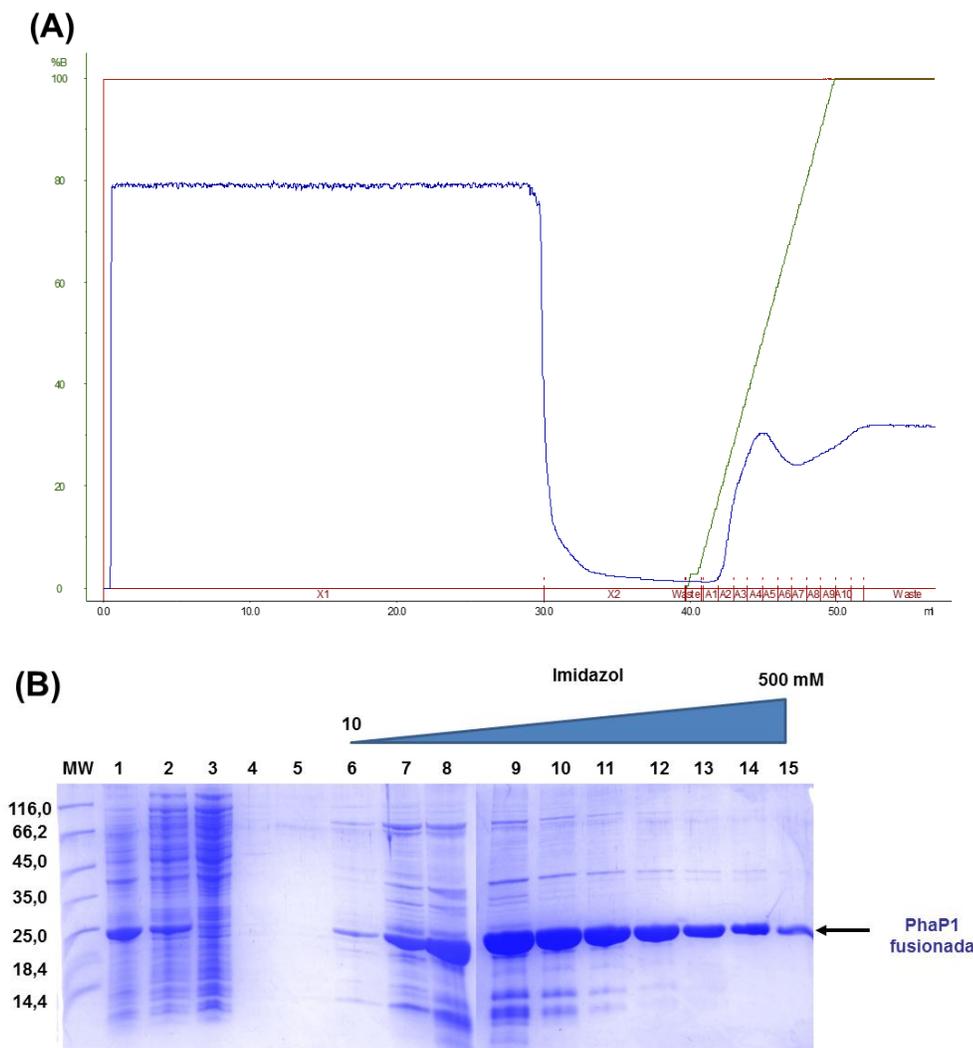


**(B)**



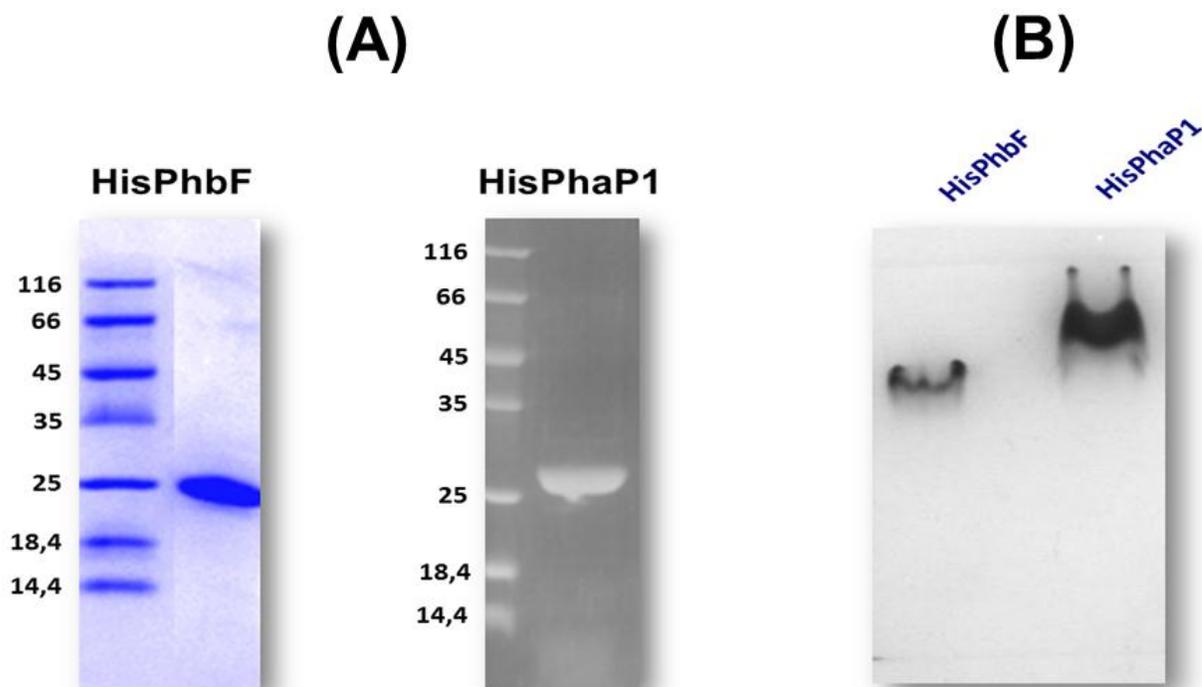
(A) A purificação foi realizada utilizando uma coluna Hi-trap chelating (5 mL). A solução de proteínas (50 mL) foi injetada, num fluxo de 1 mL/min. Após injeção a coluna foi lavada com 10 CV de tampão (100 mmol/L de NaCl; 50 mmol/L de Tris-HCl pH 7,5; 10 mmol/L de imidazol, 0,05% triton X-100). A proteína PhbF fusionada foi eluída utilizando um gradiente contínuo de 10 a 1000 mmol/L de imidazol equivalente a 20 VC. As amostras foram coletadas em alíquotas de 1mL. No eixo y, porcentagem de tampão (100 mmol/L de NaCl; 50 mmol/L de Tris-HCl pH 7,5; 1000 mmol/L de imidazol, 0,05% triton X-100). No eixo x, volume de eluição em mL. Em azul, leitura da absorção a 280 nm. Em verde, gradiente de imidazol. As amostras coletadas estão indicadas em vermelho. Waste: frações descartadas. (B) Perfil eletroforético em gel de poliacrilamida (SDS-PAGE) 12 % das frações coletadas durante a purificação. 1: fração bruta de proteínas solúveis antes da purificação. 2: fração de proteínas que não adsorveram na coluna. 3-21: frações referentes as proteínas que foram eluídas durante o gradiente de imidazol. O gel foi corado com azul de Coomassie.

**FIGURA 16 – Purificação da proteína PhaP1 fusionada a cauda de histidinas**



(A) A purificação foi realizada utilizando uma coluna Hi-trap chelating (5 mL). A solução de proteínas (50 mL) foi injetada, num fluxo de 1 mL/min. Após injeção a coluna foi lavada com 10 CV de tampão (100 mmol/L de NaCl; 50 mmol/L de Tris-HCl pH 7,5; 10 mmol/L de imidazol, 0,05% triton X-100). A proteína PhbF fusionada foi eluída utilizando um gradiente contínuo de 10 a 1000 mmol/L de imidazol equivalente a 20 VC. As amostras foram coletadas em alíquotas de 1mL. No eixo y, porcentagem de tampão (100 mmol/L de NaCl; 50 mmol/L de Tris-HCl pH 7,5; 1000 mmol/L de imidazol, 0,05% triton X-100). No eixo x, volume de eluição em mL. Em azul, leitura da absorção a 280 nm. Em verde, gradiente de imidazol. As amostras coletadas estão indicadas em vermelho. (B) Perfil eletroforético em gel de poliacrilamida (SDS-PAGE) 12 % das frações coletadas durante a purificação. 1: extrato bruto de proteínas após a lise celular por sonicação. 2: fração bruta de proteínas solúveis antes da purificação. 3: fração de proteínas que não adsorveram na coluna. 4-5: frações referentes a etapa de lavagem. 6-15: frações referentes as proteínas que foram eluídas durante o gradiente de imidazol. O gel foi corado com azul de Coomassie.

**FIGURA 17 – Pureza e homogeneidade das proteínas HisPhbF e HisPhaP1 após purificação**



(A) Eletroforese em SDS-PAGE 12% das proteínas HisPhbF e HisPhaP1 após diálise e concentração. Cada linha possui 2,5 µg de proteína purificada. (B) Eletroforese em condições não desnaturantes (NATIVE-PAGE) 12% das proteínas. Os géis foram corados com azul de Coomassie.

#### 1.4.3.5 Determinação do estado de oligomerização das proteínas PhbF e PhaP1

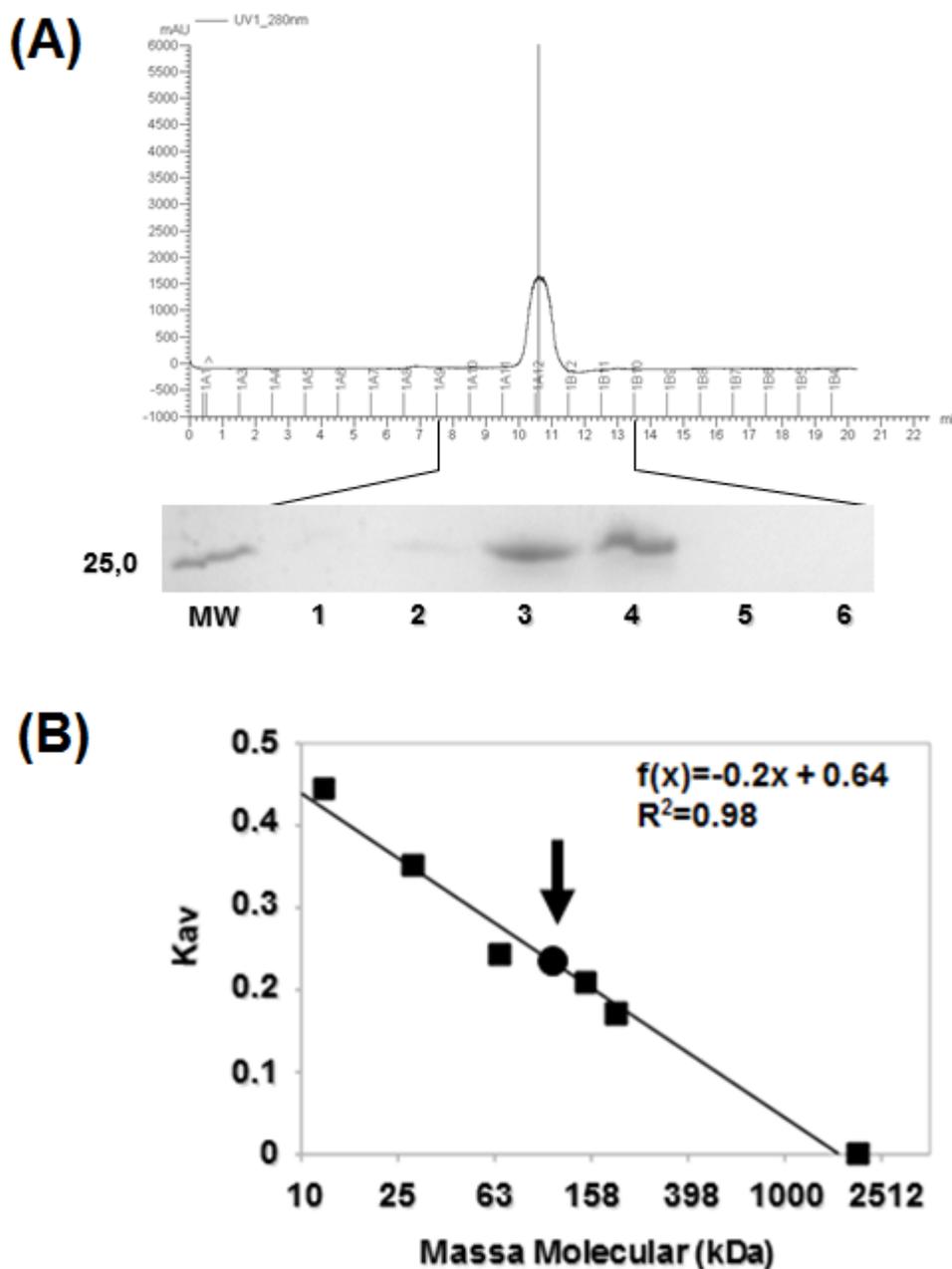
Poucos são os relatos sobre o estado de oligomerização das proteínas PhbF e PhaP1. A proteína PhaR de *P. denitrificans* é relatada como um tetrâmero (MAEHARA *et al.*, 2002) e a proteína PhaP1 de *R. eutropha* como um trîmero (NEWMANN *et al.*, 2008). As proteínas de *H. seropedicae* foram submetidas à cromatografia por gel filtração para a obtenção de informações sobre a massa molecular experimental assim como o perfil de homogeneidade. Esta última característica é importante como ponto de avaliação da qualidade do material a ser utilizado nos experimentos de espalhamento de raios-X a baixo ângulo.

As proteínas PhbF e PhaP1 possuem uma massa molecular de 104,30 e 114 kDa, respectivamente, obtidos por cromatografia em gel filtração em coluna Superose 12 HR 10/30 (Figura 18A e 19A). Estes valores foram obtidos pela submissão dos valores de  $K_{av}$  (constante de partição), calculados com base nos volumes de eluição, a equação de reta que relaciona massa molecular e  $K_{av}$  de proteínas padrão (Figura 18B e 19B).

Como HisPhbF possui uma massa monomérica calculada de 23,7 kDa estima-se que a mesma seja composta de 4,4 monômeros e desta forma compreenda um possível tetrâmero. Este experimento permitiu também confirmar a homogeneidade da amostra protéica, uma vez que não foram observados agregados e a proteína foi eluída da coluna em apenas um pico.

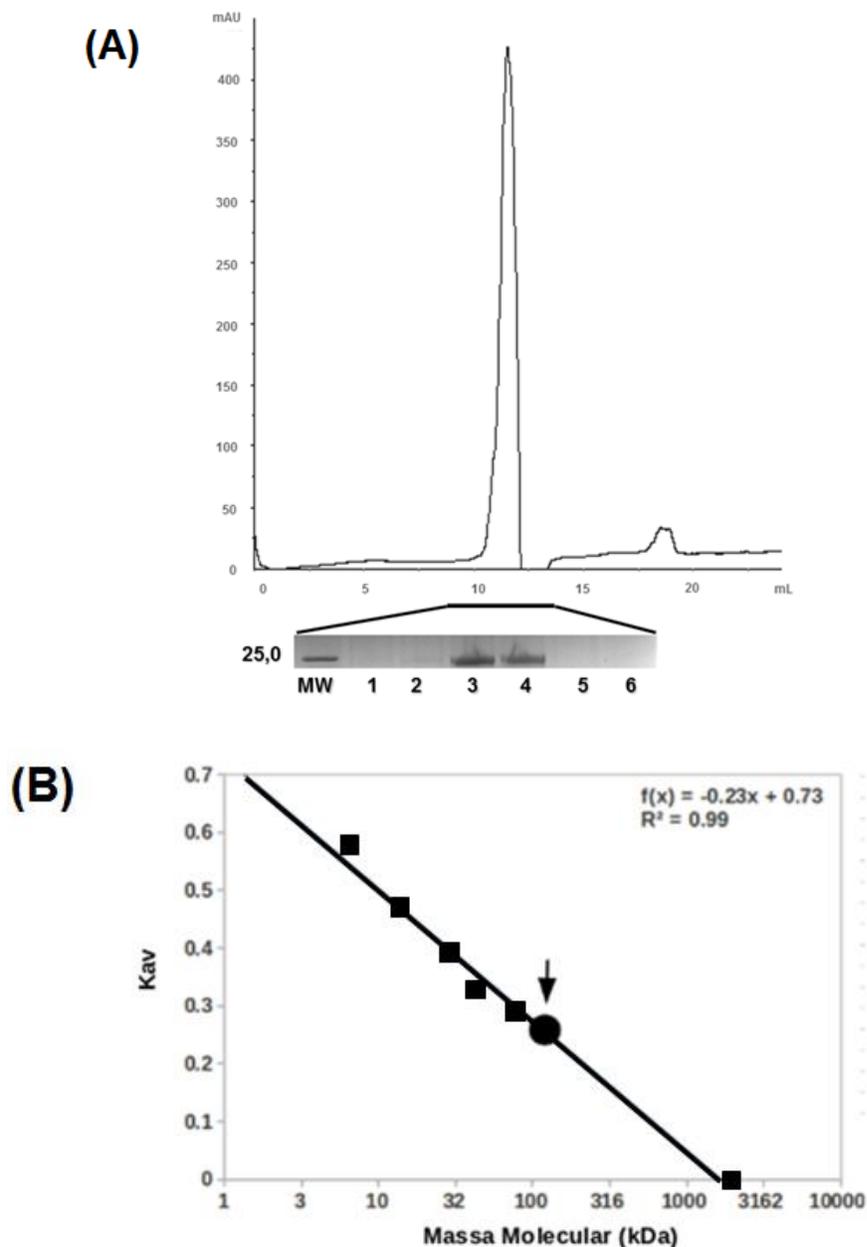
A proteína HisPhaP1 foi considerada como um possível pentâmero em solução com base na massa monomérica de 21,9 kDa. O cromatograma mostrou também que a proteína apresentou dois picos cromatográficos sendo o mais abundante correspondente à população de proteínas integras e o menos abundante correspondente aos produtos de degradação da mesma.

FIGURA 18 – Determinação do estado de oligomerização de PhbF por gel filtração



(A) Cromatograma referente à gel filtração em coluna Superose 12 HR 10/30 da proteína HisPhbF. Na ordenada, absorvância a 280nm e na abscissa, volume de eluição em ml. Abaixo do cromatograma encontra-se o perfil eletroforético em gel de poli-acrilamida (SDS-PAGE) 12% referente às frações do pico de eluição. O gel foi corado com azul de Coomassie. (B) Curva padrão  $K_{av}$  versus log da massa molecular (kDa) indicando a relação linear entre as duas grandezas representada pela equação de reta. Seta mostra o ponto de eluição da proteína HisPhbF. Os quadrados indicam as proteínas utilizadas como padrão de massa molecular (GE Healthcare).

**FIGURA 19 – Determinação do estado de oligomerização de PhaP1 por gel filtração**



(A) Cromatograma referente à gel filtração em coluna Superose 12 HR 10/30 da proteína HisPhaP. Na ordenada, absorvância a 280nm e na abscissa, volume de eluição em ml. Abaixo do cromatograma encontra-se o perfil eletroforético em gel de poliacrilamida (SDS-PAGE) 12% referente às frações do pico de eluição. O gel foi corado com azul de Coomassie. (B) Curva padrão  $K_{av}$  versus log da massa molecular (kDa) indicando a relação linear entre as duas grandezas representada pela equação de reta. Seta mostra o ponto de eluição da proteína HisPhaP1. Os quadrados indicam as proteínas utilizadas como padrão de massa molecular (GE Healthcare).

#### 1.4.3.6 Análise por microscopia de força atômica da interação da proteína PhaP1 com filmes de PHB

A proteína PhaP é encontrada, normalmente, em maior abundância sobre a superfície dos grânulos nativos de PHB (TRAINER *et al.*, 2006). O experimento de interação da proteína HisPhaP1 de *H. seropedicae* SmR1 sobre um filme de PHB artificialmente formado foi utilizado para avaliar aspectos estruturais do filme de proteínas formado e as dimensões das estruturas supramoleculares pela técnica de AFM.

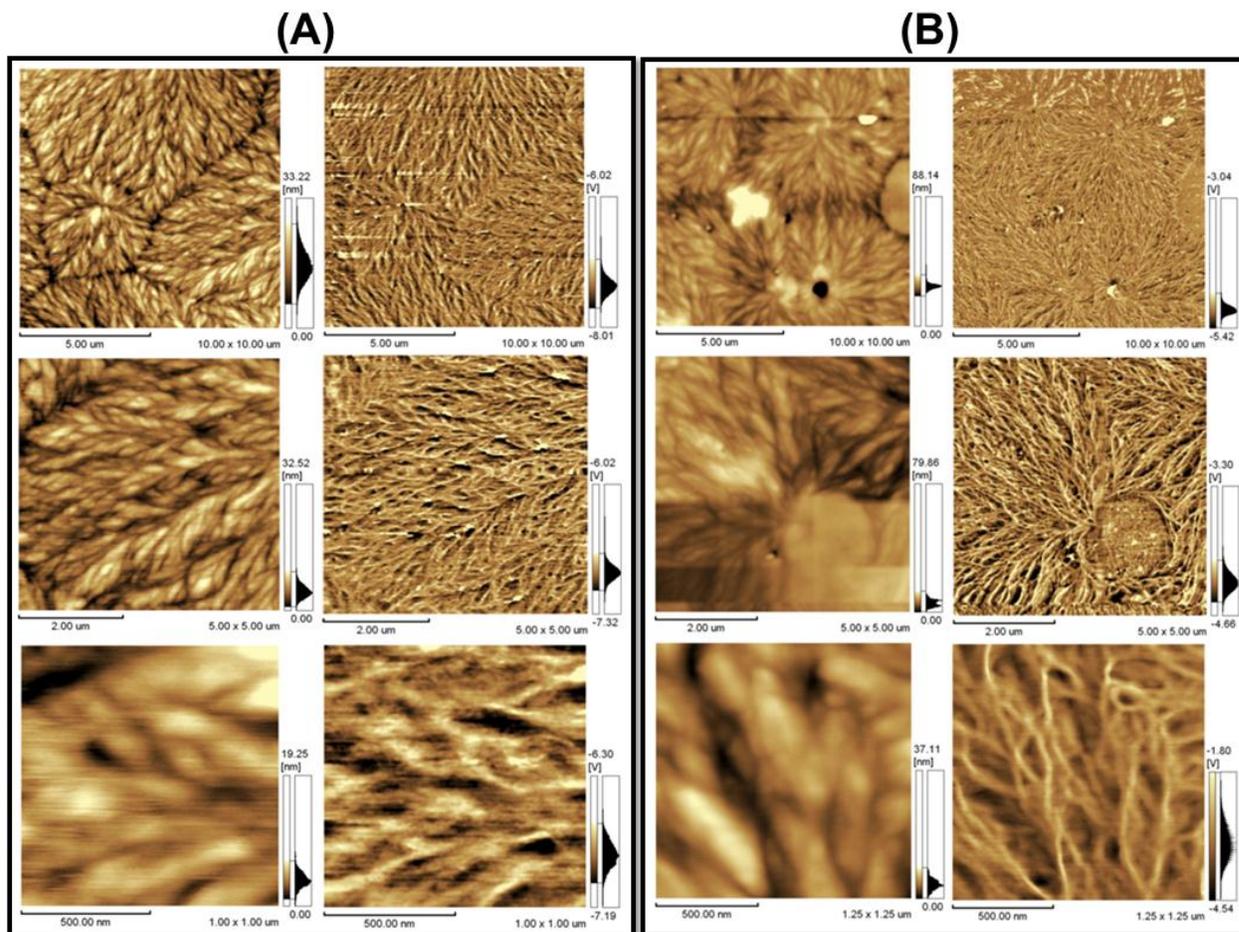
A primeira etapa do experimento consistiu na formação do filme de PHB obtido comercialmente e extraído de *H. seropedicae* SmR1. O polímero formou estruturas em forma de rede composta por microfibras com pontos centrais, comum a estruturas denominadas esferulitas (Figura 20A), o mesmo padrão foi observado para o filme formado a partir de PHB extraído de *H. seropedicae* SmR1 (Figura 20B). Este padrão de esferulitas foi observado por microscopia ótica de um filme de PHB formado por cristalização isotérmica (ABE *et al.*, 2000). Este resultado mostra que o polímero extraído possui as mesmas características morfológicas do PHB obtido comercialmente.

A segunda etapa do experimento consistiu na deposição da proteína HisPhaP1 sobre o filme de PHB extraído de *H. seropedicae* SmR1. Após a deposição, nas imagens a partir de 1  $\mu\text{m}$ , observa-se que o perfil das microfibras muda quando comparado ao material contendo somente o polímero. As microfibras tornam-se menos evidentes devido à deposição da proteína sobre a superfície (Figuras 21 – imagens de 5 e 10  $\mu\text{m}$ ). A partir da resolução de 500 nm observa-se um grande número de estruturas circulares cobrindo toda a área varrida (Figura 21). Estas estruturas possuem uma dimensão lateral média de  $177,8 \pm 11,7 \text{ \AA}$  calculadas com base em três diferentes estruturas encontradas na mesma imagem. Observa-se que elas distribuem-se por toda a área analisada.

Apesar da artificialidade deste experimento, os dados de AFM permitiram evidenciar que a proteína purificada foi capaz de adsorver sobre o filme de PHB extraído de *H. seropedicae* SmR1 na forma de estruturas que podem formar-se já em

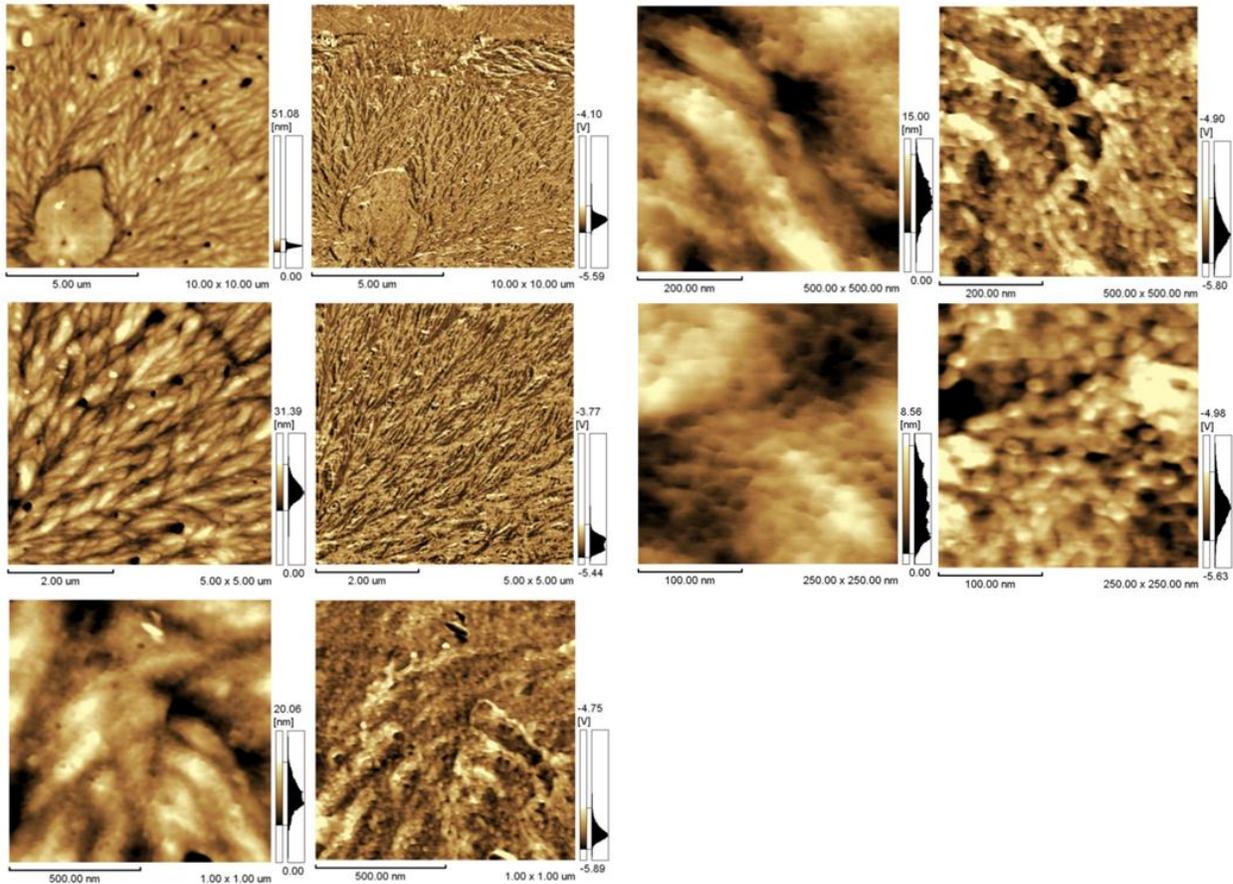
solução ou adotar as dimensões observadas somente quando em associação com o polímero PHB.

**FIGURA 20 – Imagens de AFM da deposição de PHB sobre silício**



Imagens da deposição de PHB comercial (A) e extraído de *H. seropedicae* SmR1 (B) sobre silício. Em cada quadro, as imagens a esquerda apresentam informações de topografia e as imagens a direita informações sobre a fase. A escala encontra-se na porção inferior de cada imagem e a faixa de coloração está indicada a direita. As imagens foram coletadas a temperatura ambiente e com as amostras expostas ao ar.

**FIGURA 21 – Imagens de AFM da deposição da proteína HisPhaP1 sobre filme de PHB extraído de *H. seropedicae* SmR1**



Imagens da deposição da proteína HisPhaP1 sobre filme de PHB extraído de *H. seropedicae* SmR1. Cada dupla de imagens apresentam a esquerda informação de topografia e a direita informação sobre a fase. A escala encontra-se na porção inferior de cada imagem e a faixa de coloração está indicada a direita. As imagens foram coletadas a temperatura ambiente e com as amostras expostas ao ar.

#### 1.4.3.7 Análise por dicroísmo circular

Com o objetivo de acessar a integridade estrutural das proteínas PhbF e PhaP1, na presença do detergente TritonX-100, foi utilizada a técnica do dicroísmo circular (CD). Esta técnica permite avaliar experimentalmente a estrutura secundária quanto ao conteúdo de  $\alpha$ -hélices, fitas  $\beta$  e alças.

No espectro CD das proteínas purificadas (Figuras 22A e 23A) é possível observar a presença de duas bandas negativas a 208 e 222 nm que constitui uma característica marcante de proteínas compostas predominantemente por  $\alpha$ -hélices. Este dado experimental corrobora as predições *in silico* das estruturas secundárias de PhbF e PhaP1 obtidas com o programa PsiPred (Figuras 22B e 23B). Para PhbF, os primeiros 70 resíduos de aminoácidos permitem observar a presença de um possível motivo hélice-volta-hélice (HTH) envolvido na interação da proteína com DNA (Figura 22B). Por outro lado, PhaP1 é uma proteína formada por três  $\alpha$ -hélices compostas por mais de 50 resíduos de aminoácidos cada (Figuras 23).

Estes resultados permitiram concluir que as proteínas PhbF e PhaP1 apresentaram estrutura secundária esperada, segundo a predição *in silico*, na presença do detergente na concentração de 0,05%.

FIGURA 22 – Predição da estrutura secundária da proteína PhbF

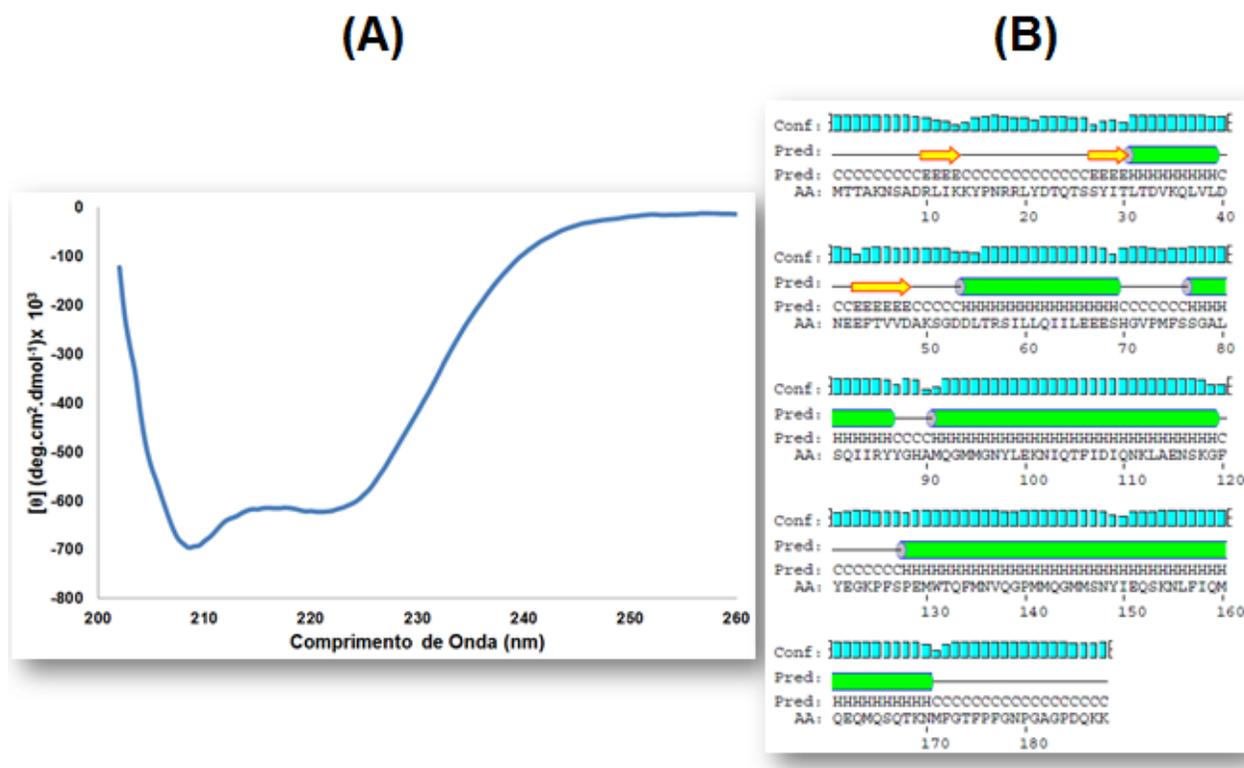
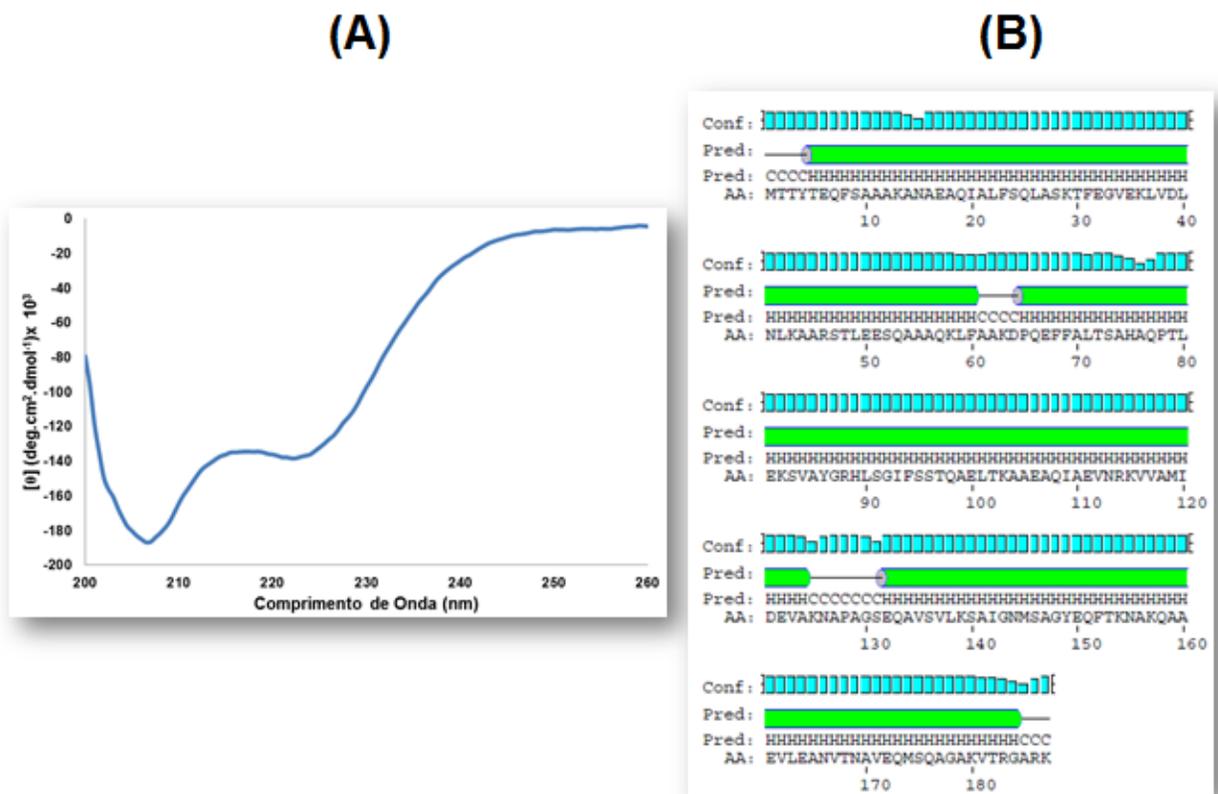


FIGURA 23 – Predição da estrutura secundária da proteína PhaP1



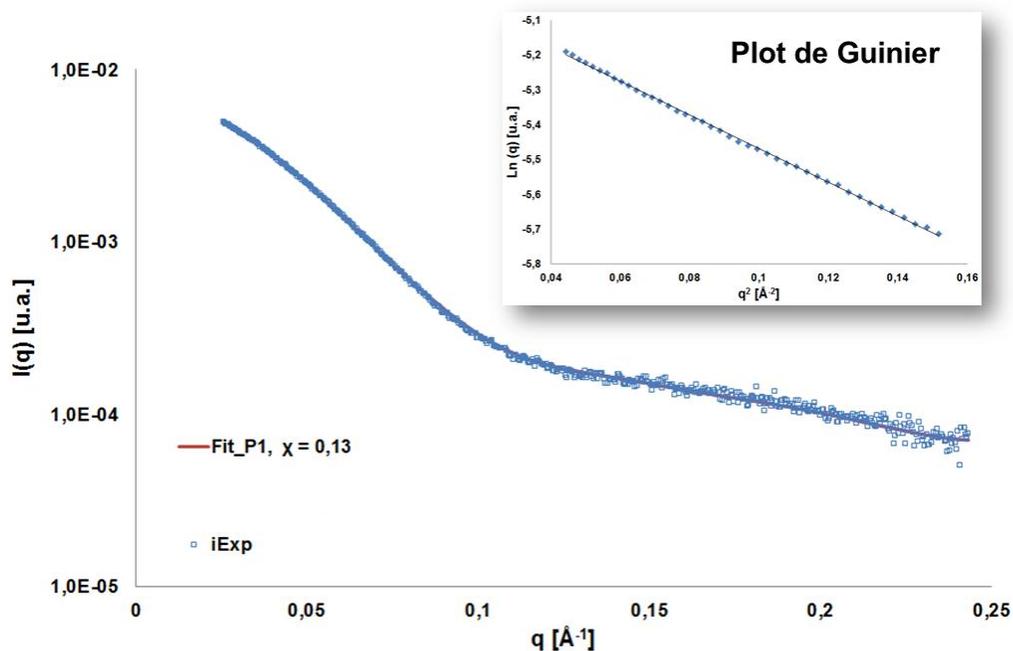
#### 1.4.3.8 Análise por espalhamento de raios-X a baixo ângulo (SAXS)

A caracterização estrutural das proteínas em estudo por SAXS foi realizada de forma complementar aos ensaios de cristalização que representou o passo limitante para a obtenção da estrutura a alta resolução. Esta técnica ainda foi utilizada para acessar o estado de oligomerização das proteínas em solução e propor um modelo estrutural experimental a baixa resolução.

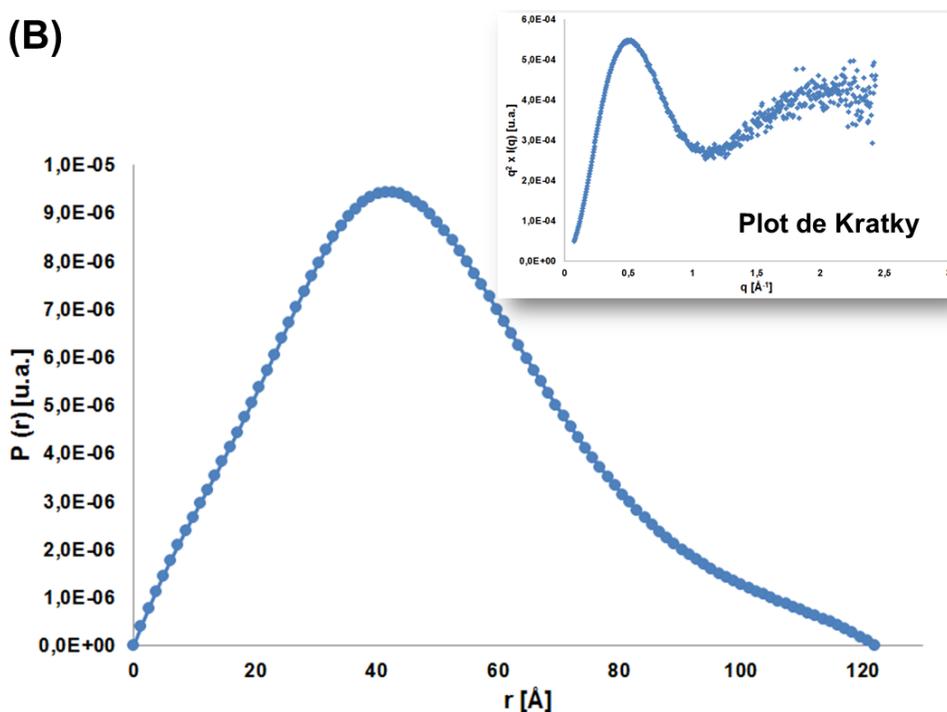
Os dados experimentais foram obtidos para as concentrações de 1, 2,5 e 5 mg/mL e mostraram padrão similar. Os dados descritos a seguir foram obtidos na concentração de 5 mg/mL e as curvas experimentais estão mostradas nas figuras 24A e 25A. A linearidade observada na região de Guinier (Figuras 24A e 25A) indica que o sistema estava monodisperso e permitiu o cálculo de um raio de giro de 38,30 e 49,8 ± 1,3 Å para PhbF e PhaP1, respectivamente. A função de distribuição de distâncias  $P(r)$ , obtida a partir dos dados experimentais, descreve PhbF como uma proteína globular com  $D_{\text{máx}}$  igual a 122 Å e PhaP1 como uma proteína alongada com  $D_{\text{máx}}$  igual a 200 Å (Figuras 24B e 25B). O plot de Kratky apresentou um pico máximo definido para as duas proteínas, contudo não foi observado uma diminuição de  $Iq^2$  com o aumento de  $q$  indicando que as proteínas encontram-se parcialmente desnaturadas ou contenham regiões com alto grau de desordem (Figuras 24B e 25B). A massa molecular foi calculada utilizando o programa SAXSMoW que baseia-se em uma função que relaciona o volume real de uma proteína em solução com o volume aparente obtido das análises de SAXS (FISCHER *et al.*, 2009). Segundo esta análise PhbF e PhaP1 possuem massas moleculares de 90,1 e 105,3 kDa. Estes dados sugerem que estas proteínas encontram-se nas formas tetramérica e pentamérica, respectivamente, sendo corroboradas pelos dados de gel filtração.

FIGURA 24 – Análises de SAXS da proteína PhbF

(A)

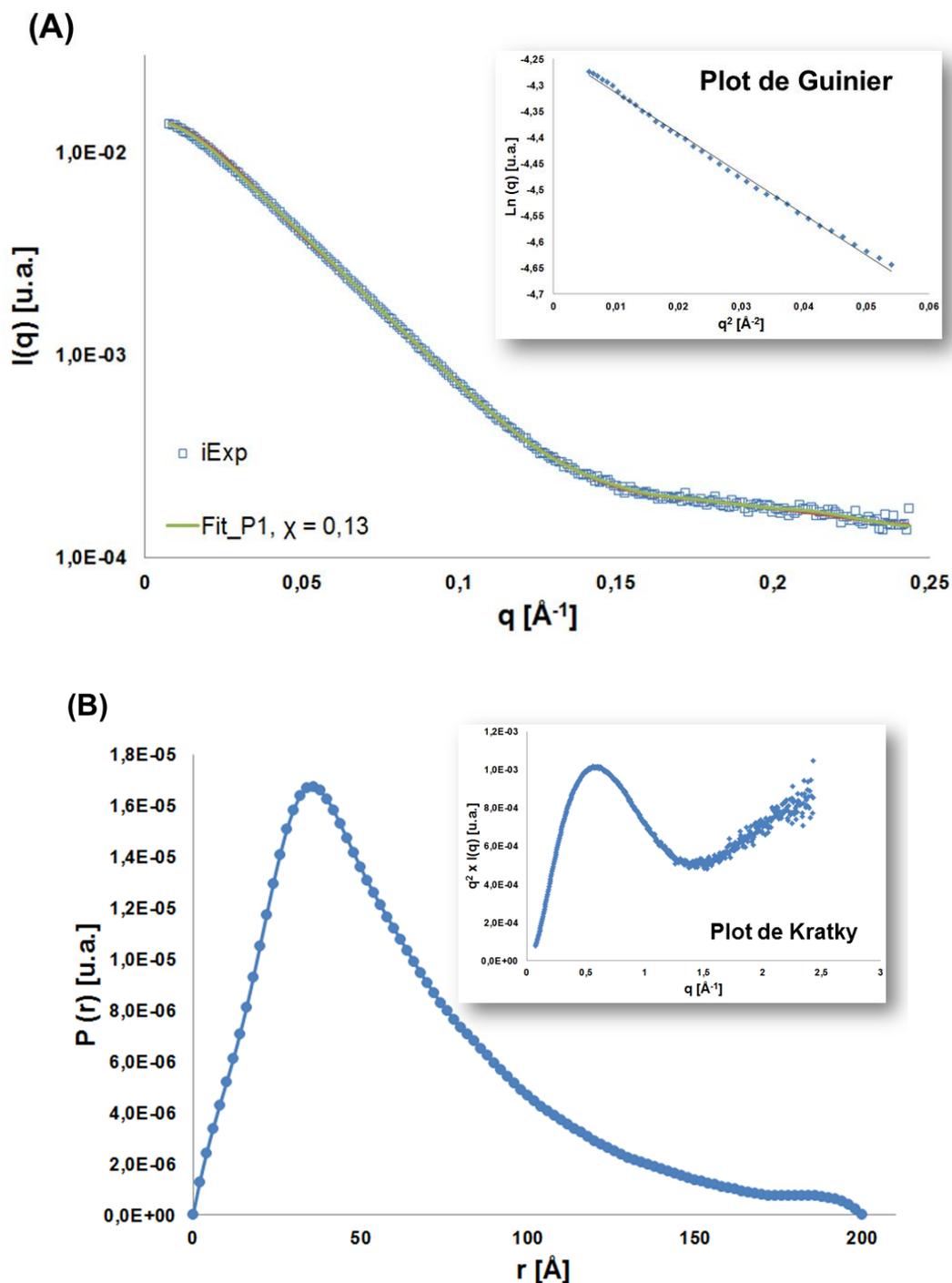


(B)



(A) Curva de espalhamento experimental da proteína HisPhbF a 5 mg/mL e o comportamento linear observado no plot de Guinier. O ajuste dos modelos *ab initio* foi calculado pelo programa DAMMIF. (B) Função de distribuição de distâncias derivados da curva experimental utilizando o programa GNOM. No quadro menor encontra-se o plot de Kratky obtido dos dados experimentais. [u.a.]: unidades arbitrárias.

FIGURA 25 – Análises de SAXS da proteína PhaP1



(A) Curva de espalhamento experimental da proteína HisPhaP1 a 5 mg/mL e o comportamento linear observado no plot de Guinier. O ajuste dos modelos *ab initio* foi calculado pelo programa DAMMIN. (B) Função de distribuição de distâncias derivados da curva experimental utilizando o programa GNOM. No quadro menor encontra-se o plot de Kratky obtido dos dados experimentais. [u.a.]: unidades arbitrárias.

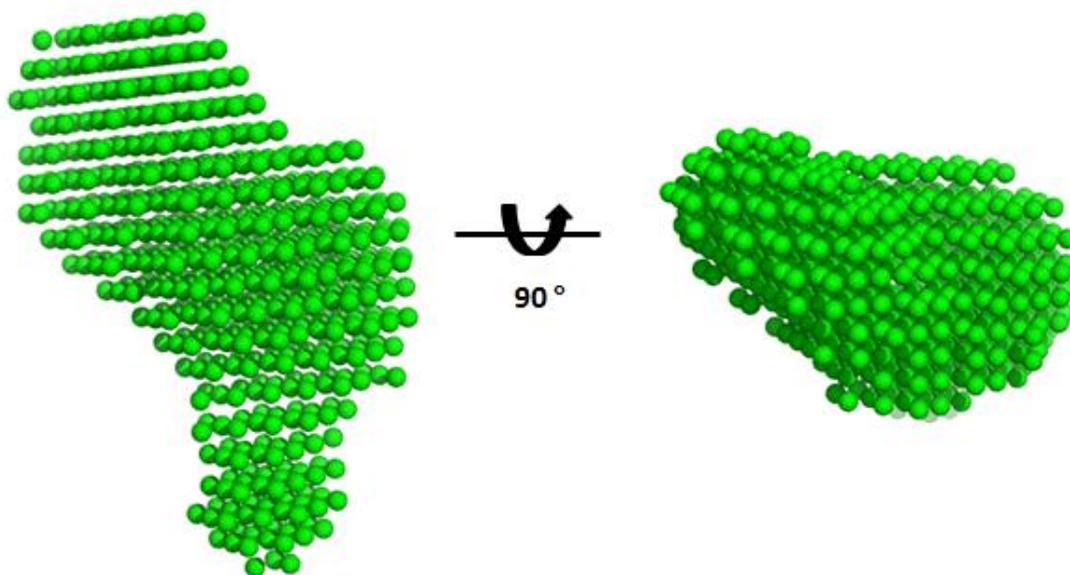
Os modelos estruturais foram gerados *ab initio* a partir dos dados de SAXS utilizando o programa DAMMIF. Este utiliza a estratégia de modelagem por esferas onde a partícula é representada por uma coleção de esferas densamente empacotadas dentro de um volume de procura. Cada esfera é classificada como pertencendo à partícula ou ao solvente. A modelagem inicia com um modelo inicial arbitrário utilizando técnicas de *annealing* simulado para construir um modelo compacto interconectado que apresente um padrão de espalhamento que melhor se ajuste aos dados experimentais (FRANKE & SVERGUN, 2009). Vinte modelos independentes foram gerados para cada proteína sem a imposição de simetria e em seguida alinhados. A estrutura média foi normalizada quanto ao volume da partícula. Todos os modelos apresentaram valores de discrepância (NSD) baixos (Tabela 2) mostrando que o programa encontrou facilmente a mesma solução partindo de parâmetros diferentes. O ajuste dos dados obtidos com os modelos *ab initio* aos dados experimentais pode ser observado nas figuras 24A e 25A. Os modelos obtidos apresentaram um bom ajuste aos dados experimentais como avaliado pelo parâmetro  $\chi$  cujos valores encontram-se na tabela 2. Os modelos finais para as proteínas PhbF e PhaP1 são mostrados na Figura 26.

O modelo *ab initio* da proteína PhbF mostrou a presença de dois possíveis domínios. A porção inferior e menor do modelo possivelmente represente o domínio de ligação ao DNA enquanto que o domínio mais volumoso corresponda ao domínio de interação da proteína aos grânulos de PHB (Figura 26A).

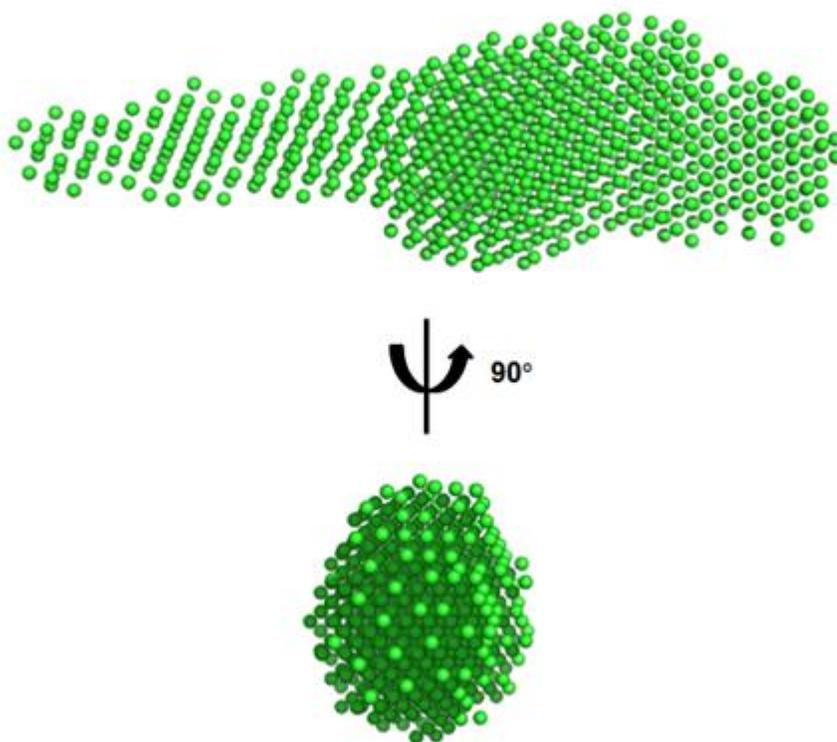
O modelo da proteína PhaP1 apresentou um formato alongado onde também é possível distinguir duas regiões. Uma região mais volumosa provavelmente envolvida na interação com o polímero (~103 Å) e uma região estendida e menor (~87Å) (Figura 26B). Esta região pode corresponder à porção desordenada da proteína e ser responsável pela parcela de desordem observada no plot de Kratky. Este modelo é diferente do proposto por NEWMANN e colaboradores (2008) que descreve PhaP1 de *R. eutropha* como um trímero com envelope molecular em formato de triângulo plano, também determinado por SAXS.

FIGURA 26 – Modelos *ab initio* das proteínas PhbF e PhaP1 construídos a partir dos dados de SAXS

(A)



(B)



Modelo do envelope molecular médio das proteínas PhbF (A) e PhaP1 (B) construído com o programa DAMMIF não impondo simetria. Figuras preparadas com o programa Pymol.

**TABELA 2 – Parâmetros estruturais obtidos com os dados de SAXS das proteínas PhbF e PhaP1**

|                         | PhbF(Exp) <sup>†</sup> | PhbF (Dam) <sup>‡</sup> | PhaP1(Exp) <sup>†</sup> | PhaP1 (Dam) <sup>‡</sup> |
|-------------------------|------------------------|-------------------------|-------------------------|--------------------------|
| MW (kDa) <sup>φ</sup>   | 90,1                   | -                       | 105,3                   | -                        |
| Oligômero estimado      | 4                      | -                       | 5                       | -                        |
| Rg (Å) <sup>*</sup>     | 38,30±0,00             | -                       | 49,8± 1,3               | -                        |
| Rg (Å) <sup>**</sup>    | 38,21±0,00             | 32,68                   | 50,94±0,0               | 49,39                    |
| Dmax (Å) <sup>**</sup>  | 122                    | 126,5                   | 200                     | 198,0                    |
| Resolução (Å)           | 25                     | 25                      | 25                      | 25                       |
| Discrepância ( $\chi$ ) | -                      | 0,130                   | -                       | 0,132                    |
| NSD                     | -                      | 0,905                   | -                       | 0,627                    |

<sup>φ</sup> Calculado a partir do programa SAXSMoW.

<sup>†</sup> Valores calculados a partir de dados experimentais.

<sup>‡</sup> Parâmetros calculados a partir dos modelos de átomos dummy.

<sup>\*</sup> Calculado a partir do plot de Guinier.

<sup>\*\*</sup> Calculado a partir da função de distribuição de distâncias.

Resolução:  $2\pi/q_{\text{máx}}$

$\chi$  Calculado com o programa DAMMIF.

NSD (*normalized spatial discrepancy*) calculado com o programa DAMAVER.

#### 1.4.3.9 Análise por modelagem estrutural utilizando o método *ab initio*

O processo de modelagem molecular consiste na obtenção de uma representação da estrutura proteica quando a mesma não pode ser obtida por métodos experimentais. Em geral existem duas categorias de modelagem: dependente de molde (*threading* e modelagem comparativa) e livre de molde ou *ab initio*. As proteínas PhbF e PhaP1 de *H. seropedicae* não podem ser modeladas corretamente pelo método comparativo, pois não possuem nenhuma estrutura no banco de dados PDB (*Protein Data Bank*) com identidade de sequência de aminoácidos superior a 35%. Desta forma o método *ab initio* é uma alternativa viável para a geração de modelos que permitam discutir alguns aspectos estruturais.

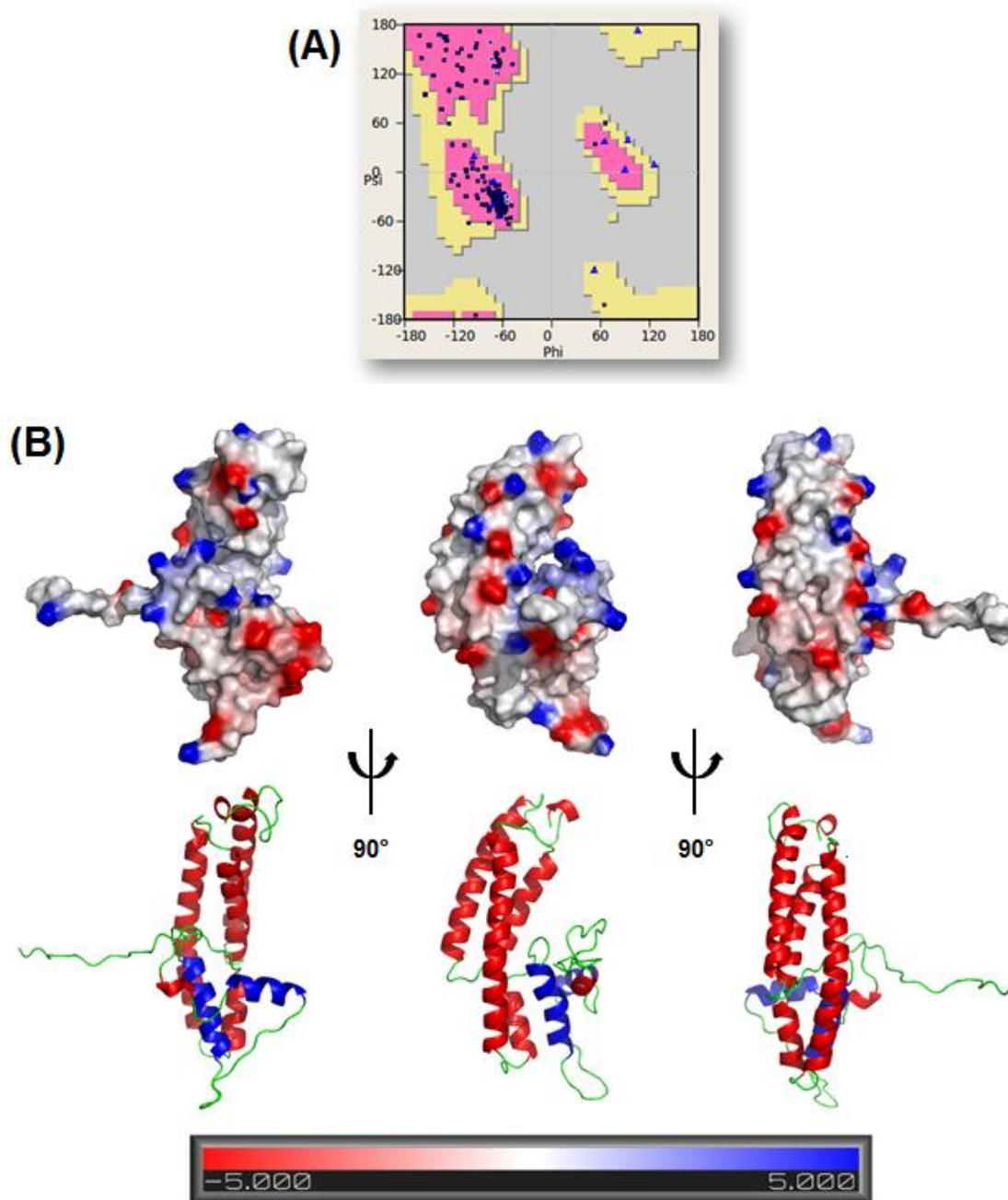
Brevemente, o método *ab initio* baseia-se na predição da estrutura proteica com base nas propriedades físicas da sequência de aminoácidos. Pequenos fragmentos de 3 e 9 aminoácidos são obtidos de proteínas com estrutura tridimensional resolvida e reunidos aleatoriamente com base na sequência de aminoácidos da proteína em estudo através da estratégia de Monte Carlo. Nesta estratégia modificações conformacionais

são realizadas no espaço torsional com o objetivo de obter-se uma estrutura de menor energia (DAS *et al.*, 2008). Rosetta tem sido classificado pelo CASP (*Critical Assessment of Structure Prediction*) como o programa que atinge um melhor desempenho com base nesta estratégia. Predições de domínios proteicos com menos de 125 aminoácidos regularmente tem alcançado desvios menores do que 5 Å em relação as estruturas determinadas experimentalmente (KAUFMANN *et al.*, 2010) mostrando a eficácia deste método.

A sequência de aminoácidos das proteínas PhbF e PhaP1 foram submetidas ao servidor Robetta (KIM *et al.*, 2004) que gerou 10000 estruturas. Estes modelos foram agrupados quanto ao parâmetro de raiz quadrada do quadrado dos desvios de C $\alpha$  (RMSD), onde valores baixos representam estruturas semelhantes. Os modelos centróides dos grupos mais populosos foram disponibilizados. Os modelos das proteínas PhbF e PhaP1 foram validados pela sua qualidade estereoquímica como pode ser observado pelos seus respectivos gráficos de Ramachandran (Figuras 27A e 28A).

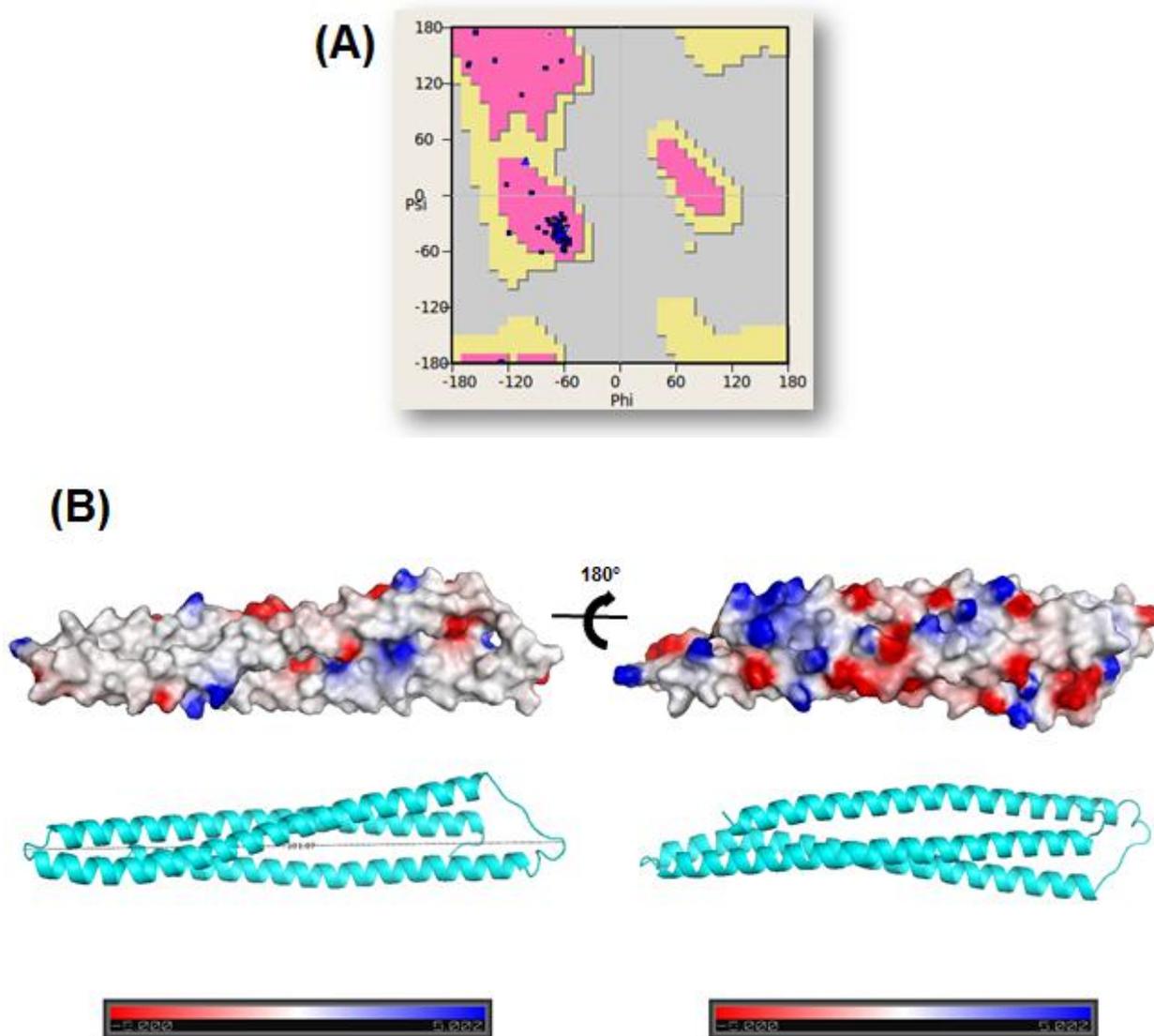
O modelo do monômero da proteína PhbF apresenta dois domínios: N-terminal composto por um possível motivo hélice-volta-hélice (HTH) relacionado com a interação da proteína com o DNA e C-terminal composto por três hélices possivelmente envolvidas na interação com o polímero PHA (Figura 27B). O motivo de HTH modelado está de acordo com a descrição geral que define este motivo como composto por um segmento de 20 aminoácidos que englobam duas  $\alpha$  hélices perpendiculares conectadas por uma volta  $\beta$  formada por 4 resíduos de aminoácido até alças longas como *loops* (LUSCOMBE *et al.*, 2000). Estes elementos podem ser observados na figura 27B (região em azul). Luscombe e colaboradores também descrevem a localização do motivo HTH em um bolsão hidrofóbico composto de 3 a 6 hélices também encontrado no modelo de PhbF. A face correspondente às hélices que formam a região C-terminal apresentou uma região diferencialmente mais hidrofóbica (Figura 27B) evidenciando uma provável região de interação com o polímero ou uma área de interação proteína-proteína importante na formação da estrutura tetramérica.

FIGURA 27 – Modelo *ab initio* do monômero da proteína PhbF



(A) Gráfico de Ramachandran mostrando a qualidade estereoquímica do modelo gerado. (B) Modelo estrutural em representação do tipo *ribbon* (parte inferior) e representação da superfície eletrostática (parte superior). Hélices possivelmente envolvidas com a ligação ao DNA (azul), hélices possivelmente envolvidas com a ligação da proteína ao PHB (vermelho), alças (verde). Escala de potencial eletrostático calculada pela equação de Poisson-Boltzmann no intervalo de -5 a +5 kbT/e. Região em vermelho representa potencial negativo e região azul potencial positivo. Superfície eletrostática gerada pelo programa Delphi e Figuras gerada com o programa Pymol.

FIGURA 28 – Modelo *ab initio* do monômero da proteína PhaP1



(A) Gráfico de Ramachandran mostrando a qualidade estereoquímica do modelo gerado. (B) Modelo estrutural em representação do tipo *ribbon* (parte inferior) e representação da superfície eletrostática (parte superior). Escala de potencial eletrostático calculada pela equação de Poisson-Boltzmann no intervalo de -5 a +5 kbT/e. Região em vermelho representa potencial negativo e região azul potencial positivo. Superfície eletrostática gerada pelo programa Delphi e Figuras gerada com o programa Pymol.

A proteína PhaP1 reveste externamente os grânulos de PHB impedindo a coalescência dos mesmos, possibilitando assim a rápida mobilização do polímero em intermediários metabólicos. O modelo *ab initio* sugere que o monômero de PhaP1 é uma proteína com característica fibrilar composta por três hélices que se alternam (Figura 28B). A análise de sua superfície eletrostática mostra a presença de uma face hidrofóbica e outra com uma distribuição de cargas mais uniforme (Figura 28B). Esta característica é concordante com estudos realizados por Hanley e colaboradores (1999) com a proteína PhaP de *R. eutropha* que sugere a presença de domínios hidrofílicos formando uma face de interação citoplasmática ou de interação com outras proteínas presentes na superfície do grânulo. Estas regiões hidrofílicas são alternadas por domínios hidrofóbicos caracterizando PhaP como uma barreira anfifílica (HANLEY *et al.*, 1999).

#### **1.4.4 Caracterização funcional das proteínas PhbF e PhaP1 de *H. seropedicae* SmR1**

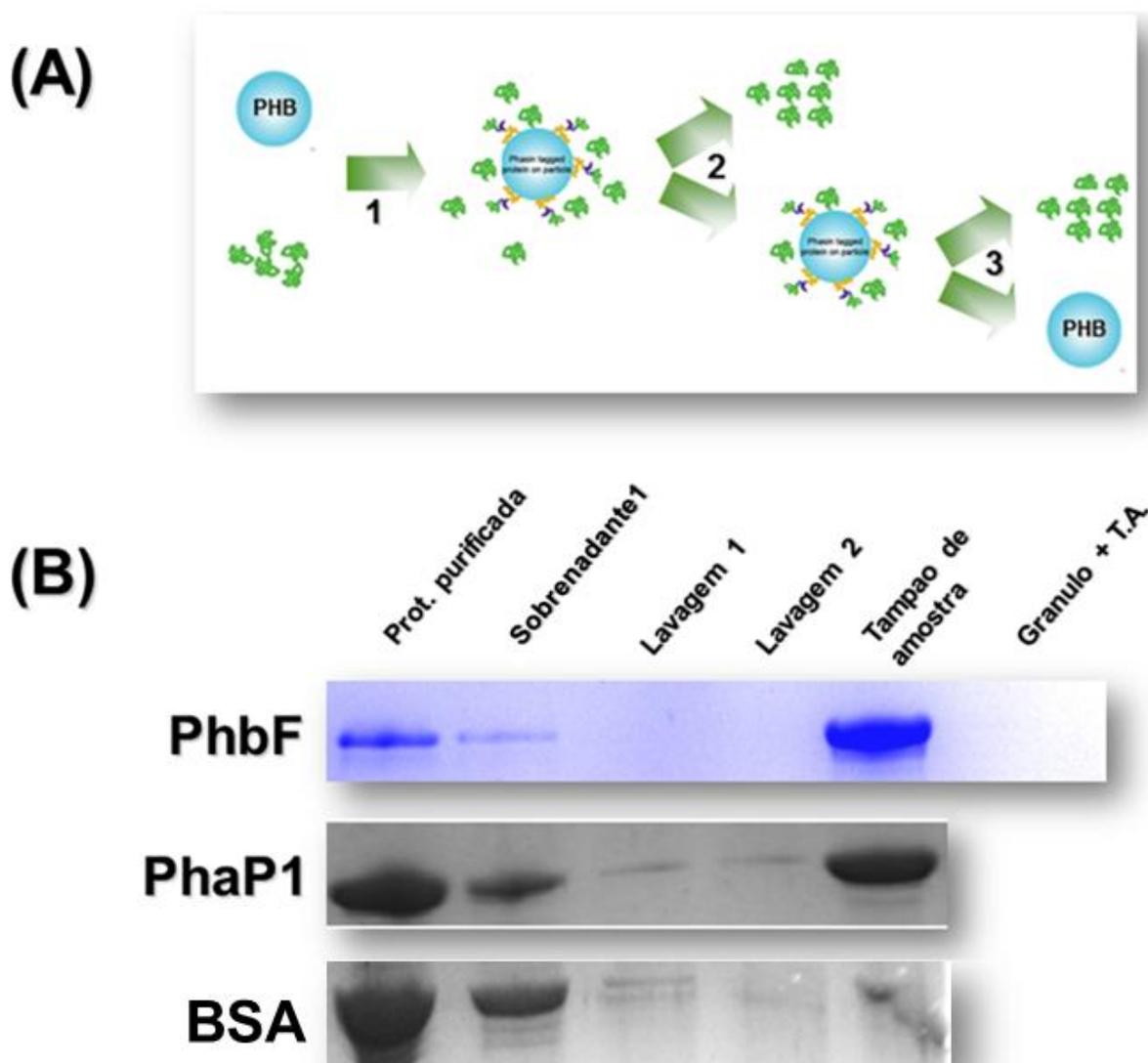
##### **1.4.4.1 Ensaio de interação das proteínas PhbF e PhaP1 com o polímero PHB**

Um ensaio de co-precipitação foi realizado com o objetivo de avaliar a atividade de interação das proteínas PhbF e PhaP1 de *H. seropedicae* SmR1 com o PHB extraído do mesmo organismo (1.3.2.4).

Um resumo do ensaio pode ser observado na figura 29A. As proteínas PhbF e PhaP1 interagiram fortemente com o polímero após duas lavagens com tampão Tris-HCl pH 7,5 e dissociação por ação dos agentes desnaturantes SDS e  $\beta$ -mercaptoetanol, sucedida de aquecimento a 95°C (Figura 29B). Como controle foi utilizado BSA que interagiu fracamente com o PHB (Figura 29B). Outro controle utilizado foi uma fração contendo só o PHB que mostrou a ausência de proteínas previamente associadas ao polímero e a eficiência do método de extração.

Este ensaio revelou a capacidade das proteínas PhbF e PhaP1 de interagirem *in vitro* com o PHB extraído de *H. seropedicae* e indicou que o mesmo ocorre *in vivo*. Este resultado também sugere que somente o polímero é capaz de funcionar como uma molécula efetora expondo sítios de interação e mobilizando as proteínas do citoplasma para o grânulo de PHB.

FIGURA 29 – Ensaio de interação entre as proteínas PhbF e PhaP1 com PHB



(A) Esquema do ensaio de interação. 1. interação. 2. lavagem e remoção do excesso proteico. 3. dissociação. (B) Perfil eletroforético em gel de poliacrilamida (SDS-PAGE) 12% das frações correspondentes ao ensaio de interação. O ensaio foi realizado em um volume final de 100  $\mu$ L contendo: 25  $\mu$ g de proteína, 1 mg de PHA, 50 mM Tris-HCl pH 7,5. A reação foi incubada a 37  $^{\circ}$ C por 10 minutos, centrifugada e a fração solúvel reservada. O material sedimentado foi lavado duas vezes em tampão 50 mM Tris-HCl pH 7,5. O grânulo de PHB extraídos de *H. seropedicae* SmR1 tratados com tampão contendo SDS e  $\beta$ -mercaptoetanol (Grânulo + T.A.). Os géis foram corados com azul de Coomassie.

#### 1.4.4.2 Interação da proteína PhbF com as regiões promotoras dos genes relacionados ao metabolismo de PHB de *H. seropedicae* SmR1

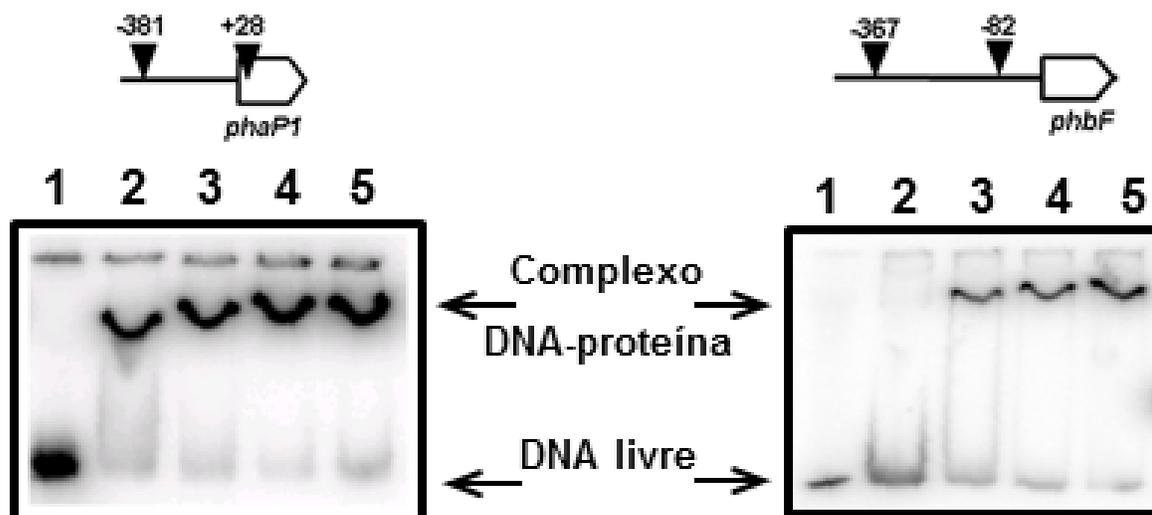
Várias evidências na literatura sugerem a ligação da proteína PhbF ao seu próprio promotor assim como aos promotores dos genes *phbC* (PHB sintase) e *phaP* (fasina) (MAEHARA *et al.*, 2001; MAEHARA *et al.*, 2002; CHOU *et al.*, 2009). Com o objetivo de estudar a atividade *in vitro* de ligação de PhbF ao DNA foram planejados e sintetizados oligonucleotídeos iniciadores para a amplificação das possíveis regiões promotoras dos genes *phbF* e *phaP1*. Foram também planejados oligonucleotídeos para os genes *phbC* (Hsero\_2999), *phaZ* (Hsero\_1622), *phaB* (Hsero\_3315), *phaC* (Hsero\_2625) e *phaZ* (Hsero\_0639) relacionados ao metabolismo de PHAs, assim como para os genes Hsero\_0262, *dksA* (Hsero\_2404) e *fadB* (Hsero\_0254) que possivelmente encontram-se no começo dos *operons* que compreendem respectivamente os genes *phaC*, *phbC* e *phaB*. Os oligonucleotídeos iniciadores planejados para os experimentos de amplificação estão listados no quadro 6.

O ensaio de retardamento de migração de banda em gel foi realizado para os promotores *phbF* e *phaP1* amplificados com oligonucleotídeos marcados radioativamente com [<sup>32</sup>P]. Este tipo de marcação proporciona maior sensibilidade ao ensaio realizado. A proteína PhbF ligou-se às duas regiões promotoras (Figura 30) com constantes de dissociação de 150 nM para o promotor *phaP1* e 450 nM para o promotor *phbF*.

As demais regiões promotoras foram amplificadas utilizando um par de oligonucleotídeos marcados nas extremidades 5' com as sondas fluorescentes 6FAM e VIC respectivamente. O tamanho de cada região promotora está listado na figura 31. Como pode ser observado na figura 32, a proteína HisPhbF foi capaz de ligar-se a todos os promotores testados e não foi capaz de ligar-se ao promotor *nifB* (REGO *et al.*, 2006), não relacionado ao metabolismo de PHA, mostrando a especificidade desta ligação.

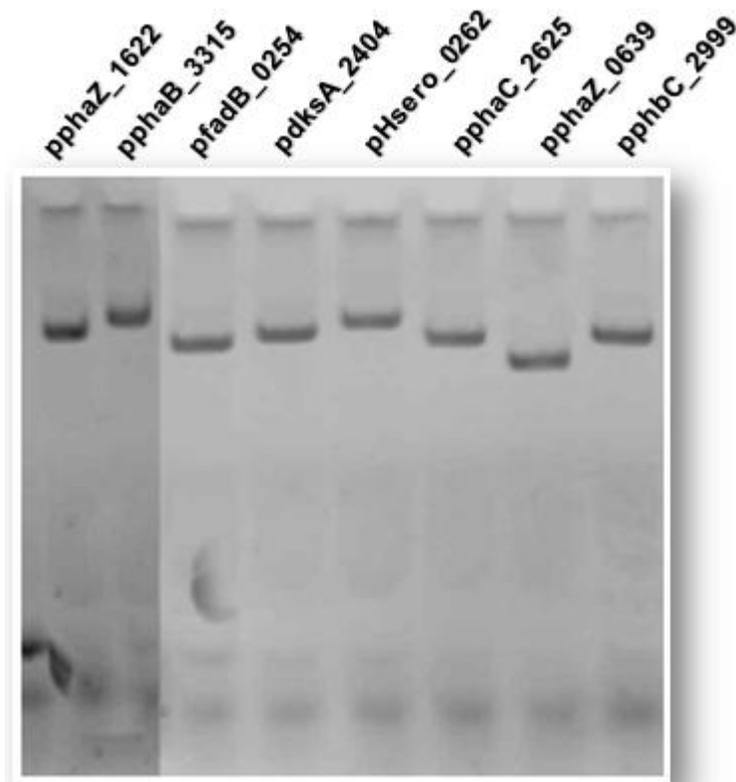
Estes ensaios mostraram que a proteína PhbF encontrava-se ativa quanto a sua atividade de interação com DNA *in vitro*.

**FIGURA 30 – Ensaio de retardamento da migração de banda do referente a interação da proteína PhbF com os promotores *phbF* e *phaP1* marcados com [<sup>32</sup>P]**



Os promotores *phbF* (325 bp) e *phaP1* (409 bp) foram marcados com [<sup>32</sup>P] e incubados a 30 °C por 5 minutos na presença de concentrações crescentes da proteína HisPhbF purificada. As linhas de 1-5 indicam as concentrações 0, 280, 570, 860 e 1100 nM considerando HisPhbF como uma proteína tetramérica. A formação do complexo DNA-proteína foi analisada por separação em gel de poliacrilamida não desnaturante 4%. DNA de timo (100 ng) foi utilizado como competidor não específico em um excesso molar de dez vezes. O esquema acima das fotos indica a porção da região promotora utilizada nos ensaios. Os números acima das setas pretas indicam a posição em relação ao códon de início de tradução.

**FIGURA 31 – Amplificação dos promotores alvos utilizando oligonucleotídeos marcados com as sondas fluorescentes VIC e 6-FAM**



| Promotor                   | pb  |
|----------------------------|-----|
| <i>PphaZ</i> (Hsero_1622)  | 512 |
| <i>PphaB</i> (Hsero_3315)  | 638 |
| <i>PfadB</i> (Hsero_0254)  | 543 |
| <i>PdksA</i> (Hsero_2404)  | 575 |
| <i>PHsero</i> (Hsero_0262) | 661 |
| <i>PphaC</i> (Hsero_2625)  | 545 |
| <i>PphaZ</i> (Hsero_0639)  | 404 |
| <i>PphbC</i> (Hsero_2999)  | 542 |

Eletrforese em gel de poliacrilamida 4 % dos produtos de amplificação dos promotores listados na tabela acima assim como os seus respectivos tamanhos em pb. O número ao lado do nome de cada região promotora corresponde ao número do gene anotado no genoma de *H. seropedicae*. O gel foi visualizado por excitação das sondas fluorescentes com UV a 252 nm e filtro para o corante SYBR green em transiluminador UVP-Biolmaging Systems.

FIGURA 32 – Ensaio de retardamento da migração de banda do complexo DNA promotor – proteína PhbF

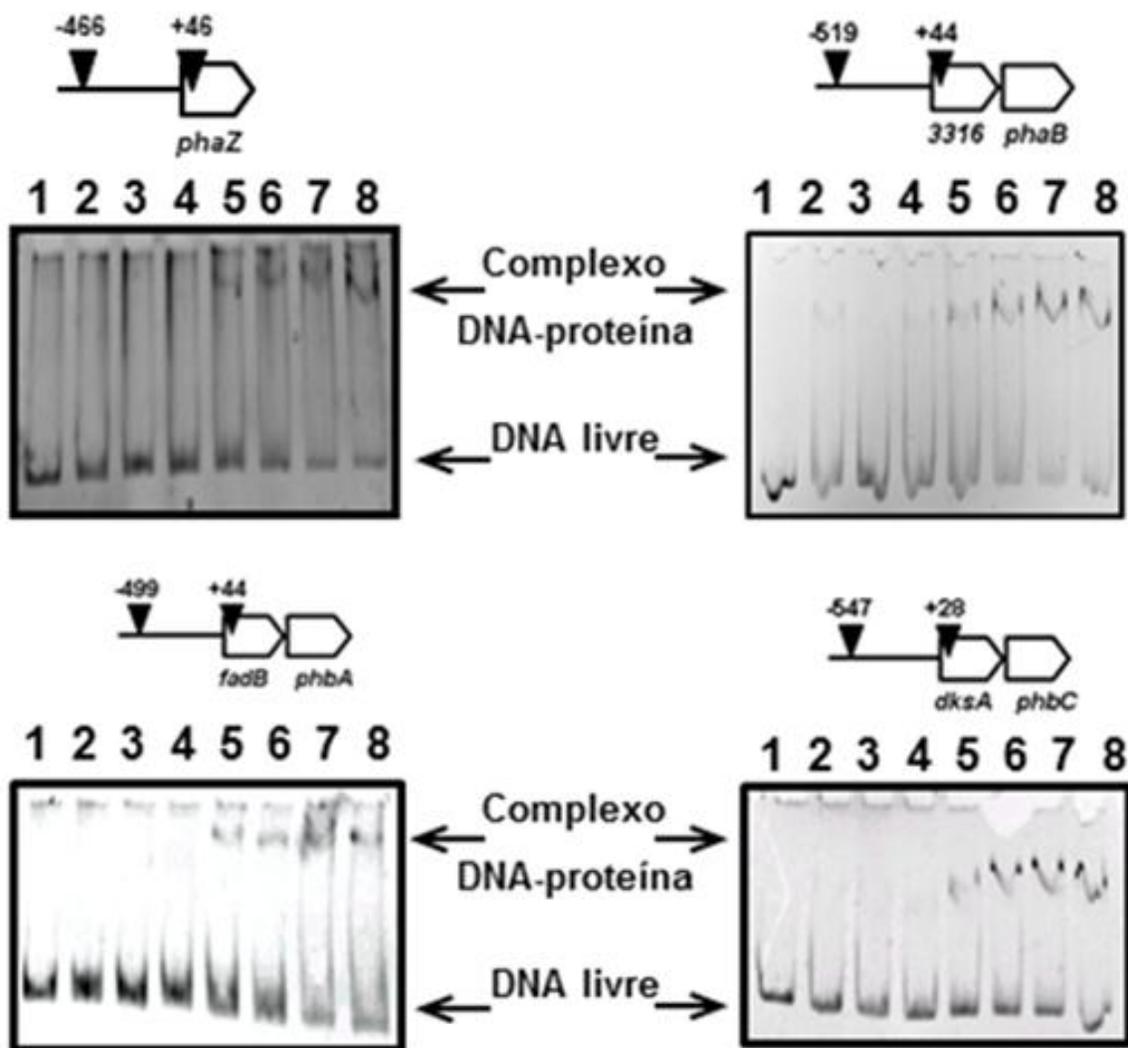
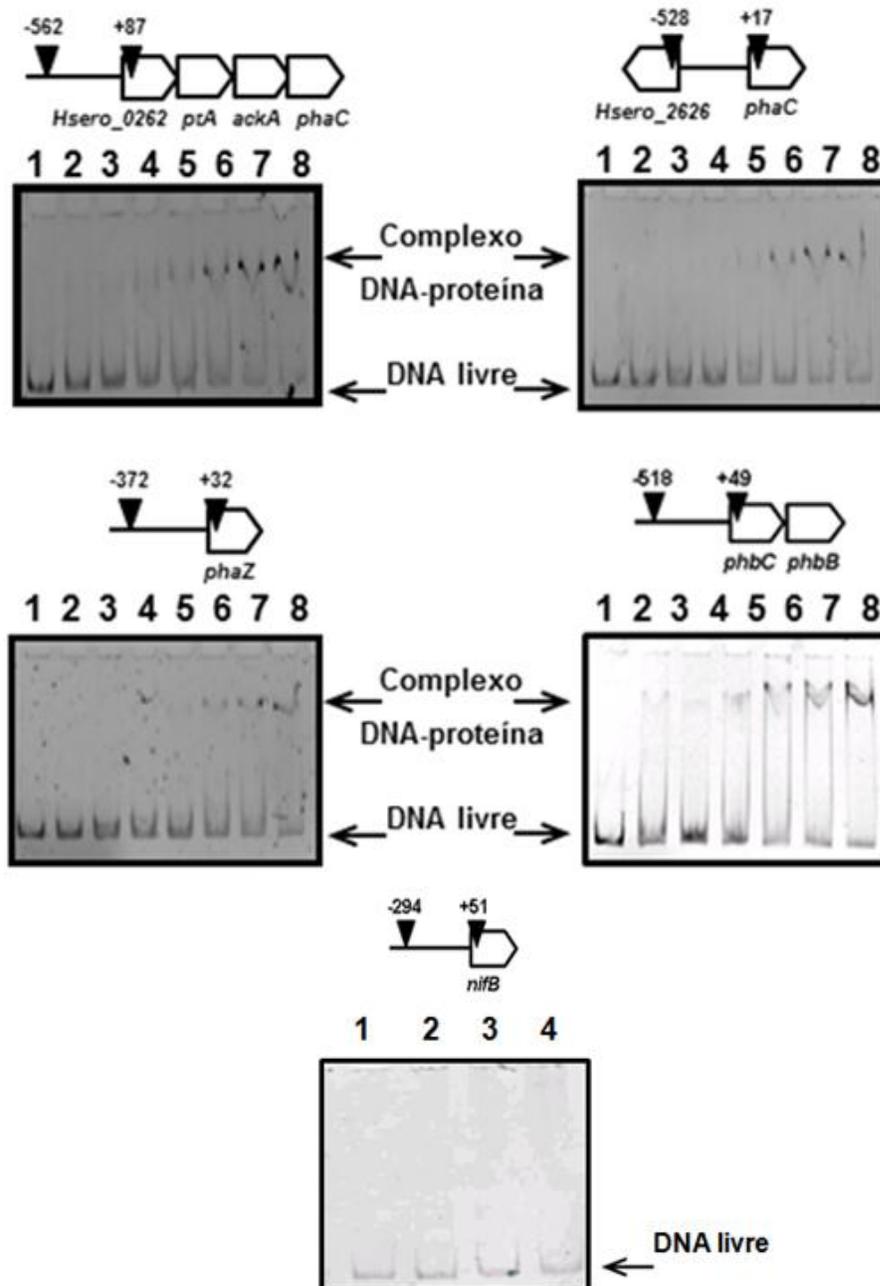


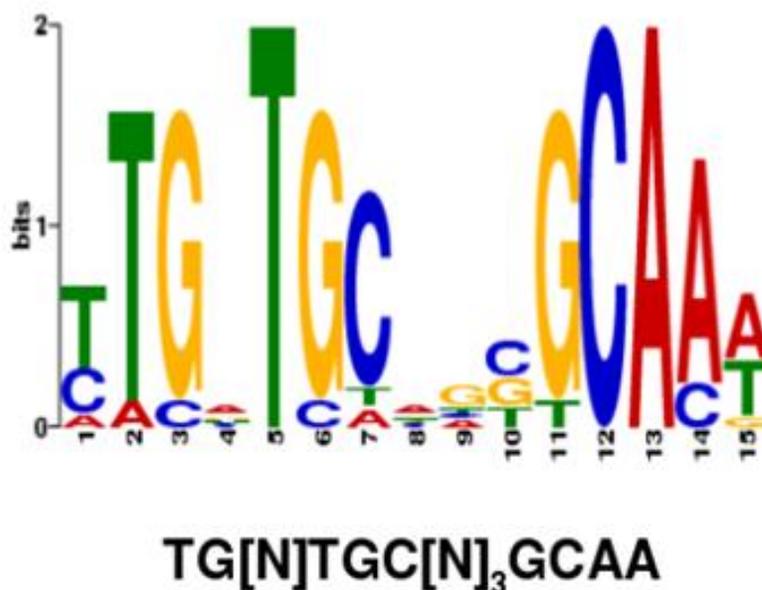
Figura continua na próxima página.



Ensaio de retardamento de migração de banda utilizando eletroforese em gel de poli-acrilamida em condição nativa 4 % (acrilamida/ bisacrilamida 29:1 (m/m)) em tampão TAE 2X. Concentrações crescentes da proteína HisPhbF tetramérica (0, 62, 125, 250, 500, 750, 1000 or 1250 nM – linhas 1-8) foram incubados com aproximadamente 40 nM de cada promotor em tampão de ligação (10 mmol/L Tris-HCl pH 7,5; 80 mmol/L NaCl, 1mM EDTA, 10mM  $\beta$ -mercaptoetanol, 5% (m/v) glicerol). A reação foi incubada cinco minutos a 30 °C. As bandas foram visualizado por excitação das sondas fluorescentes com UV a 252 nm e filtro para o corante SYBR green em transiluminador UVP-Biolmaging Systems. O promotor *nifB* foi utilizado como controle negativo. Linhas de 1-4 indicam as concentrações de 0, 62, 250 or 1250 nM de HisPhbF tetramérica. O esquema acima das fotos indica a porção da região promotora utilizada nos ensaios. Os números acima das setas pretas indicam a posição em relação ao códon de início de tradução.

Como PhbF interagiu com todas as regiões promotoras testadas, as suas sequências foram submetidas ao programa MEME (BAILEY *et al.*, 2009) com o objetivo de identificar um possível motivo conservado e que dessa forma pudesse sugerir um provável sítio de ligação da proteína. Como pode ser observado na figura 33 foi identificado o motivo 5`-TG[N]TGC[N]<sub>3</sub>GCAA-3`. As posições dos motivos em relação ao códon de início de tradução variam entre 80 pb para o gene *phbC* e 562 pb para o gene Hsero\_0262 (Figura 33). Motivo semelhante foi observado também no microrganismo *R. sphaeroides* FJ1 (CHOU & YANG, 2009) com a sequência 5`-CTGC[N]<sub>3</sub>GCAG-3`. Já em *P. denitrificans* foram encontrados dois módulos ricos no motivo TGC os quais são reconhecidos pela proteína PhaR (MAEHARA *et al.*, 2002). De forma geral o motivo TGC assim como a sequência palindrômica GCA são encontrados nas sequências reconhecidas por PhaR dos microrganismos acima citados. Estes relatos corroboram desta forma o motivo de ligação predito para a proteína PhbF de *H. seropedicae*.

**FIGURA 33 – Predição do provável motivo de ligação da proteína PhbF às regiões promotoras**

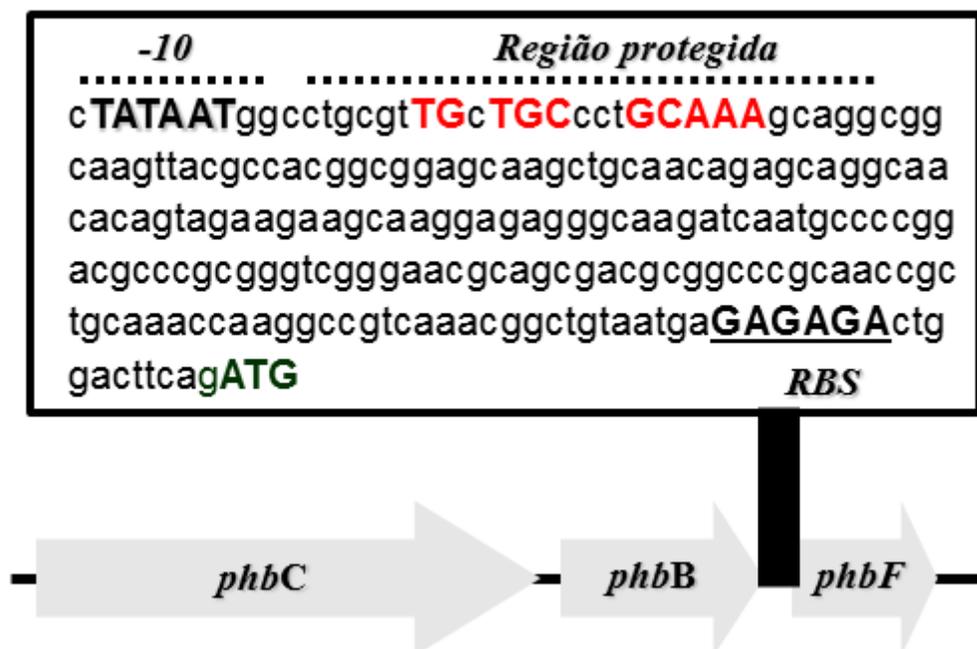


| Promotor           | Fita | Início | Sítio  |
|--------------------|------|--------|--|
| <i>PHsero_0262</i> | +    | 562    | CAATACCGCGT <b>TGATGCATCGCAA</b> TCGCGCCCGGC           |
| <i>PphaC_2625</i>  | -    | 126    | GGTCCGCAACCT <b>TGGTGC</b> TGG <b>GCAA</b> TGGCATGGTGG |
| <i>PphaZ_1622</i>  | -    | 308    | GATACTATCCT <b>TGTTCG</b> GAC <b>GCAA</b> TAGCATAGCAC  |
| <i>PdksA_2404</i>  | +    | 400    | GCTGGCAAAACT <b>TGATGC</b> AGGT <b>CAA</b> ATTCAGGGCCG |
| <i>PphbF</i>       | +    | 167    | ATGGCCTGCGT <b>TGCTGC</b> CCTG <b>GCAA</b> AAAGCAGGCCG |
| <i>PphaZ_0639</i>  | +    | 77     | GTAGTTCTTGCT <b>TGGTGA</b> TGC <b>GCAA</b> AGTCGTCTTAG |
| <i>PfadB_0254</i>  | +    | 350    | GCGTGCGGCGT <b>TGAT</b> CTAGC <b>GCAA</b> TCCGGCCCGGC  |
| <i>PphaP1</i>      | -    | 255    | TTGTTGAAATA <b>TGATGC</b> ATT <b>GCA</b> CAAGCGAATTCT  |
| <i>PphbC_2999</i>  | -    | 80     | ACTAACAAAAC <b>TGCTGC</b> CCG <b>GCAA</b> TTGAGCGGCAT  |
| <i>PphaB_3315</i>  | +    | 205    | CGGCAGCCTCTACA <b>TGCTGG</b> <b>GCAA</b> GAAAAGCCGTT   |

Esquema representando a sequência consenso possivelmente reconhecida pela proteína PhbF de *H. seropedicae*. No eixo y encontra-se a frequência de certo nucleotídeo em uma determinada posição na sequência representada em bits. Abaixo do esquema encontra-se a possível sequência consenso teórica reconhecida por PhaR de *H. seropedicae*. A tabela lista informações sobre fita, posição relativa (pb) ao códon de início de tradução e motivo encontrado a montante de cada gene. Em azul estão representados os pares de base conservados em cada motivo. Os dados foram gerados pelo programa MEME.

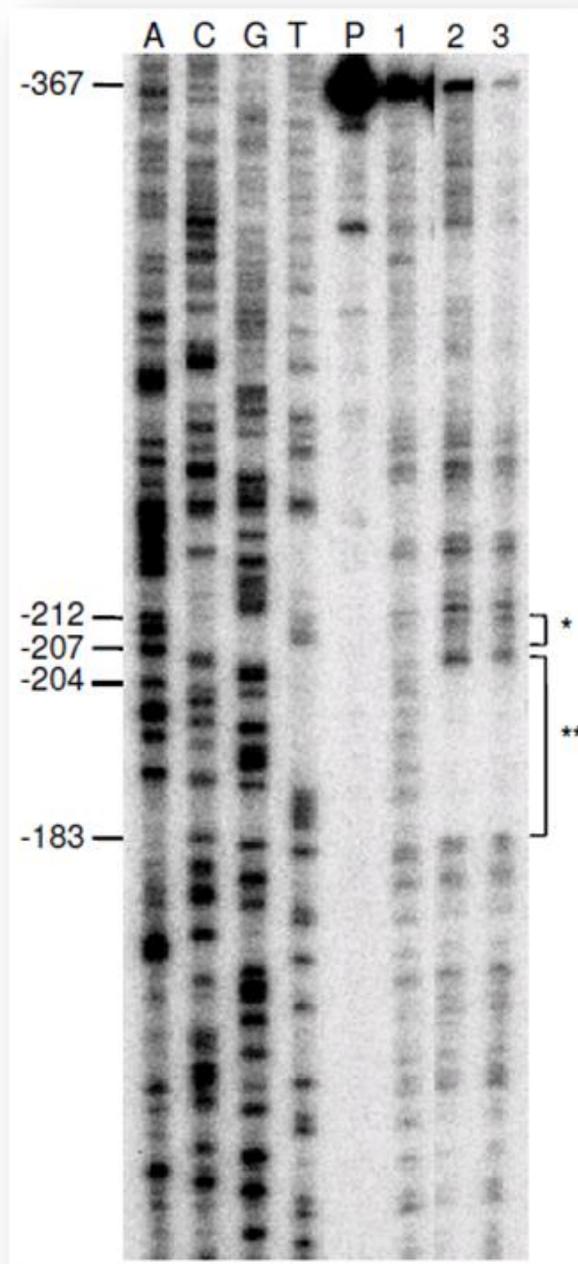
Com o objetivo de mapear *in vitro* a região de ligação da proteína PhbF de *H. seropedicae* e avaliar se a sequência consenso 5`-TG[N]TGC[N]<sub>3</sub>GCAA-3` é importante para a ligação da proteína ao DNA foi realizado o experimento de *DNaseI footprinting* utilizando a região promotora do gene *phbF* (300 pb) marcada com [<sup>32</sup>P]. Este experimento permitiu observar que PhbF protegeu uma área de 24 pb entre as posições 181 e 204 pb acima do códon de início de tradução de PhbF (Figura 35) e que engloba a sequência consenso teórica (Figura 34). Este resultado confirma que esta sequência é realmente reconhecida por PhbF. A região de proteção encontra-se a dois pares de base a jusante de um provável promotor sigma 70 (posições 208 a 212 pb) indicando ser PhbF uma proteína repressora de sua própria transcrição em *H. seropedicae*.

**FIGURA 34 – Interação da proteína HisPhbF com a região promotora *phbF***



Representação da região gênica de *phbF*. A sequência da região intergênica entre *phbB* e *phbF* está detalhada. A linha pontilhada mostra um provável módulo -10 do promotor  $\sigma^{70}$  e a região protegida no ensaio por HisPhbF respectivamente. A sequência consenso predita por bioinformática está mostrada em letras maiúsculas dentro da região protegida. RBS: sítio de ligação ao ribossomo.

**FIGURA 35 – Região promotora de *phbF* é protegida por HisPhbF da ação da DNaseI**



A fita não codificadora do gene *phbF* foi utilizada como sonda. O ensaio foi realizado na ausência (linha 1) ou na presença de 155 (linha 2) ou 312 nM (linha 3) da proteína HisPhbF purificada calculada como tetrâmero. A linha P indica a região promotora não digerida. A reação de sequenciamento está indicada nas linhas A, C, G e T. A região protegida da DNaseI está indicada por \*\*. O provável promotor sigma 70 está indicado por \*. Os números indicam a posição da base em relação ao códon de início de tradução.

#### **1.4.4.3 Caracterização do padrão de expressão dos promotores *phbF* e *phaP1* de *H. seropedicae* SmR1**

Os estudos *in vitro* da proteína PhbF permitiram a identificação do provável sítio de ligação ao DNA assim como a interação com as regiões promotoras dos genes *phbF* e *phaP1*. Os estudos *in vivo* baseados no ensaio de atividade de  $\beta$ -galactosidase foram realizados para avaliar, através do gene repórter *lacZ*, a função repressora de PhbF em *E. coli* ET8000. Esta bactéria oferece um ambiente heterólogo sem a influência da produção do polímero PHB, pois não possui as proteínas necessárias para a sua síntese. Na segunda etapa do estudo, a expressão de *phbF* e *phaP1* foram analisadas transcricionalmente e traducionalmente durante o crescimento de *H. seropedicae* SmR1. O objetivo principal foi estudar a expressão das proteínas PhbF e PhaP1 durante do crescimento celular sob a influência do metabolismo de PHB.

#### **1.4.4.4 Expressão dos promotores *phbF* e *phaP1* em *E. coli* estirpe ET8000**

As regiões promotoras dos genes *phbF* e *phaP1* (Figura 36) foram clonadas no vetor pMP220 para avaliar a expressão através do ensaio de  $\beta$ -galactosidase. Estes vetores foram transformados em *E. coli* estirpe ET8000. Estes promotores foram escolhidos por apresentarem um provável sítio de ligação a PhbF conservado e um padrão claro de ligação à proteína.

Primeiramente a atividade de  $\beta$ -galactosidase foi avaliada sem a presença da proteína PhbF. A expressão do promotor *phbF* foi aproximadamente duas vezes maior quando comparada a *phaP1* (Figura 37). Quando a proteína PhbF foi expressa a partir de um promotor do tipo *tac*, observou-se uma diminuição de 10 vezes na expressão do promotor *phbF* e de 4 vezes em *phaP1* (Figura 37). A expressão da proteína NifH não afetou a atividade de  $\beta$ -galactosidase mostrando que a repressão transcricional não foi causada pela mobilização da maquinaria celular na expressão de uma proteína exógena. Estes resultados mostram que os promotores analisados são funcionais em *E. coli* e *in vivo* a proteína PhbF pode reprimir a transcrição dos genes *phbF* e *phaP1*.

**FIGURA 36 – Sequência de nucleotídeos das regiões promotoras dos genes *phbF* e *phaP1* de *H. seropedicae* SmR1**

**(A)**

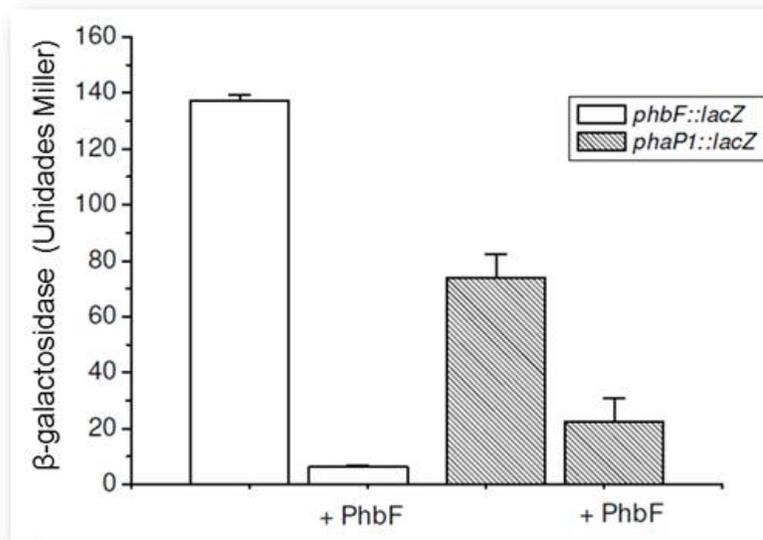
[*phbB*] **TCGGACTTCTCCC**TCAACGGCGGCCTCTACATGGGCTGACCGGCCGTAG  
**Oligo F**  
 CTGGATGGGGCCGGACGCGGTGCATCGCCGCATCCGGCTTCTCACTGGCGGGGA  
 TTTTCTCCGCCAGGCATGAATGGGCTCCCCATTTTCTTGTCTTCCCCCCT**TATAATG**  
 $\sigma^{70}$   
 GCCTGCGTT**TGCTGCCCTGCAA**AAAGCAGGCGGCAAGTTACGCCACGGCGGAGCA  
**Sítio para PhbF**  
 AGCTGCAACAGAGCAGGCAACACAGTAGAAGAAGCAAGGAGAGGGCAAGATCAAT  
 GCCCCGGACGCCCGCGGGTCGGGAACGCAGCGACGCGGCCCGCAACCGCTGCA  
 AACCAAGGCCGTCAAACGGCTGTAATGA**GAGAGA**CTGGACTTCAG**ATGACTACTG**  
**RBS**  
**CAAAAATT**CGGCAGATCGCTTGATCA**AGAAATACCCAATCGC** [*phbF*]  
**Oligo R**

**(B)**

[Hsero\_1670] **ACACAACACTCTCG**TTTTTCTCGGCAGAGCTTGCCGCACTGCAACA  
**Oligo F**  
 AACTGCCTTGATGCGGTGCGCAAACGATTGCGCGCACGGGATTGATAACCAGAA  
 TCCAGTTCCTTGCAGCAGAGTCTATCTGTCTTGAATGCCTCTCAAAAAATCACCTA  
 AGTCCTTGATTAGACGTGGATCAACTTTCCACTGCTTTTCGCCATTCCCATGTGATG  
 AAAGCT**TTGACT**TTTATGCACGATTA**AACTAGAA**TTTCGCTTGT**TGCAATGCATCA**TATT  
 $\sigma^{70}$   $\sigma^{70}$  **Sítio para PhbF**  
 CAACAAGAAATGACATCGGAATCCACCTGTTTCGTTCGAGATCGGTCCGTAAGTGTTT  
 ACTGTTCGCGCTGCCTTGTTTCATCAACCAACTAAAGA**GAAGGA**TGTTTCAAC**ATGAC**  
**RBS**  
**GACCTACACC****GAGCAATTTCCG** [*phaP1*]  
**Oligo R**

Análise das regiões promotoras dos genes *phbF* (A) e *phaP1* (B). As regiões regulatórias conservadas estão em negrito e parte das regiões codificadoras está em vermelho. As sequências sublinhadas indicam as regiões utilizadas para a construção dos oligonucleotídeos para a amplificação das regiões promotoras.

**FIGURA 37 – A proteína PhbF reprime a expressão dos promotores *phbF* e *phaP1* de *H. seropedicae* SmR1 em *E. coli* ET8000**



A atividade de  $\beta$ -galactosidase foi medida a partir de culturas de *E. coli* ET8000 cultivadas por 8 horas em meio mínimo M9 contendo as fusões pMP220::*phbF* ou pMP220::*phaP1*. A proteína PhbF foi expressa a partir de um promotor *tac* do vetor pMMS31(+PhbF). Os dados representam os desvios padrão médios de três leituras independentes. A atividade do vetor pMP220 (controle) na presença de pMMS31 foi inferior a 6 unidades Miller.

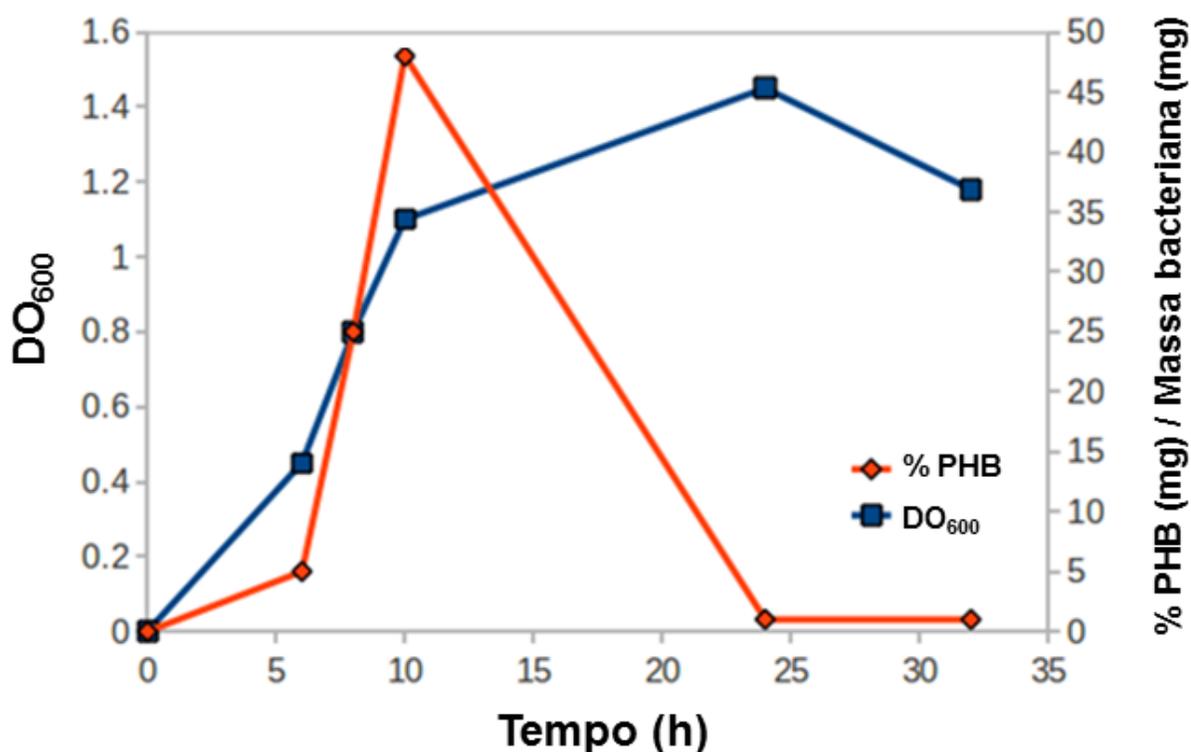
#### **1.4.4.5 Caracterização do padrão de expressão das proteínas PhbF e PhaP1 em *H. seropedicae* SmR1**

A produção de PHB ocorre quando uma fonte de carbono está presente em excesso e ao mesmo tempo o crescimento celular é limitado por elementos essenciais como ferro, magnésio, nitrogênio, enxofre, fosfato ou potássio (STEINBÜCHEL & SCHLEGEL, 1989). Assim a concentração deste polímero flutua ao longo das fases do crescimento celular em função do balanço entre síntese e mobilização desta fonte de reserva. O padrão de produção de PHB em *H. seropedicae* SmR1 quando cultivado em meio NFbHPN é mostrado na figura 38, gentilmente cedida por MÜLLER-SANTOS. Neste meio inicial a relação C/N é de aproximadamente 6,5. A quantidade de polímero produzido por *H. seropedicae* SmR1 aumentou exponencialmente quando a bactéria entrou na fase estacionária de crescimento ( $D.O_{600} = 1,1$ ).

O efeito da produção do polímero PHB sobre a expressão dos promotores dos genes *phbF* e *phaP1* foi avaliado durante 24 horas através das fusões plasmidiais *phbF::lacZ* e *phaP1::lacZ* transformadas na estirpe SmR1 de *H. seropedicae*. As expressões das proteínas foram igualmente acompanhadas através de ensaios de imunodeteção.

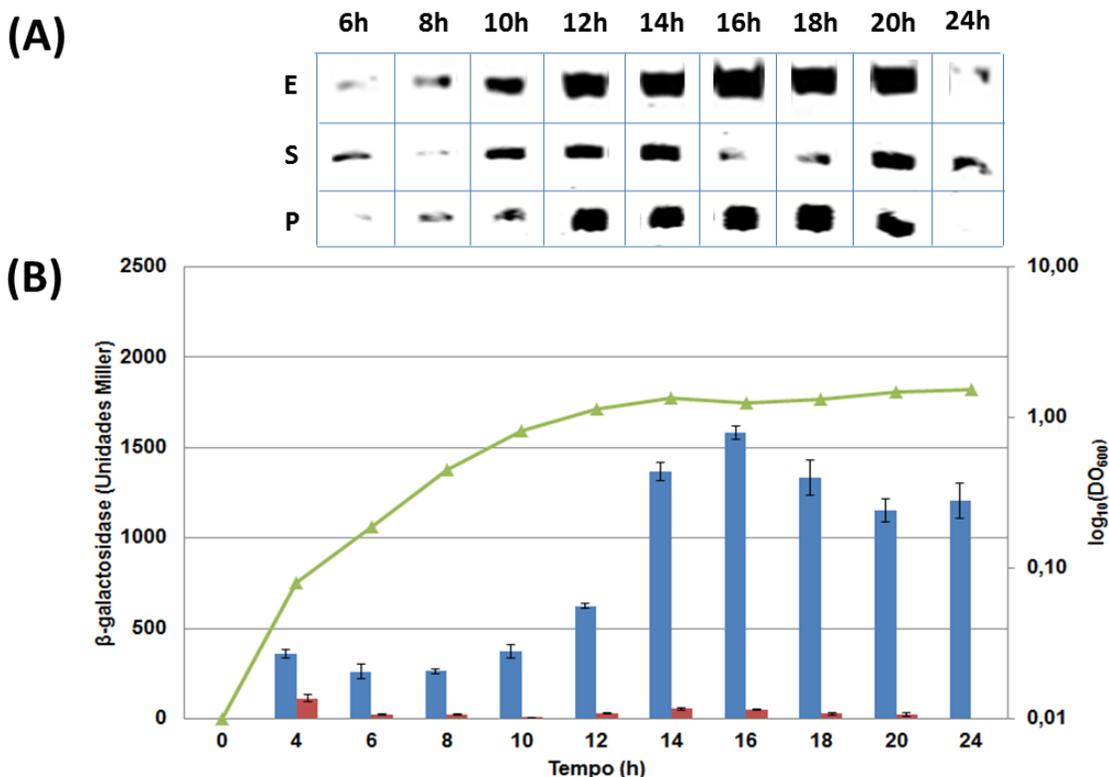
O promotor *phbF* em *H. seropedicae* apresentou um aumento na expressão a partir das 12 horas de crescimento celular atingindo um pico as 16 horas (Figura 39B). Este aumento na atividade de  $\beta$ -galactosidase a uma D.O.<sub>600</sub> entre 1,0 e 1,1 coincidiu com o aumento na produção de PHB nas condições testadas por Müller-Santos (Figura 38).

**FIGURA 38 – Perfil de produção de polihidroxibutirato (PHB) em *H. seropedicae***



*H. seropedicae* estirpe SmR1 cultivado em meio NFbHPN malato após diluição da cultura saturada para uma D.O.<sub>600</sub> de 0,01. As células foram incubadas a 30 °C sob 120 rpm de agitação. Linha azul mostra o crescimento celular monitorado a D.O.<sub>600</sub>. Linha vermelha mostra a quantidade de PHB expressa em % de polímero (mg) por biomassa bacteriana (mg). Resultado cedido pelo Prof. Dr. Marcelo Müller-Santos.

**FIGURA 39 – Padrão transcricional e traducional da expressão de PhbF em *H. seropedicae* SmR1**



*H. seropedicae* SmR1 contendo a fusão *phbF::lacZ* cultivado em meio NFbHP contendo 20 mM  $\text{NH}_4\text{Cl}$  a partir de uma  $D.O_{600}$  de 0,01. Cultivo a 30 °C sob agitação a 120 rpm. A cada 2 horas alíquotas do meio de cultura foram coletadas para os ensaios de  $\beta$ -galactosidase e para a detecção da proteína expressa pelo ensaio de imunodeteccção. (A) Análise da expressão e mobilidade da proteína PhbF. Extrato celular total (E), fração solúvel (S) e fração precipitado (P) foram submetidos a SDS-PAGE seguido de imunodeteccção utilizando o anticorpo para HisPhbF. (B) Acompanhamento da atividade de  $\beta$ -galactosidase ao longo do crescimento celular. Barras em azul representam a atividade em unidades Miller do promotor *phbF* e as barras vermelhas a atividade basal do vetor pMP220. Cada medida corresponde à média de três culturas independentes medidas em triplicata. As barras representam o desvio padrão entre as atividades das culturas independentes. A linha verde representa o crescimento bacteriano acompanhado pela medida da absorbância da cultura a 600 nm.

Esta etapa pode ser interpretada com base na dinâmica de síntese de PHB, onde o aumento no número e no tamanho dos grânulos do polímero pode causar o desligamento da proteína PhbF das regiões promotoras, permitindo assim o aumento da própria expressão. Este evento pode ser observado pelo maior aparecimento da proteína a partir das 12 horas na fração precipitado que contém as proteínas presentes no grânulo sedimentado (Figura 39A). No pico de expressão do promotor *phbF* (16 horas) observa-se que a maior parte da proteína encontra-se associada ao grânulo uma vez que uma fração muito pequena da proteína é encontrada na fração solúvel que equivale às proteínas citoplasmáticas. A partir das dezesseis horas de crescimento ocorre a diminuição na atividade do promotor *phbF* (Figura 39B). Este fato pode compreender a fase de mobilização de PHB onde a degradação leva a diminuição no tamanho dos grânulos com a concomitante dissociação da proteína PhbF. Este retorno da proteína ao citoplasma pode ser evidenciado no ponto de 20 horas onde a proteína PhbF volta a aparecer em maior concentração na fração solúvel ou citoplasmática (Figura 39A).

Um aumento na atividade de expressão de  $\beta$ -galactosidase foi observado a partir das 12 horas para o promotor *phaP1* (Figura 39B). Este padrão observado também para o promotor *phbF* indica o acoplamento da expressão destes promotores ao metabolismo de PHB em *H. seropedicae* SmR1. Este acoplamento também é relatado em *R. eutropha* (YORK *et al.*, 2002).

A proteína PhaP1 é detectada na fração solúvel ou citoplasmática somente entre 8 e 12 horas de crescimento (Figura 40A), mostrando que a partir das 14 horas passa a haver uma maior mobilização da faseína para a fração precipitado onde encontram-se as proteínas associadas aos grânulos de PHB. A partir das 18 horas não foi observado o retorno da proteína a fração solúvel (Figura 40A) como observado para a proteína PhbF (Figura 39A). Este resultado pode ser explicado em função da insolubilidade da proteína que após o tratamento de separação das frações solúvel (citoplasma) e precipitado (grânulos de PHB) manteve-se só na última fração. Outra possibilidade seria a degradação de PhaP1 quando dissociada do grânulo na fase estacionária de crescimento. A proteção à degradação de PhaP quando está associada ao grânulo foi sugerida como uma forma de regulação das concentrações celulares da proteína de *R.*

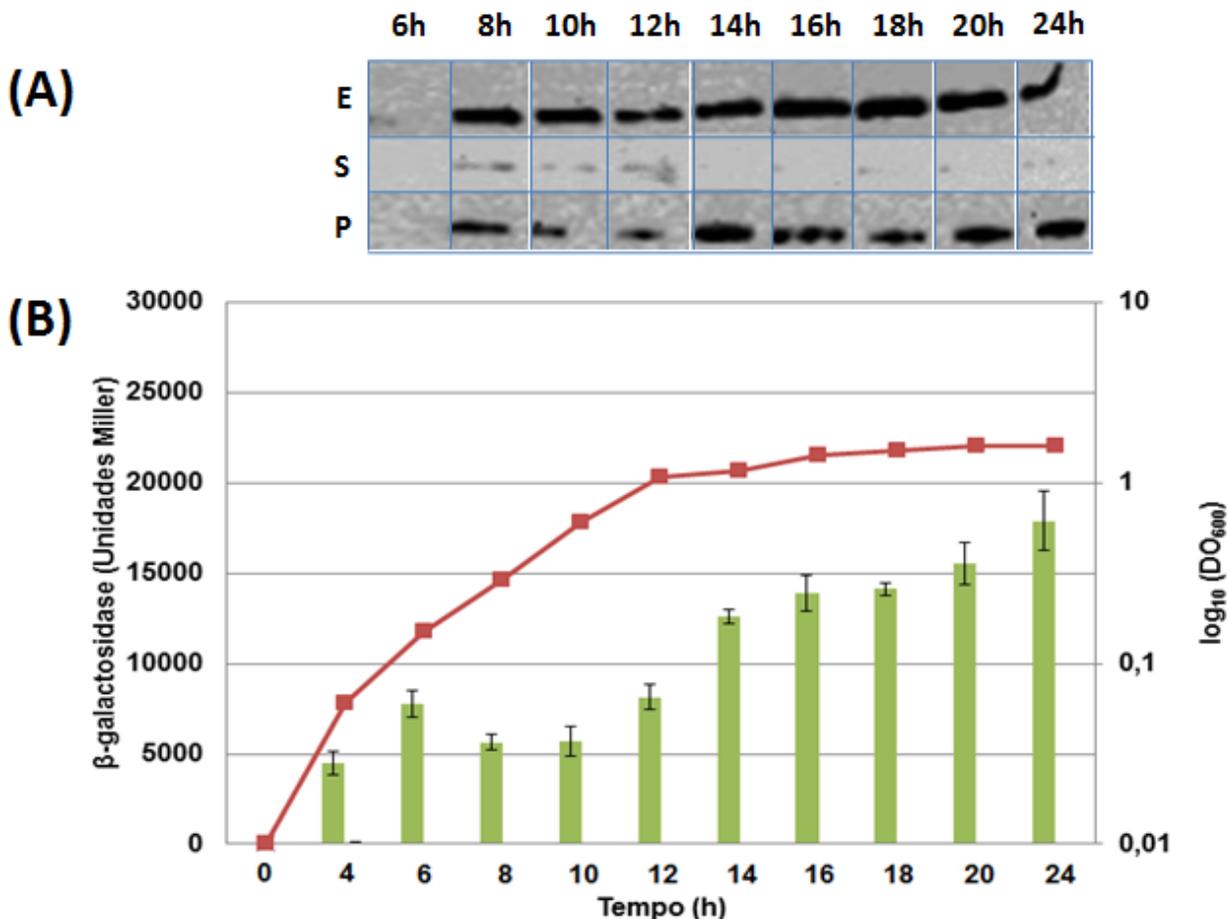
*eutropha* (YORK *et al.*, 2001). Esta suscetibilidade a degradação foi também observada para a proteína PhaP1 de *H. seropedicae* durante o processo de sua purificação.

PhaP1 foi observada em praticamente todas as frações precipitado durante o crescimento de *H. seropedicae*. Esta observação está de acordo com o sugerido para a proteína PhaP de *R. eutropha* onde a presença de quantidades mínimas de PHA na célula seriam o suficiente para desencadear a expressão de PhaP (YORK *et al.*, 2001) que, conseqüentemente, estaria associada ao grânulo.

O promotor *phaP1* apresentou níveis transcricionais mais elevados em *H. seropedicae* quando comparado ao promotor *phbF*. Esta grande expressão permitiu CHAVES e colaboradores (2007) a classificá-la como a quarta proteína mais expressa no proteoma de referência de *H. seropedicae*. Este nível de expressão poderia também ser justificado como uma forma de compensar o curto tempo de vida desta proteína em função da degradação.

Outro ponto importante a ser discutido consiste no papel regulatório de PhbF sobre a expressão do promotor *phaP1* em *H. seropedicae*. York e colaboradores (2002) descrevem PhbF como o fator acoplador da expressão de PhaP1 a produção de PHA em *R. eutropha*. Contudo em *H. seropedicae* este mecanismo não foi evidente ao observar a expressão do promotor *phaP1* de forma temporal, uma vez que não foi observado uma diminuição da expressão do mesmo após 18 horas de crescimento como evidenciado para o promotor *phbF*. Apesar de York e colaboradores (2002) terem chegado a esta hipótese por meio da construção do mutante *phaR<sup>-</sup>* em *R. eutropha* e não através da expressão do promotor durante o crescimento, não pode-se descartar a possibilidade de que o mesmo mecanismo exista em *H. seropedicae*. Este resultado pode ainda ser explicado do ponto de vista da estabilidade do RNA mensageiro. A ação exercida pela proteína PhbF sobre a expressão do promotor *phaP1* pode ter sido camuflada pela maior estabilidade do RNA mensageiro, onde a repressão a nível transcricional não se reflete na expressão traducional.

**FIGURA 40 – Padrão transcricional e traducional da expressão de PhaP1 em *H. seropedicae* SmR1**



*H. seropedicae* SmR1 contendo a fusão *phaP1::lacZ* cultivado em meio NFbHP contendo 20 mM NH<sub>4</sub>Cl a partir de uma D.O.<sub>600</sub> de 0,01. Cultivo a 30 °C sob agitação a 120 rpm. A cada 2 horas alíquotas do meio de cultura foram coletadas para os ensaios de  $\beta$ -galactosidase e para a detecção da proteína expressa pelo ensaio de imunodeteção. (A) Análise da expressão e mobilidade da proteína PhaP1. Extrato celular total (E), fração solúvel (S) e fração precipitado (P) foram submetidos a SDS-PAGE seguido de imunodeteção utilizando o anticorpo para HisPhbF. (B) Acompanhamento da atividade de  $\beta$ -galactosidase ao longo do crescimento celular. Barras em verde representam a atividade em unidades Miller do promotor *phaP1* e as barras vermelhas a atividade basal do vetor pMP220. Cada medida corresponde à média de três culturas independentes medidas em triplicata. As barras representam o desvio padrão entre as atividades das culturas independentes. A linha vermelha representa o crescimento bacteriano acompanhado pela medida da absorbância da cultura a 600 nm.

## 1.5 CONSIDERAÇÕES FINAIS

O metabolismo de polihidroxialcanoatos é uma importante via de assimilação e reserva de intermediários metabólicos, funcionando como suprimento de carbono em condições de privação dos mesmos. O estudo dos mecanismos moleculares envolvidos na regulação da mobilização deste polímero é uma etapa importante para a compreensão do papel do PHA na célula. Em *H. seropedicae*, a regulação deste processo assim como as proteínas que participam do metabolismo de PHAs são pouco conhecidas. Neste trabalho as proteínas PhbF e PhaP1 foram selecionadas com o objetivo de caracterizá-las estrutural e funcionalmente. Estas proteínas não possuem estrutura cristalográfica resolvida e poucos são os relatos que correlacionam o papel *in vivo* com características estruturais. As duas proteínas desempenham funções regulatórias relacionadas à transcrição de genes ligados ao metabolismo e ao tamanho dos grânulos. Estes são macroestruturas que funcionam como sítio para a estocagem, síntese e degradação deste polímero.

PhbF é uma proteína composta basicamente por  $\alpha$ -hélices distribuída em dois domínios. Um domínio N-terminal associado à interação ao DNA e um domínio C-terminal relacionado à interação com os grânulos de PHB. As duas atividades foram confirmadas *in vitro* mostrando a confiabilidade das análises de predição. PhbF é um tetrâmero assim como muitas proteínas que ligam-se ao DNA e desempenham papel repressor, como por exemplo, as proteínas da família TetR. Contudo PhbF não apresentou qualquer similaridade quanto a sequência de aminoácidos com membros desta família e parece possuir uma conformação diferenciada, como observado pelo envelope molecular obtido com dados de SAXS. Este modelo está de acordo com os dados de análise de sequência que sugeriram a presença de dois domínios. Esta disposição também foi observada no modelo estrutural *ab initio* que apresentou o domínio N-terminal como do tipo hélice-volta-hélice. Mais estudos são necessários para compreender melhor a organização destes monômeros para compor a estrutura quaternária final e cujo envelope foi sugerido pela técnica de SAXS. A função de ligação ao DNA foi demonstrada *in vitro* e também revelou a capacidade de PhbF de ligar-se a regiões promotoras de genes relacionados a todas as etapas centrais no

metabolismo de PHA. Esta ligação sugere que o metabolismo de PHA é finamente regulado e este processo é influenciado pela presença do polímero uma vez que PhbF também é capaz de ligar-se ao PHB *in vitro*. Através da análise de todas as regiões promotoras foi possível sugerir a sequência consenso 5`-TG[N]TGC[N]<sub>3</sub>GCAA-3` como o módulo reconhecido por PhbF durante a ligação ao DNA.

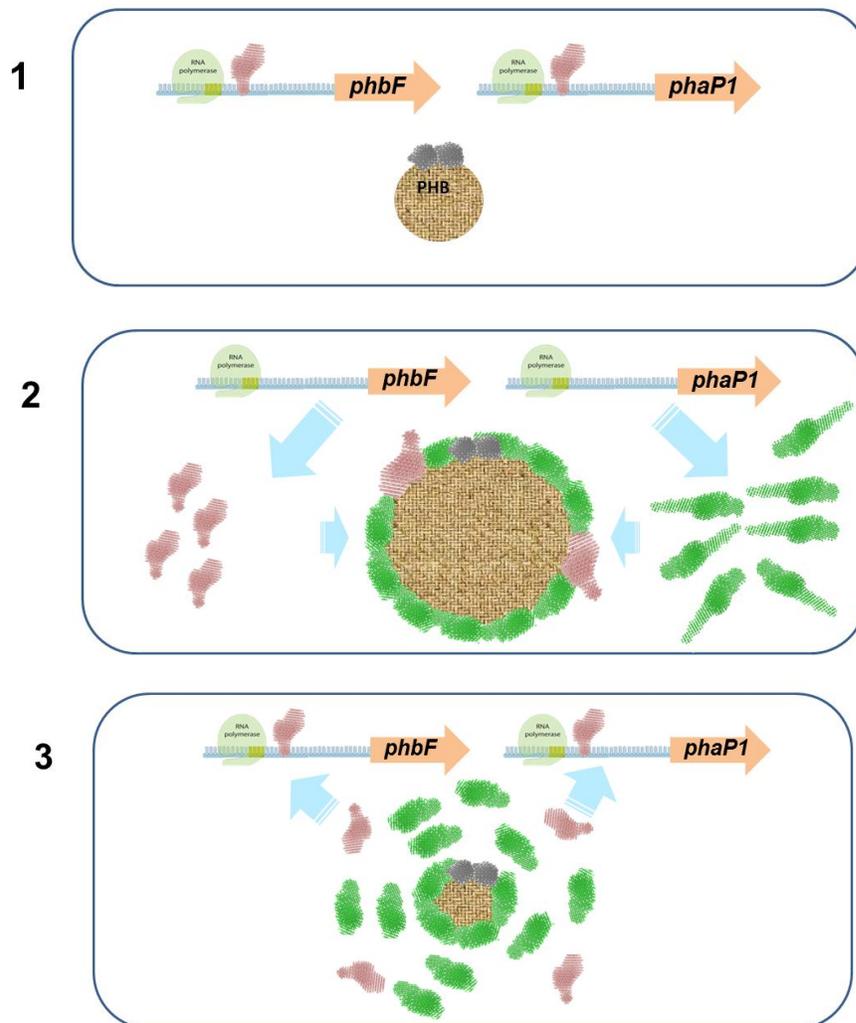
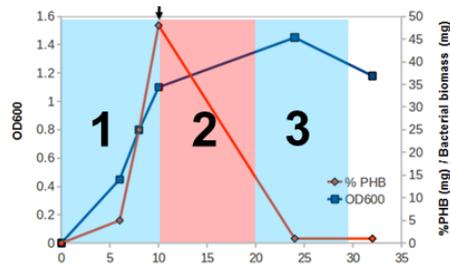
A proteína PhaP1, assim como PhbF, é composta somente por  $\alpha$ -hélices presentes possivelmente na forma de três grandes hélices, uma característica de proteínas com perfil fibrilar e como observado no modelo estrutural *ab initio*. O modelo ainda descreve o monômero de PhaP1 como sendo formado por uma face com distribuição de cargas e uma face hidrofóbica. Esta descrição é similar a de Hanley e colaboradores (1999) que se basearam apenas na sequência de aminoácidos e não em um modelo estrutural. O perfil fibrilar também é sugerido pelo envelope molecular obtido com dados de SAXS que descreve PhaP1 como uma proteína de dimensão máxima de 200 Å e raio de giro de 50,94 Å. No modelo obtido por SAXS aparecem um módulo mais volumoso com 113 Å de comprimento e uma extensão menos volumosa com 87 Å de comprimento. O primeiro módulo provavelmente é a região mais hidrofóbica uma vez que este maior volume pode estar associado à interação de moléculas de TritonX-100, responsáveis pela solubilização da proteína. Por outro lado, o segundo módulo sugere a presença de uma região desordenada que possa ser importante para a interação da proteína ao grânulo de PHB. Este modelo possui certa discrepância em relação ao modelo estrutural *ab initio* cuja dimensão máxima é de aproximadamente 101 Å. Esta observação sugere que a conformação monomérica seja diferenciada da proposta quando a proteína encontra-se na estrutura quaternária pentamérica em solução. Estruturas com dimensão máxima de 189 Å foram observadas quando PhaP1 foi adsorvida em um filme de PHB. Esta dimensão é compatível com a calculada com os dados de SAXS, contudo não permite concluir que a conformação da proteína em solução e associada ao polímero seja a mesma.

As análises estruturais realizadas durante este trabalho permitiram descrever um retrato inicial da estrutura das proteínas PhbF e PhaP1 e também levantar questões importantes a serem consideradas na continuidade dos estudos.

O papel da proteína PhbF assim como a sua influência sobre a expressão da proteína PhaP1 foi avaliada *in vivo* englobando-se os aspectos estruturais descritos acima. A proteína PhbF apresentou a atividade de repressão de sua própria transcrição assim como da transcrição do gene *phaP1* sugerindo um papel central no processo de regulação. Este mecanismo de regulação foi sugerido pelos ensaios de  $\beta$ -galactosidase realizados em *E. coli*. O sistema experimental permitiu observar de forma isolada o papel de PhbF sem a influência do polímero PHB sobre a sua função. Em *H. seropedicae* foi possível sugerir o acoplamento entre a transcrição dos genes *phbF* e *phaP1* ao metabolismo de PHB. A migração das proteínas do citoplasma aos grânulos de PHB, sugeridos pelos ensaios de imunodeteção, permitem ligar este fenômeno a atividade de ligação ao polímero observada *in vitro*. Por outro lado a mobilidade da proteína PhbF entre o citoplasma, onde desempenha o seu papel de repressor da transcrição, e o grânulo durante as fases de produção e degradação do polímero ilustram as atividades de ligação tanto ao DNA como ao PHB.

Os resultados descritos acima permitiram propor um modelo de regulação envolvendo a proteína PhbF em *H. seropedicae* SmR1 (Figura 41) e compartilhado com o modelo proposto por Pötter e colaboradores (2002) para *R. eutropha*. Neste modelo o crescimento de *H. seropedicae* pode ser dividido em três etapas. Na primeira etapa a bactéria encontra-se na fase exponencial de crescimento onde a proteína PhbF reprime a sua transcrição assim como a transcrição do gene *phaP1* (Figura 41-1). Na segunda etapa, a bactéria entra na fase estacionária de crescimento onde se inicia a síntese do polímero causando o aumento no tamanho e número de grânulos. A proteína PhbF interage com o polímero resultando na desrepressão dos genes *phbF* e *phaP1*. Nesta fase a proteína PhaP1 passa a ser necessária para evitar a coalescência dos grânulos nascentes e as proteínas expressas ligam-se ao grânulo à medida que o mesmo aumenta de tamanho (Figura 41-2). Na terceira etapa, as concentrações extracelulares da fonte de carbono diminuem e a bactéria passa a degradar o polímero PHB resultando na diminuição do tamanho dos grânulos e dissociação das proteínas PhbF e PhaP1. A migração da proteína PhbF do grânulo ao citoplasma resulta novamente na repressão da transcrição dos genes *phbF* e *phaP1* (Figura 41-3).

**FIGURA 41 – Modelo de regulação desempenhada pela proteína PhbF em *H. seropedicae* SmR1**



A primeira figura representa a curva de crescimento de *H. seropedicae* e os números indicam esquematicamente as três diferentes etapas na regulação de PhbF. 1: Fase exponencial, 2: Início da fase estacionária e 3: Fase estacionária tardia. As estruturas em verde indicam a proteína PhaP1 e em vermelho a proteína PhbF. As setas azuis indicam o sentido de mobilidade das proteínas.

## 1.6 CONCLUSÕES

- As proteínas PhbF e PhaP1 foram expressas e purificadas com grau de pureza superior a 93% sendo classificadas quanto a forma oligomérica, respectivamente, como tetrâmero e pentâmero.
- As proteínas foram capazes de ligar-se *in vitro* ao polímero PHB extraído de *H. seropedicae* integrando o papel destas dentro do metabolismo de PHA.
- A proteína PhaP1 formou estruturas com dimensão de  $177,8 \pm 11,7 \text{ \AA}$  sobre filmes formados por PHB mostrando que a proteína associa-se de forma ordenada sobre o mesmo.
- A proteína PhbF liga-se a onze regiões promotoras de genes relacionados ao metabolismo de PHA mostrando o seu papel regulatório nesta via.
- O motivo consenso  $5\text{'-TG[N]TGC[N]}_3\text{GCAA-3'}$  é a sequência reconhecida por PhbF durante a ligação ao DNA.
- PhbF é uma proteína repressora da expressão dos genes *phbF* e *phaP1* em *E. coli* ET8000.
- A expressão dos genes *phbF* e *phaP1* é acoplada à síntese de PHB em *H. seropedicae*.
- O modelo a baixa resolução de PhbF permitiu determinar as suas dimensões em solução e observar a presença de dois domínios: um domínio de ligação ao DNA e outro domínio de interação ao PHB.
- O modelo a baixa resolução de PhaP1 permitiu determinar as suas dimensões em solução e classificá-la como uma proteína alongada e parcialmente desordenada nas condições testadas.

**CAPÍTULO 2: DETERMINAÇÃO DA ESTRUTURA  
DA PROTEÍNA HFQ DE *Herbaspirillum*  
*seropedicae* SmR1 POR DIFRAÇÃO DE RAIOS-X**

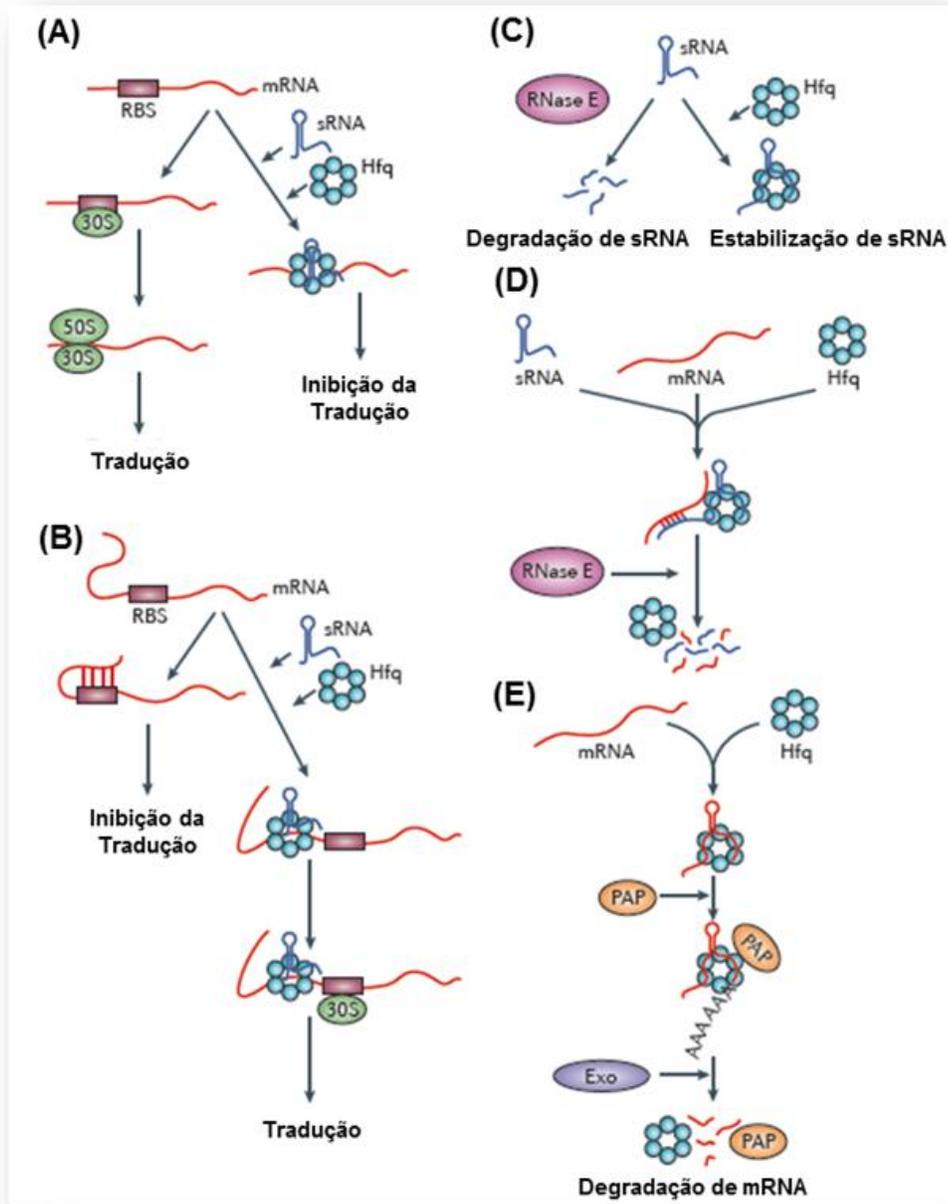
## 2.1 INTRODUÇÃO

Hfq é uma chaperona de RNA envolvida na regulação pós-transcricional em várias espécies bacterianas. A inativação do gene *hfq* em *E. coli* resultou em uma variedade de fenótipos, indicando que esta proteína é importante no controle da expressão em vários aspectos no metabolismo bacteriano (TSUI *et al.*, 1994; WACHI *et al.*, 1999; SONNLEITNER *et al.*, 2003; ANSONG *et al.*, 2009). Hfq também é descrita como um importante fator de virulência em bactérias patogênicas como *Brucella abortus*, *Legionella pneumophila*, *Pseudomonas aeruginosa*, *Vibrio cholerae* e *Yersinia enterocolitica* (ROBERTSON *et al.*, 1999; MCNEALY *et al.*, 2005; DING *et al.*, 2004; NAKAO *et al.*, 2005) e parece participar de um tipo de mecanismo de silenciamento mediado por pequenos RNAs (sRNA ou *small RNAs*) em procariotos (AIBA, 2007). sRNA são RNAs não codificadores, que contém entre 50-300 nucleotídeos, envolvidos na regulação pós-transcricional de várias funções celulares (STORZ *et al.*, 2011). Hfq pode regular a expressão gênica através de vários mecanismos: a) favorecendo a interação entre o sRNA regulatório e seu mRNA alvo, promovendo a estabilidade do mRNA (Figura 42A) ou a degradação dependendo do alvo (Figura 42D); b) mudando a acessibilidade do sRNA alvo a RNases (Figura 42C); c) promovendo o pareamento de bases entre o sRNA regulatório e o mRNA alvo, levando a repressão ou ativação da tradução (Figura 42B,A); d) estimulando a poliadenilação de mRNA desencadeando a degradação (Figura 42E) (STORZ *et al.*, 2004; VOGEL & LUISI, 2011).

Esta proteína é amplamente distribuída e pertence à família de proteínas chamada Sm e Sm-like, cujos membros são capazes de se ligar a RNA e tem uma conformação característica composta de uma  $\alpha$ -hélice N-terminal seguida de cinco fitas  $\beta$  (WILUSZ & WILUSZ, 2005) (Figura 43). As estruturas quaternárias possuem a forma de um anel cujas faces são denominadas proximal, onde se encontra a  $\alpha$ -hélice, e distal (face oposta) (Figura 43) (BRENNAN *et al.*, 2007). Dois motivos altamente conservados (Sm1 e Sm2) são compartilhados com proteínas de processamento do RNA em humanos (HERMANN *et al.*, 1995). As estruturas cristalográficas de vinte proteínas Hfq de nove organismos diferentes já foram resolvidas por cristalografia de raios-X e depositadas no banco de dados PDB. As características das estruturas destes

microrganismos, consideradas relevantes para este trabalho, estão apresentadas no quadro 7 e a estrutura representativa da Hfq de *E. coli* está mostrada na figura 43.

**FIGURA 42 – Mecanismos de ação da proteína Hfq**



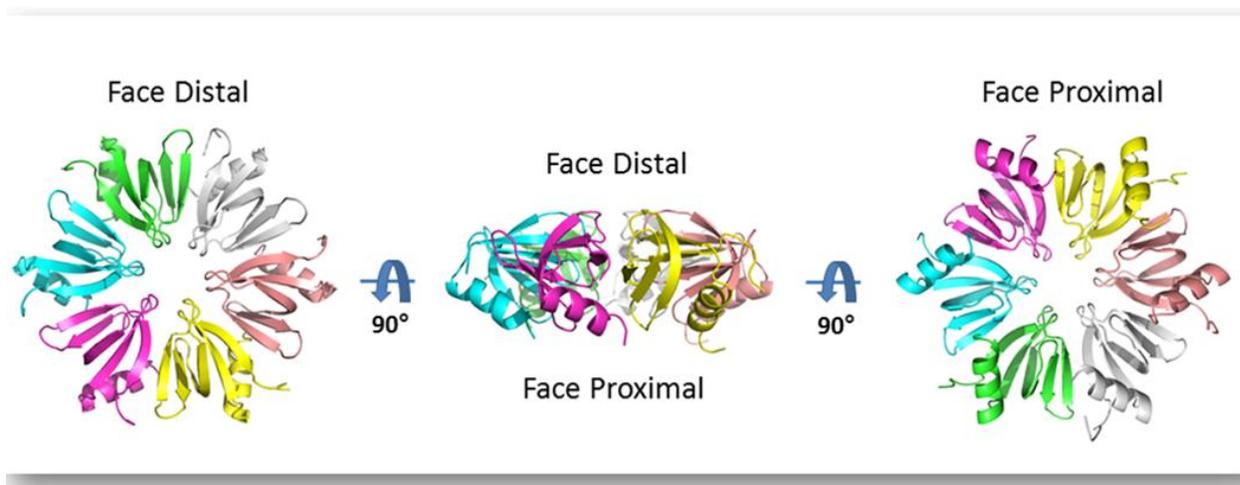
Modificado de VOGEL & LUISI (2011). (A-E) Representação esquemática dos diferentes mecanismos de ação da proteína Hfq relacionados à função de chaperona de RNA.

**QUADRO 7 – Proteínas Hfq com estrutura cristalográfica depositada no PDB**

| Código PDB | Organismo                     | Resolução (Å) | Nº de resíduos do monômero | Oligômero | Característica                             | Identidade com HsHfq (%) |
|------------|-------------------------------|---------------|----------------------------|-----------|--|--------------------------|
| * 3SB2     | <i>H. seropedicae</i>         | 2,63          | 79                         | 6         | Nativa                                     | 100                      |
| 3INZ       | <i>P. aeruginosa</i>          | 1,70          | 82                         | 6         | Mutação H57T                               | 77                       |
| 3M4G       | <i>P. aeruginosa</i>          | 2,05          | 82                         | 6         | Mutação H57A                               | 77                       |
| 3AHU       | <i>B. subtilis</i>            | 2,20          | 78                         | 6         | Complexo com o RNA (AG) <sub>3</sub>       | 49                       |
| 3HSB       | <i>B. subtilis</i>            | 2,20          | 78                         | 6         | Complexo com o RNA (AG) <sub>3</sub> A     | 49                       |
| 1U1S       | <i>P. aeruginosa</i>          | 1,60          | 82                         | 6         | Baixa concentração de sal                  | 78                       |
| 1U1T       | <i>P. aeruginosa</i>          | 1,90          | 82                         | 6         | Alta concentração de sal                   | 77                       |
| 3RES       | <i>E. coli</i>                | 2,00          | 102                        | 6         | Complexo com ADP                           | 82                       |
| 2YLC       | <i>S. typhimurium</i>         | 1,30          | 74                         | 6         | Complexo com o RNA U <sub>6</sub>          | 80                       |
| 2YLB       | <i>S. typhimurium</i>         | 1,15          | 74                         | 6         | Nativa                                     | 80                       |
| 3RER       | <i>E. coli</i>                | 1,70          | 102                        | 6         | Complexo com o RNA AU <sub>6</sub> A e ADP | 82                       |
| 1KQ2       | <i>S. aureus</i>              | 2,71          | 77                         | 6         | Complexo com o RNA AU <sub>5</sub> G       | 35                       |
| 3GIB       | <i>E. coli</i>                | 2,40          | 102                        | 6         | Complexo com o RNA poli(A)                 | 82                       |
| 3QHS       | <i>E. coli</i>                | 2,85          | 102                        | 6         | Nativa                                     | 82                       |
| 1KQ1       | <i>S. aureus</i>              | 1,55          | 77                         | 6         | Nativa                                     | 35                       |
| 1HK9       | <i>E. coli</i>                | 2,15          | 102                        | 6         | Resíduos de 73-102 deletados               | 82                       |
| 2YHT       | <i>E. coli</i>                | 2,90          | 102                        | 6         | Nativa                                     | 82                       |
| 3HFO       | <i>Synechocystis sp.</i>      | 1,30          | 70                         | 6         | Nativa                                     | Não<br>significante      |
| 3HFN       | <i>Anabaena sp.</i>           | 2,31          | 72                         | 6         | Nativa                                     | Não<br>significante      |
| 2QTX       | <i>M. jannaschii</i>          | 2,50          | 71                         | 6         | Nativa                                     | 30                       |
| 1MGQ       | <i>M. thermoautotrophicum</i> | 1,70          | 83                         | 7         | Nativa                                     | Não<br>significante      |

Informações coletadas em 28/11/2011. HsHfq: Hfq de *H. seropedicae* SmR1. Nativa: proteína Hfq sem modificações. \* Este trabalho.

**FIGURA 43 – Estrutura representativa da proteína Hfq de *E. coli***



Estrutura 1HK9 de *E.coli*. Cada monômero está representado por uma cor diferente.

Todas as estruturas são homohexaméricas com exceção da proteína Hfq de *Methanotermobacter thermoautotrophicum* (Quadro 7) que cristalizou como um homoheptâmero. Destas estruturas, seis (1KQ2 de *Staphylococcus aureus*, 3GIB e 3RER de *Escherichia coli*, 3HSB e 3AHU de *Bacillus subtilis*, e 2YLC de *Salmonella typhimurium*) estão complexadas com RNA. Através destes estudos estruturais os autores procuraram principalmente, elucidar os mecanismos de interação entre proteína e RNA. O primeiro complexo foi obtido por co-cristalização da proteína Hfq de *S. aureus* com a sequência AU<sub>5</sub>G, rica em uracilas. Este modelo mostrou a ligação deste RNA à região proximal da proteína e destacou o papel do poro central na ligação ao sRNA (SCHUMACHER *et al.*, 2002). Além disso, em *E. coli*, o complexo proteína-RNA mostrou que a face distal do homohexâmero da Hfq está envolvida na ligação a um RNA poli(A), que estaria representando uma fração de um tipo de RNA alvo (LINK *et al.*, 2009). Este último modelo de ligação está associado ao aumento da estabilidade do RNA, diferente do RNA rico em uridinas ligado à proteína Hfq de *S. aureus* (SCHUMACHER *et al.*, 2002). Estas duas estruturas mostraram que Hfq pode distinguir RNAs diferentes. O complexo Hfq-U<sub>6</sub> de *Salmonella thyphimurium* (2YLC) também revelou que a proteína reconhece preferencialmente RNAs ricos em uridinas na porção 3' (SAUER & WEICHENRIEDER, 2011). Já o complexo Hfq65-AU<sub>6</sub>A-ADP de *E. coli*

mostrou uma ligação diferencial da sequência AU<sub>6</sub>A, que corresponde a um segmento interno do sRNA DsrA. Neste caso, o RNA AU<sub>6</sub>A interage simultaneamente com a porção proximal e distal da proteína Hfq (WANG *et al.*, 2011).

As sequências de aminoácidos das proteínas Hfq da *archaea* termofílica *Methanococcus janaschii* (2QTX) e das cianobactérias *Synechocystis sp.* PCC 6803 (3HFO) e *Anabaena PCC 7120* (3HFN) são as menos conservadas quando comparadas às demais Hfqs. A estrutura de *M. janaschii* tem as características mais particulares:  $\alpha$ -hélice mais curta, diâmetro menor, face distal mais carregada negativamente e o *loop* L4 e a região C-terminal com estruturas distintas (NIELSEN *et al.*, 2007). Apesar destas diferenças, a conformação do tipo Sm e a função protéica são conservados. As proteínas Hfq das cianobactérias *Synechocystis sp.* PCC 6803 e *Anabaena PCC 7120* ligam-se fracamente aos sRNAs de *E. coli*, cuja ligação a Hfq já é bem descrita, presumivelmente devido a um sítio diferenciado de ligação ao RNA e a diferenças de carga nos sítios de ligação (BØGGILD *et al.*, 2009). Comparações entre a estrutura cristalográfica da proteína Hfq selvagem e mutante no motivo conservado YKHAI (3INZ e 3M4G) de *P. aeruginosa* (MOSKALEVA *et al.*, 2010) mostraram a importância de ligações de hidrogênio específicas nas regiões intermonômero. Outra característica estrutural amplamente discutida é a região C-terminal não conservada. Esta região flexível tem sido relatada como essencial para o mecanismo de riborregulação (OLSEN *et al.*, 2010). A importância desta região na interação com RNAs mais longos, assim como a primeira descrição estrutural sobre sua desordem em solução, foi reportada utilizando a técnica de SAXS (BEICH-FRANDSEN *et al.*, 2011).

A proteína Hfq de *H. seropedicae* SmR1 começou a ser estudada durante o curso de mestrado de KADOWAKI (2008) em função de seu potencial papel na fixação biológica de nitrogênio por meio de mecanismos de regulação pós-transcricional. O possível envolvimento desta proteína no metabolismo de nitrogênio foi sugerido em função da sua influência sobre a expressão do ativador transcricional NifA de *Azorhizobium caulinodans* (KAMINSKI *et al.*, 1994), *Rhodobacter capsulatus* (DREPPER *et al.*, 2002) e *Rhizobium leguminosarum* (ZHANG & HONG, 2009). Neste trabalho, a proteína Hfq de *H. seropedicae* foi clonada em vetor de expressão, purificada e caracterizada *in vitro* quanto a sua ligação ao sRNA DsrA de *E. coli*. A

caracterização estrutural começou a ser realizada através das técnicas de dicroísmo circular, gel filtração e espalhamento dinâmico de luz (DLS). A cristalização e coleta dos dados de difração de raios-X foram realizadas como etapa experimental para a resolução da estrutura tridimensional de Hfq. Durante o Doutorado deu-se continuidade aos estudos de caracterização estrutural de Hfq de *H. seropedicae* SmR1 por meio da resolução da estrutura tridimensional a nível atômico e a análise do envelope molecular em solução por meio da técnica de difração de raios-X a baixo ângulo (SAXS). Nesta etapa do trabalho houve a intensa colaboração do professor Dr. Jorge Iulek da Universidade Estadual de Ponta Grossa.

## **2.2 OBJETIVO**

### **2.2.1 Objetivo geral**

Caracterizar a estrutura cristalográfica e em solução da proteína Hfq de *H. seropedicae*.

### **2.2.2 Objetivos específicos**

- Resolver a estrutura cristalográfica da proteína Hfq de *H. seropedicae*.
- Realizar uma análise comparativa com as estruturas de proteínas já depositadas no banco de dados PDB.
- Obter informações estruturais em solução da proteína Hfq por SAXS e correlacioná-la a estrutura cristalográfica.

## 2.3. MATERIAIS E MÉTODOS

### 2.3.1 Cristalização, coleta de dados de difração de raios-X e análise estrutural da proteína Hfq de *H. seropedicae* SmR1

A proteína Hfq de *H. seropedicae* foi expressa em *E. coli* BL21 (DE3) e purificada na sua forma não fusionada em tampão contendo 50 mmol/L de Tris-HCl pH 8 e 100 mmol/L de NaCl como descrito por KADOWAKI (2008). O cristal da proteína formou-se em poucos dias na condição de cristalização número 16 do kit PACT (Qiagen) contendo 0,1 M PCB (propionato de sódio, cacodilato de sódio e Bis-Tris propano) pH 7,0 e 25% (m/v) PEG 1500. À gota contendo o cristal foi adicionado glicerol a uma concentração final de 5% (v/v) com o objetivo de crioproteção. Em seguida o cristal foi montado em um *loop* e acoplado a cabeça goniométrica sob um fluxo de gás nitrogênio a 100K. Os dados de difração de raios-X foram coletados em um difratômetro modelo Rigaku UltraX-18HF equipado com um detector Mar345 utilizando 2 minutos de tempo de exposição por imagem no Laboratório Nacional de Luz Síncrotron (LNLS). Trezentas e quatro imagens com ângulo de oscilação de 1° foram integradas, escalonadas e as reflexões promediadas com o programa XDS (KABSCH, 2010). As fases iniciais foram calculadas utilizando o programa Phaser (MCCOY *et al.*, 2007) e utilizando como modelo de busca a estrutura do hexâmero da proteína Hfq de *E. coli* (PDB code 1HK9) editada com o programa Chainsaw (STAIN, 2008). O modelo foi ajustado manualmente com o programa Coot (EMSLEY & COWTAN, 2004) e submetido a refinamento por vários ciclos pelo programa PHENIX.REFINE (ADAMS *et al.*, 2010). A qualidade do modelo foi avaliada pelo programa PROCHECK (LASKOWSKI *et al.*, 1993) e no servidor molprobity (DAVIS *et al.*, 2004). Os valores de RSCC e RSR foram calculados pelo programa Mapman (KLEYWEGT & JONES, 1996). Os fatores de estrutura e as coordenadas atômicas foram depositados no banco de dados PDB com o código 3SB2. As figuras dos modelos estruturais foram montadas com o programa Pymol (SCHRODINGER, 2010) e o potencial eletrostático foi calculado utilizando-se o programa DelPhi (ROCCHIA *et al.*, 2001).

### 2.3.2 Espalhamento de raios-X a baixo ângulo e construção do modelo a baixa resolução

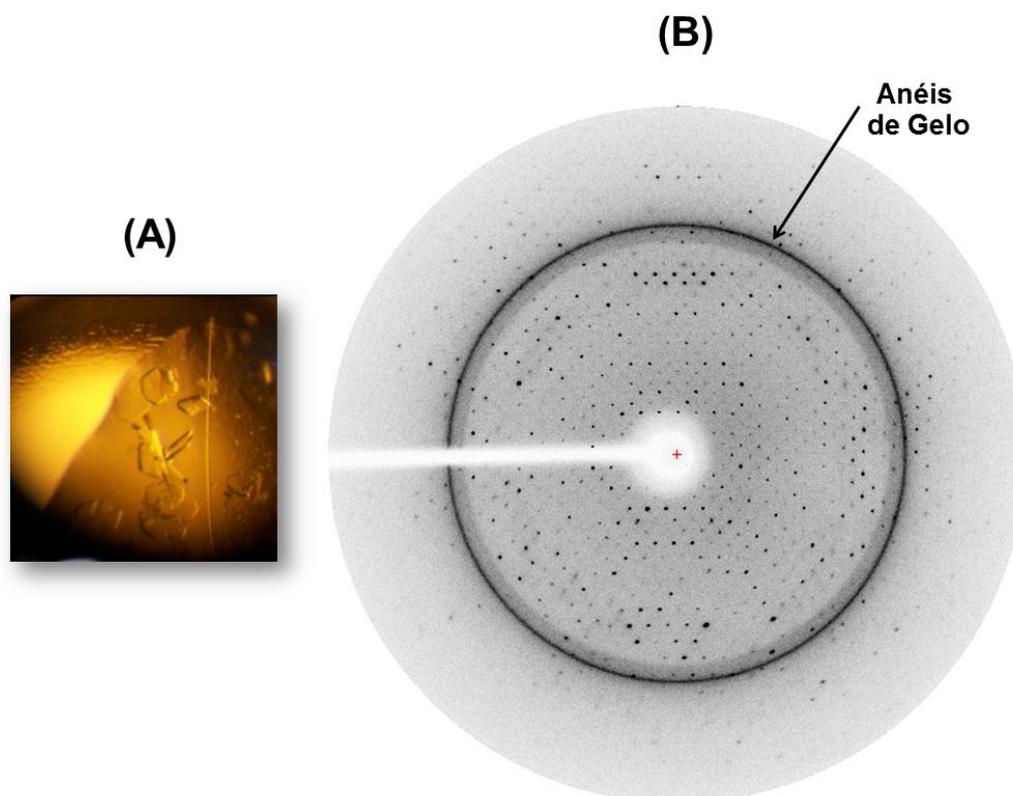
Os dados de SAXS da proteína Hfq foram coletados utilizando a linha de luz SAXS2 (comprimento de onda 1,488 Å) no LNLS com o detector 2D (MAR-165). Os experimentos foram realizados a 25 °C com uma distância entre o detector e a amostra de 1534,5 mm. A solução protéica foi submetida a um tempo de exposição de 10 minutos correspondente a 5 *frames* de 2 minutos cada. Duas concentrações protéicas foram testadas, 0,90 e 4,35 mg/mL. As medidas da menor concentração foram escolhidas e utilizadas uma vez que apresentaram uma boa relação sinal-ruído e não foi observado padrão de agregação. Os dados foram corrigidos quanto às variações na intensidade dos raios-X e subtraídos do espalhamento do tampão e da radiação de fundo utilizando o programa FIT2D (HAMMERSLEY, 1997). A função de distribuição dos pares de distâncias e o raio de giro foram calculados com o programa GNOM (SVERGUN, 1992) utilizando um diâmetro máximo de 80 Å. A massa molecular experimental foi determinada com o programa SAXS MoW (FISCHER *et al.*, 2010). Vinte envelopes moleculares a baixa resolução foram gerados com o programa GASBOR (SVERGUN *et al.*, 2001) impondo simetria do tipo P6 e o modelo médio foi calculado com o programa DAMAVER (VOLKO *et al.*, 2003). A sobreposição dos modelos cristalográficos e a baixa resolução foi realizada pelo programa SUPERCOMB (KOZIN & SVERGUN, 2001). A distância máxima e o raio de giro dos modelos foram calculados com o programa CRY SOL 2.6 (SVERGUN *et al.*, 1995).

## 2.4. RESULTADOS

### 2.4.1 Processamento das imagens e indexação

Trezentas e quatro imagens foram coletadas para o cristal (Figura 44A) obtido na condição contendo 0,1 M PCB (propionato de sódio, cacodilato de sódio e Bis-Tris propano) pH 7,0 e 25% (m/v) PEG 1500. A primeira avaliação das imagens referentes aos padrões de difração de raios-X mostrou a presença de anéis devidos à difração por gelo (Figura 44B) e quatro imagens com falhas na forma de faixas brancas cruzando o padrão de difração. Para avaliar o efeito dos fatores citados nas estatísticas dos dados, o processamento foi realizado incluindo ou excluindo as imagens com falha, assim como as reflexões que se encontravam na região dos anéis devido ao gelo. O resumo dos processamentos está mostrado na tabela 3.

**FIGURA 44 – Avaliação do padrão de difração de raios-X da proteína Hfq**



(A) Cristal utilizado na coleta de dados. (B) Padrão de difração de raios-X mostrando a presença de anéis de gelo.

**TABELA 3 – Resumo dos dados estatísticos de processamento de imagens**

|                            | 1          | 2          | 3          | 4          | 5          | 6          |
|----------------------------|------------|------------|------------|------------|------------|------------|
| <b>N° reflexões únicas</b> | 24788      | 24845      | 22167      | 22202      | 24253      | 24379      |
| <b>Rsymm (%)</b>           | 4,3(28,5)  | 3,8(27,5)  | 4,0(29,3)  | 3,6(28,4)  | 4,0(29,9)  | 3,6(27,7)  |
| <b>I/σ(I)</b>              | 21,5(4,7)  | 23,0(4,9)  | 22,1(4,8)  | 23,7(4,7)  | 22,6(4,7)  | 24,7(4,8)  |
| <b>Completeza (%)</b>      | 90,6(62,4) | 90,9(63,3) | 81,1(63,8) | 81,0(60,9) | 88,9(65,3) | 89,2(63,7) |

(1) com anéis de gelo e todas as imagens; (2) com anéis de gelo e sem 4 imagens com falhas; (3) sem anéis de gelo (3,91-3,51) e com todas as imagens; (4) sem anéis de gelo (3,91-3,51) e sem 4 imagens com falhas; (5) sem anéis de gelo (3,70-3,62) e com todas as imagens; (6) sem anéis de gelo (3,70-3,62) e sem 4 imagens com falhas. Valores entre parênteses representam os parâmetros dos dados de maior resolução (2,6 Å).

Com base nestes resultados decidiu-se utilizar os dados do processamento que levou em consideração as reflexões presentes nas regiões de difração por gelo e excluindo as imagens que apresentavam falha. Esta decisão foi tomada em função da melhor completeza e fator R deste processamento quando comparado aos demais.

A auto-indexação foi realizada com o programa XDS que sugeriu os grupos de ponto 2 e 222 como os de maior simetria para o processamento (Tabela 4).

**TABELA 4 – Grupos espaciais sugeridos para o processamento dos dados**

|           | Rede Bravais | Qualidade do ajuste | a    | b    | c    | $\alpha$ | $\beta$ | $\gamma$ | Grupos espaciais |
|-----------|--------------|---------------------|------|------|------|----------|---------|----------|------------------|
| <b>44</b> | aP           | 0,0                 | 59,2 | 60,5 | 64,8 | 90,0     | 90,0    | 105,5    | P1               |
| <b>31</b> | aP           | 0,1                 | 59,2 | 60,5 | 64,8 | 90,0     | 90,0    | 74,5     | P1               |
| <b>34</b> | mP           | 0,5                 | 59,2 | 64,8 | 60,5 | 90,0     | 105,5   | 90,0     | P2, P2(1)        |
| <b>14</b> | mC           | 22,2                | 72,4 | 95,2 | 64,8 | 90,0     | 90,0    | 88,7     | C2               |
| <b>13</b> | oC           | 22,4                | 72,4 | 95,2 | 64,8 | 90,0     | 90,0    | 88,7     | C222, C222(1)    |

As dimensões da unidade de cela a, b e c são expressas em angstroms e os ângulos  $\alpha$ ,  $\beta$  e  $\gamma$  em graus.

Os dados foram inicialmente indexados e promediados nos grupos espaciais P2 e C222 que apresentaram  $R_{\text{symm}}$  global de 4,9 e 51,3 % respectivamente. Estas estatísticas indicaram em um primeiro momento o grupo pontual 2 como mais provável, mesmo após várias tentativas realizadas para a solução no grupo pontual 222. Ausências sistemáticas indicaram um eixo de ordem 2 helicoidal. Por fim, o grupo espacial  $P2_1$  mostrou-se ser o correto durante os refinamentos, pois a estrutura hexamérica não é composta por dois trímeros iguais, o que justificaria uma simetria superior.

#### **2.4.2 Resolução da estrutura cristalográfica, refinamento e validação**

A resolução da estrutura cristalográfica de proteínas não depende somente dos dados extraídos das imagens referentes aos padrões de difração. É necessário também se ter conhecimento do parâmetro conhecido como “fases” para, juntamente com as amplitudes dos fatores de estrutura obtidos a partir dos dados de difração, determinar a posição dos átomos que compõem a proteína (DRENTH, 1993). Dentre os métodos existentes foi escolhido o de Substituição Molecular que se baseia na similaridade previamente conhecida entre a proteína em estudo e outra com estrutura tridimensional resolvida.

A estimativa do conjunto inicial de fases para os fatores de estrutura da proteína Hfq de *H. seropedicae* foi realizada utilizando como molde a estrutura da Hfq de *E. coli* (1HK9), uma vez que estas proteínas compartilham 82% de identidade (Quadro 7). A substituição molecular e alguns ciclos de refinamento confirmaram o grupo espacial como  $P2_1$  com um hexâmero por unidade assimétrica (conteúdo de solvente de 43,3 %). A resolução dos dados utilizados foi de 2,6 Å convergindo para um R-work e um R-free de 0,20 e 0,28, respectivamente. As estatísticas de processamento de dados e de refinamento estão mostradas na tabela 5. Os valores de R-work e de R-free correspondem a índices globais da concordância entre o modelo estrutural e os dados experimentais e com o objetivo de analisar esta concordância de forma local (por resíduo) foram utilizados os índices RSCC (*Real Space Correlation Factor*) e RSR (*Real Space Factor*). Estes índices permitem quantificar o ajuste de cada resíduo na densidade eletrônica. Na figura 45 observa-se que de forma geral as regiões dos

resíduos 1-10, 48-52 e 65-69 possuem valores baixos para RSCC e altos para RSR que indicam uma grande diferença entre o mapa de densidade observado e o calculado. Estas regiões correspondem ao N-terminal, a volta entre as fitas  $\beta$ 3 e 4 e C-terminal concordando com a alta flexibilidade já esperada para estas regiões.

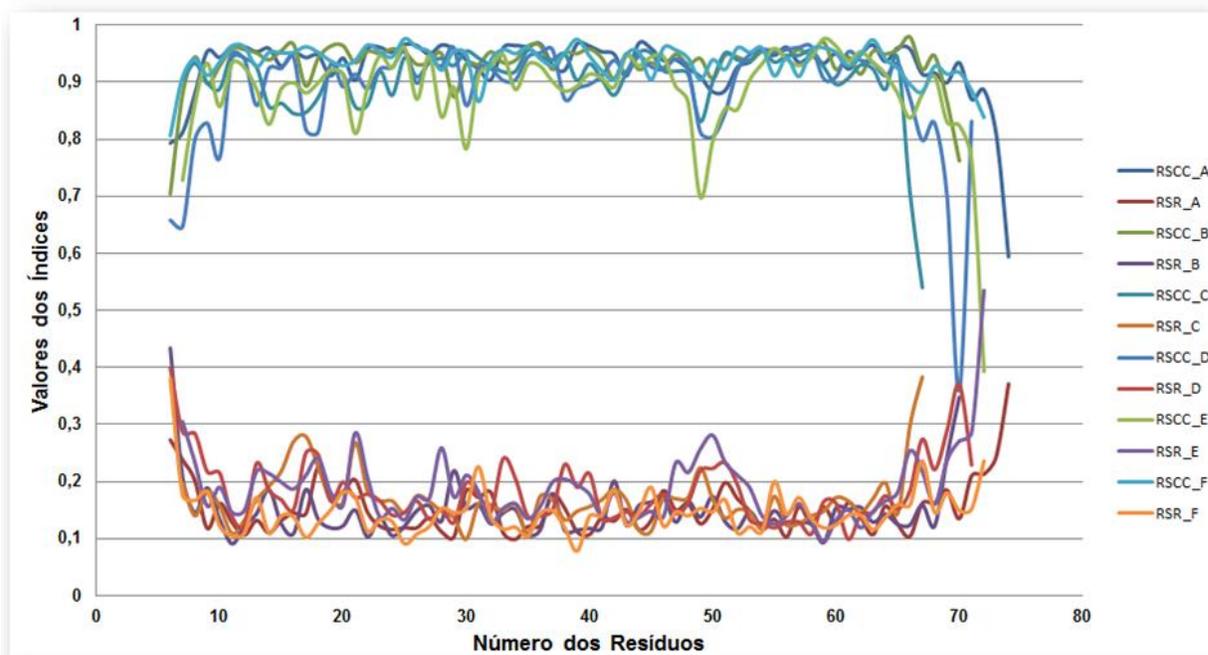
A qualidade estereoquímica do modelo cristalográfico foi avaliada utilizando-se o programa Procheck que gerou também o diagrama de Ramachandran (Figura 46). Na estrutura depositada, 98,3% dos resíduos de aminoácidos encontram-se em regiões permitidas atestando a qualidade do modelo. Uma característica que chama a atenção são os resíduos de asparagina 49 de todas as cadeias que se encontram em regiões generosamente permitidas e não permitidas. Esta conformação não usual será discutida abaixo.

**TABELA 5 – Estatísticas de coleta de dados e refinamento da estrutura cristalográfica**

| <b>Código de depósito no PDB</b>                 | <b>3SB2</b>  |
|--|--|
| <b>Estatísticas de coleta de dados</b>           |  |
| Grupo espacial                                   | P2 <sub>1</sub>  |
| Parâmetros de cela unitária                      | a = 59,64, b = 65,32, c = 60,99 Å, $\beta = 105,5^\circ$ |
| Conteúdo da unidade assimétrica                  | Um homohexâmero  |
| Comprimento de onda (Å)                          | 1,54 (anôdo rotatório)                                   |
| Resolução (Å)                                    | 66,00 – 2,63 (2,70 – 2,63)                               |
| Nº. de reflexões únicas                          | 13212 (977)  |
| Redundância                                      | 6,02 (5,78)  |
| Completeza (%)                                   | 97,2 (94,9)  |
| I/ $\sigma$ (I)                                  | 29,21 (6,52)   |
| R <sub>sym</sub> (%)                             | 4,5 (30,5)   |
| R-meas(%)  | 4,9 (33,5)   |
| Mosaicidade (°)                                  | 0,402  |
| Fator B pelo gráfico de Wilson (Å <sup>2</sup> ) | 59,9   |
| <b>Estatísticas de refinamento</b>               |  |
| Resolução (Å)                                    | 23,47 – 2,63   |
| R <sub>work</sub> (%)                            | 0,20   |
| R <sub>free</sub> (%)                            | 0,28   |
| RMSD – distância de ligação (Å)                  | 0,002  |
| RMSD – ângulo de ligação (°)                     | 0,541  |
| Átomos da proteína                               | 3024   |
| Moléculas de água                                | 210  |
| Fator B médio (proteína) (Å <sup>2</sup> )       | 57,1   |
| Fator B médio (águas) (Å <sup>2</sup> )          | 48,6   |
| <b>Regiões de Ramachandran (%)</b>               |  |
| Mais favorável                                   | 85,5   |
| Adicionalmente permitida                         | 12,8   |
| Generosamente permitida                          | 1,4  |
| Não permitidas                                   | 0,3  |

As estatísticas de Ramachandran foram calculadas com o programa PROCHECK. Valores entre parênteses indicam as estatísticas dos dados da faixa de maior resolução.

**FIGURA 45 – Qualidade do ajuste da estrutura da proteína Hfq na densidade eletrônica**



Os índices RSCC (*Real Space Correlation Factor*) estão registrados na parte superior do gráfico e os índices RSR (*Real Space Factor*) na parte inferior em função do resíduo de aminoácido. As cores dos índices correspondentes a cada cadeia estão indicadas no lado direito do gráfico. Os valores foram calculados pelo programa Mapman.

**FIGURA 46 – Qualidade estereoquímica da estrutura cristalográfica de Hfq**

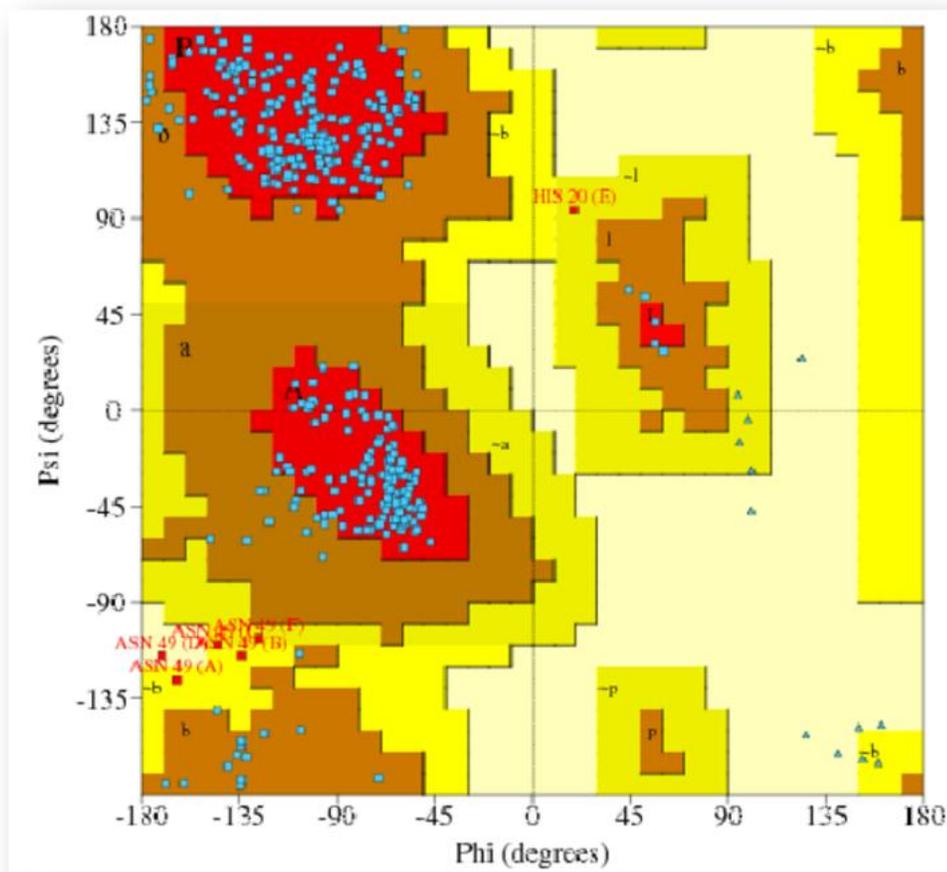
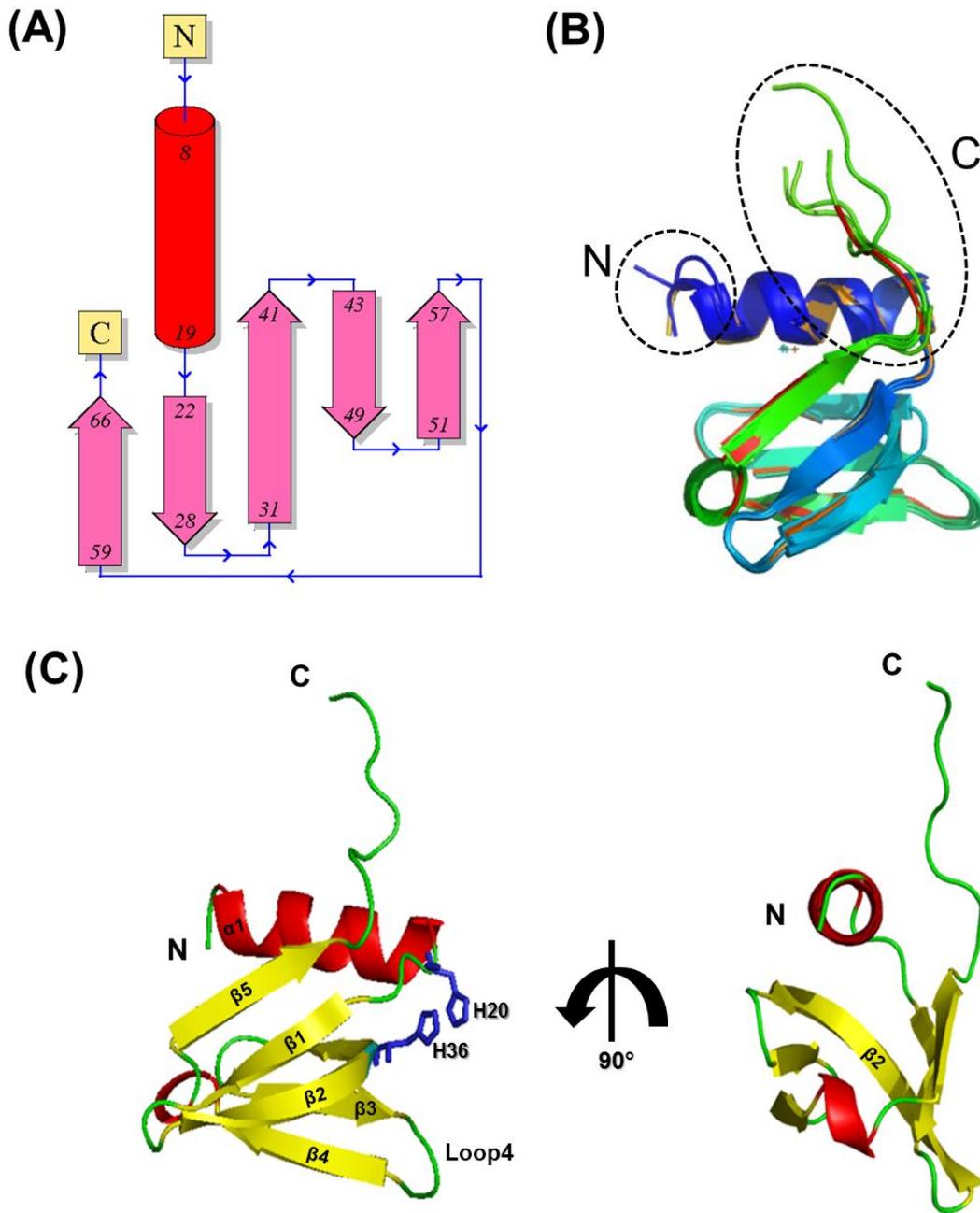


Diagrama de Ramachandram. Regiões permitidas (A, a, B, b, L e l), regiões generosamente permitidas (~a, ~b e ~l) e não permitidas (bege). Os quadrados vermelhos destacam o número, a cadeia e o tipo de resíduo de Histidina e Asparagina encontrado nas regiões generosamente permitidas e não permitidas. Figura gerada pelo programa PROCHECK.

### 2.4.3 Descrição da estrutura geral da proteína Hfq de *H. seropedicae* SmR1

A topologia conservada do tipo Sm de cada monômero compreende uma  $\alpha$ -hélice (resíduos 8-19) seguida de cinco fitas  $\beta$  ( $\beta$ 1, resíduos 22-28;  $\beta$ 2, resíduos 31-41;  $\beta$ 3, resíduos 43-49;  $\beta$ 4, resíduos 51-57;  $\beta$ 5, resíduos 59-66) (Figura 47A). A fita  $\beta$  mais longa ( $\beta$ 2) forma um arco que passa entre as fitas  $\beta$ 1 e  $\beta$ 3. Esta topologia é encontrada em todas as estruturas disponíveis de Hfq. Um centro hidrofóbico é formado por fitas  $\beta$  sobre o qual se acomoda a  $\alpha$ -hélice (Figura 47C). Cinco *loops* conectam os elementos de estrutura secundária, sendo o quinto o mais longo (Figura 47C).

FIGURA 47 – Estrutura monomérica da proteína Hfq de *H. seropedicae* SmR1



(A) Representação esquemática da topologia da proteína Hfq. Cilindro representa estrutura secundária do tipo  $\alpha$ -hélice e as setas representam fitas  $\beta$ . Os números representam o começo e fim dos elementos de estrutura secundária referente à sequência de aminoácidos. (B) Sobreposição dos monômeros realizada com o programa Pymol. As regiões flexíveis estão indicadas dentro dos círculos. (C) Monômero A da proteína Hfq com os resíduos de histidinas 20 e 36 destacados em azul. Figura preparada com o programa Pymol. N: extremidade N-terminal. C: extremidade C-terminal.

A sobreposição dos monômeros que compõem a unidade assimétrica revelou uma região C-terminal flexível e desordenada (Figura 47B), como relatadas por Beich-Frandsen e colaboradores (2011) para a proteína Hfq de *E. coli*, que pode ser igualmente importante na interação com RNAs longos. Esta análise também permitiu avaliar similaridade estrutural calculada pelo programa SuperPose (MAITI *et al.*, 2004). De forma geral os monômeros são similares estruturalmente entre si como observado pelos valores de desvio médio quadrático (RMSD) em torno de 1 Å.

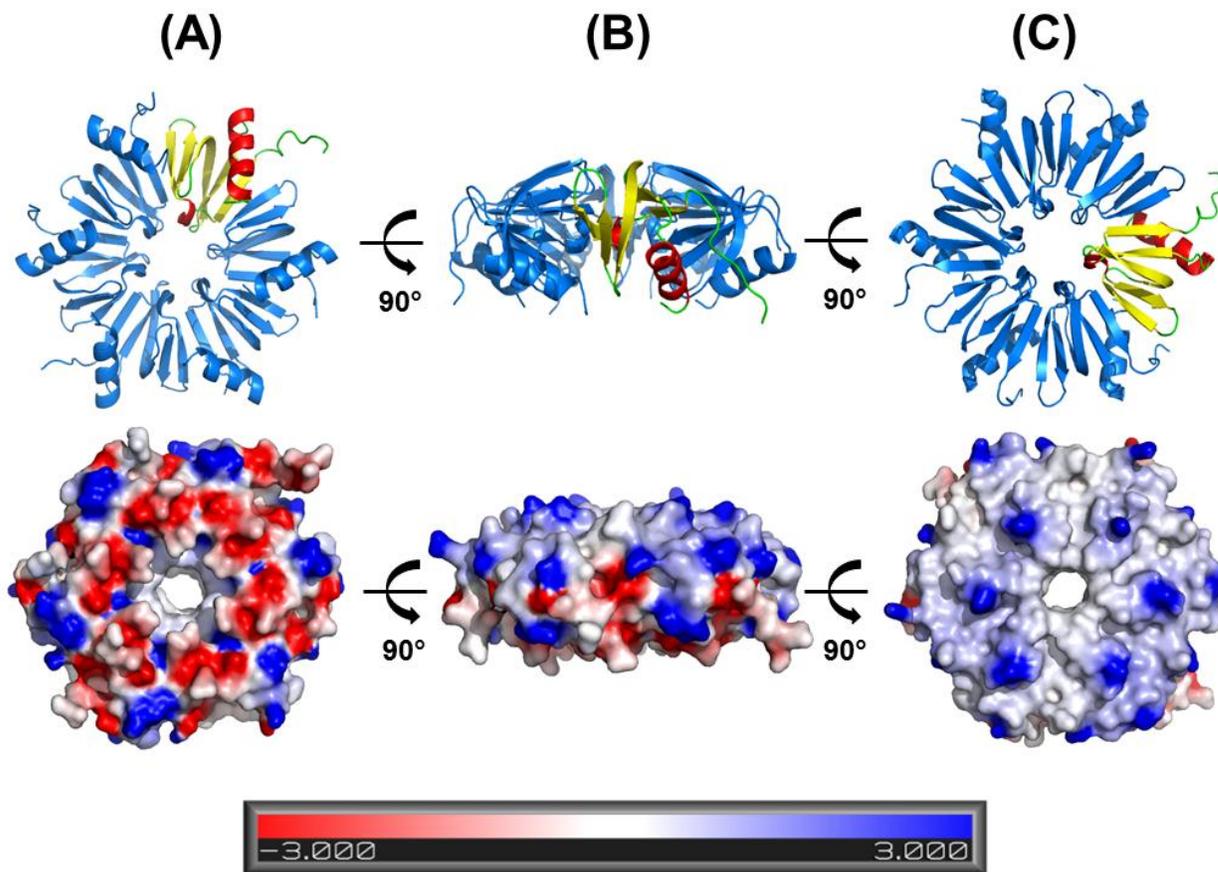
**TABELA 6 – Similaridade estrutural entre os monômeros**

|   | A              | B              | C              | D              | E              | F |
|---|----------------|----------------|----------------|----------------|----------------|---|
| A | -              | -              | -              | -              | -              | - |
| B | 0,85<br>(0,75) | -              | -              | -              | -              | - |
| C | 0,91<br>(0,53) | 0,94<br>(0,40) | -              | -              | -              | - |
| D | 1,21<br>(1,13) | 1,02<br>(0,86) | 0,77<br>(0,46) | -              | -              | - |
| E | 1,29<br>(1,17) | 0,99<br>(0,67) | 0,96<br>(0,57) | 1,13<br>(0,90) | -              | - |
| F | 1,27<br>(1,11) | 1,04<br>(0,81) | 1,06<br>(0,59) | 1,29<br>(1,07) | 1,08<br>(0,69) | - |

A sobreposição e as estatísticas de RMSD foram calculadas pelo programa SuperPose (MAITI *et al.*, 2004). Os valores fora dos parênteses levam em consideração todos os átomos da proteína e os valores entre parênteses correspondem aos cálculos levando em consideração somente os C $\alpha$ .

A proteína Hfq de *H. seropedicae* SmR1 tem o formato clássico de anel hexamérico, com diâmetro máximo de 72 Å e um diâmetro mínimo de 64 Å, apresentando uma conformação levemente distorcida na estrutura cristalográfica (Figura 48). A espessura do anel é de 34 Å com um poro central de 13 Å de diâmetro. A análise da superfície eletrostática mostra claramente que a face proximal (Figura 48) possui mais cargas do que a região distal (Figura 48) que é mais hidrofóbica e possivelmente está envolvida em mecanismos de interação proteína-proteína. A região do poro central é composta de cargas positivas indicando os sítios de interação com RNA (Figura 48).

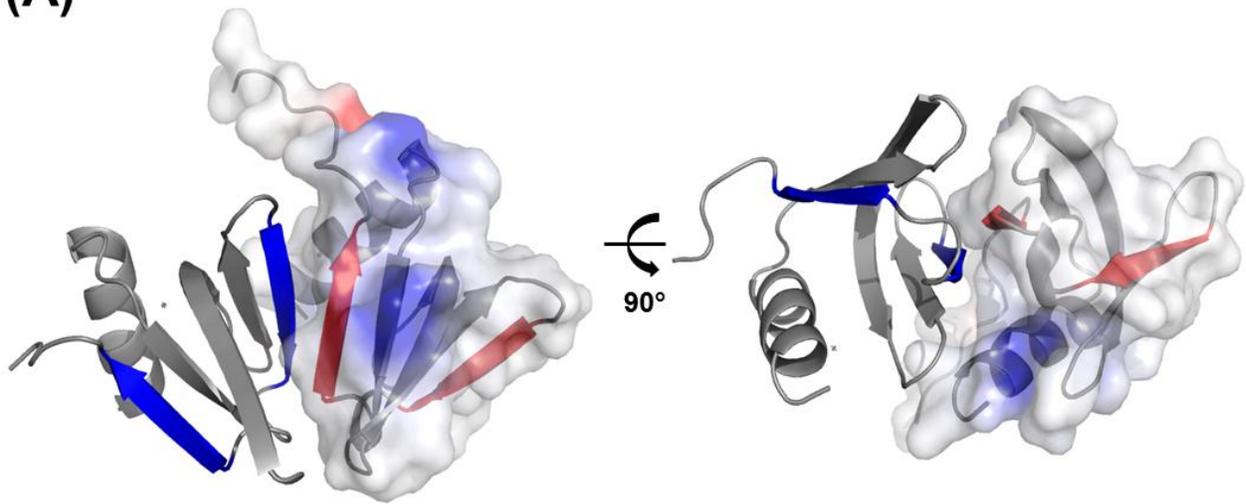
FIGURA 48 – Estrutura geral da proteína Hfq



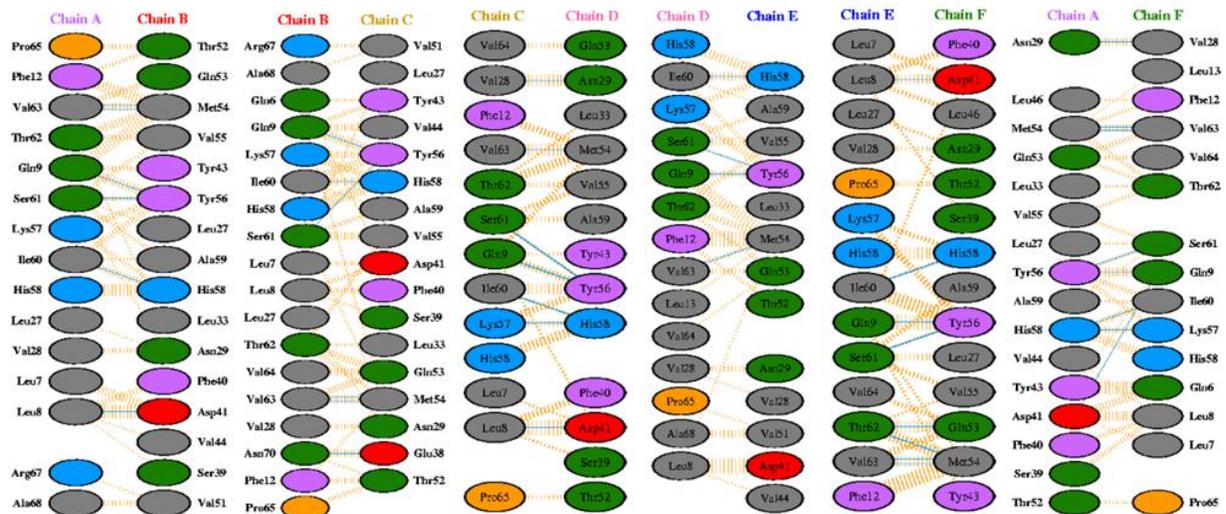
Estrutura cristalográfica da proteína Hfq de *H. seropedicae* Hfq a 2.6 Å de resolução. Na parte superior da figura a estrutura está representada na forma de *cartoon* e na parte inferior representação da superfície eletrostática. Escala de potencial eletrostático calculada pela equação de Poisson-Boltzmann no intervalo de -5 a +5 kbT/e. Região em vermelho representa potencial negativo e região azul potencial positivo. (A) face proximal, (B) visão lateral e (C) face distal. Superfície eletrostática gerada pelo programa Delphi e figuras geradas com o programa Pymol.

FIGURA 49 – Interações intermoleculares na proteína Hfq de *H. seropedicae* SmR1

(A)



(B)



(A) As interações intermoleculares existentes na proteína Hfq estão representadas para o par de monômeros A e F, na forma de *cartoon*. O envelope molecular do monômero A é mostrado na forma de transparência. As fitas  $\beta$  em azul e vermelho indicam as regiões de maior contato entre os monômeros. Figuras geradas com o programa Pymol. (B) Diagrama mostrando os resíduos de aminoácido que participam das interações intermoleculares. O nome das cadeias está indicado na porção superior de cada fileira. As linhas tracejadas em laranja indicam interações não ligadas e as linhas azuis ligações de hidrogênio. Os resíduos estão representados por elipses e coloridos de acordo com o tipo de aminoácido: alifáticos em cinza, aromáticos em lilás, polar não carregado em verde, carregados positivamente em azul, carregados negativamente em vermelho e prolina destacado em laranja.

**TABELA 7 – Interações intermoleculares na proteína Hfq de *H. seropedicae* SmR1**

| Cadeias | N° resíduos na interface | Área de interface (Å <sup>2</sup> ) | N° pontes salinas | N° pontes dissulfeto | N° ligações de H | N° contatos de não-ligação |
|---------|--------------------------|-------------------------------------|-------------------|----------------------|------------------|----------------------------|
| A & B   | 15:16                    | 730:736                             | --                | --                   | 6                | 79                         |
| B & C   | 18:17                    | 812:798                             | --                | --                   | 6                | 78                         |
| C & D   | 13:13                    | 666:657                             | --                | --                   | 7                | 63                         |
| D & E   | 14:13                    | 689:711                             | --                | --                   | 4                | 67                         |
| E & F   | 14:14                    | 675:702                             | --                | --                   | 8                | 80                         |
| F & A   | 16:15                    | 720:702                             | --                | --                   | 8                | 83                         |

Informações obtidas com o programa PDBSUM.

O hexâmero encontrado na unidade assimétrica da estrutura cristalográfica da proteína Hfq consiste também na estrutura quaternária observada pelos experimentos de gel filtração (KADOWAKI, 2008) e por SAXS como discutido na seção 2.4.5. Os monômeros possuem uma superfície média de interface de 716 Å<sup>2</sup> variando provavelmente em função do empacotamento cristalino (Tabela 7). Os resíduos de aminoácidos encontrados nas fitas β4 e 5 (Figura 49A) participam em maior número das interações intermonômero. A estrutura é estabilizada principalmente por interações hidrofóbicas e em menor importância por interações de hidrogênio (Figura 49B e Tabela 7).

#### 2.4.4 Análise estrutural comparativa

Estudos comparativos foram realizados com o objetivo de revisitar a estrutura da família de proteínas Hfq a fim de afirmar a conservação de sua conformação assim como discutir elementos estruturais não ressaltados e discutidos anteriormente na literatura.

A primeira questão é o empacotamento cristalino que é diretamente relacionado à distribuição de cargas superficiais que indicam particularidades que podem influir na interação proteína-proteína, importante na cristalização. Neste contexto, o primeiro mapa de densidade eletrônica revelou o perfil de empacotamento cristalino da proteína Hfq de *H. seropedicae*. Dois monômeros de um hexâmero interagem com a região de poro de outro hexâmero através da face distal (Figura 50), provavelmente foi este empacotamento que permitiu o refinamento de uma região C-terminal mais longa e ordenada de um dos dois monômeros que foi denominado monômero A. Este empacotamento é compartilhado com a proteína Hfq de *P. aeruginosa*, contudo difere da proteína de *E. coli* onde o empacotamento leva a um empilhamento dos hexâmeros que apresentam maior superfície de contato (Figura 50). O menor contato entre as proteínas Hfq de *H. seropedicae* reflete-se em uma menor compactação cristalina constatada por um conteúdo de solvente de 43%, superior aos 33% calculado para o empacotamento cristalino da proteína Hfq de *E. coli* (1HK9).

Outra questão foi a comparação entre os cinco *loops* que conectam os elementos de estrutura secundária. Dentre eles o *loop5*, que liga as fitas  $\beta 4$  e  $\beta 5$ , possui o motivo altamente conservado YKHA1 (Figura 51A) que contribui para a formação do poro central. Por outro lado, o resíduo Asn49, presente no *loop4* que possui uma variabilidade de aminoácidos entre as estruturas analisadas, possui uma conformação não usual como visto anteriormente no diagrama de Ramachandran. Esta conformação é estabilizada por uma ligação de hidrogênio entre o grupo amida de Asn49 e o grupo carbonil de Val51 (Figura 51B). Este resíduo é conservado em *E. coli* (códigos PDB: 1HK9 and 3GIB) e *P. aeruginosa* (código PDB: 1U1T) e encontra-se igualmente em uma conformação não usual, sugerindo a sua importância para a atividade e ou estabilidade de Hfq. Outra particularidade deste *loop* é a composição por

três resíduos, comum às estruturas citadas anteriormente, mas contrária à composição por quatro resíduos encontrada nas outras Hfq.

A estrutura cristalográfica de Hfq de *H. seropedicae* revelou a interação entre os resíduos His20 e His36 que não está presente nas demais estruturas de Hfq depositadas no PDB (Figura 51C). Estes resíduos localizados no *loop1* e na fita  $\beta 2$ , respectivamente, interagem através de um empilhamento do tipo  $\pi$ - $\pi$ , podendo ser um fator estabilizador da conformação da Hfq (Figura 47C e 51C).

**FIGURA 50 – Empacotamento cristalino das proteínas Hfq de *H. seropedicae* e de *E.coli***

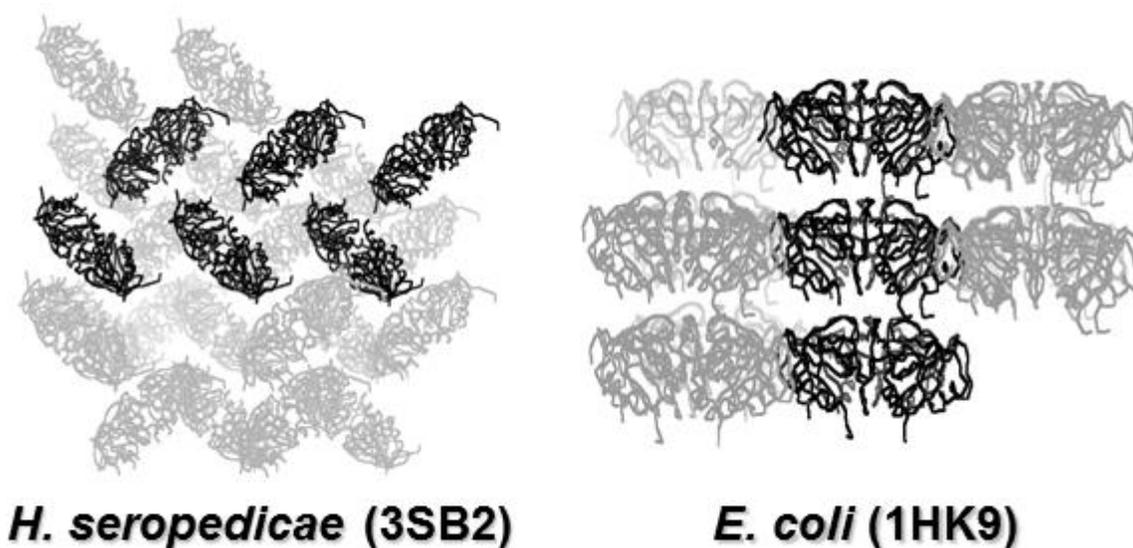
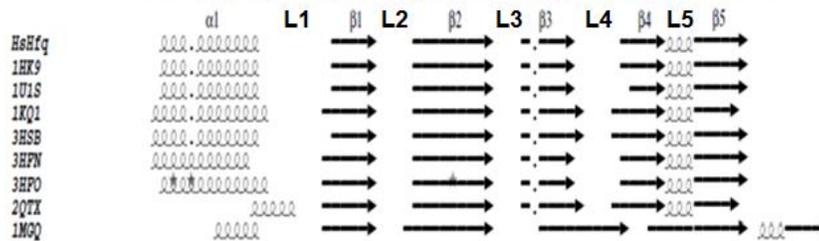
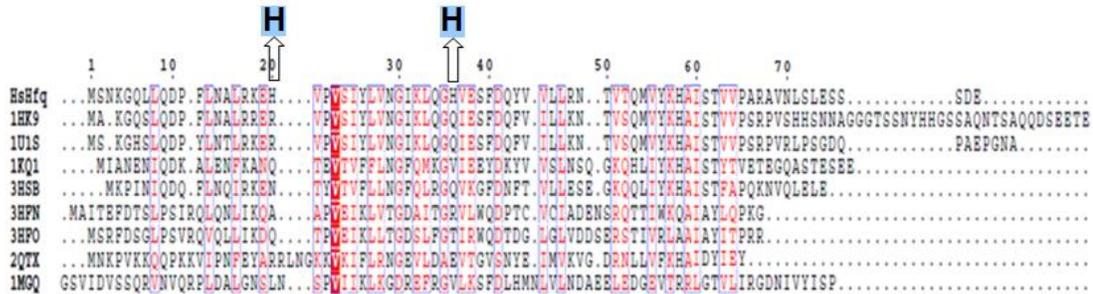


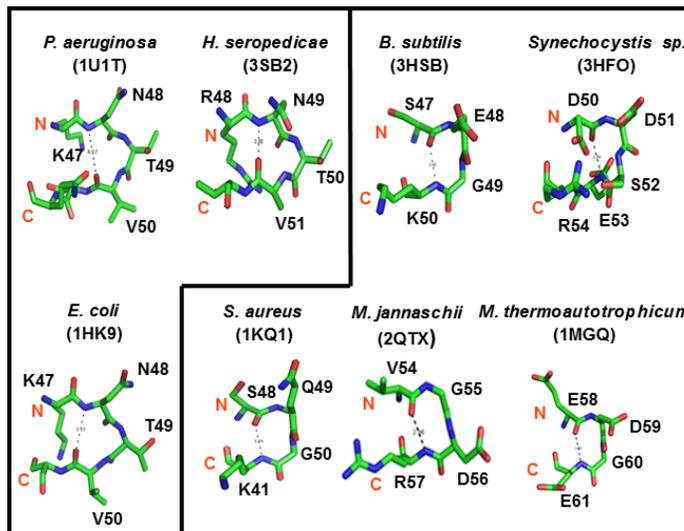
FIGURA 51 – Análise comparativa de proteínas Hfq de diferentes microrganismos

(A)

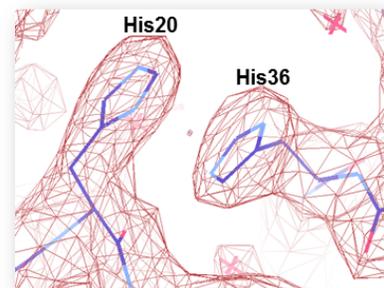


(B)

3 resíduos de aminoácidos      4 resíduos de aminoácidos



(C)



(A) Alinhamento das seqüências de aminoácidos e estruturas secundárias das proteínas Hfq com estruturas tridimensionais disponíveis no banco de dados PDB. Os aminoácidos conservados estão dentro de caixas e destacados de acordo com a intensidade da cor vermelha. Os resíduos de histidina 20 e 36 estão indicados por setas largas e pretas. (B) Comparação entre os loops 4. As estruturas estão agrupadas pelo número de resíduos no loop4 que estão representados na forma de bastões coloridos de acordo com o tipo de átomo. As linhas tracejadas representam as interações que estabilizam o loop. N: N-terminal. C: C-terminal. (C) Mapa 2mFo-DFc com um contorno de 1,0  $\sigma$  em torno dos resíduos 20 e 36 de histidinas.

Continuação da legenda na próxima página.

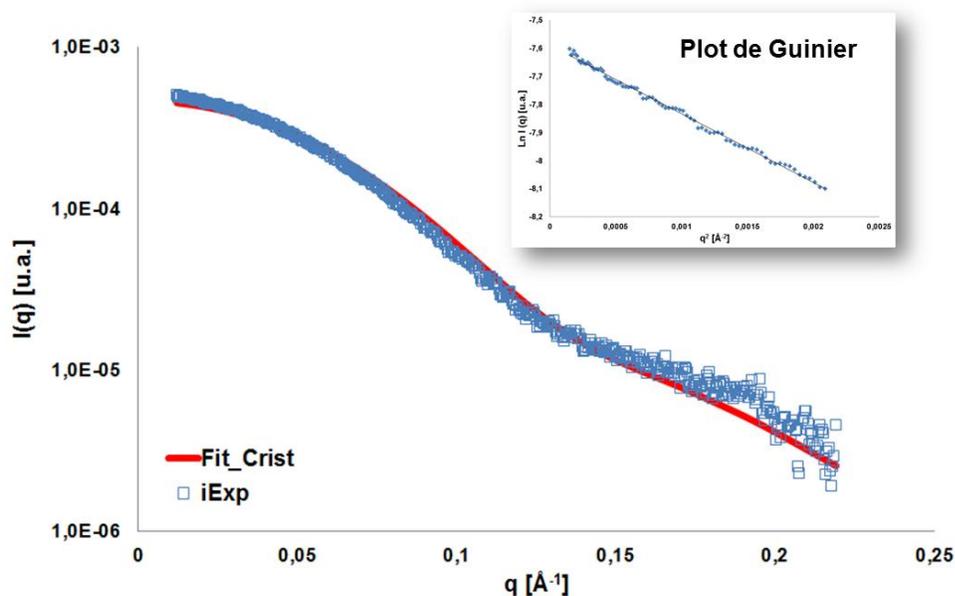
Os códigos de depósito das estruturas cristalográficas são 1HK9 para *Escherichia coli*, 1U1S para *Pseudomonas aeruginosa*, 1KQ1 para *Staphylococcus aureus*, 3HSB para *Bacillus subtilis*, 3HFN para *Anabaena PCC 7120*, 3HFO para *Synechocystis sp. PCC 6803*, 2QTX para *Methanococcus janaschii* e 1MGQ para *Methanothermobacter thermoautotrophicum*. O alinhamento das estruturas secundárias está mostrado abaixo do alinhamento da sequência de aminoácidos. As sequências foram alinhadas pelo programa ClustalW e representadas pelo programa ESPript.

#### 2.4.5 Estrutura a baixa resolução da proteína Hfq em solução

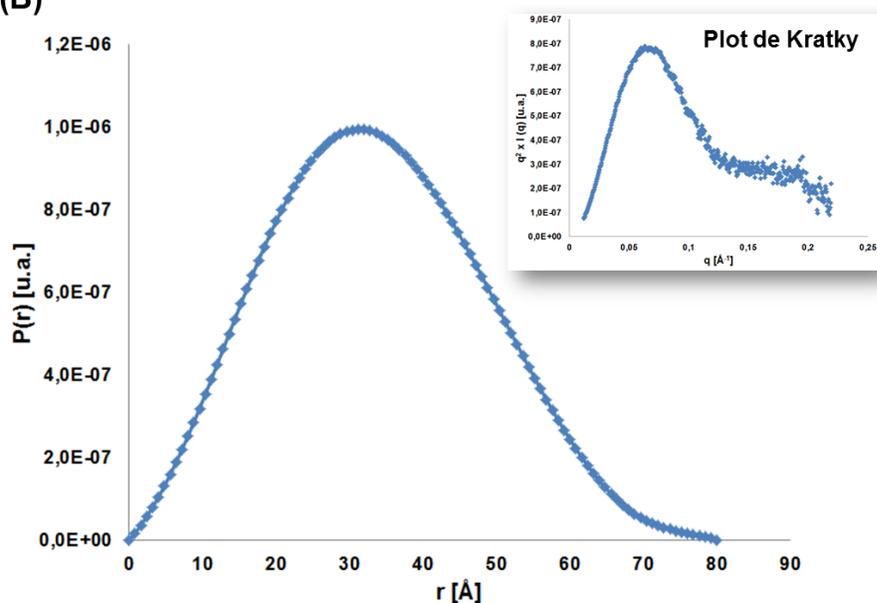
Experimentos de SAXS foram realizados com o objetivo de caracterizar estruturalmente a proteína em solução. O perfil da curva de espalhamento com as correções apropriadas foi obtida a uma concentração de 0,9 mg/mL (Figura 52A). A linearidade na região de Guinier (Figura 52A) indicou um sistema monodisperso e o plot de Kratky (Figura 52B) evidenciou um pico definido e um perfil referente a proteínas completamente estruturadas em solução. O raio de giro ( $R_g$ ) calculado, obtido através das análises de Guinier dos dados experimentais (Figura 52A), e a distância máxima ( $D_{\text{máx}}$ ), obtida com a função de distribuição de distâncias ( $p(r)$ ) (Figura 52B), foram  $26,8 \pm 0,1 \text{ \AA}$  e  $80 \text{ \AA}$ , respectivamente.  $D_{\text{máx}}$  está de acordo com a estrutura cristalográfica ( $78,2 \text{ \AA}$ ), contudo  $R_g$  mostrou-se diferente (Tabela 8) devido ao ajuste pobre da curva de espalhamento predita para a estrutura cristalográfica em relação aos dados experimentais (Figura 52A), que levou a uma discrepância de 5,3. Esta discrepância pode ser reflexo da falta de informação estrutural dos resíduos encontrados nas regiões N e C-terminal desordenadas e suas orientações, assim como a protuberância não ocupada pela estrutura cristalográfica observada na região superior do modelo *ab initio*. A massa molecular de 55,1 kDa foi estimada baseando-se na curva de SAXS utilizando o método implementado no programa SAXSMow e que concorda com os resultados de gel filtração e a estrutura cristalográfica. Os parâmetros estruturais extraídos dos dados experimentais obtidos com o experimento de SAXS estão sintetizados na tabela 8.

FIGURA 52 – Análises de SAXS da proteína Hfq

(A)



(B)

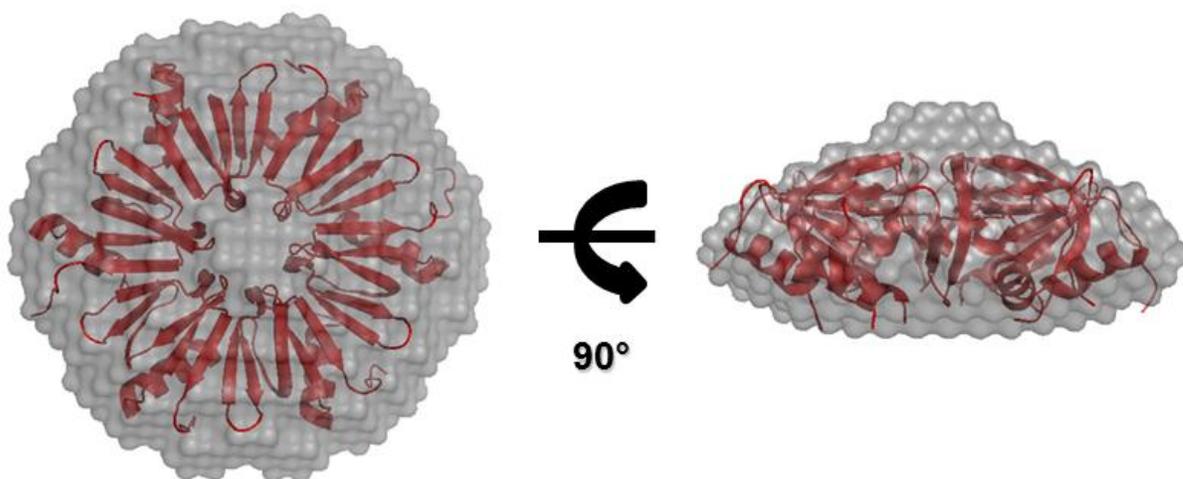


(A) Curva de espalhamento experimental da proteína Hfq a 0,9 mg/mL e o comportamento linear observado no plot de Guinier. O ajuste dos modelos *ab initio* foi calculados pelo programa DAMMIF. (B) Função de distribuição de distâncias derivados da curva experimental utilizando o programa GNOM. No quadro menor encontra-se o plot de Kratky obtido dos dados experimentais.

O melhor modelo *ab initio* foi obtido através da imposição de simetria de ordem 6 nos cálculos realizados pelo programa DAMMIN. Vinte modelos independentes foram gerados com baixa discrepância, indicada pelo índice NSD (*normalized spatial discrepancy*) de 0,6. A curva do modelo ajustou-se adequadamente à curva experimental de SAXS como atestado pelo valor de  $\chi$  de 1,43. As estruturas a baixa e alta resolução (por SAXS e cristalografia) foram sobrepostas e mostraram um bom ajuste (Figura 53). Estes resultados mostram que a estrutura cristalográfica da proteína Hfq de *H. seropedicae* é mantida em solução.

Este é o segundo relato na literatura sobre a caracterização estrutural da proteína Hfq em solução pelo método de SAXS. O primeiro relato foi para a proteína Hfq de *E. coli* com o objetivo de descrever a função da região C-terminal desta proteína (BEICH-FRANDBSEN *et al.*, 2011). Estes autores realizaram medidas com a proteína nativa e com uma deleção dos últimos trinta e sete resíduos de aminoácidos que revelaram uma flexibilidade e desordem na porção C-terminal. Este trabalho mostrou que a estrutura cristalográfica era menor do que o modelo de envelope molecular justificado pela presença da região C-terminal cuja contribuição espacial foi observada somente por SAXS. Esta característica também foi observada em menor escala para a proteína Hfq de *H. seropedicae* SmR1 e sugerida em função da discrepância entre a estrutura cristalográfica e o envelope molecular *ab initio*. Este menor efeito pode ser justificado em função da região C-terminal ser mais curta para a proteína Hfq de *H. seropedicae* quando comparada a de *E. coli*. Os dois modelos mostram que a proteína Hfq mantém a conformação hexamérica e em forma de anel em solução assim como destacaram a flexibilidade da porção C-terminal.

**FIGURA 53 – Estrutura a baixa resolução da proteína Hfq em solução**



Envelope *ab initio* a baixa resolução (cinza) sobreposto à estrutura cristalográfica (vermelho). A sobreposição foi realizada pelo programa SUPERCOMB. A figura foi preparada com o programa Pymol.

**TABELA 8 – Parâmetros estruturais obtidos com os dados de SAXS da proteína Hfq**

|                        | Hfq(Exp) <sup>†</sup> | Hfq(Dam) <sup>‡</sup> | Hfq(Crys) <sup>Δ</sup> |
|------------------------|-----------------------|-----------------------|------------------------|
| MW (kDa) <sup>φ</sup>  | 55,1                  | -                     | -                      |
| Oligômero              | 6                     | -                     | 6                      |
| Rg (Å) <sup>*</sup>    | 25,99 ± 0,07          | -                     | -                      |
| Rg (Å) <sup>**</sup>   | 26,80 ± 0,10          | 23,73                 | 22,3                   |
| Dmax (Å) <sup>**</sup> | 80                    | 76,03                 | 78,2                   |
| Resolução (Å)          | 28                    | 28                    | 2,63                   |
| Discrepância (χ)       | -                     | 1,43                  | 5,31                   |
| NSD                    | -                     | 0,6                   | -                      |

<sup>φ</sup> Calculado a partir do programa SAXSMoW.

<sup>†</sup> Valores calculados a partir de dados experimentais.

<sup>‡</sup> Parâmetros calculados a partir dos modelos de átomos dummy.

<sup>\*</sup> Calculado a partir do plot de Guinier.

<sup>\*\*</sup> Calculado a partir da função de distribuição de distâncias.

<sup>Δ</sup> Valores calculados com o programa CRY SOL.

Resolução:  $2\pi/q_{\text{máx}}$

χ Calculado com o programa DAMMIN.

NSD (*normalized spatial discrepancy*) calculado com o programa DAMAVER.

## 2.5 CONCLUSÕES

- A estrutura da proteína Hfq de *Herbaspirillum seropedicae* estirpe SmR1 foi resolvida por cristalografia a uma resolução de 2,6 Å, e corresponde a primeira estrutura de uma chaperona de RNA de uma bactéria diazotrófica e endofítica.
- A proteína Hfq apresentou a topologia Sm conservada.
- A estrutura permitiu a discussão de duas diferentes conformações para o loop4 sugerindo que esta região é importante para a função protéica.
- Dois resíduos de histidinas encontrados somente na estrutura de Hfq de *H. seropedicae* interagem através de um empilhamento do tipo  $\pi$ - $\pi$  e podem ser importante na estabilização da estrutura.
- O envelope *ab initio* calculado a partir dos experimentos de SAXS sugere que a conformação observada na estrutura cristalográfica é mantida em solução.

### 3. REFERÊNCIAS BIBLIOGRÁFICAS

ABE, H.; KIKKAWA, Y.; IWATA, T.; AOKI, H.; AKEHATA, T.; DOI, Y. Microscopic visualization on crystalline morphologies of thin films for poly[(R)-3-hydroxybutyric acid] and its copolymer. **Polymer**, v. 41, p. 867–874, 2000.

ADAMS, P.D.; AFONINE, P.V.; BUNKOCZI, G.; CHEN, V.B.; DAVIS, I.W.; ECHOLS, N.; HEADD, J.J.; HUNG, L.; KAPRAL, G.J.; GROSSE-KUNSTLEVE, R.W.; MCCOY, A.J.; MORIARTY, N.W.; OEFFNER, R.; READ, R.J.; RICHARDSON, D.C.; RICHARDSON, J.S.; TERWILLIGERE, T.C.; ZWARTA, P.H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. **Acta Cryst. D**, v. 66, p. 213-221, 2010.

ADLER, A.J.; GREENFIELD, N.J.; FASMAN, G.D. **Met. Enzymol.**, London and New York: Academic Press, vol. XXVII D, pp. 675, 1973.

AIBA, H. Mechanism of RNA silencing by Hfq-binding small RNAs. **Curr. Opin. Microbiol.**, v. 10, p. 1-6, 2007.

ANDERSON, A.J.; DAWES, E.A. Occurrence, metabolism, metabolic rate, and industrial uses of bacterial polyhydroxyalkanoates. **Microbiol. Ver.**, v. 54, p. 450–472, 1990.

ANSONG, C.; YOON, H.; PORWOLLIK, S.; MOTTAZ-BREWER, H.; PETRITIS, B.O.; JAITLY, N.; ADKINS, J.N.; MCCLELLAND, M.; HEFFRON, F.; SMITH, R.D. Global Systems-Level Analysis of Hfq and SmpB Deletion Mutants in *Salmonella*: Implications for Virulence and Global Protein Translation. **PLoS ONE**, v. 4, p. 1-13, 2009.

BALDANI J.I.; BALDANI V.L. History on the biological nitrogen fixation research in graminaceous plants: special emphasis on the Brazilian experience. **An Acad. Bras. Cienc.**, v. 77, p. 549-79, 2005.

BALDANI, J.I.; BALDANI, V.L.D.; SELDIN, L.; DÖBEREINER, J. Characterization of *Herbaspirillum seropedicae* gen. nov., sp. nov., a new root-associated nitrogen-fixing bacterium. **Intl. J. Sys. Bact.**, n. 36, p. 86-93, 1986.

BALSANELLI, E.; SERRATO, R.V.; DE BAURA, V.A.; SASSAKI, G.; YATES, M.G.; RIGO, L.U.; PEDROSA, F.O.; SOUZA, E.M.; MONTEIRO, R.A. *Herbaspirillum seropedicae* *rfbB* and *rfbC* genes are required for maize colonization. **Environ. Microbiol.**, v. 12, p. 2233–2244, 2010.

BASTIÁN, F.; COHEN, A.; PICCOLI, P.; LUNA, V.; BARALDI, R.; BOTTINI, R. Production of indole-3-acetic acid and gibberellins A<sub>1</sub> and A<sub>3</sub> by *Acetobacter diazotrophicus* and *Herbaspirillum seropedicae* in chemically-defined culture media. **Plant Growth Regulation**, v. 24, p. 7-11, 1998.

BAILEY, T.L.; WILLIAMS, N.; MISLEH, C.; LI, W.W. MEME: discovering and analyzing DNA and protein sequence motifs. **Nucl. Acids Res.**, v. 34, 2006.

BEICH-FRANDSEN, M.; VECEREK, B.; KONAREV, P.V.; SJÖBLOM, B.; KLOIBER, K.; HÄMMERLE, H.; RAJKOWITSCH, L.; MILES, A.J.; KONTAXIS, G.; WALLACE, B.A.; SVERGUN, D.I.; KONRAT, R.; BLÄSI, U.; DJINOVIC-CARUGO, K. Structural insights into the dynamics and function of the C-terminus of the *E. coli* RNA chaperone Hfq. **Nucl. Acids Res.**, v. 39, p. 4900-4915, 2011.

BENELLI, E.M.; BUCK, M.; POLIKARPOV, I.; SOUZA, E.M.; CRUZ, L.M.; PEDROSA, F.O. *Herbaspirillum seropedicae* signal transduction protein PII is structurally similar to the enteric GlnK. **Eur. J. Biochem.**, v. 269, p. 3296-3303, 2002.

BENELLI, E.M.; BUCK, M.; SOUZA, E.M.; YATES, M.G.; PEDROSA, F.O. Uridylylation of the PII protein from *Herbaspirillum seropedicae*. **Can. J. Microbiol.**, v. 47, p. 309-314, 2001.

BENELLI, E.M.; SOUZA, E.M.; FUNAYAMA, S.; RIGO, L.U.; PEDROSA, F.O. Evidences for two possible glnB-type genes in *Herbaspirillum seropedicae*. **J. Bacteriol.**, v. 179, p. 4623-4626, 1997.

BERNADO, P.; SVERGUN, D.I. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. **Mol. BioSyst.**, 2011

BINNIG, G.; QUATE, C.F.; GERBER, CH. Atomic force microscope. **Phys. Rev. Lett.**, v. 56, p. 930-933, 1986.

BODDEY, R.M.; OLIVEIRA, O. C.; URQUIAGA, S.; REIS, V. M.; OLIVARES, F. L.; BALDANI, V. L. D.; DÖBEREINER, J. Biological nitrogen fixation associated with sugar cane and rice: Contributions and prospects for improvement. **Plant and Soil**, v. 174, p. 195-209, 1995.

BØGGILD, A.; OVERGAARD, M.; VALENTIN-HANSEN, P.; BRODERSEN, D.E. Cyanobacteria contain a structural homologue of the Hfq protein with altered RNA-binding properties. **FEBS J.**, v. 276, p. 3904-3915, 2009.

BOHMERT, K.; BALBO, I.; STEINBÜCHEL, A.; TISCHENDORF, G.; WILLMITZER, L. Constitutive expression of the  $\beta$ -ketothiolase gene in transgenic plants. A major obstacle for obtaining polyhydroxybutyrate-producing plants. **Plant Physiol.**, v. 128, p. 1282–1290, 2002.

BONATTO, A.C.; COUTO, G.H.; SOUZA, E.M.; ARAUJO, L.M.; PEDROSA, F.O.; NOINDORF, L.; BENELLI, E.M. Purification and characterization of the bifunctional uridylyltransferase and the signal transducing proteins GlnB and GlnK from *Herbaspirillum seropedicae*. **Prot. Express. Purif.**, v. 55, p. 293-299, 2007.

BONATTO, A.C.; SOUZA, E.M.; PEDROSA, F.O.; YATES, M.G.; BENELLI, E.M. Effect of T- and C-loop mutations on the *Herbaspirillum seropedicae* GlnB protein in nitrogen signalling. **Res. Microbiol.**, v. 156, p. 634-640, 2005.

BRADFORD, M.M. A rapid and sensitive method for the quantification of microgram quantities of protein utilization: the principle of protein-dye binding. **Anal. Biochem.**, v. 72, p. 248-254, 1976.

BRENNAN, R.G.; LINK, T.M. Hfq structure, function and ligand binding. **Curr. Opin. Microbiol.**, v. 10, p. 125-133, 2007.

BYROM, D. (1994). Polyhydroxyalkanoates, In: Mobley DP (ed) **Plastic from microbes: microbial synthesis of polymers and polymer precursors**. Hanser Munich, pp. 5-33.

CATALAN, A.I.; FERREIRA, F.; GILL, P.R.; BATISTA, S. Production of polyhydroxyalkanoates by *Herbaspirillum seropedicae* grown with different sole carbon sources on lactose when engineered to express the *lacZlacY* genes. **Enzyme Microbial. Technol.**, v. 20, p. 1352-1357, 2007.

Cell-Porator Voltage Booster – Instruction Manual – Life Technologies, Inc. 1991.

CHAVES, D.F.S.; FERRER, P.P.; SOUZA, E.M.; CRUZ, L.M.; MONTEIRO, R.A.; PEDROSA, F.O. A two-dimensional proteome reference map of *Herbaspirillum seropedicae* proteins. **Proteomics**, v. 7, p. 3759-3763, 2007.

CHOU, M.E.; YANG, M.K. Analyses of binding sequences of the PhaR protein of *Rhodobacter sphaeroides* FJ1. **FEMS Microbiol. Lett.**, v. 302, p. 138-143, 2010.

CHOU, M.E.; CHANG, W.T.; CHANG, Y.C.; YANG, M.K. Expression of four *pha* genes involved in poly- $\beta$ -hydroxybutyrate production and accumulation in *Rhodobacter sphaeroides* FJ1. **Mol. Genet. Genomics**, 2009.

CHUBATSU, L.S.; MONTEIRO, R.A.; SOUZA, E.M.; OLIVEIRA, M.A.S.; YATES, M.G.; WASSEM, R.; BONATTO, A.C.; HUERGO, L.F.; STEFFENS, M.B.R.; RIGO, L.U.; PEDROSA, F.O. Nitrogen fixation control in *Herbaspirillum seropedicae*. **Plant and Soil**, 2011.

CHUNG, C.T., NIEMELA, S.L., MILLER, R.H. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. **Proc. Nat. Acad. Sci. USA**, v. 86, p. 2172-5, 1989.

COUTO, G.H. **Expressão e Purificação da Proteína GlnD de *Herbaspirillum seropedicae***. Curitiba, Universidade Federal do Paraná, Setor de Ciências Biológicas, 99p., Dissertação de Mestrado, 2005.

DAS, R.; BAKER, D. Macromolecular Modeling with Rosetta. **Annu. Rev. Biochem.**, v. 77, p. 36-82, 2008.

DAVIS, I.W.; MURRAY, L.W.; RICHARDSON, J.S.; RICHARDSON, D.C. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. **Nucl. Acids Res.**, v. 32, p. 615-619, 2004.

DE EUGENIO, L.I.; GALAN, B.; ESCAPA, I.F.; MAESTRO, B.; SANZ, J.M.; GARCIA, J.L.; PRIETO, M.A. The PhaD regulator controls the simultaneous expression of the *pha* genes involved in polyhydroxyalkanoate metabolism and turnover in *Pseudomonas putida* KT2442. **Environ. Microbiol.**, v. 12, p. 1591-1603, 2010.

DENNIS, D.; SEIN, V.; MARTINEZ, E.; AUGUSTINE, B. PhaP Is Involved in the Formation of a Network on the Surface of Polyhydroxyalkanoate Inclusions in *Cupriavidus necator* H16. **J. Bacteriol.**, v. 190, p. 555–563, 2008.

DING, Y.; DAVIS, B.M.; Waldor, M.K. Hfq is essential for *Vibrio cholera* virulence and downregulates sigma expression. **Mol. Microbiol.**, v. 53, p. 345-354, 2004.

DRENTH, J. **Principles of Proteins X-Ray Crystallography**. Springer-Verlag. New York, USA, 1994.

DREPPER, T.; RAABE, K.; GIAOURAKIS, D.; GENDRULLIS, M.; MASEPOHL, G.; KLIPP, W. The Hfq-like protein NrfA of the phototrophic purple bacterium *Rhodospirillum rubrum* controls nitrogen fixation via regulation of *nifA* and *anfA* expression. **FEMS**, v. 215, p. 221-227, 2002.

EGGINK, G.; DEWAARD, P.; HUIJBERTS, G.N.M. Formation of novel poly(hydroxyalkanoates) from long-chain fatty acids. **Can. J. Microbiol.**, v. 41, p. 14-21, 1995.

EMSLEY, P.; COWTAN, K. Coot: model-building tools for molecular graphics. **Acta Cryst. D**, v. 60, p. 2126-2132, 2004.

FEIGIN, L.A.; SVERGUN, D.I. Structure Analysis by Small-Angle X-ray and Neutron Scattering. Plenum Press, New York, 1987.

FISCHER, H.; NETO, M.O.; NAPOLITANO, H.B.; POLIKARPOV, I.; CRAIEVICH, A.F. Determination of the molecular weight of proteins in solution from a single small-angle x-ray scattering measurement on a relative scale. **J. Appl. Cryst.**, v. 43, p. 101-109, 2010.

FRANKE, D.; SVERGUN, D.I. DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering. **J. Appl. Cryst.**, v. 42, p. 342-346, 2009.

FULLER, R.C.; O'DONNELL, J.P.; SAULNIER, J.; REDLINGER, T.E.; FOSTER, J.; LENZ, J.W. The supramolecular architecture of the polyhydroxyalkanoate inclusions in *Pseudomonas oleovorans*. **FEMS Microbiol. Rev.**, v. 103, p. 279-288, 1992.

GALÁN B.; DINJASKI, N.; MAESTRO, B.; DE EUGENIO, L.I.; ESCAPA, I.F.; SANZ, J.M.; GARCÍA, J.L.; PRIETO, M.A. Nucleoid-associated PhaF phasin drives intracellular location and segregation of polyhydroxyalkanoate granules in *Pseudomonas putida* KT2442. **Mol. Microbiol.**, v. 79, p. 402-418, 2011.

GALVÃO, C.W.; PEDROSA, F.O.; SOUZA, E.M.; YATES, M.G.; CHUBATSU, L.S.; STEFFENS, M.B. The *recX* gene product is involved in the SOS response in *Herbaspirillum seropedicae*. **Can. J. Microbiol.**, v. 49, p. 145-50, 2003.

GOUET, P.; ROBERT, X.; COURCEL, E. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. **Nucl. Acids Res.**, v. 31, p. 3320–3323, 2003.

GRIEBEL, R.; SMITH, Z.; MERRICK, M. Metabolism of poly-beta-hydroxybutyrate. Purification, composition, and properties of native poly-beta-hydroxybutyrate granules from *Bacillus megaterium*. **Biochemistry**, v. 7, p. 3676–3681, 1968.

GUINIER, A.; FOURNET, G. **Small Angle Scattering of X-Rays**. New York, Wiley, 1955.

HALL, T.A. **BioEdit 4.8**. Raleigh, 1997-2001. 1 arquivo (11,5 Mb); Disponível em: <<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>> BioEdit: a userfriendly biological sequence alignment editor and analysis program for Windows 95/98/NT.

HAMMERSLEY, A.P. **FIT2D: An Introduction and Overview**, ESRF Internal Report. ESRF97HA02T, 1997.

HANLEY, S.Z.; PAPPIN, D.J.C.; RAHMAN, D.; WHITE, A.J.; ELBOROUGH, K.M.; SLABAS, A.R.. Re-evaluation of the primary structure of *Ralstonia eutropha* phasin and implications for polyhydroxyalkanoic acid granule binding. **FEBS Letters**, 1999.

HAYWOOD, G.W.; ANDERSON, A.J.; CHU, L.; DAWES, A.E. The role of NADH- and NADPH-linked acetoacetyl-CoA reductases in the poly-3-hydroxybutyrate synthesizing organism *Alcaligenes eutrophus*. **FEMS Microbiol. Lett.**, v. 52, p.259–264, 1988.

HERMANN, H.; FABRIZIO, P.; RAKER, V.A.; FOULAKI, K.; HORNIG, H.; BRAHMS, H.; LUHRMANN, R. snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. **EMBO J.**, v. 14, p. 2076–2088, 1995.

HERNANDEZ-ELIGIO, A.; CASTELLANOS, M.; MORENO, S.; ESPÍN, G. Transcriptional activation of the *Azotobacter vinelandii* polyhydroxybutyrate biosynthetic genes *phbBAC* by PhbR and RpoS. **Microbiology**, 2011.

HISANO, T.; KASUYA, K.; TEZUKA, Y.; ISHII, N.; KOBAYASHI, T.; SHIRAKI, M.; OROUDJEV, E.; HANSMA, H.; IWATA, T.; DOI, Y.; SAITO, T.; MIKI, K. The Crystal Structure of Polyhydroxybutyrate Depolymerase from *Penicillium funiculosum* Provides Insights into the Recognition and Degradation of Biopolyesters. **J. Mol. Biol.**, v. 356, p. 993–1004, 2006.

HUERGO, L.F.; NOINDORF, L.; GIMENES, C.; LEMGRUBER, R.S.; CORDELLINI, D.F.; FALARZ, L.J.; CRUZ, L.M.; MONTEIRO, R.A.; PEDROSA, F.O.; CHUBATSU, L.S.; SOUZA, E.M.; STEFFENS, M.B. Proteomic analysis of *Herbaspirillum seropedicae* reveals ammonium-induced AmtB dependent membrane sequestration of PII proteins. **FEMS Microbiol. Lett.**, v. 308, p. 40–7, 2010.

HUIJBERTS, G.N.; EGGINK, G.; de WAARD, P.; HUISMAN, G.W.; WITHOLT, B. *Pseudomonas putida* KT2442 cultivated on glucose accumulates poly(3-hydroxyalkanoates) consisting of saturated and unsaturated monomers. **Appl. Environ. Microbiol.**, v. 58, p. 536-544, 1992.

JENDROSSEK, D. Polyhydroxyalkanoate Granules are Complex Subcellular Organelles (Carbonosomes). **J. Bacteriol.**, v. 191, p. 3195–3202, 2009.

JENDROSSEK, D.; HANDRICK, R. Microbial degradation of polyhydroxyalkanoates. **Annu. Rev. Microbiol.**, v. 56, p. 403–432, 2003.

JIA, Y.; KAPPOCK, T.J.; FRICK, T.; SINSKEY, A.J.; STUBBE, J. Lipases Provide a New Mechanistic Model for Polyhydroxybutyrate (PHB) Synthases: Characterization of the Functional Residues in *Chromatium Vinosum* PHB Synthase. **Biochemistry**, v. 39, p. 3927-3936, 2000.

JONES, D.T. Protein secondary structure prediction based on position-specific scoring matrices. **J. Mol. Biol.**, v.292, p.195-202, 1999.

JURASEK, L.; MARCHESSAULT, R.H. The Role of Phasins in the Morphogenesis of Poly(3-hydroxybutyrate) Granules. **Biomacromolecules**, v. 3, p. 256-261, 2002.

KABSCH, W. XDS. **Acta Cryst. D**, v. 66, p. 125-132, 2010.

KADOWAKI, M.A.S. **Caracterização estrutural e funcional da proteína Hfq de *Herbaspirillum seropedicae***. Tese de Mestrado em Ciências (Bioquímica). Editora da UFPR- Curitiba-PR, 2008.

KADOURI, D.; JURKEVITH, E.; OKON, Y. Involvement of the reserve material poly-β-hydroxybutyrate in azospirillum brasiliense stress endurance and root colonization. **Appl. Environ. Microbiol.**, v. 6, p. 3244-3250, 2003.

KADOWAKI, M.A.; MÜLLER-SANTOS, M.; REGO, F.G.; SOUZA, E.M.; YATES, M.G.; MONTEIRO, R.A.; PEDROSA, F.O.; CHUBATSU, L.S.; STEFFENS, M.B. Identification and characterization of PhbF: A DNA binding protein with regulatory role in the PHB metabolism of *Herbaspirillum seropedicae* SmR1. **BMC Microbiol.**, v. 14, 2011a.

KADOWAKI, M.A.; IULEK, J.; BARBOSA, J.A.R.G.; PEDROSA, F.O.; SOUZA, E.M.; CHUBATSU, L.S.; MONTEIRO, R.A.; OLIVEIRA, M.A.S.; STEFFENS, M.B. Structural characterization of the RNA chaperone Hfq from the nitrogen-fixing bacterium *Herbaspirillum seropedicae* SmR1. **Bioch. Biophys. Acta.**, 2011b.

KAMINSKI, P.A.; ELMERICH, C. The Control of *Azorhizobium caulinodans nifA* expression by oxygen, ammonia and by the HF-I-like protein, NrfA. **Mol. Microbiol.**, v. 28, p. 603-613, 1998.

KAUFMANN, K.W.; LEMMON, G.H.; DELUCA, S.L.; SHEEHAN, J.H.; MEILER, J. Practically Useful: What the ROSETTA Protein Modeling Suite Can Do for You. **Biochemistry**, v. 49, p. 2987-2998, 2010.

KESSLER, B.; WESTHUIS, R.; WITHOLT, B.; EGGINK, G. Production of microbial polyesters: fermentation and downstream processes. In: W. Babel, A. STEINBÜCHEL, (Eds.), *Advances in Biochemical Engineering Biotechnology: Biopolyesters*, Vol. 71, Springer, Berlin, 2001, pp. 159-182.

KESSLER, B.; WITHOLT, B. Factors involved in the regulatory network of polyhydroxyalkanoate metabolism. **J. Biotechnol.**, v.86, p. 97-104, 2001.

KIM, D.E.; CHIVIAN, D.; BAKER, D. Protein structure prediction and analysis using the Robetta server. **Nucl. Acids Research**, v. 32, 2004.

KLASSEN, G.; PEDROSA, F.O.; SOUZA, E.M.; YATES, M.G.; RIGO, L.U. Sequencing and functional analysis of the *nifENXorf1orf2* gene cluster of *H. seropedicae*. **FEMS Microbiol. Lett.**, v. 181, p. 165-170, 1999.

KLASSEN, G.; PEDROSA, F.O.; SOUZA, E.M.; FUNAYAMA, S.; RIGO, L.U. Effect of nitrogen compounds on nitrogenase activity in *Herbaspirillum seropedicae* strain SmR1. **Can. J. Microbiol.**, n. 43, p. 841–846, 1997.

KLEYWEGT, G.J.; JONES, T.A. xdIMAPMAN and xdIDATAMAN - programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. **Acta Cryst. D**, v.52, p.826-828, 1996.

KOZIN, M.B.; SVERGUN, D.I. Automated matching of high- and low-resolution structural models. **J. Appl. Cryst.**, v. 34, p. 33-41, 2001.

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

LAGEVEEN, R. G.; HUISMAN, G.W.; PREUSTING, H.; KETELAAR, P.; EGGINK, G.; WITHOLT, B. Formation of polyesters by *Pseudomonas oleovorans*: effect of substrates on formation and composition of poly(R)-3-hydroxyalkanoates and poly(R)-3-hydroxyalkenoates. **Appl. Environ. Microbiol.**, v. 54, p. 2924–2932, 1988.

LASKOWSKI, R.A. PDBsum new things. **Nucl. Acids Res.**, v. 37, p. 355-359, 2009.

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

LEMOIGNE, M. Produits de deshydratation et de polymerisation de l'acide b-oxobutyrique. **Bull. Soc. Chem. Biol.** v. 8, p. 770-782, 1926.

LINK, T.M.; VALENTIN-HANSEN, P.; BRENNAN, R.G. Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. **Proc. Nat. Acad. Sci. USA**, v. 106, p. 1-6, 2009.

LUSCOMBE, N.M.; AUSTIN, S.E.; BERMAN, H.M.; THORNTON, J.M. An overview of the structures of protein-DNA complexes. **Genome Biology**, v. 1, 2000.

MACNEIL, T.; MACNEIL, D.; TYLER, B. Fine-structure deletion map and complementation analysis of the *glnA-glnL-glnG* region in *Escherichia coli*. **J. Bacteriol.**, v. 150, p. 1302-1313, 1982.

MADISON, L.L.; HUISMAN, G.W. Metabolic Engineering of Poly(3-Hydroxyalkanoates): From DNA to Plastic. **Microbiol. Mol. Biol. Rev.** v. 63, n. 1, p. 21-53, 1999.

MAEHARA, A.; UEDA, S.; NAKANO, H.; YAMANE, T. Analyses of a polyhydroxyalkanoic acid granule-associated 16-kilodalton protein and its putative regulator in the *pha* locus of *Paracoccus denitrificans*. **J. Bacteriol.**, v. 181, p. 2914-2921, 1999.

MAEHARA, A.; DOI, Y.; NISHIYAMA, T.; TAKAGI, Y.; UEDA, S.; NAKANO, H.; YAMANE, T. PhaR, a protein of unknown function conserved among short-chain-length polyhydroxyalkanoic acid-producing bacteria, is a DNA-binding protein and represses *Paracoccus denitrificans* *phaP* expression in vitro. **FEMS Microbiol. Lett.**, v. 200, p. 9-15, 2001.

MAEHARA, A.; TAGUCHI, S.; NISHIYAMA, T.; YAMANE, T.; DOI, Y. A Repressor Protein, PhaR, Regulates Polyhydroxyalkanoate (PHA) Synthesis via Its Direct Interaction with PHA. **J. Bacteriol.**, v. 184, n. 14, p. 3992-4002, 2002.

MAITI, R.; VAN DOMSELAAR, G.H.; ZHANG, H.; WISHART, D.S. SuperPose: a simple server for sophisticated structural superposition. **Nucl. Acids Res.**, v. 32, p. 590-594, 2004.

MCCOOL, G.J.; CANNON, M.C. Polyhydroxyalkanoate inclusion body-associated proteins and coding region in *Bacillus megaterium*. **J. Bacteriol.**, v. 181, p. 585–592, 1999.

MCCOY, A.J.; GROSSE-KUNSTLEVE, R.W.; ADAMS, P.D.; WINN, M.D.; STORONIA, L.C.; READA, R.J. Phaser crystallographic software. **J. Appl. Cryst.**, v. 40, p. 658-674, 2007.

MCNEALY, T.L.; FORSBACH-BIRK, V.; SHI, C.; MARRE, R. The Hfq homolog in *Legionella pneumophila* demonstrates regulation by LetA and RpoS and interacts with the global regulator CsrA. **J. Bacteriol.**, v. 187, p. 1527-1532, 2005.

MILLER J. H. **Experiments in Molecular Genetics**, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1992.

MIYAZAKI, S.; TAKAHASHI, K.; SHIRAKI, M.; SAITO, T.; TEZUKA, Y.; KASUYA, K. Properties of a Poly(3-hydroxybutyrate) Depolymerase from *Penicillium funiculosum*. **Journal of Polymers and the Environment**, v. 8, n. 4, p. 175-182, 2002.

MONTEIRO, R.A.; SCHMIDT, M.A.; BAURA, V.A.; BALSANELLI, E., WASSEM, R.; YATES, M.G.; RANDI, M.A.F.; PEDROSA, F.O.; SOUZA, E.M. Early colonization pattern of maize (*Zea mays* L. *Poales*, *Poaceae*) roots by *Herbaspirillum seropedicae* (Burkholderiales, Oxalobacteraceae). **Genet. Mol. Biol.**, v. 31, p. 932-937, 2008.

MONTEIRO, R.A.; SOUZA, E.M.; YATES, M.G.; PEDROSA, F.O.; CHUBATSU, L.S. Fnr is involved in oxygen control of *Herbaspirillum seropedicae* N-truncated NifA protein activity in *Escherichia coli*. **Appl. Environ. Microbiol.**, v. 69, p. 1527-1531, 2003.

MONTEIRO, R. A.; SOUZA, E. M.; YATES, M.G.; PEDROSA, F. O.; CHUBATSU, L.S. Use of lactose to induce expression of soluble NifA protein domains of *Herbaspirillum seropedicae* in *Escherichia coli*. **Can. J. Microbiol.**, n. 46, p. 1-4, 2000.

MONTEIRO, R.A.; SOUZA, E.M.; FUNAYAMA, S.; YATES, M.G.; PEDROSA, F.O.; CHUBATSU, L.S. Expression and functional analysis of an N-truncated NifA protein of *Herbaspirillum seropedicae*. **FEBS Lett.**, v. 447, p. 283-286, 1999a.

MONTEIRO, R.A.; SOUZA, E.M.; YATES, M.G.; PEDROSA, F.O.; CHUBATSU, L.S. In-trans regulation of the N-truncated NifA protein of *Herbaspirillum seropedicae* by the N-terminal domain. **FEMS Microbiol. Lett.**, v. 180, p. 157-161, 1999b.

MOSKALEVA, O.; MELNIK, B.; GABDULKHAKOV, A.; GARBER, M.; NIKONOV, S.; STOLBOUSHKINA, E.; NIKULIN, A. The structures of mutant forms of Hfq from *Pseudomonas aeruginosa* reveal the importance of the conserved His57 for the protein hexamer organization. **Acta Cryst. F**, v. 66, p. 760-764, 2010.

MÜNCH, R.; HILLER, K.; GROTE, A.; SCHEER, M.; KLEIN, J.; SCHOBERT, M.; JAHN, D. Virtual Footprinting and PRODORIC: an integrative framework for regulon prediction in prokaryotes. **Bioinformatics**, v. 21, p. 4187-4189, 2005.

NAKAO, H.; WATANABE, H.; NAKAYAMA, S.; TAKEDA, T. *yst* gene expression in *Yersinia enterocolitica* is positively regulated by a chromosomal region that is highly homologous to *Escherichia coli* host factor 1 gene (*hfq*). **Mol. Microbiol.**, v. 18, p. 859-865, 2005.

NEUMANN, L.; SPINOZZI, F.; SINIBALDI, R.; RUSTICHELLI, F.; PÖTTER, M.; STEINBÜCHEL, A. Binding of the Major Phasin, PhaP1, from *Ralstonia eutropha* H16 to Poly(3-Hydroxybutyrate) Granules. **J. Bacteriol.**, v. 190, p. 2911–2919, 2008.

NIELSEN, J.S.; BØGGILD, A.; ANDERSEN, C.B.F.; NIELSEN, G.; BOYSEN, A.; BRODERSEN, D.E.; VALENTIN-HANSEN, P. An Hfq-like protein in *archaea*: Crystal structure and functional characterization of the Sm protein from *Methanococcus jannaschii*. **RNA**, v. 13, p. 2213-2223, 2007.

NOINDORF, L.; BONATTO, A.C.; MONTEIRO, R.A.; SOUZA, E.M.; RIGO, L.U.; PEDROSA, F.O.; STEFFENS, M.B.; CHUBATSU, L.S. Role of PII proteins in nitrogen fixation control of *Herbaspirillum seropedicae* strain SmR1. **BMC Microbiol.**, v. 11, p. 11-8, 2011.

NOINDORF, L.; REGO, F.G.M.; MONTEIRO, R.A.; WASSEM, R.; CRUZ, L.M.; RIGO, L.U.; SOUZA, E.M.; STEFFENS, M.B.R.; PEDROSA, F.O.; CHUBATSU, L.S. Characterization of the *orf1glnKamtB* operon of *Herbaspirillum seropedicae*. **Arch. Microbiol.**, v. 185, p. 55-62, 2006.

OBST, M.; STEINBÜCHEL, A. Cyanophycin – an Ideal bacterial Nitrogen Storage Material with Unique Chemical Properties. **Microbiol. Monogr.**, p. 167-193, 2006.

OJUMU, T.V.; YU, J.; SOLOMON, B.O. Production of Polyhydroxyalkanoates, a bacterial biodegradable polymer. **Afr. J. Biotechnol.**, v. 3, p. 18-24, 2004.

OLSEN, A.S.; MØLLER-JENSEN, J.; BRENNAN, R.G.; VALENTIN-HANSEN, P. C-Terminally Truncated Derivatives of *Escherichia coli* Hfq Are Proficient in Riboregulation. **J. Mol. Biol.**, v. 404, p. 173-182, 2010.

PACE, C.N.; VAJDOS, F.; FEE, L.; GRIMSLEY, G.; GRAY, T. How to measure and predict the molar absorption coefficient of a protein. **Protein Science**, v. 4, p. 2411-2423, 1995.

PAPAGEORGIU, A.C.; HERMAWAN, S.; SINGH, C.B.; JENDROSSEK, D. Structural Basis of Poly(3-Hydroxybutyrate) Hydrolysis by PhaZ7 Depolymerase from *Paucimonas lemoignei*. **J. Mol. Biol.**, v. 382, p. 1184–1194, 2008.

PARRY, D.A.D.; BAKER, E.N. Biopolymers. **Rep. Prog. Phys.**, v. 47, p. 1133-1232, 1984.

PEDROSA, F.O.; MONTEIRO, R.A.; WASSEM, R.; CRUZ, L.M.; AYUB, R.A.; COLAUTO, N.B.; FERNANDEZ, M.A.; FUNGARO, M.H.; GRISARD, E.C.; HUNGRIA, M. *et al.* Genome of *Herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grasses. **PLoS Genet.**, v. 7, 2011.

PEDROSA, F.O.; BENELLI, E.M.; YATES, M.G.; WASSEM, R.; MONTEIRO, R.A.; KLASSEN, G.; STEFFENS, M.B.R.; SOUZA, E.M.; CHUBATSU, L.S.; RIGO, L.U. Recent developments in the structural organization and regulation of nitrogen fixation genes in *Herbaspirillum seropedicae*. **J. Biotechnol.**, v. 91, p. 189-195, 2001.

PEDROSA, F.O.; TEIXEIRA, K.R.S.; MACHADO, I.M.P.; STEFFENS, M.B.R.; KLASSEN, G.; BENELLI, E.M.; FUNAYAMA, S.; RIGO, L.U.; ISHIDA, M.L.; YATES, M.G.; SOUZA, E.M. Structural organization and regulation of the *nif* genes of *Herbaspirillum seropedicae*. **Soil Biol. Biochem.**, v. 29, p. 843-846, 1997.

PERALTA-GIL, M.; SEGURA, D.; GUZMAN, J.; SERVIN-GONZALEZ, L.; ESPIN, G. Expression of the *Azotobacter vinelandii* poly-beta-hydroxybutyrate biosynthetic *phbBAC* operon is driven by two overlapping promoters and is dependent on the transcriptional activator PhbR. **J. Bacteriol.**, v.184, p.5672-5677, 2002.

PERSUHN, D.C.; SOUZA, E.M.; STEFFENS, M.B.; PEDROSA, F.O.; RIGO, L.U. The transcriptional activator NtrC controls the expression and activity of glutamine synthetase in *Herbaspirillum seropedicae*. **FEMS Microbiol. Lett.**, v. 192, p. 217-221, 2000.

PIEPER-FÜRST, U.; MADKOUR, M.H.; MAYER, F.; STEINBÜCHEL, A. Identification of the region of a 14-kilodalton protein of *Rhodococcus ruber* that is responsible for the binding of this phasin to polyhydroxyalkanoic acid granules. **J. Bacteriol.**, v. 177, p. 2513–2523, 1995.

PIEPER-FURST, U.; MADKOUR, M.H.; MAYER, F.; STEINBUCHHEL, A. Purification and Characterization of a 14-Kilodalton Protein That Is Bound to the Surface of Polyhydroxyalkanoic Acid Granules in *Rhodococcus ruber*. **J. Bacteriol.**, v. 176, p. 4328-4337, 1994.

PIMENTEL, J.P.; OLIVARES, F.; PITARD, R.M.; URQUIAGA, S.; AKIBA, F.; DOBEREINER, J. Dinitrogen fixation and infection of grass leaves by *Pseudomonas rubrisulalbicans* and *Herbaspirillum seropedicae*. **Plant and Soil**, v. 137, p. 61-65, 1991.

POTTER, M.; MADKOUR, M.H.; MAYER, F.; STEINBÜCHEL, A. Regulation of phasin expression and polyhydroxyalkanoate (PHA) granule formation in *Ralstonia eutropha* H16. **Microbiology**, v. 148, p. 2413-2426, 2002.

POTTER, M.; MULLER, H.; REINECKE, F.; WIECZOREK, R.; FRICKE, F.; BOWIEN, B.; FRIEDRICH, B.; STEINBÜCHEL, A. The complex structure of polyhydroxybutyrate (PHB) granules: four orthologous and paralogous phasins occur in *Ralstonia eutropha*. **Microbiology**, v. 150, p. 2301-2311, 2004.

POTTER, M.; STEINBÜCHEL, A. Poly(3-hydroxybutyrate) Granule-Associated Proteins: Impacts on Poly(3-hydroxybutyrate) Synthesis and Degradation. **Biomacromolecules**, v. 6, p. 552-560, 2005.

POTTER, M.; STEINBÜCHEL, A. Biogenesis and Structure of Polyhydroxyalkanoate Granules. **Microbiol. Monogr.**, p. 109-136, 2006.

PRIETO, M.A.; BUHLER, B.; JUNG, K.; WITHOLT, B.; KESSLER, B. PhaF, a polyhydroxyalkanoate-granule-associated protein of *Pseudomonas oleovorans* GPo1 involved in the regulatory expression system for *pha* genes. **J. Bacteriol.**, v. 181, p. 858-868, 1999.

PUTNAM, C.D.; GREG, M.H.; HURA, L.; TAINER, J.A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures. **Quarterly Reviews of Biophysics**, v. 40, p. 191-285, 2007.

RATCLIFF, W.C.; KADAM, S.V.; DENISON, R.F. Poly-3-hydroxybutyrate (PHB) supports survival and reproduction in starving rhizobia. **FEMS Microbiol. Ecol.**, v. 65, p. 391-399, 2008.

REGO, F. G. M.; CHUBATSU, L. S.; PEDROSA, F. O.; YATES, M. G. ; WASSEM, R.; STEFFENS, M. B. R.; RIGO, L.; SOUZA, E. M.. The expression of *nifB* gene from *Herbaspirillum seropedicae* is dependent upon the NifA and RpoN proteins. **Can. J. Microbiol.**, 2006.

REHM, B.H.A. Polyester synthases: natural catalysts for plastics. **Biochem. J.**, v. 376, p. 15–33, 2003.

REHM, B.H.A. Genetics and biochemistry of polyhydroxyalkanoate granule self-assembly: The key role of polyester synthases. **Biotechnol. Lett.**, v. 28, p. 207-213, 2006.

ROBERTSON, G.T.; ROOP, R.M.J. The *Brucella abortus* host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice. **Mol. Microbiol.**, v. 34, p. 690-700, 1999.

ROCCHIA, W.; ALEXOV, E.; HONIG, B. Extending the Applicability of the Nonlinear Poisson-Boltzmann Equation: Multiple Dielectric Constants and Multivalent Ions. **J. Phys. Chem. B**, v. 28, p. 6507-6514, 2001.

RODRIGUES, T.E.; SOUZA, V.E.; MONTEIRO, R.A.; GERHARDT, E.C.; ARAUJO, L.M.; CHUBATSU, L.S.; SOUZA, E.M.; PEDROSA, F.O.; HUERGO, L.F. In vitro interaction between the ammonium transport protein AmtB and partially uridylylated forms of the P(II) protein GlnZ. **Bioch. Biophys. Acta.**, v. 1814, p. 1203-9, 2011.

RONCATO-MACCARI, L.D.B.; RAMOS, H.J.O.; PEDROSA, F.O.; ALQUINI, Y.; CHUBATSU, L.S.; YATES, M.G.; RIGO, L.; STEFFENS, M.B.R.; SOUZA, E.M. Endophytic *Herbaspirillum seropedicae* expresses nif genes in gramineous plants. **FEMS Microbiol. Ecol.**, v. 45, p. 39-47, 2003.

RUTHERFORD, K.; PARKHILL, J.; CROOK, J.; HORSNELL, T.; RICE, P.; RAJANDREAM, M. A.; BARRELL, B. Artemis: sequence visualization and annotation. **Bioinformatics**, v. 16, p. 944-945, 2000.

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

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-termination inhibitors. **Proc. Nat. Acad. Sci. USA**, v. 74, p. 5463-5467, 1977.

SANTOS, N.; CASTANHO, M. An Overview of the Biophysical Applications of Atomic Force Microscopy. **Biophysical Chemistry**, v. 107, p. 133–149, 2004.

SAUER, E.; WEICHENRIEDER, O. Structural basis for RNA 3'-end recognition by Hfq. **Proc. Nat. Acad. Sci. USA**, v. 108, p. 13065–13070, 2011.

SCHMIDT, M.A.; SOUZA, E.M.; BAURA, V.; WASSEM, R.; YATES, M.G.; PEDROSA F.O.; MONTEIRO, R.A. Evidence for the endophytic colonization of *Phaseolus vulgaris* (common bean) roots by the diazotroph *Herbaspirillum seropedicae*. **Braz. J. Med. Biol. Res.**, v. 44, p. 182-185, 2011.

SCHRODINGER, L.L.C. The PyMOL Molecular Graphics System, Version 1.3r1, 2010.

SCHUBERT, P.; STEINBÜCHEL, A.; SCHLEGEL, H.G. Cloning of the *Alcaligenes eutrophus* genes for synthesis of poly- $\beta$ -hydroxybutyric acid (PHB) and synthesis of PHB in *Escherichia coli*. **J. Bacteriol.**, v.170, p. 5837-5847, 1988.

SCHUMACHER, M.A.; PEARSON, R.F.; MOLLER, T.; VALENTIN-HANSEN, P.; BRENNAN, R.G. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. **EMBO J.**, v. 21, p. 3546-3556, 2002.

SERRATO, R.V.; SASSAKI, G.L.; CRUZ, L.M.; CARLSON, R.W.; MUSZYŃSKI, A.; MONTEIRO, R.A.; PEDROSA, F.O.; SOUZA, E.M.; IACOMINI, M. Chemical composition of lipopolysaccharides isolated from various endophytic nitrogen-fixing bacteria of the genus *Herbaspirillum*. **Can. J. Microbiol.**, v. 56, p. 342-7, 2010.

SILVA, L.G.; MIGUENS, F.C.; OLIVARES, F.L. *Herbaspirillum seropedicae* and sugarcane endophytic interaction investigated by using high pressure freezing electron microscopy. **Brazilian Journal of Microbiology**, v. 34, p. 69-71, 2003.

SIMON, R.; PRIEFFER, U.; PUHLER, A. A broad host range mobilization system for in vitro genetic engineering: transposon mutagenesis in Gram negative bacteria. **BioTechnology**, v. 1, p. 784-791, 1983.

SONNLEITNER, E.; HAGENS, S.; ROSENAU, F.; WILHELM, S.; HABEL, A.; JÄGER, K.; BLÄSI, U. Reduced virulence of a *hfq* mutant of *Pseudomonas aeruginosa* O1. **Microb. Pathog.**, v. 35, p. 217-228, 2003.

SOUZA, A.L.; INVITTI, A.L.; REGO, F.G.; MONTEIRO, R.A.; KLASSEN, G.; SOUZA E.M.; CHUBATSU, L.S.; PEDROSA, F.O.; RIGO, L.U. The involvement of the *nif*-associated ferredoxin-like genes *fdxA* and *fdxN* of *Herbaspirillum seropedicae* in nitrogen fixation. **J. Microbiol.**, v. 48, p. 77-83, 2010.

SOUZA, E.M.; PEDROSA, F.O.; RIGO, L.U.; MACHADO, H.B.; YATES, M.G. Expression of the *nifA* gene of *Herbaspirillum seropedicae*: role of the NtrC and NifA binding sites and of the -24/-12 promoter element. **Microbiology**, v. 146, p. 1407-1418, 2000.

SOUZA, E.M.; PEDROSA, F.O.; DRUMMOND, M.; RIGO, L.U.; YATES, M.G. Control of *Herbaspirillum seropedicae* NifA activity by ammonium ions and oxygen. **J. Bacteriol.**, v. 181, p. 681-684, 1999.

SOUZA, E.M.; PEDROSA, F.O.; MACHADO, H. B.; DRUMMOND, M.; YATES, M.G. The N-terminus of the NifA protein of *H. seropedicae* is probably involved in sensing of ammonia. In: **Nitrogen Fixation: Fundamentals and Applications**. TIKHONOVICH, I.A.; NIKOLAI, A.P.; VASSILY, I.R.; WILLIAM, E.N. (eds.) X International Congress in Nitrogen Fixation. St. Petersburg, p. 260, 1995.

SOUZA, E.M.; FUNAYAMA, S.; RIGO, L.U.; PEDROSA, F.O. Cloning and characterization of the *nifA* gene from *Herbaspirillum seropedicae* strain Z78. **Can. J. Microbiol.**, v. 37, p. 425-429, 1991a.

SOUZA, E.M.; FUNAYAMA, S.; RIGO, L.U.; YATES, M.G.; PEDROSA, F.O. Sequence and structural organization of a *nifA*-like gene and part of a *nifB*-like gene of *Herbaspirillum seropedicae* strain Z78. **J. Gen. Microbiol.**, v. 137, p. 1511-1522, 1991b.

SPAINK, H.P.; OKKER, R.J.H.; WIJFFELMAN, C.A.; PEES, E.; LUGTENBERG, B.J.J. Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRLJI. **Plant Mol. Biol.** v. 9, p. 27-39, 1987.

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

STEFFENS, M.B.; RIGO, L.U.; FUNAYAMA, S.; SOUZA, E.M.; MACHADO, H.B.; PEDROSA, F.O. Cloning of a *recA*-like gene from the diazotroph *Herbaspirillum seropedicae* strain Z78. **Can. J. Microbiol.**, v. 39, p. 1096-102, 1993.

STEINBÜCHEL, A. Polyhydroxyalkanoic acids. In: **Byrom D (ed) Biomaterials: novel materials from biological sources**. Stockton, New York, pp. 124-213, 1991.

STEINBÜCHEL, A.; VALENTIN, H.E. Diversity of bacterial polyhydroxyalkanoic acids. **FEMS**, v. 128, p.219-228, 1995.

STEINBÜCHEL, A.; SCHLEGEL, H.G. Excretion of pyruvate by mutants of *Alcaligenes eutrophus*, which are impaired in the accumulation of poly( $\beta$ -hydroxybutyric acid) (PHB), under conditions permissive for synthesis of PHB. **Appl. Microbiol. Biotechnol.**, v. 31, p. 168–175, 1989.

STORZ, G.; VOGEL, J.; WASSARMAN, K.M. Regulation by Small RNAs in Bacteria: Expanding Frontiers. **Mol. Cell**, v. 43, 2011.

STORZ, G.; OPDYKE, J.A.; ZHANG, A. Controlling mRNA stability and translation with small, noncoding RNAs. **Curr. Opin. Microbiol.**, v. 7, p. 140-144, 2004.

STUBBE, J.; TIAN, J. Polyhydroxyalkanoate (PHA) homeostasis: the role of the PHA. **Synthese Nat. Prod. Rep.**, v. 20, p. 445–457, 2003.

STUBBE, J.; TIAN, J.; HE, A.; SINSKEY, A.J.; LAWRENCE, A.G.; LIU, P. NONTEMPLATE-DEPENDENT POLYMERIZATION PROCESSES: Polyhydroxyalkanoate Synthases as a Paradigm. **Annu. Rev. Biochem.** v. 74, p. 433-480, 2005.

SUN, J.; PENG, X.; VAN IMPE, J.; VAN DER LEYDEN, J. The *ntrB* and *ntrC* genes are involved in the regulation of poly-3-hydroxybutyrate biosynthesis by ammonia in *Azospirillum brasilense* Sp7. **Appl. Environ. Microbiol.**, v. 66, p. 113–117, 2000.

SVERGUN, D.I.; PETOUKHOV, M.V.; KOCH, M.H.J. Determination of Domain Structure of Proteins from X-Ray Solution Scattering. **Biophys. J.**, v. 80, p. 2946-2953, 2001.

SVERGUN, D.I.; BARBERATO, C.; KOCH, M.H. CRY SOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. **J. Appl. Cryst.**, v. 28, p. 768-773, 1995.

SVERGUN, D.I. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. **J. Appl. Cryst.**, v. 25, p. 495-503, 1992.

TADRA-SFEIR, M.Z.; SOUZA, E.M.; FAORO, H.; MULLER-SANTOS, M.; BAURA, V.A.; TULESKI, T.R.; RIGO, L.U.; YATES, M.G.; WASSEM, R.; PEDROSA, F.O.; MONTEIRO R.A. Naringenin regulates expression of genes involved in cell wall synthesis in *Herbaspirillum seropedicae*. **Appl. Environ. Microbiol.**, v. 77, p. 2180-3, 2011.

TAMURA, K.; PETERSON, D.; PETERSON, N.; STECHER, G.; NEI, M.; KUMAR, S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. **Mol. Biol. Evol.**, v. 28, p. 2731–2739, 2011.

THOMPSON, J.D.; HIGGINS, D.G.; GIBSON, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucl. Acids Res.**, v. 11, p. 4673–4680, 1994.

TIAN, S.J.; LAI, W.J.; ZHENG, Z.; WANG, H.X.; CHEN, G.Q. Effect of over-expression of phasin gene from *Aeromonas hydrophila* on biosynthesis of copolyesters of 3-hydroxybutyrate and 3-hydroxyhexanoate. **FEMS Microbiol. Lett.**, v. 244, p. 19–25, 2005.

TOKIWA, Y.; CALABIA, B.P.; UGWU, C.U.; AIBA, S. Biodegradability of Plastics. **Int. J. Mol. Sci.**, v. 10, p. 3722-3742, 2009.

TRAINER, M.; CHARLES, T. The role of PHB metabolism in the symbiosis of rhizobia with legumes. **Appl. Microbiol. Biotechnol.**, v. 71, p. 377–386, 2006.

TREWHELLA, J. Insights into biomolecular function from small-angle scattering. **Curr. Opin. Struct. Biol.**, v. 7, p. 702-708, 1997.

TSUI, H.C.; LEUNG, H.C.E.; WINKLER, M.E. Characterization of broadly pleiotrophic phenotypes caused by an Hfq insertion mutation in *Escherichia coli* K-12. **Mol. Microbiol.**, v. 13, p. 35-49, 1994.

TSUTAKAWA, S.E.; HURA, G.L.; FRANKEL, K.A.; COOPER, P.K.; TAINER, J.A. Structural analysis of flexible proteins in solution by small angle X-ray scattering combined with crystallography. **Journal of Structural Biology**, v. 158, p. 214-223, 2007.

VOGEL, J. ; LUISI, B.F. Hfq and its constellation of RNA. **Nature Microbiol.**, v. 9, p. 578-589, 2011.

VOLKOV, V.V., SVERGUN, D. I. Uniqueness of *ab initio* shape determination in small-angle scattering. **J. Appl. Cryst.**, v. 36, p. 860–864, 2003.

WACHI, M.; TAKADA, A.; NAGAI, K. Overproduction of the outer-membrane proteins FepA and FhuE responsible for iron transport in *Escherichia coli* *hfq::cat* mutant. **BBRC**, v. 264, p. 525-529, 1999.

WANG, W.; WANG, L.; ZOU, Y.; ZHANG, J.; GONG, Q.; WU, J.; SHI, Y. Cooperation of *Escherichia coli* Hfq hexamers in DsrA binding. **Genes Dev.**, v. 25, p. 2106-2117, 2011.

WASSEM, R.; SOUZA, E.M.; YATES, M.G.; PEDROSA, F.O.; BUCK, M. Two roles for integration host factor at an enhancer-dependent *nifA* promoter. **Mol. Microbiol.**, v. 35, p. 756-764, 2000.

WILUSZ, C.J.; WILUSZ, J. Eucaryotic Lsm proteins: lessons from bacteria. **Nat. Struct. Mol. Biol.**, v. 12, p. 1031-1036, 2005.

WESTERMEIER, R.; NAVEN, T. Part III: Course Manual, Step 9: In-gel digestion. In: WESTERMEIER, R.; NAVEN, T. (Ed.). **Proteomics in Practice. A laboratory Manual of Proteome Analysis**. Wiley-VCH, p.261, 2004.

WIECZOREK, R.; PRIES, A.; STEINBÜCHEL, A.; MAYER, F. Analysis of a 24-Kilodalton Protein Associated with the Polyhydroxyalkanoic Acid Granules in *Alcaligenes eutrophus*. **J. Bacteriol.**, v. 177, p. 2425–2435, 1995.

WILLIAMSON, D.H.; WILKINSON, J.F. The isolation and estimation of the  $\beta$ -hydroxybutyrate inclusions of *Bacillus* species. **J. Gen. Microbiol.**, v. 19, p. 198–209, 1958.

YOUNG, J.P.W. Phylogenetic classification of nitrogen-fixing organisms. In: STACEY, G.; BURRIS, R.H.; EVANS, H.J. (eds.). **Biological nitrogen fixation. New York: Chapman & Hall**, p. 43–88, 1992.

YORK, G.M.; STUBBE, J.; SINSKEY, A.J. The *Ralstonia eutropha* PhaR protein couples synthesis of the PhaP phasin to the presence of polyhydroxy-butyrate in cells and promotes polyhydroxybutyrate production. **J. Bacteriol.**, v. 184, p. 59–66, 2002.

YORK, G.M.; JUNKER, B.H.; STUBBE, J.; SINSKEY, A.J. Accumulation of the PhaP Phasin of *Ralstonia eutropha* Is Dependent on Production of Polyhydroxybutyrate in Cells. **J. Bacteriol.**, v. 183, p. 4217-4226, 2001.

ZHANG, Y.; HONG, G. Post-transcriptional regulation of NifA expression by Hfq and RNase E complex in *Rhizobium leguminosarum* bv. *viciae*. **Acta Biochim. Biophys. Sin.**, v.41, p. 719-730, 2009.

ZINN, M.; WITHOLT, B.; EGLI, T. Occurrence, synthesis and medical application of bacterial polyhydroxyalkanoate. **Advan. Drug. Deliv. Rev.**, v. 53, p. 5-21, 2001.