# Design of Bacterial Polyester Beads for Recombinant Protein Production, Biomolecule Separation and Detection

A thesis presented in partial fulfilment of the
requirements of the degree of
Doctor of Philosophy
in
Microbiology

at Massey University, Palmerston North,
New Zealand.

Jinping Du
2018

# Abstract

Protein recovery and biomolecule detection are commonly required for scientific research as well as industrial activities. However, it is generally complicated and costly either to produce and purify recombinant proteins (especially therapeutic proteins) from engineered *Escherichia coli* cells, or to directly separate proteins or detect other biomolecules from natural sources. Here the PHA synthase (PhaC) mediated polyhydroxyalkanoate (PHA) bead display technology was explored as a solution to these problems by developing streamlined processes with less complex steps to achieve protein recovery and biomolecule detection.

Firstly, by fusing a target protein to PhaC via a self-cleavable linker tag of either sortase (sortase A from *Staphylococcus aureus*) or intein (DnaB mini intein from *Synechocystis sp.* PCC 6803), new self-cleavable recombinant protein production and purification resins were developed. It was shown that the PhaC fusion could mediate *in vivo* production of PHA beads displaying the target protein. Functional target protein could be obtained at high purity from isolated PHA beads by incubation with $CaCl_2$ and triglycine (in the case of the self-cleavable sortase tag) or by a pH shift to 6 (in the case of the self-cleavable intein tag). Six recombinant proteins were successfully produced and purified via the intein approach, including 3 model proteins (*Aequorea victoria* green fluorescent protein (GFP), *Mycobacterium tuberculosis* vaccine candidate Rv1626, and the synthetic immunoglobulin G (IgG) binding ZZ domain of protein A derived from *Staphylococcus aureus*) and 3 therapeutic proteins (human tumour necrosis factor alpha (TNFα), human interferon alpha-2b (IFNα2b), and human granulocyte colony-stimulating factor (G-

CSF)). Of these, TNFα and IFNα2b were also successfully produced and purified via the sortase approach.

Secondly, *in vivo* one-step production of PHA affinity resins was achieved by fusing to PhaC differently customised OBody ligands. These ligands were previously engineered by other groups from the OB-fold domain of aspartyl-tRNA synthetase (aspRS) from *Pyrobactulum aerophilum*, by using phage display technology, to have specific binding affinities to biomolecules of interest. The resulting recombinant OBody beads were used for lysozyme sepration from a complex substrate, and for progesterone (P4) binding. Further optimisation of the P4 binding condition is necessary before the OBody bead system can be used for P4 detection in bovine milk. However, recombinant immobilisation of OBody ligands on the surface of PHA beads expands not only the attractiveness of these emerging OBody scaffolds, but also the utility scope of PHA beads as affinity resins.

# Acknowledgements

To start a PhD study at my own expense is not a choice without pressure, especially for one lasting over 6 years. All who have accompanied me through this journey are greatly appreciated.

First I would like to thank my supervisor Professor Bernd Rehm. I could not have gone so far without his continuous support and endless patience. Bernd has kept looking for resources to relieve my financial burden, he kindly reimbursed my tuition fees for the first year, offered me a research assistant position at the final stage, and allowed my one-year suspension in between for working with Polybatics Ltd..

I would also like to thank my co-supervisor Professor Rosie Bradshaw for providing invaluable quick feedback and countless careful proofreading cycles on this thesis, regardless of early mornings, late nights, weekends or holidays. The thesis could never come into birth without her lasting encouragement.

I value the expert guidance on my project from my previous co-supervisor Dr Mark Patchett who has just retired from Massey. I specially thank Trevor Loo for his technical help on protein mass spectrometry analysis. I thank Dr Zoe Jordens and Professor Kathryn Stowell for their valuable advice on thesis writing.

I thank Tracy Thompson and Polybatics Ltd. for providing me the working opportunity as a research technician from February 2012 to August 2013, which has gained me much research experience as well as financial support towards my PhD completion.

I would also like to take this opportunity to thank Professor Ipsita Roy, Associate Professor Evelyn Sattlegger and Dr Nigel Larsen - my thesis examiners, for their very helpful comments and suggestions.

I am grateful to all the Rehm lab members and former Polybatics staff for their support and assistance. Special thanks go to Shuxiong Chen, Dr Anika Jahns, Dr Iain Hay, Dr

Katrin Grage, Dr David Hooks, Dr Jenny Drapes, Dr Jason Lee, Dr Majela Gonzalez-Miro, Dr Patricia Rubio Reyes and Ogura Kampachiro for their precious time and generous help.

I thank the help and support from Ann Truter, Cynthia Cresswell, Debra Cresswell, Paul Hocquard and Peter Lewis for making my life in the IFS family a lot easier.

I treasure the tremendous help from Hongping Jin, Jing Wang and Shuguang Zhang when I first came to New Zealand. I appreciate the spiritual inspiration from Lynette Spencer, Isaac Fung and Jun Zhou when I was at a loss. I thank my old friends Shicong Jia and Yizhe Cheng for all the long-distance support from China.

My special gratitude goes to my parents and my mother-in-law for their constant love. My deepest thoughts go to my deceased father-in-law who had been longing to but could not witness my graduation now.

I owe a lot to my husband Zhujun Wang for being my on call driver to and from Massey at countless weekends and late nights, and for taking more responsibility in looking after the family when I have been occupied by experiments or writing up. I would not have the courage to begin nor complete my PhD study without his unconditional love and support.

My final thanks go to my two lovely daughters, Jiaqi and Anya, for being considerate when mum has to take care of those naughty *E. coli* and bead kids in the lab, for believing in and being patient for mum's completion of her PhD study. You are angels shining in my life with your bright smiles.

# Table of contents

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| °C | Degrees Celsius |
| ABC solution | 50 mM ammonium bicarbonate (ABC) solution |
| ACP | Acyl carrier protein |
| AGE | Agarose gel electrophoresis |
| $Amp^r$ | Ampicillin resistance |
| ANOVA | Analysis of variance |
| APS | Ammonium persulfate |
| aspRS | Aspartyl-tRNA synthetase |
| B7 | Progesterone (P4) binding OBody B7, a synthetic peptide engineered based on the OB-fold domain of aspRS from *Pyrobaculum aerophilum* |
| bp | Base pair(s) |
| Bis-Tris | Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane |
| BSA | Bovine serum albumin |
| BSA-P4 | BSA conjugated progesterone |
| $Ca^{2+}$ | Calcium ion |
| CAT | Chloramphenicol acetyltransferase |
| CBD | Chitin binding domain derived from pTwin1 vector (NEB) |
| CLSM | Confocal laser scanning fluorescence microscopy |
| CD | Circular dichroism |
| $Cm^r$ | Chloramphenicol resistance |
| CoA | Coenzyme A |
| D7 | P4 binding OBody P4013-D7, a synthetic peptide engineered based on the OB-fold domain of aspRS from *P. aerophilum* |
| DABCYL | 4-((4-(dimethylamino)phenyl)azo)benzoic acid as a FRET receptor quencher |
| DDA | Dimethyldioctyldecyl ammonium bromide as a vaccine adjuvant |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxynucleotide triphosphate |
| DsbC | Disulphide isomerase |
| DTT | Dithiothreitol |
| EDANS | 5-((2-Aminoethyl)amino)naphthalene-1-sulfonic acid as a FRET donor fluorophore |
| EDTA | Ethylenediaminetetraacetic acid |
| EGFP | Enhanced green fluorescent protein |
| ELISA | Enzyme-linked immunosorbent assay |
| Erv1p | Sulfhydryl oxidase |
| FIA | Fluorescence immunoassay |
| FLAG tag | DYKDDDDK octapeptide used used as a tag for affinity chromatography |
| FM | Fluorescent Microscopy |
| FRET | Förster resonance energy transfer |
| g | Gram |

| | |
|---|---|
| *g* | G force(s) |
| GAPs | Granule Associated Proteins |
| G-CSF | Human granulocyte colony-stimulating factor, short isoform without signal peptide or VSE after the QEKL residue |
| GFP | Green fluorescent protein from *Aequorea victoria* |
| GMOs | Genetically modified organisms |
| *gor* | Gene coding for glutaredoxin reductase |
| GRAS | Generally recognized as safe |
| G-scar | The extra glycine residue left on a protein upon self-cleavage of sortase tag |
| GST tag | Glutathione S-transferase used as a tag for affinity chromatography |
| h | Hour(s) |
| HA | (*R*)-3-hydroxyacyl |
| 3-HB | 3-hydroxybutyrate |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| His tag | Polyhistidine used as a tag for immobilized metal affinity chromatography |
| HRP | Horse radish peroxidise |
| *Hsa* PolII intein | Salt-dependent PolII intein derived from *Halobacterium salinarum* |
| *Hut* MCM2 intein | Salt-dependent MCM2 intein derived from *Halorhabdus utahensis* |
| IB / IBs | Inclusion body / Inclusion bodies |
| IFNα2b | Human interferon alpha 2b, without signal peptide |
| IgG | Immunoglobulin G |
| IPTG | Isopropyl-β-D-thiogalactopyranoside |
| kb | Kilobase(s) |
| $K_D$ | Equilibrium dissociation constant |
| kDa | Kilo Daltons |
| LacZ | β-galactosidase |
| LB | Luria-Bertani broth |
| LC-MS/MS | Liquid Chromatography with tandem mass spectrometry analysis |
| LFI | Lateral flow immunoassay |
| Linker | A flexible DNA linker inserted between 3’ end of *phaC* gene and 5’ end of a target coding region to facilitate the folding of the target protein or peptide |
| LPXTG / LPETG | Sortase recognises this five amino acid signal and cleaves between the T and G in the presence of $Ca^{2+}$ +/- triglycine |
| µg | Microgram |
| µl | Microlitre |
| µM | Micromolar |
| M | Molar |
| MBP | Maltose binding protein |
| mg | Milligram |
| min | Minute(s) |
| ml | Millilitre |
| MOPS | 3-(N-morpholino) propanesulfonic acid |
| *Mth* RIR1 intein | RIR1 intein derived from *Methanobacterium thermoautotrophicum* (pTWIN2 vector, NEB), which self-cleaves upon addition of thiol reagents |

| | |
|---|---|
| *Mtu* ΔI-CM intein | Engineered pH-senstive RecA inteins derived from *Mycobacterium tuberculosis* recA (*Mtu* RecA) |
| *Mxe* GyrA intein | GyrA mini intein derived from *Mycobacterium xenopi* (pTWIN1 vector, NEB), which self-cleaves upon addition of thiol reagents |
| ng | Nanogram |
| *Npu* DnaE intein | pH-senstive DnaE intein derived from *Nostoc punctiforme* |
| NusA | N utilisation substance protein A |
| O6 | Lysozyme binding OBody L200EP-06, a synthetic peptide engineered based on the OB-fold domain of aspRS from *P. aerophilum* |
| OBody | A synthetic non-antibody protein scaffold engineered based on the OB-fold domain of aspRS from *P. aerophilum* |
| OPD | *o*-Phenylenediamine dihydrochloride |
| ori | Origin of replication |
| PAb | Polyclonal antibody |
| P4 | Progesterone |
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| PHA | Polyhydroxyalkanoate |
| PhaA | β-ketothiolase |
| $PHA_{MCL}$ | Medium chain length PHA |
| $PHA_{SCL}$ | Short chain length PHA |
| $PHA_{SCL-MCL}$ | PHA containing mixtures of C3-C5 and C6-C14 monomers |
| PhaB | Acetoacetyl-CoA reductase |
| *phaCAB* | PHA biosynthesis gene operon |
| PhaC | PHA synthase |
| PhaE | Type III PHA synthase subunit |
| PhaM | Regulatory protein |
| PhaP | Phasin protein |
| PhaR | Transcriptional regulator that control the transcription of phasin and the synthesis of PHA beads; Or Type IV PHA synthase subunit |
| PhaZ | Intracellular PHA depolymerase |
| PHB | Polyhydroxybutyrate, poly(3-hydroxybutyric acid) |
| RBS | Ribosome binding site |
| RIA | Radioimmunoassay |
| rpm | Revolutions per minute |
| SD | Standard deviation |
| SDS | Sodium dodecyl sulphate |
| SDS-PAGE | Sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| SG-linker | A triplicate SGGGG linker introduced at C-terminus of PhaC to maintain its hydrophobic environment for a proper PHA synthase functionality |
| S-S bond | disulphide bond |
| *Ssp* Intein | DnaB helicase mini intein from *Synechocystis* sp. strain PCC6803, derived from pTwin1 vector (NEB), which self-cleaves when pH drops to 6 |
| SrtA | Sortase A from *Staphylococcus aureus* minus the N-terminal membrane anchor region |

| | |
|---|---|
| SrtAc / $SrtA_{\Delta 59}$ | catalytic core of SrtA (amino acids 60-206) |
| Strep tag | AWRHPQFGG or WSHPQFEK peptide used as a tag for affinity chromatography |
| T7 promoter | Promoter for bacteriophage T7 RNA polymerase |
| T7 terminator | Transcription terminator for bacteriophage T7 RNA polymerase |
| TB | Terrific Broth |
| TBE | Tris-Borate-EDTA buffer |
| TEM | Transmission Electron Microscopy |
| TEMED | N, N, N', N'-tetramethylethyl-endiamine |
| $Tet^r$ | Tetracycline resistance |
| TEV | Tobacco Etch Virus protease |
| TNFα | Human tumour necrosis factor alpha, soluble form |
| Tris | Trishydroxymethylaminomethane |
| *trxB* | Gene coding for thioredoxin reductase |
| v/v | Volume / volume ratio |
| WT | Wild Type |
| w/v | Weight / volume ratio |
| ZZ | Two copies of a IgG binding Z domain of protein A derived from *Staphylococcus aureus* |

Abbreviations (one and three letter codes) for amino acids

| | |
|---|---|
| A / Ala | Alanine |
| C / Cys | Cysteine |
| D / Asp | Aspartic acid |
| E / Glu | Glutamic acid |
| F / Phe | Phenylalanine |
| G / Gly | Glycine |
| H / His | Histidine |
| I / Ile | Isoleucine |
| K / Lys | Lysine |
| L / Leu | L / Leu |
| M / Met | Methionine |
| N / Asn | Asparagine |
| P / Pro | Proline |
| Q / Gln | Glutamine |
| R / Arg | Arginine |
| S / Ser | Serine |
| T / Thr | Threonine |
| V / Val | Valine |
| W / Trp | Tryptophan |
| Y / Tyr | Tyrosine |

# Chapter 1: Introduction

It is of great importance in both academic and commercial sectors to either produce and purify recombinant proteins from engineered *Escherichia coli* cells, or to directly separate proteins or detect other biomolecules from natural sources (Lightfoot & Moscariello 2004; Oliveira & Domingues 2018). However, these processes are usually complicated and costly, and urgently need improvement. Here polyhydroxyalkanoate (PHA) bead display technology was explored for developing streamlined processes with less complicated steps toward protein recovery and biomolecule detection.

## 1.1 Polyhydroxyalkanoate (PHA) and PHA bead display technology

### 1.1.1 Polyhydroxyalkanoate (PHA)

Polyhydroxyalkanoate (PHA), a polymer made up of (*R*)-3-hydroxy fatty acid units joined through ester bonds, is a naturally occurring polyester synthesised by both Gram-negative bacteria (such as some species belonging to the genera *Cupriavidus* and *Pseudomonas*) and Gram-positive bacteria (such as some species of *Bacillus*) (Lu *et al.* 2009; Tan *et al.* 2014), some halophilic Archaea (members of the family *Halobacteriaceae*) (Han *et al.* 2010; Poli *et al.* 2011), as well as some eukaryotic yeast cells such as *Candida tropicalis* (Priji *et al.* 2013; Priji *et al.* 2016). These microorganisms produce intracellular PHA granules as carbon / energy reservoirs. This generally occurs under imbalanced nutrient conditions such as excessive carbon but can also occur under limited nitrogen, phosphorus and oxygen conditions. The PHAs are also mobilised or

degraded under carbon starvation conditions (Grage *et al.* 2009; Rehm 2010; Obruca *et al.* 2017). On average 5 to 10 PHA granules (or beads) with sizes ranging from 50 to 500 nm are accumulated in each cell (Rehm 2007; Grage *et al.* 2009; Koller *et al.* 2010), reaching up to 80% of cell dry weight in fed-batch bacterial fermentations (Madison & Huisman 1999; Chen 2009; Tan *et al.* 2014).

In addition to native PHA producers, an expanding set of genetically modified organisms (GMOs) harbouring appropriate PHA biosynthesis genes (section 1.2.1) have been created to recombinantly produce PHA. These non-native PHA producers include not only prokaryotes such as Gram-negative *Escherichia coli* (Li *et al.* 2007) and Gram-positive *Bacillus subtilis* (Law *et al.* 2003; Wang *et al.* 2006; Singh *et al.* 2009; Lin & Chen 2017), but also eukaryotes including yeasts such as *Saccharomyces cerevisiae* and *Pichia pastoris* (Breuer *et al.* 2002; Poirier *et al.* 2002), insects such as *Spodoptera frugiperda* (Williams *et al.* 1996) and plants such as *Arabidopsis thaliana*, tobacco and sugarcane (Bohmert *et al.* 2000; Bohmert-Tatarev *et al.* 2011; McQualter *et al.* 2016).

The most well-known and well-studied form of PHA, poly(3-hydroxybutyric acid) (PHB), which is based on 3-hydroxybutyrate (3-HB) monomers, was also the first recorded bacterial PHA polymer, and was identified by Maurice Lemoigne from *Bacillus megaterium* more than 90 years ago (Albuquerque & Malafaia 2018). So far, over 150 different (*R*)-3-hydroxy fatty acids have been reported as PHA constituents (Parlane *et al.* 2017). This number is growing through supply of tailor-made monomers or monomer precursors as carbon substrates for bacterial PHA production (Jia *et al.* 2016), through genetic modification or metabolic engineering of PHA producing strains (Chen *et al.* 2015), or through chemical or physical modification of existing PHA polymers (Levine *et al.* 2016). The monomer diversity and vast modification possibilities lead to a range of

PHAs with different thermal and mechanical properties. PHAs, occurring as either homopolymers or copolymers, contain typically 600 - 35,000 (*R*)-3-hydroxy fatty acid monomers (Khanna & Srivastava 2005), and fall into two major groups depending on the monomer size. The two groups are (1) short chain length PHA ($PHA_{SCL}$), the classic and the most-investigated, where the monomer contains 3-5 carbon atoms (Wang *et al.* 2016), and (2) medium chain length PHA ($PHA_{MCL}$), typically produced by *Pseudomonas* species where the monomer contains 6-14 carbon atoms (Rai *et al.* 2011). Generally $PHA_{SCL}$ is hard and brittle with a high level of crystallinity (40-80%) and a high melting temperature (80 - 180°C). In contrast, $PHA_{MCL}$ is more elastomeric with low crystallinity (20 - 40%) and lower melting point (30 - 80°C) (Zinn & Hany 2005). Recently, $PHA_{SCL-MCL}$ copolymers or terpolymers (PHA containing two or three mixtures of C3-C5 and C6-C14 monomers) are attracting industrial interest due to better thermal and physical properties than either $PHA_{SCL}$ or $PHA_{MCL}$ (Nomura *et al.* 2004; Bhubalan *et al.* 2010; Balakrishna Pillai & Kumarapillai 2017).

Due to their diversity as well as their biodegradability, biocompatibility and thermoplastic nature, PHA polymers could be potentially used as renewable plastics, biomedical materials (such as sutures, repair patches, stents, tissue engineering scaffolds or drug delivery carriers), biochemical precursors or even feed supplements (Chen 2009; Somleva *et al.* 2013; Ali & Jamil 2016; Michalak *et al.* 2017). However, all these applications are based on extracted PHAs themselves that are no longer in granule form, and thus without any granule associated proteins (GAPs).

PHAs obtained via fermentation are said to be 5-10 times more expensive than traditionaly oil based or chemically synthesised plastics, thus PHA usage is largely limited to biomedical / biotech areas (Parlane *et al.* 2016). Granule form can justify the

high production cost, because it can add additional protein functionalities of high value. Only when beads are in gruanue form, the intact GAPs could be potentially used to tag various protein ligands of interest onto beads through genetic recombinant technologies, thus add up bead functionality and broaden their usage in ligand related applications, such as protein recovery, biomolecule detection, disease diagnosis and vaccine delivery.

In native PHA-producing cells, hydrophobic granular PHA polymers are surrounded by proteins that are responsible for PHA metabolism, including enzymes such as PHA synthases (PhaC) that catalyse polymerisation of PHA, PHA depolymerases (PhaZ) responsible for PHA degradation / mobilisation, structural protein phasins (PhaP) that control the size and number of PHA beads, as well as regulatory proteins (PhaR) that control the synthesis of PHA beads and the transcription of PhaP (Rehm 2010; Parlane *et al.* 2017).

These granule associated proteins (GAPs), especially PHA synthase (PhaC), have been used extensively through recombinant DNA technologies as anchors to display functional proteins on the surface of either recombinantly, natively or even chemically produced PHA beads. The resulting PHA beads could have many uses, depending on the functionality of the surface displayed proteins. For example, protein immobilisation (Moldes *et al.* 2004; Peters & Rehm 2006; Rasiah & Rehm 2009; Blatchford *et al.* 2012; Hooks *et al.* 2013; Robins *et al.* 2013; Chen *et al.* 2014; Lee *et al.* 2014; Jahns & Rehm 2015; Seo *et al.* 2016; Bello-Gil *et al.* 2017; Hafizi *et al.* 2017; Ran *et al.* 2017), recombinant protein production and purification (Banki *et al.* 2005; Barnard *et al.* 2005; Wang *et al.* 2008; Geng *et al.* 2010; Zhang *et al.* 2010; Grage *et al.* 2011; Zhou *et al.* 2011), biomolecule recognition such as protein separation, endotoxin removal or antigen detection (Brockelbank *et al.* 2006; Grage & Rehm 2008; Peters & Rehm 2008; Lewis

& Rehm 2009; Li *et al.* 2011), diagnostic testing (Lee *et al.* 2005; Bäckström *et al.* 2007; Atwood & Rehm 2009; Chen *et al.* 2014), bioimaging (Jahns *et al.* 2008), insecticide carrier (Moldes *et al.* 2006), vaccine delivery (Parlane *et al.* 2009; Parlane *et al.* 2011; Martinez-Donato *et al.* 2016; Rubio Reyes *et al.* 2016) and targeted drug delivery (Yao *et al.* 2008; Lee *et al.* 2011).

## 1.1.2 Biosynthesis of PHA and self-assembly of PHA beads

### 1.1.2.1 PHA Biosynthesis

There are three major natural pathways leading to PHA biosynthesis. For the most classical natural pathway of PHA biosynthesis, represented by PHB ($PHA_{SCL}$) biosynthesis in *Cupriavidus necator* (with earlier synonyms of *Wautersia eutropha, Ralstonia eutropha* or *Alcaligenes eutrophus*) (Vandamme & Coenye 2004), three key enzymes are required. These are *β*-ketothiolase (PhaA, which condenses two molecules of acetyl-CoA to form acetoacetyl-CoA), acetoacetyl-CoA reductase (PhaB, which reduces the acetoacetyl-CoA to form (*R*)-3-hydroxybutyryl-CoA), and PHA synthase (PhaC, which polymerizes (*R*)-3-hydroxybutyryl-CoA to form PHA with the release of CoA) (Normi *et al.* 2005; Rehm 2007). Corresponding biosynthesis genes quite frequently are clustered together, like the *phaCAB* operon in *C. necator* (Rehm 2003; Rehm 2007; Grage *et al.* 2009).

Two other major natural pathways are represented by $PHA_{MCL}$ biosynthesis found in the genus *Pseudomonas*, where either the fatty acid β-oxidation cycle (via enoyl-CoA intermediate) or fatty acid *de novo* synthesis cycle (via (*R*)-3-hydroxyacyl-ACP

intermediate) could be independently utilised to provide (*R*)-3-hydroxyacyl-coA thioester monomers for PHA synthesis (Rehm *et al.* 1998; Tsuge *et al.* 2003; Chen *et al.* 2015).

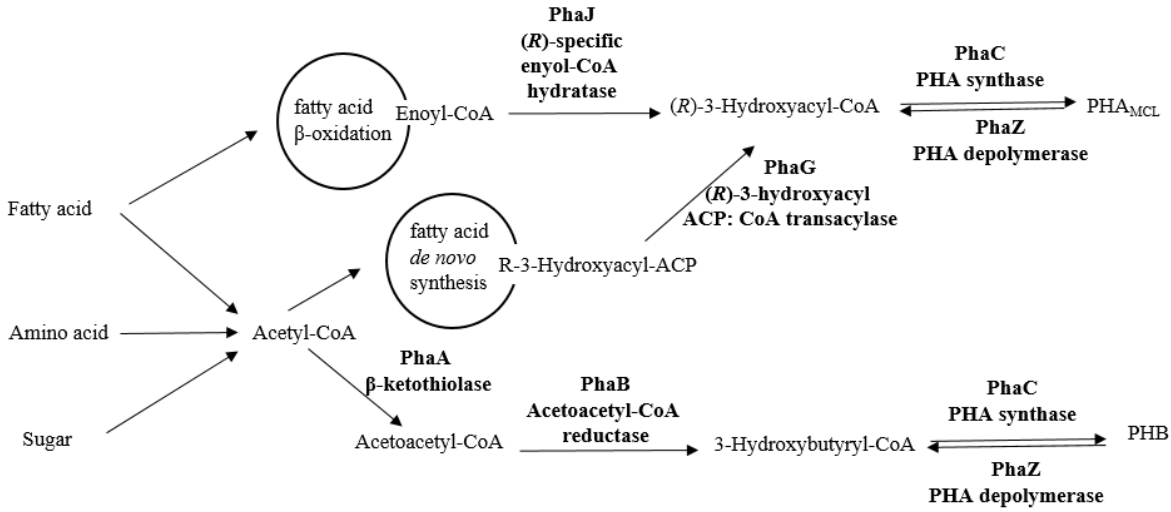



**Figure 1.1 Major pathways of PHA biosynthesis and major enzymes involved in PHA metabolism.**

With extensive metabolic engineering efforts in either native or non-native PHA producers, so far about 10 engineered pathways leading to PHA biosynthesis have been reported (Meng *et al.* 2014; Chen *et al.* 2015). Generally starting with different carbon sources, these pathways use different metabolic routes to supply various kinds of PHA precursors (namely (*R*)-3-hydroxyacyl-coA thioester monomers) for corresponding PHA synthases to catalyse the polymerisation reactions (Meng *et al.* 2014; Chen *et al.* 2015).

Thirty years ago, recombinant PHB production by introducing the whole PHA biosynthesis gene operon from *C. necator* into *E. coli* was published separately by two groups at the same time (Schubert *et al.* 1988; Slater *et al.* 1988). Since then, more precised genetic and metabolic engineering has been achieved in a wide range of organisms due to increased knowledge of genes and pathways related to PHA biosynthesis. For instance, by transforming different PHA biosynthesis genes into different plant compartments, transgenic *Arabidopsis thaliana* could be engineered to

produce either $PHA_{SCL}$, $PHA_{MCL}$ or $PHA_{SCL-MCL}$ polymers (Mittendorf *et al.* 1998; Bohmert *et al.* 2000; Matsumoto *et al.* 2009). The $PHA_{MCL}$ production in *A. thaliana* was straightforward as it simply involved expressing and targeting the PHA synthase *phaC1* gene from *Pseudomonas aeruginosa* into leaf peroxisome and glyoxysomes in cotyledons where fatty acid β-oxidation pathway is extremely active (Mittendorf *et al.* 1998). Whereas $PHA_{SCL}$ production in *A. thaliana* was achieved by coexpressing and targeting the *phaA*, *B,* and *C* genes (encoding *β*-ketothiolase, acetoacetyl-CoA reductase and PHA synthase, respectively) from *C. necator* into leaf chloroplasts where acetyl-CoA is abundant (Bohmert *et al.* 2000).

In contrast to $PHA_{SCL}$ or $PHA_{MCL,}$ the production of $PHA_{SCL-MCL}$ in the plastids of *A. thaliana* was the most complex, as it required both short and medium chain length monomers to be supplied through different metabolic pathways, along with a PHA synthase with broad substrate specificity to enable incorporation of these monomers into the final PHA product (Matsumoto *et al.* 2009). Briefly, 4 differently originated wild type or mutated genes were coexpressed and targeted to the plastids of *A. thaliana* in that study. These included *phaA* and *phaB* genes from *C. necator* encoding *β*-ketothiolase and acetoacetyl-CoA reductase for converting acetyl-CoA to (*R*)-3-hydroxybutyryl-CoA monomers; along with a gene encoding a mutated (F87T) 3-ketoacyl-acyl carrier protein (ACP) synthase III (FabH) from *E. coli* to supply medium chain length monomers via the fatty acid *de novo* synthesis pathway. According to the study of Matsumoto *et al.* (2009) the F87T mutant was able to recognise medium chain length 3-ketoacyl-ACP substrates in addition to short chain length 3-ketoacyl-ACP substrates that are recognised by wild type FabH. Lastly a gene coding for a mutant PhaC1 (ST/QK) (a Ser325Thr and Gln481Lys double mutants) PHA synthase from *Pseudomonas* sp. 61-3 for PHA polymerisation reaction (Matsumoto *et al.* 2009).

Therefore, factors affecting the final PHA composition include not only the PHA biosynthesis pathways involved, but also the key PHA synthases whose substrate specificities ultimately decide the PHA structure (Rehm 2003; Meng *et al.* 2014; Chen *et al.* 2015).

### 1.1.2.2 PHA synthases

Though all PHA synthases polymerise the acyl moiety of the (*R*)-3-hydroxyacyl-CoA substrates by releasing free CoA, they vary a lot in subunit compositions and substrate specificities, which form the basis for their classification (Rehm 2003).

There are four major groups of PHA synthases (Rehm 2003; Rehm 2006; Rehm 2007). In terms of subunit compositions, Class I and II PHA synthases, represented respectively by the synthases from *C. necator* and *P. aeruginosa,* consist of a single subunit of PhaC (molecular mass (MW) = 60 - 73 kDa) (Timm & Steinbüchel 1992; Rehm 2003; Hoffmann & Rehm 2004; Rehm 2006; Zou *et al.* 2017). Class III and IV PHA synthases, represented respectively by the synthases from *Allochromatium vinosum* and *Bacillus megaterium*, comprise a secondary subunit PhaE (~40 kDa) in the case of Class III synthases, or PhaR (~20 kDa) in the case of Class IV synthases, in addition to the catalytic subunit PhaC synthases (~40 kDa) in both (Rehm 2003; Rehm 2006; Zou *et al.* 2017). In terms of substrate specificities, generally Class I, III and IV PHA synthases prefer monomers of short chain length (C3-C5) to form $PHA_{SCL}$, whereas Class II PHA synthases polymerise monomers of medium chain length (C6-C14) to form $PHA_{MCL}$ (Rehm 2007; Zou *et al.* 2017), with PhaC1 or PhaC2 showing similar substrate specificities (Qi *et al.* 1997; Guo *et al.* 2013).

Class I and II PHA synthases exist in equilibrium as monomers and dimers, and the dimeric forms have been shown to be the active forms (Zhang *et al.* 2003; Rehm 2007). Though amino acid sequences of all four types of PhaCs are quite diversified with highly variable N-terminal regions (Rehm 2007), they do share in the C-terminal catalytic domains an essential PhaC box sequence ([G/S]-X-C-X-[G/A]-G) (Tsuge *et al.* 2015), as well as a crucial Cys-His-Asp catalytic triad (Rehm 2003). This includes, for example, Cys319-His508-Asp480 of class I PHA synthase from *C. necator* (Gerngross *et al.* 1994; Jia *et al.* 2001), Cys296-His453-Asp452 of class II PHA synthase from *P. aeruginosa* (Amara & Rehm 2003), Cys149-His331-Asp302 of class III PHA synthase from *A. vinosum* (formally *Chromatium vinosum*) (Jia *et al.* 2000), and Cys151-His335-Asp306 of class IV PHA synthase from *Bacillus cereus* (Tsuge *et al.* 2015). All these features are similar to hydrolases such as lipases bearing a core α/β hydrolase fold (an α/β sheet composed of eight β-sheets connected by six α-helices) (Ollis *et al.* 1992), except that the serine in the catalytic site is replaced by cysteine in PHA synthases (Jaeger *et al.* 1995; Jia *et al.* 2000; Jendrossek & Handrick 2002; Rehm 2003; Tsuge *et al.* 2015).

Crystal structures became available only very recently for two Class I PHA synthases, one from *C. necator* (Wittenborn *et al.* 2016; Kim *et al.* 2017; Kim *et al.* 2017) (PhaC$_{Cn}$ hereinafter) and the other from *Chromobacterium* sp. USM2 (Chek *et al.* 2017) (PhaC$_{Csp}$ hereinafter). The PhaC$_{Cn}$ structures published by two individual groups are similarly in a partially open form with an obvious channel or tunnel (in the vicinity of the active site) that is proposed for substrate entrance or binding (Wittenborn *et al.* 2016; Kim *et al.* 2017; Kim *et al.* 2017), while the PhaC$_{Csp}$ structure is in a closed form where the active site is covered by a CAP subdomain and the catalytic residues are facing a water-filled large channel inside the protein (Chek *et al.* 2017). It is suggested that these structural

differences might represent two dimeric states of the Class I PHA synthase which raises the possibility of rearrangement via conformational change of the CAP subdomain (Chek *et al.* 2017). Nevertheless, the above mentioned monomer-dimer equilibrium and active dimeric forms, the Cys-His-Asp catalytic triad, and the core α/β hydrolase fold (though with 4-5 more β-sheets and 2 more α-helices as compared to a canonical lipase α/β hydrolase fold) were all confirmed in these crystal structures (Wittenborn *et al.* 2016; Chek *et al.* 2017; Kim *et al.* 2017).

Despite the structural knowledge, the exact mechanisms for Class I PHA synthase catalytic reaction remain unknown except that a single active site from the PhaC dimer is required for PHA biosynthesis (Wittenborn *et al.* 2016; Chek *et al.* 2017; Kim *et al.* 2017). For example for the PhaC$_{Cn}$ structure, two different mechanisms have been proposed by the two individual groups (Wittenborn *et al.* 2016; Kim *et al.* 2017). The first suggests a substrate entrance channel which is also used for the release of free CoA, and a separate egress route for the elongation of the nascent PHA chain (Wittenborn *et al.* 2016). In this mechanism, the catalytic Cys319 keeps bonding with the growing PHA chain while repeatedly accepting the (*R*)-3-hydroxyacyl (HA) moiety and releasing CoA moiety (Wittenborn *et al.* 2016), but the proposed egress route was too narrow and would require substantial conformational changes for the growing PHA chain to pass through (Chek *et al.* 2017). Whereas the second suggested mechanism is in favour of a ping-pong mechanism in which the substrate and the growing PHA chain take turns to enter and leave the catalytic site through a single tunnel, with the catalytic Cys319 acting to transfer one HA monomer from HA-CoA to the PHA chain every cycle (Kim *et al.* 2017), which is inconsistent with the detection of a product covalently bound to the catalytic Cys319 (Wodzinska *et al.* 1996). Furthermore, comments on the role of N-terminal region of PhaC are different: one is localising PhaC$_{Cn}$ onto PHA granule and stabilizing the growing

PHA polymer (Kim *et al.* 2017), while the other is stabilizing dimeric $PhaC_{Csp}$ (Chek *et al.* 2017). Clear elucidation might be reached in the near future with more knowledge accumulated on crystal structures of PHA synthases.

### 1.1.2.3 Self-assembly of PHA beads

*In vitro* synthesis of PHAs and self-assembly of spherical beads was first demonstrated by Gerngross and Martin with purified PHA synthases and (*R*)-3-hydroxybutyryl-CoA substrates (Gerngross & Martin 1995). However, the exact *in vivo* PHA bead assembly mechanism remains unclear. A total of three models have been proposed for *in vivo* PHA granule assembly: (i) the currently favoured micelle model, (ii) the controversial membrane budding model, and (iii) the relatively new scaffolding model.

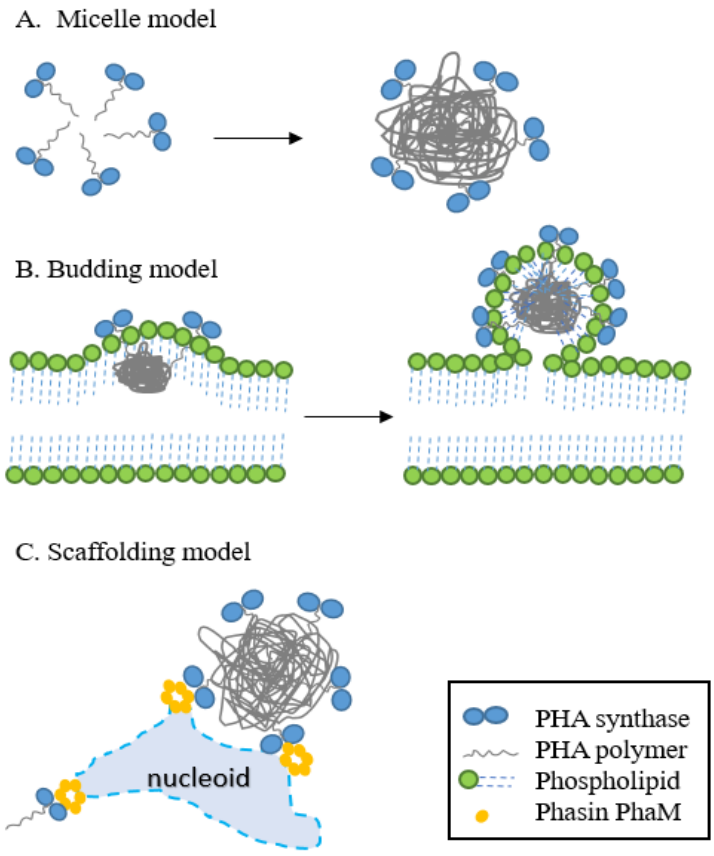



Figure 1.2 **PHA bead self-assembly models.**

In the micelle model, soluble dimerised PHA synthase interacts with the substrate randomly in the cytoplasm to initiate the polymerisation process, during which the growing hydrophobic PHA chain remains covalently attached to the PHA synthase (Gerngross *et al.* 1993). Several of these amphipathic synthase-polymer molecules self-aggregate to form energetically-favourable micelle-like structures with a granular hydrophobic core surrounded by PHA synthase (Gerngross *et al.* 1993). In native PHA producers, these intracellularly formed nascent PHA granules are further coated by other granule-associated proteins (GAPs) that are involved in PHA metabolism, such as Phasins (PhaP), depolymerases (PhaZ) or regulatory proteins (PhaR) (Thomson *et al.* 2010; Parlane *et al.* 2017). The micelle model is well supported by the above-mentioned *in vitro* PHA formation study using only purified PHA synthase and substrate (Gerngross & Martin 1995).

The membrane budding model is based on early observations of the lipid membrane-like matter surrounding either isolated PHA beads or beads inside intact cells (Lundgren *et al.* 1964; Jensen & Sicko 1971; Dunlop & Robards 1973; Mayer & Hoppert 1997). In this model, soluble PHA synthases associate with or stay close to the inner cell membrane, and the polymerisation reaction occurs in the inter-membrane space, where the elongation of PHA chains makes the membrane fold around them, leading to granules surrounded by a phospholipid monolayer with PHA synthases and other proteins attached, which eventually bud from the membrane into the cytoplasm (Jendrossek *et al.* 2007; Thomson *et al.* 2010).

Evidence in favour of the membrane budding model includes data from transmission electron microscopy (TEM), confocal laser scanning fluorescence microscopy (CLSM) or fluorescence microscopy (FM) demonstrating that nascent PHA beads are not

randomly distributed in the cytosol but rather frequently located at or near the cytoplasmic membranes (such as the cell poles, cell wall or cell division sites) in either native or recombinant PHA producers (Jendrossek 2005; Peters & Rehm 2005; Jendrossek *et al.* 2007; Xiao *et al.* 2015). However, evidence against this model is accumulating. For example, an electron cryotomography technique used to examine near-native state PHA biogenesis in the native producer *C. necator* revealed that, along the length of the cell, PHA granules of different sizes are distributed towards the centre of the cytoplasm and are coated with only a discontinuous surface layer, which is more like a partial protein coat rather than a continuous phospholipid layer (Beeby *et al.* 2012). The absence of a phospholipid layer for *in vivo* formed PHA beads is further evidenced by fluorescence microscopy (FM) colocalisation analysis of phospholipid-targeted fluorescent proteins in three native PHA producers (*C. necator*, *Pseudomonas putida* and *Magnetospirillum gryphiswaldense*) (Bresan *et al.* 2016). In that study, DsRed2EC and other fluorescent proteins were separately fused with the phospholipid-binding domain (LactC2) of lactadherin, and the fusion proteins were found to only colocalise with the cytoplasmic membrane but not with PHA granules in all circumstances (Bresan *et al.* 2016).

The scaffolding model arose due to the granule localisation close to dark "mediation elements" in the cell centre as observed first in *C. necator* and later in *Comamonas* sp. EB172 via transmission electron microscopy (TEM) (Tian *et al.* 2005; Mumtaz *et al.* 2011). Undefined at that time, these mediation elements were proposed to serve as scaffolding initiation sites for granule formation (Tian *et al.* 2005; Mumtaz *et al.* 2011). These mediation elements are now believed to be bacterial nucleoids (Jendrossek & Pfeiffer 2014). In the model PHA producer *C. necator*, a regulatory protein PhaM that can bind not only PHA synthase PhaC but also DNA is proposed to be responsible for anchoring the PhaM-PhaC-PHB initiation complex to a bacterial nucleoid for granule

formation (Jendrossek & Pfeiffer 2014). However, this PhaM mediated scaffolding event could be unnecessary for PHA granule formation, as an earlier study showed that even a Δ*phaM* mutant *C. necator* strain formed PHA granules (Pfeiffer *et al.* 2011).

#### 1.1.2.4 PHA granule associated proteins (GAPs)

In addition to PHA synthases, three other major types of proteins are also known to associate with PHA beads in a diverse range of native PHA producers; these proteins include depolymerases, phasins and regulatory proteins. Except for PHA synthases that are essential for PHA synthesis and covalently bound to PHA granules, all other granule associated proteins (GAPs) are nonessential for PHA accumulation and only attached to the granules through hydrophobic interactions (Draper *et al.* 2013). Provided below is a brief summary of important GAPs discovered in the model strain *C. necator*.

##### 1.1.2.4.1 Depolymerases (PhaZ)

Intracellular PHA depolymerases bind to PHA granules and act through thiolysis reactions to degrade or mobilise native PHA (Uchino *et al.* 2007). So far in the model PHA producer *C. necator*, nine enzymes have been reported or predicted to mediate PHA degradation, including seven PHA depolymerases (PhaZ1 to PhaZ7) and two 3-hydroxybutyrate-oligomer hydrolases (PhaY1 and PhaY2) (Sznajder & Jendrossek 2014; Arikawa *et al.* 2016). Alternative names have been suggested by Kobayashi et al. (2005) that better reflect their amino acid sequence similarities (Kobayashi *et al.* 2005): PhaZa1 to PhaZa5 (for PhaZ1 to PhaZ5) are a group of isoenzymes with active cysteine sites; PhaZb and PhaZc (for oligohydrolases PhaY1 and PhaY2 respectively) both contain active sites resembling the lipase box (Gly-X-Ser-X-Gly) (Kobayashi *et al.* 2005); while

PhaZd1 and PhaZd2 (for PhaZ6 and PhaZ7) are another group showing high activity as well as the Ser-Asp-His catalytic triad (Abe *et al.* 2005; Sznajder & Jendrossek 2014). Several *in vivo* PHA mobilisation studies have been carried out using *phaZ* gene deletion mutants of *C. necator* in different combinations, but except for PhaZa1 whose *in vivo* PHA mobilisation function has been proved (Uchino *et al.* 2008), a clear role for any other enzyme is yet to be elucidated. However one recent study showed that PhaZd1 affects the molecular weight of PHA accumulated in *C. necator* (Arikawa *et al.* 2016).

#### 1.1.2.4.2 Phasins (PhaP)

Phasins are the most abundant GAPs, and are of small size (MW = 11-25 kDa) (Grage *et al.* 2009). So far in the model PHA producer *C. necator*, seven phasins (PhaP1 to PhaP7) have been identified (Pfeiffer & Jendrossek 2012). Of these, PhaP1 has been most extensively studied and among the 6 other variants, only three (PhaP2 to PhaP4) are homologous to PhaP1 with 38-70% identity (Pfeiffer & Jendrossek 2012). Nevertheless, all 7 phasin proteins contain a common characteristic "phasin 2 motif" which is a region enriched in hydrophobic residues (Pfeiffer & Jendrossek 2012). The phasins consist of a hydrophobic domain responsible for association with the surface of the PHB granules, and a predominantly hydrophilic/amphiphilic domain exposed to the cytoplasm of the cell (Pötter & Steinbüchel 2005). Among them, the most studied, PhaP1, has been shown by secondary structure analysis to occur as a planar triangular homotrimer (Neumann *et al.* 2008) and is bound to the hydrophobic surface of PHA polymers as soon as the polymer is available (Cho *et al.* 2012). PhaP1 functions to control the amount, size and number of PHA granules, to prevent PHA granule aggregation and inhibit nonspecific association of other proteins with the PHA granules (Grage *et al.* 2009; Draper *et al.* 2013; Maestro & Sanz 2017; Parlane *et al.* 2017).

#### 1.1.2.4.3 Regulatory proteins (PhaR and PhaM)

In *C. necator*, the two identified regulatory proteins PhaR and PhaM both are capable of binding not only to PHA granules but also to DNA (or to the nucleoid) (Pötter *et al.* 2002; York *et al.* 2002; Pfeiffer *et al.* 2011; Wahl *et al.* 2012; Bresan & Jendrossek 2017).

The transcriptional regulator PhaR functions through binding to PHA granules as well as DNA promoter regions upstream of the *phaP* (actually the most studied *phaP1* gene) and *pha*R genes (Pötter *et al.* 2002; York *et al.* 2002). In their proposed regulation model, PhaR binds the *phaP* promoter so as to inhibit *phaP* transcription under non-PHA accumulating conditions. But once PHA accumulation begins, PhaR binds to the hydrophobic surfaces of newly formed PHA granules, leading to a reduced level of free PhaR until it is insufficient to repress the transcription of *phaP* gene. Then PhaP protein (actually the most studied PhaP1 protein) is produced and immediately attaches to the PHA granules. After the granules have reached maximum size and are completely coated by PhaP along with PhaR, the excess free PhaR binds to the promoter regions of both *phaP* and *pha*R genes, and prevents transcription of both (Pötter *et al.* 2002; York *et al.* 2002).

As mentioned in the scaffolding model (section 1.1.2.3), regulatory protein PhaM (26.6 kDa) is able to locate the PhaM-PhaC-PHB complex to the bacterial nucleoid through binding to both PHA granule and DNA/nucleoid (Jendrossek & Pfeiffer 2014). PhaM also interacts with PhaP5 and affects the number, size and distribution of PHA granules (Pfeiffer *et al.* 2011). One study indicated that PhaM has little or no direct interaction with the PHA polymer chain (Ushimaru & Tsuge 2016). In addition, a very recent study

suggested that PhaM associates with PhaC through interaction with N-terminal domain of the synthase, and activates PhaC by providing a more extensive surface area for PhaC to interact with the growing PHA polymer (Kim *et al.* 2017).

### 1.1.3 Functionalised PHA beads mediated via GAP fusion

As mentioned in section 1.1.1, granule associated proteins (GAPs), especially PHA synthase (PhaC), due to their ability to bind to PHA beads, have been extensively explored to display functional proteins on the surface of PHA beads by designing appropriate GAP-target protein fusions. The resulting PHA beads find applications in areas such as recombinant protein production and purification as well as biomolecule recognition.

#### 1.1.3.1 PHA beads as recombinant protein production and purification platforms

Previously, granule associated proteins (GAPs) including PHA synthase PhaC, phasin protein (PhaP, actually the most studied PhaP1) and regulatory protein PhaR from *C. necator* each have been used for recombinant protein production and purification. By fusing a target protein to one of these GAPs through a linker consisting of either a protease cleavage site or a self-cleavage Intein tag ((see section 1.2.1 for a detailed explanation), engineered from an appropriate protein splicing element known as Intein, this kind of tag is cleavable at one of its termini by pH shift or thiol addition), the target protein could be anchored by its GAP fusing partner onto the surface of PHA beads that serve as purification resins. The resulting PHA beads carrying the target protein could be easily

separated from soluble cell fractions, and finally release the target protein from beads either via protease treatment or self-cleavage of the Intein tag.

In one study involving the PHA synthase PhaC, a plasmid expressing PhaC-EK-X (EK represents enterokinase cleavage site, and X the target protein) was transformed into a recombinant *E. coli* strain harbouring *phaA* and *phaB* genes (encoding *β*-ketothiolase (PhaA) and acetoacetyl-CoA reductase (PhaB) that are key enzymes for PHA biosynthesis) (Grage *et al.* 2011). This enabled production of PHA beads displaying the fusion protein of PhaC-EK-X. Then after bead isolation and washing steps, the beads were treated with enterokinase to release pure target protein such as fluorescent protein HcRed and anti-β-galactosidase antibody single chain variable fragment scFv13R4 (Grage *et al.* 2011).

In a different study, phasin protein PhaP (actually the most studied PhaP1) was used in a similar approach involving thrombin cleavage. By transforming a plasmid expressing PhaP-THR-tPA (THR represents thrombin cleavage site, and tPA represents recombinant human tissue plasminogen activator) into a recombinant *E. coli* strain harbouring *phaCAB* genes (encode all three key enzymes for PHA biosynthesis), recombinant human tissue plasminogen activator could be successfully produced and purified (Geng *et al.* 2010).

A collaborative group based in the USA was the first to combine the self-cleaving tag technology with PHA bead display technology in the production and purification of recombinant proteins (Banki *et al.* 2005; Barnard *et al.* 2005). In their pioneering work, a plasmid expressing PhaP-PhaP-PhaP-intein-X or PhaP-PhaP-intein-X (PhaP was used in mulitiple copies to increase hydrophobic association with PHA beads, the intein used was a pH inducible *Mtu* ΔI-CM mini-intein, and X represents the target protein) was

transformed into engineered *E. coli* cells harbouring the PHA biosynthesis genes *phaABC*. PHA beads attached with corresponding fusion proteins could be produced and isolated, then the target protein was released by lowering pH. Four recombinant proteins were purified using this method, including maltose binding protein (MBP), β-galactosidase (LacZ), chloramphenicol acetyltransferase (CAT) and N utilisation substance protein A (NusA) (Banki *et al.* 2005).

Next, the same group tested a different intein in native PHA producers by using *C. necator* to produce PhaP-intein-X (the intein used was a thiol controllable *Mxe* GyrA intein, and X represents the target protein) protein respectively. Similarly PHA beads with corresponding attached fusion proteins could be produced and isolated, then the target protein (GFP or LacZ) was released by addition of thiol reagents (Barnard *et al.* 2005). Later Wang's group was successful in using exactly the same PhaP-PhaP-PhaP-intein-X and recombinant *E. coli* system from the pioneer study (Banki *et al.* 2005) to produce and purify porcine interferon alpha (PoIFNα) (Zhou *et al.* 2011).

Chen's group from Tsinghua University extended this recombinant protein purification platform mediated by self-cleaving PHA beads further by using *in vitro* chemically synthesised PHA beads and using PhaR tags (Wang *et al.* 2008; Zhang *et al.* 2010). In their studies, the PHA polymer was chemically produced separately, while the recombinant protein was produced by *E. coli* in the form of either PhaP-intein-X or PhaR-intein-X (the intein used was a pH inducible *Ssp* DnaB intein, and X represents the target protein). Then the PHA polymer and the cleared cell lysate were mixed together to pull down the fusion proteins, and finally via pH shift induced intein cleavage, the target protein (including EGFP (enhanced green fluorescent protein), MBP and LacZ) could be released from the beads (Wang *et al.* 2008; Zhang *et al.* 2010).

### 1.1.3.2 PHA beads as affinity resins

Granule associated proteins (GAPs) including PHA synthase PhaC and phasin protein could also be used to functionalise PHA beads with appropriate affinity ligands in order to separate, remove or detect a target biomolecule based on molecular recognition.

For example, the group led by Professor Bernd H. A. Rehm has developed a variety of PHA bead affinity resins by using engineered PHA synthase (PhaC) fusions (Brockelbank *et al.* 2006; Grage & Rehm 2008; Peters & Rehm 2008; Lewis & Rehm 2009). This included IgG purification resins of PHA beads displaying the immunoglobulin G (IgG) binding ZZ domain of protein A from *Staphylococcus aureus* (Brockelbank *et al.* 2006; Lewis & Rehm 2009); β-galactosidase (antigen) detection resins of PHA beads displaying anti-β-galactosidase antibody single chain variable fragment scFv13R4 (Grage & Rehm 2008); and biotin-recognising affinity resins of PHA beads displaying different variants of streptavidin that are applicable for ELISA, DNA purification, enzyme immobilisation and flow cytometry (Peters & Rehm 2008).

In addition, the above-mentioned Tsinghua group developed endotoxin removal resins of PHA beads by mixing chemically prepared PHA particles with recombinantly produced and purified rhLBP-PhaP fusion protein (rhLBP = human lipopolysaccharide binding protein, PhaP was actually the most studied PhaP1) (Li *et al.* 2011).

## 1.2 Existing technologies related to protein recovery and biomolecular detection

### 1.2.1 Recombinant protein production and purification from *E. coli* cells

In a conventional recombinant protein production scheme, upstream production and downstream purification are two separate processes, and the latter often requires expensive chromatography columns and considerable process optimisation (Asenjo & Andrews 2009). Although affinity chromatography using affinity tags such as His, FLAG, GST or Strep can facilitate the recombinant protein purification process (Young *et al.* 2012), their industrial scale up is however limited by the additional protease treatment and chromatography required for tag removal.

Amongst various recombinant proteins, therapeutic proteins are especially hard to express in *E. coli*, because overexpression of heterologous eukaryotic genes in *E. coli* tends to result in proteins as insoluble inclusion bodies (IBs). For example, one recent study reported that human tumour necrosis factor alpha (TNFα) production in *E. coli* resulted in an IB format product of up to 50% (Zhang *et al.* 2014). Similarly human interferon alpha-2b (IFNα2b) production in *E. coli* was frequently reported to be in the form of IB aggregates (Rabhi-Essafi *et al.* 2007). In addition, human granulocyte colony-stimulating factor (G-CSF) production in *E. coli* tends to aggregate and forms inclusion bodies (IBs) (Do *et al.* 2014). Indeed, several recently published studies on production and purification of recombinant TNFα, IFNα2b and G-CSF focus on optimising protein refolding from inclusion bodies as well as optimising chromatography operations (Vemula *et al.* 2015; Wang *et al.* 2015; Romanov *et al.* 2017). Many are trying to avoid tedious refolding by adopting solubility-enhancing tags or affinity tags in combination with specialized

affinity resins plus additional protease treatment efforts in tag removal (Rabhi-Essafi *et al.* 2007; Do *et al.* 2014; Alizadeh *et al.* 2015).

Several types of self-cleaving tags, including sortase and intein, are emerging as alternatives to protease treatment for tag removal. For example, sortase A (SrtA) from *Staphylococcus aureus* functions as a transpeptidase by identifying proteins with a C-terminal five amino acid LPXTG signal (X represents any amino acid), cleaving between the T and G, and ligating the T to peptidoglycan of the cell wall (Clancy *et al.* 2010). *In vitro* studies revealed that cleavage activity of recombinant SrtA was independent of triglycine, but the cleavage rate was increased in the presence of triglycine (Ton-That *et al.* 2000). Previously, the catalytic core of SrtA (SrtAc or $SrtA_{\Delta 59}$, amino acids 60-206) was used for protein purification in the format of a $His_6$-SrtA-LPETG-target protein fusion (Mao 2004) or a similar biotin reactive bls-SrtA-LPETG-target protein / blsbls-SrtA-LPETG-target protein (bls represents biotin label signal) fusion (Matsunaga *et al.* 2010), which upon binding to respective affinity resins, released free target proteins into elution fractions via SrtA activation in the presence of $Ca^{2+}$ +/- triglycine (Mao 2004; Matsunaga *et al.* 2010). Nevertheless, these processes still relied on the use of expensive and specialized affinity resins.

Another well-studied self-cleaving tag is intein, an intervening protein element that is involved in protein splicing, namely the self-excision and ligation of its flanking peptides (exteins) (Perler *et al.* 1994). Various inteins have been engineered to abolish the ligation but retain cleavage at only one terminus (either N- or C- terminus) controllable by pH / temperature / thiol agents (Shi *et al.* 2013). The DnaB mini intein from *Synechocystis sp.* PCC 6803 has been commercialised and widely used in the pTWIN1 vector (New England Biolabs, Hitchin, UK); this mini intein is activated at pH 6-7 at 25°C and

undergoes self-cleavage at the C-terminus. Recently, this intein has been employed in combination with polyhydroxyalkanoate (PHA) beads (as detailed in sections 1.1) for target protein production and purification (Wang *et al.* 2008; Zhang *et al.* 2010). These strategies rely on phasin (Wang *et al.* 2008) or a regulatory protein (PhaR) (Zhang *et al.* 2010) that non-covalently associates with PHA beads as "affinity" tags and fusion partners of target proteins. By incubating PHA beads with cell lysate containing PhaP-intein-target protein or PhaR-intein-target protein, the target protein can be anchored to PHA beads via PhaP / PhaR and thus separated from cell debris and host proteins. The pH inducible intein as linker subsequently allows the release of target protein by a pH drop. However, the non-covalent anchoring of the target protein to PHA beads can cause leakage of the respective PhaP-intein-target protein or PhaR-intein-target protein during the PHA bead wash cycles (Wang *et al.* 2008; Zhang *et al.* 2010).

### 1.2.2 Protein separation or biomolecule detection from natural sources

There are often industrial needs to separate or detect biologically important molecules from natural sources, like for instance commercial lysozyme separation from hen egg white in the food and pharmaceutical industry (Shahmohammadi 2017), and progesterone (P4) detection in the dairy industry (Jang *et al.* 2017). So far affinity separation and affinity detection techniques exploiting the biorecognition between a biomolecule and its ligand (or binding partner) are commonly regarded as the most efficient techniques. Biorecognition, or molecular recognition, is defined as the specific noncovalent interaction between a biomolecule and its ligand (or binding partner), which plays a vital role in cellular activities (McCammon 1998). Examples of such biomolecule - ligand pairs include enzyme - substrate, protein - cofactor, antibody - antigen. Affinity separation and detection technologies are mostly exemplified by affinity resins with immobilised

antibodies that are capable of recognising antigenic epitopes of the molecules of interest (Tozzi *et al.* 2003; Crivianu-Gaita & Thompson 2016). Manufacture of this type of affinity resin generally requires three main processes: (1) the preparation of a support matrix (either through chemical treatment of existing natural resources such as agarose, dextrose and cellulose, or through *de novo* chemical synthesis such as polyacrylamide and polymethacrylate derivatives) (Vařilová *et al.* 2006), (2) the recombinant production and purification of an antibody or antibody fragment which is usually costly (Dias & Roque 2017) and (3) chemical cross-linking of the antibody or antibody fragment to the support matrix to avoid random orientation or denaturation of the antibody or antibody fragment, which can be a complex process (Shen *et al.* 2017). Therefore, developing alternative non-antibody affinity ligands or resins that enable simplified ligand immobilisation is necessary, such as in the cases of lysozyme separation and progesterone detection.

Lysozyme (E.C.3.2.1.17, N-acetyl-muramic-hydrolase, ~14.3 kDa) is widely used as a preservative in food and is in high demand in the pharmaceutical industry due to its bacteriostatic and bactericidal effects (Cegielska-Radziejewska *et al.* 2008). Lysozyme is usually obtained from hen egg white but separation methods such as crystallization, adsorption, chromatography, and ultrafiltration are often complicated and inefficient. Lysozyme can also be separated using substrate (chitin or chitosan)-based affinity chromatography (Yuan *et al.* 2009; Wolman *et al.* 2010), but even this method suffers from slow flow rates and high resin costs (Abeyrathne *et al.* 2013). Furthermore, affinity chromatography methods based on metal ions (Ergün *et al.* 2007; Derazshamshir *et al.* 2008; Baydemir *et al.* 2013; Liu *et al.* 2013) or dyes (Arica *et al.* 2004; Altintaş & Denizli 2006) have been developed for lysozyme separation. However, these methods often require significant process optimisation to ensure selectivity, as metal ions and dyes are

not as specific as biological ligands like antibodies or substrates (El Khoury *et al.* 2015). There are only sparse reports of successful lysozyme separation using affinity chromatography based on engineered antibodies. Actually only a synthetic octapeptide analogous to a lysozyme antibody light chain hypervariable loop (L3) (Hahn *et al.* 2001), as well as variable fragment (Fv) or single chain Fv (scFv) derived from lysozyme antibodies have been documented (Berry *et al.* 1991; Welling *et al.* 1991; Berry & Pierce 1993; Fong *et al.* 2002). Therefore, there is an urgent need for new types of affinity resins carrying selective non-antibody ligands that can enable efficient and streamlined lysozyme separation.

Progesterone (P4, ~ 314.46 g/mol) is a 21-carbon steroid hormone secreted mainly by the corpus luteum (CL, which forms after ovulation of the oocyte in the ovary) that affects the menstrual cycle, pregnancy, and embryogenesis in female mammals (Jang *et al.* 2017). The P4 level in cow's milk is typically within a range of 1-10 ng/ml, and has long been used as an indicator of reproductive status (Daems *et al.* 2017). Bovine milk P4 is minimal on Day 0 of oestrus (ovulation) at 1-2 ng/ml, increases within 2 days as the CL starts to grow and peaks at ~ 3.5 ng/ml around Day 10. It remains stable for about one week and sharply decreases from Day 17 as the CL begins to regress in the absence of fertilization till Day 21 when the next oestrus occurs. On the contrary, for a pregnant cow where the CL is maintained, milk P4 level is usually higher than 7 ng/ml and remains high throughout gestation (Simersky *et al.* 2007; Samsonova *et al.* 2015; Daems *et al.* 2017). Regular milk P4 level detection contributes to accurate ovulation prediction for timely artificial insemination, to optimise herd reproductive performance and maintain farm profitability, which is of great importance in the dairy industry. However, existing chromatography or mass spectrometry technology used for P4 detection is usually complicated and requires expensive equipment (Gao *et al.* 2016; Goyon *et al.* 2016).

There are various reports on P4 antibody based immunoassay methods such as radioimmunoassay (RIA) (Byszewska-Szpocińska & Markiewicz 2006), electrochemical immunosensor (Zhang *et al.* 2013), fluorescence immunoassay (FIA) (Käppel *et al.* 2007), lateral flow immunoassay (LFI) (Samsonova *et al.* 2015) and enzyme-linked immunosorbent assay (ELISA) (Wu *et al.* 2014). However, due to the inherent high production costs and low stability related to antibodies during or after immobilisation, as well as the limited numbers of tests per assay kit, they are still costly for daily use on farms (Posthuma-Trumpie *et al.* 2009). Research on new types of detection involving non-antibody ligands bound to a matrix would help to bring down the related bovine milk P4 detection cost.

Thanks to advances in protein engineering technologies (Grönwall & Ståhl 2009; Ståhl *et al.* 2013), various engineered non-antibody protein scaffolds are emerging as alternative elements for molecular recognition (Skerra 2007; Zhao *et al.* 2013; Škrlec *et al.* 2015; Dias & Roque 2017). One such promising candidate scaffold is the OB-fold, which was originally identified in four different proteins that were able to bind oligonucleotides or oligosaccharides, including staphylococcal nuclease, the anticodon binding domain of yeast asp-tRNA synthetase, as well as B-subunits of heat-labile enterotoxin and verotoxin-1 from *E. coli* (Murzin 1993). The OB-fold comprises a 5-stranded closed β-barrel that presents a concave binding face (Murzin 1993). The OB-fold is now known to be a common domain in a growing protein superfamily within all three kingdoms of life, with various binding affinities to oligonucleotides, oligosaccharides, proteins, metal ions or catalytic substrates (Arcus 2002). Those natural OB-fold harbouring proteins have no significant sequence similarity but most of them recognise their respective binding partners through the fold-related binding face (Steemson *et al.* 2014). It was thus hypothesised that the easily-adaptable binding face of the ancient OB-fold domain is

tolerant to mutation (Murzin 1993). Previously, several residues on the binding face of the OB-fold domain of aspartyl-tRNA synthetase (aspRS) from *Pyrobactulum aerophilum* were artificially randomized to make a library by using phage display technology, and the resulting library was screened for binding to a molecule of interest (Steemson 2011; Steemson *et al.* 2014). For example, an engineered OB-fold (termed an OBody) with a 3 nM affinity for hen egg-white lysozyme has been demonstrated (Steemson *et al.* 2014). It would be of great interest to immobilise such engineered biomolecular recognising OBodies on a proper support matrix for affinity separation or detection.

## 1.3 Proposed solution and the significance of this research

There is great potential to simplify processes related to protein recovery and biomolecular detection by integrating the emerging self-cleavable tag technology or OBody technology with the established PHA bead display technology via a PhaC fusion approach.

The proposed research is of great research and industrial importance. To properly design self-cleaving PhaC fusions to gain controlled release of functional target proteins would expand knowledge of the PHA bead display technology, and provide a streamlined process for recombinant protein production and purification. Simpler procedures and optimisation would benefit the protein industry economically. Furthermore, this study has the potential to enable large scale production and purification of medically important therapeutic proteins, which would help to bring down their production cost and retail price.

In addition, immobilisation of OBody ligands onto PHA beads is aimed to provide a "proof of concept" of new types of cheaper affinity resins (OBody beads) useful for separation or detection of industrially important enzymes or other biomolecules. This would also broaden applications of both OBody technology and PHA bead display platform.

## 1.4 Aims, objectives and hypotheses

In view of the generally complicated and costly processes related to protein recovery and biomolecular detection, the aims of this research were to develop (1) a streamlined process with less complicated steps toward the production and purification of recombinant proteins, especially therapeutic proteins, and (2) a simplified process for preparation of affinity resins with non-antibody ligands that could be used for separation and detection of industrially important biomolecules. Specifically, the PHA synthase (PhaC) mediated PHA bead display technology was explored to develop (1) self-cleavable recombinant protein production and purification resins, and (2) affinity resins carrying non-antibody ligands useful for biomolecular separation and detection.

The first aim of this research was to design PHA beads as self-cleavable recombinant protein production and purification resins, and was achieved by the following objectives:

**1. Determine whether sortase tags can be combined with PHA bead display technology to develop self-cleavable recombinant protein production and purification resins**

First, the *phaC-srtA* hybrid gene (encoding PHA synthase (PhaC) and sortase A from *Staphylococcus aureus*) was transformed into *E. coli* cells along with *phaAB* genes (encoding *β*-ketothiolase (PhaA) and acetoacetyl-CoA reductase (PhaB) that are key enzymes for PHA biosynthesis) to assess the production of bead fusion proteins and the self-cleavage activity of sortase in the presence of $CaCl_2$.

Then, the *phaC-srtA-x* hybrid gene (there was a LPETG coding region immediately before *x*, and *x* = coding gene for a target protein) was co-transformed with *phaAB* genes into *E. coli* cells to assess the production of bead fusion proteins and the self-cleavage of the sortase (in the presence of $CaCl_2$ and triglycine) so as to release a G-tagged target protein. Human tumour necrosis factor alpha (TNFα) and human interferon alpha-2b (IFNα2b) were tested as therapeutic target proteins.

**2. Determine whether intein tags can be combined with PHA bead display technology to develop self-cleavable recombinant protein production and purification resins**

First, the *phaC-intein-x* hybrid gene (*intein* = coding gene for the pH inducible *Ssp* DnaB mini intein, and *x* = coding gene for a model protein) was co-transformed with *phaAB* genes into *E. coli* cells to assess the production of bead fusion proteins and the pH drop induced self-cleavage of the intein so as to release the model protein. Model proteins tested for this object were *Aequorea victoria* green fluorescent protein (GFP), *Mycobacterium tuberculosis* vaccine candidate Rv1626, and the synthetic immunoglobulin G (IgG) binding ZZ domain of protein A derived from *Staphylococcus aureus*.

Then, this self-cleaving system was similarly tested for the production and purification of therapeutic targets by replacing the *x* gene (in *phaC-intein-x* construct) to one coding for TNFα, IFNα2b or human granulocyte colony-stimulating factor (G-CSF).

The hypothesis behind this aim was that the intergration of self-cleavable tags and PhaC synthase mediated PHA bead display platform would simplify the production and purification of recombinant proteins.

The second aim of this research was to design PHA beads as affinity resins carrying non-antibody ligands useful for biomolecular separation and detection, and was achieved by the following objectives:

**1. Determine whether a lysozyme recognising OBody ligand can be immobilised on PHA beads and used for lysozyme separation**

To this end, the coding region for lysozyme-recognising OBody ligand was fused to the *phaC* gene, and co-transformed with *phaAB* genes into *E. coli* cells to assess the production of bead fusion proteins, as well as the lysozyme separation usage of the resulting OBody beads as affinity resins.

**2. Determine whether a progesterone recognising OBody ligand can be immobilised on PHA beads and used for progesterone detection**

The coding region for two progesterone (P4)-recognising OBody ligands was each fused to the *phaC* gene, and respectively co-transformed with *phaAB* genes into *E. coli* cells to assess the production of bead fusion proteins, as well as the P4 detection usage of the

resulting OBody beads as affinity resins. Parameters such as P4 binding capacity and P4 binding affinity were carefully investigated.

The hypothesis behind this aim was that the immoblisation of OBody ligands on PHA beads would maintain the functionality of respective OBody ligands in separation and detection of industrially important biomolecules .

# Chapter 2: Materials and Methods

## 2.1 Bacterial strains and plasmids

The *Escherichia coli* strains and plasmids used in this study are listed in Tables 2.1 and 2.2 below.

### 2.1.1 *E. coli* strains

**Table 2.1** *E. coli strains used in this study*

| Strain | Genotype * | References |
|---|---|---|
| XL1-Blue | *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI*$^{q}$Z$\Delta$*M15* Tn*10* (Tet$^{r}$)] | Stratagene, La Jolla, USA |
| BL21 (DE3) | F$^{-}$ *dcm ompT hsdS* ($r_{B}^{-}$ $m_{B}^{-}$) gal $\lambda$(DE3) | Novagen, Madison, USA |
| SHuffle® T7 express | *fhuA2 lacZ::T7 gene1* [lon] *ompT ahpC gal* $\lambda$*att::*:pNEB3-r1-*cDsbC* (Spec$^{R}$, *lacI*$^{q}$) $\Delta$*trxB sulA11 R(mcr-73::miniTn10*--Tet$^{S}$)2 [dcm] *R(zgb-210::Tn10* --Tet$^{S}$) *endA1 ⊿gor* $\Delta$*(mcrC-mrr)114::IS10* | New England BioLabs (NEB), Hitchin, UK |

* Tet$^{r}$, tetracycline resistance.

XL1-Blue was used for plasmid propagation, BL21 (DE3) was used for standard bead production, and SHuffle® T7 express was used for bead production where a target protein contains disulphide bond.

### 2.1.2 Plasmids

**Table 2.2** *Plasmids used in this study*

| Plasmid | Description* | Sources or references |
|---|---|---|
| **Plasmid for expression of PhaC-sortase or PhaC-sortase-LPETG-target protein fusions** | | |
| pET14b | Amp$^{r}$; T7 promoter | Novagen |

| Plasmid | Description* | Sources or references |
|---|---|---|
| pETC[a] | pET14b derivative encoding PhaC | (Peters & Rehm 2008) |
| pMCS69[b] | $Cm^r$; pBBR1MCS derivative encoding PhaA and PhaB | (Amara & Rehm 2003) |
| pMCS69E[b] | pBBR1MCS derivative encoding Erv1p in addition to PhaA and PhaB | Iain Hay |
| pET14b:phaC-linker-MalE | pET14b derivative encoding PhaC-linker-MalE | (Jahns & Rehm 2009) |
| pET14b-PhaC-SrtA[a] | pET14b derivative encoding PhaC-linker-$SrtA_{\Delta N59}$ | This study |
| pET14b-PhaC-SrtA-TNFα[a] | pET14b derivative encoding PhaC-linker-$SrtA_{\Delta N59}$-LPETG-TNFα | Iain Hay |
| pET14b-PhaC-SrtA-IFNα2b[a] | pET14b derivative encoding PhaC-linker-$SrtA_{\Delta N59}$-LPETG-IFNα2b | Iain Hay |
| **Plasmid for expression of PhaC-intein-target protein fusions** | | |
| pTWIN1 | $Amp^r$; pBR322 derivative encoding two mini-inteins (*Ssp* DnaB and *Mxe* Gyr) | NEB |
| pET14b | $Amp^r$; T7 promoter | Novagen |
| pMCS69[b] | $Cm^r$; pBBR1MCS derivative encoding PhaA and PhaB | (Amara & Rehm 2003) |
| pMCS69E[b] | pBBR1MCS derivative encoding Erv1p in addition to PhaA and PhaB | Iain Hay |
| pET14b-PhaC-linker-ZZ | pET14b derivative encoding PhaC-linker-ZZ | (Jahns *et al.* 2013) |
| pET14b-ZZ(−)PhaC | pET14b derivative encoding ZZ-PhaC | (Brockelbank *et al.* 2006) |
| pET14b-PhaC-linker-GFP | pET14b derivative encoding PhaC-linker-GFP | (Jahns & Rehm 2009) |
| pPOLY-C-phaC-rv1626 | pET14b derivative encoding PhaC-RV1626 | (Rubio Reyes *et al.* 2016) |
| pET14b-PhaC-SrtA-TNFα | pET14b derivative encoding human TNFα | Iain Hay |
| pET14b-PhaC-SrtA-IFNα2b | pET14b derivative encoding human IFNα2b | Iain Hay |
| pET14b-PhaC-SrtA-G-CSF | pET14b derivative encoding human G-CSF | Iain Hay |
| pET14b-PhaC-Intein-GFP[a] | pET14b derivative encoding PhaC-Intein-GFP | This study |
| pET14b-PhaC-Intein-Rv1626[a] | pET14b derivative encoding PhaC-Intein-Rv1626 | This study |
| pET14b-PhaC-Intein-ZZ[a] | pET14b derivative encoding PhaC-Intein-ZZ | This study |
| pET14b-PhaC-Intein-TNFα[a] | pET14b derivative encoding PhaC-Intein-TNFα | This study |
| pET14b-PhaC-Intein-IFNα2b[a] | pET14b derivative encoding PhaC -Intein-IFNα2b | This study |
| pET14b-PhaC-Intein-G-CSF[a] | pET14b derivative encoding PhaC -Intein-GCSF | This study |

| Plasmid | Description* | Sources or references |
|---|---|---|
| **Plasmid for expression of fusion proteins between PhaC and lysozyme binding OBody** | | |
| pET14b | Amp$^r$; T7 promoter | Novagen |
| pETC[a] | pET14b derivative encoding PhaC | (Peters & Rehm 2008) |
| pMCS69[b] | Cm$^r$; pBBR1MCS derivative encoding PhaA and PhaB | (Amara & Rehm 2003) |
| pProEx HTb-L200EP-06 | Amp$^r$; pProEx HTb derivative encoding OBody L200EP-06 | (Steemson 2011) |
| pET14b-GFP-PhaC | pET14b derivative encoding GFP-PhaC | (Jahns *et al.* 2013) |
| pET14b-PhaC-linker-SG linker-GFP | pET14b derivative encoding PhaC-linker-SG linker-GFP | (Jahns & Rehm 2009) |
| pET14b-O6-PhaC[a] | pET14b derivative encoding OBody L200EP-06-PhaC | This study |
| pET14b-PhaC-O6[a] | pET14b derivative encoding PhaC-linker-SG linker-OBody L200EP-06 | This study |
| **Plasmid for expression of fusion proteins between PhaC and progesterone binding OBody** | | |
| pET14b | Amp$^r$; T7 promoter | Novagen |
| pETC[a] | pET14b derivative encoding PhaC | (Peters & Rehm 2008) |
| pMCS69[b] | Cm$^r$; pBBR1MCS derivative encoding PhaA and PhaB | (Amara & Rehm 2003) |
| pProEx-B7 | Amp$^r$; pProEx derivative encoding OBody B7 | Vickery Arcus |
| pET14b-GFP-PhaC | pET14b derivative encoding GFP-PhaC | (Jahns *et al.* 2013) |
| pPOLY-N | pETC derivative containing *Nde*I, *Xma*I, *Sma*I and*Spe*I sites upstream of *phaC* | (Hay *et al.* 2014) |
| pPOLY-C | pETC derivative containing containing *Stu*I, *Xho*I, *Xma*I, *Sma*I and *Bam*HI sites downstream of *phaC* | (Hay *et al.* 2014) |
| pET14b-PhaC-D7[a] | pET14b derivative encoding PhaC-OBody P4013-D7 | This study |
| pET14b-3xD7-PhaC[a] | pET14b derivative encoding 3xOBody P4013-D7-PhaC | This study |
| pET14b-3xD7-PhaC-D7[a] | pET14b derivative encoding 3xOBody P4013-D7-PhaC-OBody P4013-D7 | This study |
| pET14b-D7-PhaC[a] | pET14b derivative encoding OBody P4013-D7-PhaC | This study |
| pET14b-B7-PhaC[a] | pET14b derivative encoding OBody B7-PhaC | This study |
| pET14b-3xB7-PhaC[a] | pET14b derivative encoding 3xOBody B7-PhaC | This study |

[a] Plasmid A encoding the PHA synthase PhaC or PhaC fusions necessary for PHA polymerisation.

[b] Helper plasmid B encoding PhaA (β-ketothiolase) and PhaB (acetoacetyl-CoA reductase) preparing *R*-(3)-hydroxybutyryl-CoA substrate required for PhaC to produce PHA; or additionally encoding Erv1p (sulfhydryl oxidase) helpful for disulphide bond formation.

[*] Amp$^r$, Ampicillin resistance; [*] Cm$^r$, chloramphenicol resistance.

## 2.2 Medium and cultivation conditions

### 2.2.1 *E. coli* growth medium

Two types of *E. coli* growth media were used in this study. Luria-Bertani (LB) medium was used for plasmid propagation and standard bead production. Terrific Broth (TB) medium was used for shorter periods of bead production to minimize premature cleavage of intein or sortase caused by cytosolic $H^+$ or $Ca^{2+}$ level. This is because TB, as compared to LB, contains increased concentrations of peptone and yeast extract such that sufficient amount of cell biomass and beads could be accumulated in shorter periods; and TB also contains a medium pH buffer solution which could counteract media acidification and the cytosolic pH drop effect that occurs during cell cultivation.

All media were autoclaved at 121°C for 20 min. Unless stated, all supplements were sterilised either by autoclaving or filtration through a 0.22 µm filter before being added to sterile autoclaved media.

#### 2.2.1.1 Luria-Bertani (LB) medium

LB medium was prepared by adding 20 g of LB (Lennox) powder (Acumedia, Lansing, USA) per litre of distilled water.

LB-agar medium was made by adding 16 g of agar (Acumedia) per litre of LB medium.

#### 2.2.1.2 Terrific Broth (TB) medium

TB medium was prepared by adding 12 g of tryptone (BD, Franklin Lakes, USA), 24 g of yeast extract (Merck, Kenilworth, USA) and 4 ml of glycerol (Thermo Fisher Scientific, Waltham, USA) per 900 ml of distilled water.

Then before use, 100 mL of sterile medium pH buffer solution was added; this was either 10×phosphate buffer (0.17 M $KH_2PO_4$, 0.72 M $K_2HPO_4$, pH 7.4) for production of beads displaying PhaC-sortase or PhaC-sortase-LPETG-target protein fusions, or 10×HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)) buffer (250 mM, pH 8.6) for production of beads displaying PhaC-intein-target protein fusions.

### 2.2.2 *E. coli* cultivation conditions

*E. coli* cultivation conditions such as antibiotic addition, IPTG (isopropyl-β-D-thiogalactopyranoside) induction, glucose supplementation, growth temperature and incubation time were as described below.

#### 2.2.2.1 Antibiotic, IPTG and glucose stocks and working concentrations

Antibiotics solutions used in this study are listed in Table 2.3. All solutions were prepared as previously described (Hay *et al.* 2014) and were passed through a 0.22 µm filter for sterilization, aliquoted and stored at -20°C for further use. Addition of antibiotics into autoclaved LB-agar media was done after cooling down the media to approximately 50°C.

**Table 2.3** *Antibiotic, IPTG and glucose stocks and working concentrations*

| Antibiotic | Stock solution | Working concentration (dilution from stock solution) |
|---|---|---|
| Ampicillin Sodium salt (Applichem, Darmstadt, Germany) | 100 mg/ml in Milli-Q water | 100 μg/ml (1:1000) |
| Chloramphenicol (Gold Biotechnology St. Louis, USA) | 50 mg/ml in 100% EtOH | 50 μg/ml (1:1000) |
| Tetracycline (Sigma, St. Louis, USA) | 12.5 mg/ml in 70% EtOH | 12.5 μg/ml (1:1000) |
| IPTG (Gold Biotechnology) | 1 M in Milli-Q water | 1 mM (1:1000) |
| Glucose (Merck) | 25% w/v in Milli-Q water | 1% (1:25) |

### 2.2.2.2 Cultivation conditions for plasmid propagation

*E. coli* strains containing plasmids (Table 2.2) were grown overnight at 37°C in an incubator oven on LB-agar plates (2.2.1.1, supplemented with antibiotics as appropriate (Table 2.3)), or in shaken (200 rpm) Erlenmeyer flasks containing about 1/5 flask volume of liquid LB medium (with antibiotics as appropriate), as previously described (Blatchford *et al.* 2012).

### 2.2.2.3 Cultivation conditions for bead production

For standard bead production, overnight pre-cultures were prepared similarly as mentioned above for plasmid propagation in liquid LB medium (2.2.2.2). Then sufficient amount of the resulting overnight pre-cultures were inoculated into 1 litre of LB media (2.2.1.1) supplemented with 1% w/v glucose and antibiotics as appropriate (in 5 litre Erlenmeyer flasks) to give main cultures with a starting $OD_{600}$ of 0.1. The main cultures were grown at 37°C for about 3 h to reach an $OD_{600}$ of 0.5 to 0.8, induced with 1 mM IPTG, and allowed to grow at 25°C for additional 45 h, as previously described (Hay *et al.* 2014).

For production of beads displaying PhaC-sortase or PhaC-sortase-LPETG-target protein fusions, similar conditions were used, except that main cultures were grown in TB media with 1×phosphate buffer as medium pH buffer solution (2.2.1.2), and after IPTG induction allowed to grow only for additional 20 h.

For production of beads displaying PhaC-intein-target protein fusions, similar conditions were used, except that main cultures were grown in TB media with 1×HEPES buffer as medium pH buffer solution (2.2.1.2), and after IPTG induction allowed to grow for additional 20 h at 22°C (to avoid premature intein cleavage at 25°C which is the optimum temperature for intein activity). In addition, in order to counteract media acidification and the cytosolic pH drop effect that occurs during cell cultivation, an extra 25 mM HEPES (pH 8.6) was added manually every 3 h for the first 12 h.

### 2.2.3 Short and long term storage of *E. coli* strains

For short term storage of *E. coli* strains, overnight cultures as described in 2.2.2.2 were streaked onto LB-agar plates (with antibiotics as appropriate). After an overnight incubation at 37°C, the plates were sealed with Parafilm and stored at 4°C for up to one month.

For long term storage of *E. coli* strains, 1 ml of the overnight culture from liquid LB medium (2.2.2.2) was mixed with 70 µl of sterile dimethylsulfoxide (DMSO) in a 2 ml cryovial tube and stored at -80°C. To revive *E. coli* strains, a small chip of frozen stock was removed via a sterile pipette tip and inoculated into sterile liquid LB medium (2.2.1.1) containing appropriate antibiotics.

## 2.3 Preparation of chemically competent *E. coli* cells

Competent *E. coli* cells were prepared as described elsewhere (Hanahan 1983). 50 ml of liquid LB medium (2.2.1.1) was inoculated with 0.5 ml inoculum of an overnight culture and incubated at 37°C for 2-3 h until the $OD_{600}$ reached approximately 0.3-0.5. The cell culture was left on ice for about 15 min and then harvested by centrifugation at 8,000×g for 15 min. The cell pellets were re-suspended in 16 ml of cold RF1 solution and left on ice for 30 min. Cells were spun again at 8,000×g for 15 min and then re-suspended in 4 ml of cold RF2 solution. Finally, 200 µl aliquots of the resulting competent cells were transferred into 1.5 ml sterile microcentrifuge tubes, snap frozen in liquid nitrogen and stored at -80°C for future use. The recipes for RF1 and RF2 solutions are detailed as below, and both were sterilized by filtration through a 0.22 µm filter:

**RF1 solution:**

100 mM RbCl, 50 mM $MnCl_2$, 30 mM KAc, 10 mM $CaCl_2 \cdot 6H_2O$, pH 5.8 adjusted with acetic acid.

**RF2 solution:**

10 mM RbCl, 10 mM MOPS, 75 mM $CaCl_2 \cdot 6H_2O$, 15 mM Glycerol, pH 5.8 adjusted with NaOH.

## 2.4 DNA manipulation and molecular cloning

### 2.4.1 Polymerase chain reaction (PCR)

PCR was used to amplify target DNA fragments either for verification (standard PCR) or subcloning (high fidelity PCR) (Sambrook *et al.* 1989). All PCRs were performed in 200

µl PCR vials and stored at -20°C for further use. Sequences of PCR primers are provided in Table 2.4.

**Table 2.4** *PCR primers used in this study*

| Primer name | Sequence from 5' to 3' (with restriction site underlined) | Used for |
|---|---|---|
| **Primer for standard PCR or DNA sequencing** | | |
| 5pET14b | GTAGTAGGTTGAGGCCGTTGA | Insert verification |
| N-phaC_R | CGATCTTGACGCCTGCCAGC | Insert verification |
| C-phaC_F | AGCCACTGGACTAACGATGC | Insert verification |
| T7_terminator | GCTAGTTATTGCTCAGCGG | Insert verification |
| T7_promoter | TAATACGACTCACTATAGGG | Insert verification |
| M13_F | CCCAGTCACGACGTTGTAAAACG | Insert verification |
| M13_R | AGCGGATAACAATTTCACACAGG | Insert verification |
| **Primer for high fidelity PCR** | | |
| O6_*Spe*I_F | CCGACTAGTGTGTATCCTAAAAAGACCCACTGGACC | pET14b-O6-PhaC |
| O6_*Spe*I_R | ATAACTAGTGTCTATTGGAAGCGGCTTGGCCTTG | pET14b-O6-PhaC |
| O6_*Sma*I_F | GATACCCGGGGTGTATCCTAAAAAGACCCACTGGACC | pET14b-PhaC-O6 |
| O6_*Bam*HI_R | TATGGATCCGTCTATTGGAAGCGGCTTGGCCTTG | pET14b-PhaC-O6 |
| D7_ *Spe*I_F | TCACTAGTATGGCTACGCATTGGACC | pET14b-D7-PhaC |
| D7_ *Spe*I_R | AGACTAGTATGGTGATGGTGGTGGTG | pET14b-D7-PhaC |
| B7_*Nde*I_F | TTCATATGGCCACCCACTGGACC | pET14b-B7-PhaC |
| B7_*Spe*I_R | AGACTAGTATGGTGATGGTGGTGGTGTTCCAGAGCGG CAGCGTCTATTGGAAGCGGC | pET14b-B7-PhaC |

### 2.4.1.1 Standard PCR

Standard PCR was performed with *Taq* DNA polymerase (Fisher Scientific International, Pittsburgh, USA). The reaction mixture (per 100 µl) contained 10 µl of 10×*Taq* reaction buffer without $MgCl_2$, 10 µl of $MgCl_2$ (at 25 mM), 5 µl of DMSO; 10 µl of each primer (10 pmoles/µl), 10 µl of dNTPs (10 mM of dATP, dTTP, dCTP and dGTP), 5-10 ng of template DNA and 1 µl of *Taq* polymerase. The following conditions were used: one cycle at 94°C for 2 min (for hot start); 35 cycles at 94°C for 30 s (for denaturing), 45-68°C (generally 5°C below the lowest $T_m$ of the primer pair) for 30 s (for annealing) and

72°C for 1 min per kb (for extension); one cycle at 72°C for 10 min (for completion of extension).

#### 2.4.1.2 High fidelity PCR

For high fidelity PCR, the hot start Platinum® *Pfx* DNA polymerase (Invitrogen, Carlsbad, USA) was used. The reaction mixture (per 100 µl) contained 10 µl of 10×*Pfx* reaction buffer, 10 µl of $MgSO_4$ (at 25 mM), 10 µl of PCR enhancer buffer; 10 µl of each primer (10 pmoles/µl), 10 µl of dNTPs (10 mM of dATP, dTTP, dCTP and dGTP), 5-10 ng of template DNA and 1 U of *Pfx* polymerase. The following conditions were used: one cycle at 94°C for 2 min (for hot start); 30 cycles at 94°C for 15 s (for denaturing), 45-68°C (generally 5°C below the lowest $T_m$ of the primer pair) for 30 s (for annealing) and 68°C for 1 min per kb (for extension); one cycle at 68°C for 10 min (for completion of extension).

### 2.4.2 Plasmid isolation and quantification

For plasmid isolation, 3-5 ml of overnight *E. coli* cultures (2.2.2.2) were collected by centrifugation at 8000×g for 1 min, and plasmid isolation was performed by alkaline lysis and extraction using the High Pure Plasmid Isolation Kit (Roche, Basel, Switzerland) according to the manufacturer's instructions.

Plasmids were quantified using a NanoDrop 1000 (Thermo Scientific, Waltham, USA) according to the manufacturer's instructions.

### 2.4.3 DNA digestion with restriction endonucleases

DNA digestion were performed according to laboratory protocols (Sambrook *et al.* 1989). Generally, 1-5 µg of plasmid DNA was digested at 37°C (or 25°C in rare cases) for 2 h with 20 U of restriction endonucleases purchased from Roche, Invitrogen or NEB. Double restriction enzyme digestion could be performed in a single reaction in compatible buffers (ideally 100% activity for both enzymes) at a temperature of choice; otherwise, a stepwise enzyme digestion was necessary due to different optimum temperature or buffer conditions. In that case, DNA Clean and Concentrator Kit (Zymo Research, Irvine, USA) was applied according to the manufacturer's instructions to recover enough digested DNA product from the first enzyme digestion reaction.

### 2.4.4 Agarose gel electrophoresis and gel purification of DNA fragment

Separation of DNA fragments was achieved with agarose gel electrophoresis (AGE) (Sambrook *et al.* 1989). Typically, 1-3% agarose gels were made in 1×TBE Buffer (50 mM Tris-HCl, 50 mM Boric acid, 2.5 mM EDTA (Ethylenediaminetetraacetic acid), pH 8.0) depending on fragment sizes being separated (1% for DNA fragments >1000 bp, 2% for DNA fragments <1000 bp and 3% for DNA fragments <100 bp). DNA samples were mixed with 6×loading dye (60% (v/v) glycerol, 20 mM Tris-HCl pH 8.0, 60 mM EDTA, 0.03% (w/v) bromophenol blue, 0.03% (w/v) xylene cyanol FF) prior to well loading, along with a suitable molecular size standard. Gel electrophoresis was run in 1×TBE buffer at about 100-150 V for 30-60 min depending on the gel electrophoresis chamber (5-8 V/cm of the distance between anode and cathode) and degree of separation required. Gels were stained for about 30 min in ethidium bromide solution (2 µg/ml) and destained

for 1 min in distilled water. Gels were visualised using a UV transilluminator (Bio-Rad, Gel Doc™ EZ system, Hercules, USA).

In order to extract and purify specific PCR products or restriction digest fragments, SYBR safe DNA gel stain (Invitrogen) was used according to manufacturer's instructions. Corresponding DNA bands were excised from gels using sterile scalpel blades under blue light (Safe Imager™ 2.0 Transilluminator, Invitrogen). The gel slice was transferred to a clean microcentrifuge tube, and the DNA fragment purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research) according to the manufacturer's instructions.

## 2.4.5 DNA ligation

Desired inserts and vectors were digested with appropriate restriction endonucleases (2.4.3) and gel purified (2.4.4). For ligation (Sambrook *et al.* 1989), inserts were combined with vectors in a 3:1 molar ratio, together with 1 µl of T4 DNA ligase (Invitrogen), and 3 µl of 5×DNA ligase buffer in a final volume of 15 µl. The reaction mixture was incubated overnight at 4°C (or at room temperature for 2 h).

### 2.4.5.1 A-tailing

Where pGEM-T Easy vector (Promega, Madison, USA), an intermediate cloning vector, was necessary for AT ligation with a PCR product generated by a proofreading DNA polymerase, the PCR product was first A-tailed by *Taq* DNA polymerase to facilitate the cloning (Trower & Elgar 1996). The PCR product was first gel purified (2.4.4) and then incubated at 72°C for 30 min with 0.5 µl of *Taq* DNA polymerase, 1 µl of 10×*Taq* reaction buffer without $MgCl_2$, 1 µl of $MgCl_2$ (at 25 mM), and 0.2 mM dATP in a final volume of

10 µl. An aliquot of the A-tailed PCR product was then incubated overnight at 4°C (or at room temperature for 2 h) with 0.5 µl pGEM-T Easy vector, 5 µl pGEM-T Easy 2×ligase buffer and 1 µl pGEM-T Easy ligase in a final volume of 10 µl.

#### 2.4.5.2 Alkaline phosphatase treatment of vectors

To prevent re-ligation of linearized plasmids created by restriction endonucleases (2.4.3), Antarctic phosphatase (NEB) was used according to manufacturer's instructions for the removal of the 5' phosphate group.

### 2.4.6 Transformation of *E. coli* cells

The transformation of *E. coli* cells has been described elsewhere (Sambrook *et al.* 1989). 200 µl aliquots of frozen *E. coli* competent cells (2.3) were thawed on ice before the addition of 1-3 µl of purified plasmid DNA or 7.5 µl ligation mix, tapped briefly and left on ice for 45 min. Cells were heat-shocked at 42°C for 90 s then immediately put back on ice for 5 min. 800 µl of liquid LB media was added and cells were incubated at 37°C with shaking (200 rpm) for 1 h. 100 µl of the cells were spread onto a LB-agar plate containing selective antibiotics as required and, once dry, the plate was incubated overnight at 37°C.

### 2.4.7 DNA sequencing

DNA sequencing of recombinant plasmids was performed by the Massey University Genome Service using a capillary ABI3730 Genetic Analyzer (Applied Biosystems Inc.). DNA sequencing samples were prepared in sterile 0.2 ml thin-walled PCR tube (Axygen,

Union City, USA) containing 400 ng of DNA template and 4 pmols of sequencing primer (Table 2.4) in a final volume of 20 µl. Sequence data were assembled and analysed by Vector NTI Advance® 11.5.3 (Invitrogen).

## 2.4.8 Preparation of plasmid constructs for this study

Plasmids for expression of PhaC-sortase or PhaC-sortase-LPETG-target protein fusions were prepared as below:

In order to test whether Class A sortase from *Staphylococcus aureus* could be functionally immobilised on PHA beads, its catalytic region (namely the coding region minus the N-terminal membrane anchor region) (Ilangovan *et al.* 2001) (amino acids 60-206, GenBank accession number WP_053875978) ($SrtA_{\Delta N59}$) was codon optimised against *E. coli* and synthesized by Genscript (Piscataway, USA) with flanking *Xho*I and *Bam*HI sites. The product was digested with *Xho*I and *Bam*HI and ligated into the corresponding sites on the plasmid pET14b-phaC-linker-MalE (Jahns & Rehm 2009), resulting in the plasmid pET14b-PhaC-SrtA as shown in Figure 7.1, Appendix 7.1.

Plasmids pET14b-PhaC-SrtA-TNFα and pET14b-PhaC-SrtA-IFNα2b (as respectively shown in Figure 7.2 and 7.3, Appendix 7.1) were constructed by Iain Hay for corresponding therapeutic protein purification via self-cleavage of sortase displayed on bead surface. The *tnfα* gene encodes a soluble form of human tumour necrosis factor alpha (TNFα) (amino acids 77-233, GenBank accession number NP_000585) and the *ifnα2b* gene encodes human interferon alpha 2b (IFNα2b) without signal peptide (amino acids 24-188, GenBank accession number NP_000596).

Plasmids for expression of PhaC-intein-target protein fusions were prepared as below:

This group of plasmids was designed for corresponding target protein purification via self-cleavage of intein displayed on bead surface.

For plasmid pET14b-PhaC-Intein-GFP as shown in Figure 7.4, Appendix 7.1, the *phaC* gene flanked by *Xba*I and *Nde*I sites was excised from plasmid pET14b-PhaC-linker-GFP (Jahns & Rehm 2009) and inserted into the corresponding *Xba*I and *Nde*I sites on the plasmid pTWIN1 (NEB); then the resulting *phaC-intein* fusion gene between *Xba*I and *Xho*I sites was excised and ligated back into the corresponding *Xba*I and *Xho*I sites of the original plasmid pET14b-PhaC-linker-GFP. The *gfp* gene encodes green fluorescent protein (GFP) minus the starting methionine from *Aequorea victoria* (amino acids 2-238, GenBank accession number P42212).

Then, for plasmid pET14b-PhaC-Intein-RV1626 as shown in Figure 7.5, Appendix 7.1, the *rv1626* gene flanked by *Xho*I and *Bam*HI sites was excised from plasmid pPOLY-C-phaC-Rv1626 (Rubio Reyes *et al.* 2016) and inserted into the corresponding *Xho*I and *Bam*HI sites of plasmid pET14b-PhaC-Intein-GFP. The *rv1626* gene encodes a full length putative transcriptional antiterminator Rv1626 from *Mycobacterium tuberculosis* (amino acids 1-205, GenBank accession number 1S8N_A).

Similarly, for plasmid pET14b-PhaC-Intein-ZZ as shown in Figure 7.6, Appendix 7.1, the *zz* coding region flanked by *Xho*I and *Bam*HI sites was excised from plasmid pET14b-PhaC-linker-ZZ (Jahns *et al.* 2013) and inserted into the corresponding *Xho*I and *Bam*HI sites of plasmid pET14b-PhaC-Intein-GFP. The *zz* coding region encodes a synthetic IgG

binding ZZ domain of protein A derived from *Staphylococcus aureus* (amino acids 1-58, GenBank accession number M74186).

To construct plasmid pET14b-PhaC-Intein-TNFα as shown in Figure 7.7, Appendix 7.1, the *tnfα* gene-containing plasmid pET14b-PhaC-SrtA-TNFα was first digested by *Age*I and self-ligated to remove the *phaC-srtA* part; then the resulting plasmid was digested by *Nco*I, at which site a *phaC-intein* fragment flanked by *Nco*I sites (resulted from *Nco*I digestion of pET14b-PhaC-Intein-GFP) was inserted.

Similarly, for plasmid pET14b-PhaC-Intein-IFNα2b as shown in Figure 7.8, Appendix 7.1, the *ifnα2b* gene-containing plasmid pET14b-PhaC-SrtA-IFNα2b was digested by *Age*I and self-ligated to remove the *phaC-srtA* part; then the resulting plasmid was digested by *Nco*I, at which site a *phaC-intein* fragment flanked by *Nco*I (resulted from *Nco*I digestion of pET14b-PhaC-Intein-GFP) was inserted.

While for plasmid pET14b-PhaC-Intein-G-CSF as shown in Figure 7.9, Appendix 7.1, a *g-csf* gene-containing fragment flanked by *Age*I restriction recognition sites was digested from pET14b-PhaC-SrtA-G-CSF (prepared by Iain Hay), and ligated with a *phaC-intein fragment* flanked also by *Age*I sites (resulted from *Age*I digestion of pET14b-PhaC-Intein-TNF). The *g-csf* gene encodes a short isoform of human granulocyte colony-stimulating factor (G-CSF) without signal peptide or VSE after the QEKL residue (amino acids 31-65 and 69-207, GenBank accession number NP_000750).

Plasmids for expression of fusion proteins between PhaC and lysozyme binding OBody were prepared as below:

This group of plasmids was designed to produce beads displaying OBodies that can be used for affinity purification of hen white egg lysozyme.

Plasmid pProEx Htb-L200EP-06 (Steemson 2011) harbouring a coding region for a lysozyme binding OBody L200EP-06 (abbreviated as O6 hereinafter) (a synthetic peptide that was engineered based on OB-fold domain of aspartyl-tRNA synthetase (aspRS) from *Pyrobaculum aerophilum*, with amino acid sequence shown in Figure 7.10, Appendix 7.1) was used as a template for high fidelity PCR (2.4.1.2). The PCR product obtained using primer set O6_*Spe*I_F / O6_*Spe*I_R (Table 2.4) was digested with *Spe*I and ligated into the corresponding *Spe*I site of pET14b-GFP-phaC (Jahns *et al.* 2013) to obtain pET14b-O6-PhaC as shown in Figure 7.11, Appendix 7.1. The PCR product obtained using primer set O6_*Sma*I_F / O6_*Bam*HI_R (Table 2.4) was digested with *Sma*I and *Bam*HI, and ligated into the corresponding sites of pET14b-phaC-linker-SG linker-GFP (Jahns & Rehm 2009) to obtain pET14b-PhaC-O6 as shown in Figure 7.12, Appendix 7.1.

Plasmids for expression of fusion proteins between PhaC and progesterone (P4) binding OBody were prepared as below:

This group of plasmids was designed to produce beads displaying OBodies that have affinity towards P4 thus can be used for P4 detection.

Coding region for a P4 binding OBody P4013-D7 (abbreviated as D7 hereinafter) (a synthetic peptide that was engineered based on OB-fold domain of aspRS from *P. aerophilum*, with amino acid sequence information from the collaborator (Vickery Arcus and the company OBodies Limited (Hamilton, New Zealand)) as shown in Figure 7.13,

Appendix 7.1) was codon optimised against *E. coli* and synthesized by Genscript with flanking *Xho*I and *Bam*HI sites. The product was digested with *Xho*I and *Bam*HI and ligated into the corresponding sites on the plasmid pPOLY-C (Hay *et al.* 2014), resulting in the plasmid pET14b-PhaC-D7 as shown in Figure 7.14, Appendix 7.1.

This plasmid pET14b-PhaC-D7 was then used as a template for high fidelity PCR (2.4.1.2) using primer set D7_*Spe*I_F / D7_*Spe*I_R (Table 2.4). The resulting PCR product was digested with *Spe*I and ligated into the corresponding *Spe*I site of pET14b-GFP-phaC (Jahns *et al.* 2013) with an intention to obtain plasmid pET14b-D7-PhaC. Unexpectedly, a plasmid pET14b-3xD7-PhaC with a triplet insertion as shown in Figure 7.15, Appendix 7.1 was obtained. This was mistaken as a normal single insertion due to poor preliminary sequencing result (covering the restriction site and a further 200 bp or so from each direction) and further cloning work was continued based on it.

Then plasmid pET14b-3xD7-PhaC was digested with *Xho*I and *Bam*HI, and ligated with the *Xho*I and *Bam*HI digested DNA fragment from Genscript product, resulting the plasmid pET14b-3xD7-PhaC-D7 as shown in Figure 7.16, Appendix 7.1.

When later a single copy insertion was found necessary, pET14b-D7-PhaC as shown in Figure 7.17, Appendix 7.1 was obtained by ligating the *Spe*I digested PCR product as mentioned above with a *Spe*I digested pPOLY-N (Hay *et al.* 2014).

When the coding region for a new generation of P4 binding OBody B7 (a synthetic peptide that was engineered based on OB-fold domain of aspRS from *P. aerophilum*, with amino acid sequence shown in Figure 7.18, Appendix 7.1) became available from plasmid pProEx-B7 (gifted by Vickery Arcus from Waikato University), primer set B7_*Nde*I_F /

B7_*Spe*I_R (Table 2.4) was designed to perform high fidelity PCR (2.4.1.2). The resulting PCR product was digested with *Nde*I and *Spe*I and ligated into the corresponding sites of pPOLY-N (Hay *et al.* 2014) so as to obtain plasmid pET14b-B7-PhaC as shown in Figure 7.19, Appendix 7.1.

Furthermore, to prepare a counterpart plasmid for pET14b-3xD7-PhaC (harbouring a triplet coding regions of the first generation of P4 binding OBody P4013-D7), sequence coding for a triplet coding regions of the OBody B7 was synthesized by Genewiz (South Plainfield, USA) with flanking *Nde*I and *Spe*I sites. The product was digested with *Nde*I and *Spe*I and ligated into the corresponding sites on the plasmid pPOLY-N (Hay *et al.* 2014), resulting in the plasmid pET14b-3xB7-PhaC as shown in Figure 7.20, Appendix 7.1.

All plasmids as mentioned above were subjected to DNA sequencing (2.4.7) which confirmed their sequences and maps to be as shown in Appendix 7.1.

## 2.5 Fluorescence microscopy analysis of *E. coli* cells producing PHA beads and isolation of PHA beads

*E. coli* cells for PHA bead production were cultured as described in 2.2.2.3. Prior to cell harvesting and bead isolation, 1 ml of cell culture was sampled for Nile Red staining and fluorescence microscopy analysis in order to assess intracellular accumulation of PHA beads, as previously described (Blatchford *et al.* 2012).

### 2.5.1 Fluorescence microscopy analysis of cells producing PHA beads

Cell pellets were collected by centrifugation at 6000×g for 1 min, then 1 ml of PBS buffer (137 mM NaCl, 2.7 mM KCl, 10.0 mM $Na_2HPO_4$, 1.76 mM $KH_2PO_4$, pH 7.4) and 10 µl of Nile Red solution (0.25 mg/ml in DMSO) were added and mixed thoroughly. The mixture was left in the dark at room temperature for 15 min. Then cells were pelleted again and washed once with 1 ml of PBS buffer. The resulting cell pellet was re-suspended in 1 ml of PBS buffer, of which 3-4 µl was spotted onto a glass slide, and covered with a coverslip. The slide was examined using a fluorescent light microscope (Olympus BX51, Japan) with a PI-41005 filter (excitation = HQ 535/50, emission = HQ 645/75) against Nile Red (excitation, 450-500 nm; emission > 528 nm) under 1000× magnification. For PHA beads carrying GFP (Green fluorescent protein) (excitation, 395/475 nm; emission, 509 nm), a U-MNIB2 Filter (Excitation = 480/20, Emission = 510LP) was also used for fluorescence microscopy. All images were captured using Magnafire$^{TM}$ 2.1 (Optronics International, Muskogee, USA).

### 2.5.2 Isolation of PHA beads

*E. coli* cells were harvested by centrifugation at 8000×g for 15 min, re-suspended via a homogenizer (MICCRA D-9 45132, Müllheim, Germany) to a 10% slurry in a tailored Lysis Buffer (as detailed below) based on a published patent application (Thompson *et al.* 2013), and then mechanically disrupted using a microfluidizer (Microfluidics M-110P, Westwood, USA). Beads were recovered by centrifugation at 6000×g for 30 min at 4°C then washed twice at 8000×g for 30 min with lysis buffer, and stored as a 20% (w/v) or 200 mg/ml (w/v) slurry in Storage Buffer (as detailed below).

**Lysis Buffer**

for standard beads:

25 mM Tris-Cl, 5 mM EDTA, 0.04% (w/v) SDS, pH 9.0

for beads displaying PhaC-sortase-LPETG-target protein fusions:

50 mM Tris-Cl, 150 mM NaCl, 10 mM EDTA, 0.04% (w/v) SDS, pH 8.8

**Storage Buffer**

for standard beads:

PBS, pH 7.4

for beads displaying PhaC-sortase-LPETG-target protein fusions:

50 mM Tris-Cl, 150 mM NaCl, 10 mM EDTA, pH 7.8

for beads displaying PhaC-intein-target protein fusions:

20 mM Tris, 500 mM NaCl, 1 mM EDTA, pH 8.6

## 2.6 Protein manipulation and analysis

In this section, analysis methods related to proteins (being immobilised on PHA beads, or cleaved off beads, or affinity purified by beads) are described.

### 2.6.1 Cleavage of target proteins from isolated PHA beads

For beads displaying PhaC-sortase-LPETG-target protein fusions or PhaC-intein-target protein fusions, after bead isolation according to 2.5.2, activation of the self-cleavage tags (sortase or intein) was performed as below in order to release target protein as soluble fractions.

Before bead activation, 1 ml of the 20% bead slurry was placed into a pre-weighed 1.5 ml tube, pelleted at 8000×g for 4 min, then washed with 1 ml of Urea Washing Buffer (detailed as below) three times to remove any residual impurities. The resulting pellet was

washed once with Pre-Cleavage Washing Buffer (detailed as below), weighed again, re-suspended to a 40% slurry in Cleavage Buffer (detailed as below), and sonicated for 2 min with Elmasonic S 15H unit (Elma Schmidbauer GmbH, Singen, Germany). The beads were incubated on a rotary mixer (Labnet Mini LabRoller, Edison, USA) for 16 h at either 37°C (for beads displaying PhaC-sortase-LPETG-target protein fusions) or 25°C (for beads displaying PhaC-intein-target protein fusions). To isolate the released soluble target protein, the mixture was centrifuged at 17,000×g for 10 min, and the supernatant was removed into a clean tube, and analysed by SDS-PAGE. Note that supernatant resulted from intein cleavage was neutralised with high pH Storage Buffer (2.5.2, but with a pH of 9.1) for beads displaying PhaC-intein-target protein fusions. For PhaC-intein-GFP beads, after supernatant collection and bead sampling, the remaining post-cleavage beads were re-suspended to a 40% slurry in Cleavage Buffer and subjected to a second round of cleavage. In addition, for beads displaying PhaC-intein-therapeutic target protein fusions, an additional 0.2% v/v Tween 20 was included in the Cleavage Buffer to improve solubility of cleaved therapeutic target proteins.

**Urea Washing Buffer**

for beads displaying PhaC-sortase-LPETG-target protein fusions:

50 mM Tris, 10 mM EDTA, 1 M urea, 2% v/v Triton X-100, pH 8.5

for beads displaying PhaC-intein-target protein fusions:

100 mM Tris, 5 mM EDTA, 1 M urea, 2% v/v Triton X-100, pH 8.6

**Pre-Cleavage Washing Buffer**

for beads displaying PhaC-sortase-LPETG-target protein fusions:

50 mM Tris-Cl, 150 mM NaCl, 0.2% v/v Tween 20, pH 7.8

for beads displaying PhaC-intein-target protein fusions:

20 mM Tris, 500 mM NaCl, 1 mM EDTA, pH 6.0

**Cleavage Buffer**

for beads displaying PhaC-sortase-LPETG-target protein fusions:

50 mM Tris-Cl, 150 mM NaCl, 0.2% v/v Tween 20, 5 mM $CaCl_2$, 10 mM triglycine, pH 7.8

for beads displaying PhaC-intein-target protein fusions:

20 mM Tris, 500 mM NaCl, 1 mM EDTA, pH 6.0

## 2.6.2 Protein resolution and identification

Sodium dodecylsulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) (Sambrook *et al.* 1989) was used to resolve proteins of interest, and western blotting or mass spectrometry was used to confirm their identity.

### 2.6.2.1 SDS-PAGE

The denaturing SDS-PAGE conditions used in this study were Bis-Tris (Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane) gel in combination with MOPS (3-(N-morpholino) propanesulfonic acid) running buffer.

#### 2.6.2.1.1 Preparation of Bis-Tris gels

Each Bis-Tris gel was prepared between two clean glass plates separated with a 1.0 mm integrated spacer (Mini PROTEAN® system, Bio-Rad). Each gel consisted of a lower separating gel layer (10-15% w/v) and an upper stacking gel layer (4% w/v), and was prepared from about 4.5 ml of separating gel mixture and about 1.5 ml of stacking gel mixture, respectively:

**Separating gel mixture (5 ml)**

| 10% w/v | 12% w/v | 15% w/v | |
|---|---|---|---|
| 1.43 ml | 1.43 ml | 1.43 ml | 3.5×Bis-Tris gel buffer (1.25 M Bis-Tris, pH 6.5-6.8) |
| 1.67 ml | 2.00 ml | 2.50 ml | 30% Acrylamide/Bis solution 37.5:1 (Bio-Rad) |
| 1.90 ml | 1.57 ml | 1.07 ml | Milli-Q water |

Approximately 3-5 mg of $Na_2SO_3$ were added while mixing the separating gel mixture in a beaker to prevent formation of air bubbles. The polymerisation reaction was started by the addition of 5 µl of N, N, N', N'-tetramethylethyl-endiamine (TEMED) and 10 µl of ammonium persulfate (APS) (40% w/v). 4.5 ml of this solution was gently poured between the two glass plates with a 5 ml pipette tip and a layer of isopropanol (about 0.5 ml) was immediately placed on top of the separating gel layer. The gel was left to set for 0.5-1 hr.

**4% Stacking gel mixture (2 ml)**

| | |
|---|---|
| 0.57 ml | 3.5×Bis-Tris gel buffer |
| 0.26 ml | 30% Acrylamide/Bis solution 37.5:1 (Bio-Rad) |
| 1.17 ml | Milli-Q water |

Once the separating gel had set, the isopropanol was washed out with distilled water. Similarly, 3-5 mg of $Na_2SO_3$ was added while mixing the stacking gel mixture in a beaker to degas, followed by the addition of 5 µl TEMED and 10 µl of APS (40% w/v) to start the polymerisation reaction. 1.5 ml of this solution was gently poured between the two glass plates with a 5 ml pipette tip on top of the separating gel layer and a comb was inserted for the formation of the wells and left to set for 30 min.

#### 2.6.2.1.2 Preparation of protein samples and electrophoresis conditions

For preparation of protein samples, 5 volumes of protein sample were mixed with 1 volume of 6×SDS loading dye (0.375 M Tris, 12% SDS, 60% glycerol, 0.6 M dithiothreitol (DTT), 0.06% bromophenol blue, pH 6.8) in a 1.5 ml microcentrifuge and incubated on a heating block at 95°C for 10 min. The denatured sample was centrifuged at 17,000×g for 5 min before loading 2-20 µl as required into wells. Either Mark12™ Unstained Standard (2.5-200 kDa) (Novex®, Invitrogen™, Thermo Fisher Scientific) or GangNam-STAIN™ Prestained Protein Ladder (10-245 kDa) (iNtRON Biotechnology, Sungnam, Korea) were used as the molecular weight standards for protein size determination.

Standard electrophoresis conditions were 15 mA through the stacking gel layer and 25 mA through the separating gel layer until samples had reached the end of the gel. Electrode buffer (500 ml per gel) was prepared with 100 ml of 5×MOPS running buffer (250 mM MOPS, 250 mM Tris, 0.5% SDS and 5 mM EDTA) and 2.5 ml of 200×reducing agent (1 M sodium bisulphite).

#### 2.6.2.1.3 Protein staining and destaining

After electrophoresis, the gel was carefully removed from the gel plate and transferred to staining solution (4 g of Coomassie blue R-250, 300 ml of ethanol, 100 ml of Acetic acid and 600 ml of distilled water) and stained for 15-30 min with slow shaking.

After staining, the gel was rinsed with distilled water and left in destaining solution (of the same recipe for staining solution but without the dye) until protein bands were visible

and background colour removed. Gels were visualized using a Gel Doc™ EZ system (Bio-Rad).

### 2.6.2.2 Western blotting

Protein bands separated by SDS-PAGE were transferred to a nitrocellulose membrane using an iBlot™ Dry Blotting System (Invitrogen) according to the manufacturer's instructions. The membrane was blocked overnight at 4ºC with 2% w/v BSA in PBST (PBS, pH 7.4, 0.1% v/v Tween-20%, pre-filtered), and then washed with PBST (3×10 min). The membrane was incubated with primary antibody (1:20,000 dilution) in PBST (with 1% w/v BSA) for 1 h at room temperature. After washing with PBST (3×10 min), the membrane was incubated with a horse radish peroxidase (HRP) conjugated secondary antibody (1:20,000 dilution) in PBST (with 1% w/v BSA) for 1 h at room temperature. After washing with PBST (3×10 min), the membrane was incubated with 2 ml of enhancer solution and 2 ml of peroxide solution from SuperSignal® West Pico Chemiluminescent Substrate kit (Thermo Scitentific) for 5 min at room temperature. After excess ECL solution was drained, the membrane was laminated between two plastic sheets, to which X-ray film (Kodak Cat# 165-1454, Rochester, USA) was exposed (10 s to 1 min) in a dark room. The film was then developed using an automated X-ray developer (ALLPRO Imaging, Melville, USA).

### 2.6.2.3 Mass spectrometry

Protein bands of interest separated by SDS-PAGE were subjected to an in-gel digestion with trypsin to obtain tryptic arginine-ending and / or lysine-ending peptide fragments (Shevchenko *et al.* 2006), before submitting for the in-house Liquid Chromatography

with tandem mass spectrometry analysis (LC-MS/MS) (Thermo QExactive Plus) which was kindly conducted by Trevor Loo. Briefly, it consisted of four major steps: gel destaining (with 50% v/v methanol in 50 mM ammonium bicarbonate (ABC) solution), reduction and alkylation of the cysteine / cystine residues (with 10 mM DTT in 50 mM ABC solution and 20 mM iodoacetamide in 50 mM ABC solution, respectively), tryptic cleavage of the protein (with 20 ng/µL trypsin (Sigma Proteomics grade T6567) in 50 mM ABC solution) and extraction of the resulting peptides (collection of the overnight digestion mixture, as well as two extractions: once with 5% v/v formic acid in 50% v/v acetonitrile and once with 0.1% formic acid in 80% acetonitrile).

## 2.6.3 Protein quantification

Total protein concentration was measured by Bradford assay, while concentration in a specific protein band was determined by densitometry analysis.

### 2.6.3.1 Bradford assay for total protein concentration

A Bradford assay (Bradford 1976) was used to quantify protein either immobilised on or purified by PHA beads. 100 µl of serially diluted protein samples (or bead samples) were prepared in a low-binding flat bottom microtitre plate (Greiner Bio-One 655101, Frickenhausen, Germany) along with known BSA (bovine serum albumin) or IgG (immunoglobulin G, GE life sciences) standards. To each well containing a sample or standard, 200 µl of filtered Bradford reagent (Bio-Rad) was loaded and incubated for 5 min in the dark at room temperature for colour development. After incubation, the absorbance was measured at 595 nm using an ELx808iu ultra microtiter plate reader (BIO-TEK Instruments Inc.). Protein concentrations in bead samples were determined by

a standard curve prepared based on absorbance readings of known protein concentrations of BSA (0.05-0.4 mg/ml).

#### 2.6.3.2 Densitometry analysis for specific protein bands

Densitometry was used to quantify a specific protein band based on a known BSA standard by analysing the SDS-PAGE gel image (2.6.2.1) through Image Lab™ Software (Version 3.0, Bio-Rad) according to their user guide.

### 2.6.4 Protein conformation / function assessment

Enzyme-linked immunosorbent assay (ELISA) was performed to measure the specific recognition of a protein antigen by a corresponding antibody. A sortase assay was used to assess the function of sortase displayed on PHA beads, while different types of affinity binding assay were designed to demonstrate protein affinities for their corresponding ligands.

#### 2.6.4.1 ELISA

As a general procedure, a high-binding flat bottom microtitre plate (Greiner Bio-One 655061) was coated at 4°C overnight with 100 µl of serially diluted protein samples (or bead samples) along with proper negative / positive controls. The plate was washed three times with 370 µl of PBST (PBS, pH 7.4, 0.05% v/v Tween-20) and then blocked with 3% w/v BSA in PBST or PBST for 1 h at room temperature. After washing three times with 370 µl of PBST, the plate was then incubated with primary antibody diluted as appropriate in 100 µl of PBS containing 1% BSA for 1 h at room temperature. After

washing three times with 370 µl of PBST, the plate was then incubated with HRP conjugated secondary antibody diluted as appropriate in 100 µl of PBS containing 1% BSA for 1 h at room temperature. After further washing, 100 µl of substrate, SIGMA*FAST*™ OPD (*o*-Phenylenediamine dihydrochloride) tablet (P9187, Sigma), as prepared according to manufacturer's instructions, was added and incubated for 30 min at room temperature. The reaction was stopped by adding 50 µl of 1 N $H_2SO_4$, and the absorbance was measured at 490 nm on an ELx808iu ultra microtiter plate reader (BIO-TEK Instruments Inc., Winooski, USA). Results were presented as optical density units at 490 nm.

For ELISA using therapeutic target proteins, a smaller 80 µl system was used due to the insufficient quantities obtained. TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7.8) was used as blank, TBST (TBS, pH 7.8, 0.05% v/v Tween 20) was used for washing, and TBS containing 1% BSA was used for dilution of primary and secondary antibodies. Respective protein standards and antibodies were all from Sino Biological Inc. (Beijing, China), namely human TNFα protein (10602-HNAE), TNFα Antibody Rabbit PAb (10602-T16), human G-CSF protein (10007-HNCE), G-CSF Antibody Rabbit PAb (10007-T16), human interferon alpha 2 protein (13833-HNAY) and IFNα2 antibody Rabbit PAb (13833-T16). Protein concentration used was 0.2 µg/ml for proteins cleaved from PHA beads, protein standards, as well as primary antibodies. Goat anti-rabbit IgG HRP-conjugate (Abcam ab6721) was used as secondary antibody at 1:3000 dilution. Note these primary antibodies from Sino Biological are conformation-specific antibodies that had been thoroughly tested to recognise only conformationally folded epitopes presented by correctly folded protein antigens in ELISA assays.

For ELISA investigating the specific binding between Rv1626 and its antibody, as well as IgG binding function of the synthetic ZZ domain of protein A derived from *Staphylococcus aureus*, ELISA plates were incubated overnight at 4°C with 50 µl of Rv1626 or ZZ at 1 ng/ul concentration in PBS buffer. And as negative and positive controls, 50 µl of wild type PhaC beads (with an equivalent 50 ng of protein amount in terms of PhaC) or ZZ-PhaC beads ((Brockelbank *et al.* 2006), with an equivalent 50 ng of protein amount in terms of ZZ alone) respectively were included. Then a primary mouse polyclonal anti-Rv1626 antibody (Rubio Reyes *et al.* 2016) was added only for Rv1626 and blank, while a non-specific primary mouse polyclonal anti-DDA antibody (Rubio Reyes *et al.* 2016) only for Rv1626 as a negative control, and a secondary goat anti-mouse IgG HRP-conjugate antibody (Abcam ab6789, UK) to all sample wells for detection of bound IgG antibodies.

### 2.6.4.2 Sortase assay

To assess the function of sortase displayed on PHA beads, a synthetic DABCYL-LPETG-EDANS substrate (AnaSpec, Fremont, USA) was used. This is a fluorescently self-quenched peptide FRET (Förster resonance energy transfer) substrate, the fluorophore EDANS (excitation, 336 nm; emission, 490 nm), a U-MNIB2 Filter (Excitation = 480/20, Emission = 510LP) and the quencher DABCYL is separated by LPETG, the 5 amino acid sortase sorting signal. If the sorting signal is cleaved, then the fluorophore is separated from the quencher and its fluorescence can be detected. Reaction was performed with a black non-binding flat bottom microplate (Greiner Bio-One 655900) by addition of 180 µl of reaction mix per well. The FRET substrate was dissolved in DMSO and added at a final concentration of 5 µM to a 5% slurry of beads in TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7.8) with 5 mM $CaCl_2$. As it was reported that triglycine was not essential for

sortase cleavage activity on TG peptide bond (Ton-That *et al.* 2000), the sortase assay was performed in the absence of triglycine. Plates were incubated at 37°C in dark for 4 h. Fluorescence was monitored over time using a FLUOstar Omega (BMG labtech, Offenburg, Germany) microplate reader with the sample shaking between readings, filters used were 340 nm for excitation, and 520/10 nm for emission.

### 2.6.4.3 GFP fluorescence measurement

To assess the functionality of GFP protein resulted from intein cleavage reaction, 100 µl of GFP-containing soluble fraction samples were mixed with 100 µl of TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7.8) and added into wells of black non-binding flat bottom microplate (Greiner Bio-One 655900). A FLUOstar Omega (BMG labtech, Offenburg, Germany) microplate reader was used to measure the fluorescence of GFP (excitation, 395/475 nm; emission, 509 nm), with a filter setup of 380/10 nm for excitation, and 520/10 nm for emission.

### 2.6.4.4 Lysozyme binding assay

Beads were prewashed once in TBST (150 mM NaCl, 50m M Tris-HCl, 0.1% v/v Tween 20, pH 8.4). Approximately 50 mg of beads were added to 1 ml of TBST containing 2 mg/ml BSA, 2 mg/ml skim milk powder, and 1 mg/ml lysozyme. The beads were resuspended and incubated on a rotary mixer (Labnet Mini LabRoller) for 20 min, after which the beads were sedimented by centrifugation at 6,000×g for 4 min. The supernatant was removed, and the beads were washed three times (each wash consisting of resuspension of the beads, centrifugation, and removal of supernatant) in TBST with 0.05% v/v Tween 20. To elute the bound protein, the beads were resuspended in 500 µl

of 50 mM glycine (pH 2.0); the beads were removed by centrifugation at 16,000×g, and the supernatant was neutralized with 50 µl of 1 M $K_2HPO_4$. The protein content of the resulting elution fractions was assessed by Bradford assay (2.6.3.1).

### 2.6.4.5 Progesterone binding capacity test

A commercial Progesterone EIA Kit (Cayman Chemical, Ann Arbor, USA) was used to assess progesterone (P4) binding capacity of P4 binding OBody beads. The EIA kit assay is based on competitive antibody binding between free P4 in liquid sample and progesterone-acetylcholinesterase (P4 tracer) included with kit, thus colour intensity upon addition of Ellman's reagent is inversely proportional to free P4 amount. Therefore, the EIA kit cannot directly measure the amount of P4 bound by beads, but works indirectly by detecting P4 levels in the solution before and after bead incubation. Thus the amount of P4 reduced after bead incubation could be deemed as that bound by beads (either through the specific biorecognition of P4 by the OBody ligand D7, or through non-specific physical attachment of P4 to PHA beads). Therefore, to consider only the specific binding between P4 and OBody ligand D7, a simple wash step was introduced for beads after P4 incubation to wash off any physically attached P4, which could be detected by the EIA kit, and should be subtracted as well in the calculation of bead P4 binding capacity.

In detail, a P4 stock solution was prepared in absolute ethanol at a concentration of 2 mg/ml. Then 1 volume of the stock solution was mixed with 4 volumes of PBS buffer (137 mM NaCl, 2.7 mM KCl, 10.0 mM $Na_2HPO_4$, 1.76 mM $KH_2PO_4$, pH 7.4), resulting a P4 solution of 400 µg/ml, which is defined as a feed fraction for later quantification of total P4 added. About 20 mg of beads were mixed with 1.5 ml of the feed fraction at 37ºC

for 30 min on a rotary mixer (Labnet Mini LabRoller), after which beads were sedimented by centrifugation at 17,000×g for 10 min. The supernatant was collected as an unbound fraction for later quantification of soluble P4 left. Then beads were washed once with 1.5 ml of PBS and the wash fraction was collected as well for later quantification of P4 that was non-specifically attached to beads and washed off. The feed, unbound and wash fractions were diluted as appropriate to be within the concentration ranges of P4 standard provided by the kit and respectively assayed for P4 content. The P4 binding capacity of beads was calculated as:

$$\frac{Feed\ concentration \times Feed\ volume - Unbound\ concentration \times Unbound\ volume - Wash\ concentration \times Wash\ volume}{grams\ of\ beads\ used}$$

### 2.6.4.6 Progesterone binding assay to obtain apparent equilibrium dissociation constant ($K_D$)

In order to reveal the binding strength (or binding affinity) between progesterone (P4) and OBody beads that is represented by an equilibrium dissociation constant ($K_D$), a P4 binding assay was carried out with fixed amount of beads (or soluble OBody where appropriate) but varying amount of a biotin labelled P4 (progesterone 3-PEG11-biotin (Cayman Chemical), gifted by Vickery Arcus). As compared to the progesterone EIA kit assay in 2.6.4.5, the use of a biotin labelled P4 in combination with HRP-Conjugated Streptavidin (Thermo Fisher Scientific, gifted by Vickery Arcus) would enable a sensitive direct quantification of P4 bound on beads. Here the SIGMA*FAST*™ OPD (*o*-Phenylenediamine dihydrochloride) tablet (Sigma) was used as HRP substrate, reaction was stopped by addition of $H_2SO_4$ and colour intensity was monitored at 490 nm.

Bead samples were assayed at a concentration of 0.5, 0.1 or 0.01 µM bead fusion protein where appropriate (in terms of respective target fusion protein for OBody beads or PhaC itself for WT PhaC beads) based on densitometry analysis. And P4 tested was serially diluted with a concentration range from 0.02 to 43.84 µM. Soluble OBody samples were assayed at a concentration of 0.5 µM.

This P4 binding assay was performed similarly to ELISA described in 2.6.4.1, but in a smaller 50 µl system. The primary antibody was replaced by serially diluted progesterone 3-PEG11-biotin (Cayman Chemical, Ann Arbor, USA), and the secondary antibody was replaced by HRP-Conjugated Streptavidin (Thermo Fisher Scientific). PBS was included as blank, WT PhaC beads were included as negative controls, and wells coated with beads only (without addition of P4) were included as bead background control. Before graphing, PBS blank reading was subtracted from soluble OBody sample reading where OBody sample was assayed, and the bead background reading was subtracted from each bead sample reading.

## 2.7 Statistical Analysis

Analyses were carried out in duplicate (n=2) or triplicate (n=3) and were presented as means ± SD. A one way ANOVA (Analysis of variance) (Minitab® 18) was used to determine statistical significance, and differences were considered significant when $P<0.05$, which means greater than 95% of likelihood or probability that the difference observed in sample means is real and not due to chance.

# Chapter 3: Results

## 3.1 Design of PHA beads as self-cleavable protein purification resins

Recombinant protein production and purification from *Escherichia coli* often involves expensive chromatography equipment and complicated procedures, especially for therapeutic proteins (Jozala *et al.* 2016; Oliveira & Domingues 2018). Even the most efficient affinity chromatography technique that is based on the usage of affinity tags has the disadvantages of additional protease treatment and chromatography for tag removal (Pina *et al.* 2014).

As alternatives to protease treatment for tag removal, self-cleaving tags such as sortase and intein are gaining popularity. For example, the catalytic core of sortase A (SrtA) from *Staphylococcus aureus* has been developed as a self-cleaving tag that is able to recognise a LPXTG signal (X represents any amino acid) and cleaves between the T and G in the presence of $Ca^{2+}$ +/- triglycine (Ton-That *et al.* 2000; Mao 2004; Clancy *et al.* 2010; Matsunaga *et al.* 2010). Furthermore, differently engineered self-clevable inteins that are controllable by pH or thiols have been widely used for recombinant protein purification (Lahiry *et al.* 2017). Recently, a variety of inteins (including pH inducible *Mtu* ΔI-CM mini-intein and *Ssp* DnaB mini intein, as well as thiol inducible *Mxe* GyrA intein) have been used in combination with polyhydroxyalkanoate (PHA) beads for recombinant protein purification (Banki *et al.* 2005; Barnard *et al.* 2005; Wang *et al.* 2008; Zhang *et al.* 2010; Zhou *et al.* 2011). These strategies rely on phasin (PhaP, actually the most studied PhaP1) (Banki *et al.* 2005; Barnard *et al.* 2005; Wang *et al.* 2008; Zhou *et al.*

2011) or a regulatory protein (PhaR) (Zhang *et al.* 2010) that non-covalently associates with PHA beads as "affinity" tags in the form of PhaP-intein-X or PhaR-intein-X fusions; X represents a target protein, and in two studies two or three copies of PhaP were used to increase PhaP hydrophobic association with PHA beads (Banki *et al.* 2005; Zhou *et al.* 2011). Corresponding PhaP or PhaR fusions attached to PHA beads (that were produced either recombinantly (Banki *et al.* 2005; Zhou *et al.* 2011), natively (Barnard *et al.* 2005) or even chemically (Wang *et al.* 2008; Zhang *et al.* 2010)) could be easily separated from other cellular components through centrifugation and washing cycles, and the target protein could be released from beads via subsequent inducible self-cleavage of the inteins (Banki *et al.* 2005; Barnard *et al.* 2005; Wang *et al.* 2008; Zhang *et al.* 2010; Zhou *et al.* 2011). However, the non-covalent anchoring of the target protein to PHA beads can cause leakage of the respective PhaP-intein-X or PhaR-intein-X during the PHA bead wash cycles.

PHA beads are naturally occurring nanometre-scale polyester granules which are accumulated intracellularly under imbalanced nutrient conditions such as excessive carbon (Grage *et al.* 2009). The most well-known type of PHA is composed of poly-*β*-hydroxybutyrate (PHB) homopolymer (Ke *et al.* 2016). Taking PHB biosynthesis as an example, formation of PHA requires three key enzymes, namely, *β*-ketothiolase (PhaA, which condenses two molecules of acetyl-CoA to form acetoacetyl-CoA), acetoacetyl-CoA reductase (PhaB, which reduces the acetoacetyl-CoA to form (*R*)-3-hydroxybutyryl-CoA), and PhaC the PHA synthase (which polymerizes (*R*)-3-hydroxybutyryl-CoA to form PHA) (Normi *et al.* 2005). PhaA and PhaB do not attach to PHA beads, but PhaC remains covalently attached to the nascent PHA chain, which self assembles into beads having a polyester core that is surrounded by the unique covalently bound PhaC. Along with this are various non-covalently bound granule associated proteins (GAPs), such as

the PHA depolymerase (PhaZ) responsible for PHA metabolism, the structure protein phasin (PhaP) that controls the size and number of PHA beads, and the regulatory protein (PhaR) that controls the synthesis of PHA beads and the transcription of PhaP (Rehm 2006; Rehm 2007; Rehm 2010; Draper *et al.* 2013). The unique bead covalent attachment nature of PhaC makes it an ideal anchor to display a target protein on the bead surface mediated via translational fusions between PhaC and the target protein in engineered microbial cells (Grage *et al.* 2009; Jahns & Rehm 2009). Previously, by genetic engineering of PhaC, PHA beads have been designed to display proteins with diverse functions for different purposes such as affinity separation, protein production, enzyme immobilisation, diagnostic testing and vaccine delivery (Grage *et al.* 2009; Hooks *et al.* 2014).

In order to develop a time-efficient and streamlined process of recombinant protein production and purification, the first objective of this PhD study was to display the self-cleaving tag sortase A (SrtA) from *S. aureus* on the surface of PHA beads which serve as purification resins. As depicted in Figure 3.1, it was hypothesised that by introducing a PhaC-sortase-LPETG-target protein fusion expressed by plasmid A, along with PhaA and PhaB proteins expressed by helper plasmid B, *E. coli* cells could be engineered to produce PHA beads displaying PhaC-sortase-LPETG-target protein fusions. Therefore, the target protein could be first produced and sequestered on the natural PHA resins, then separated from contaminating host proteins via simple PHA bead isolation / washing steps, and finally cleaved with a minimal G scar by specific release into the soluble fraction via sortase activation triggered by $Ca^{2+}$ +/- triglycine.

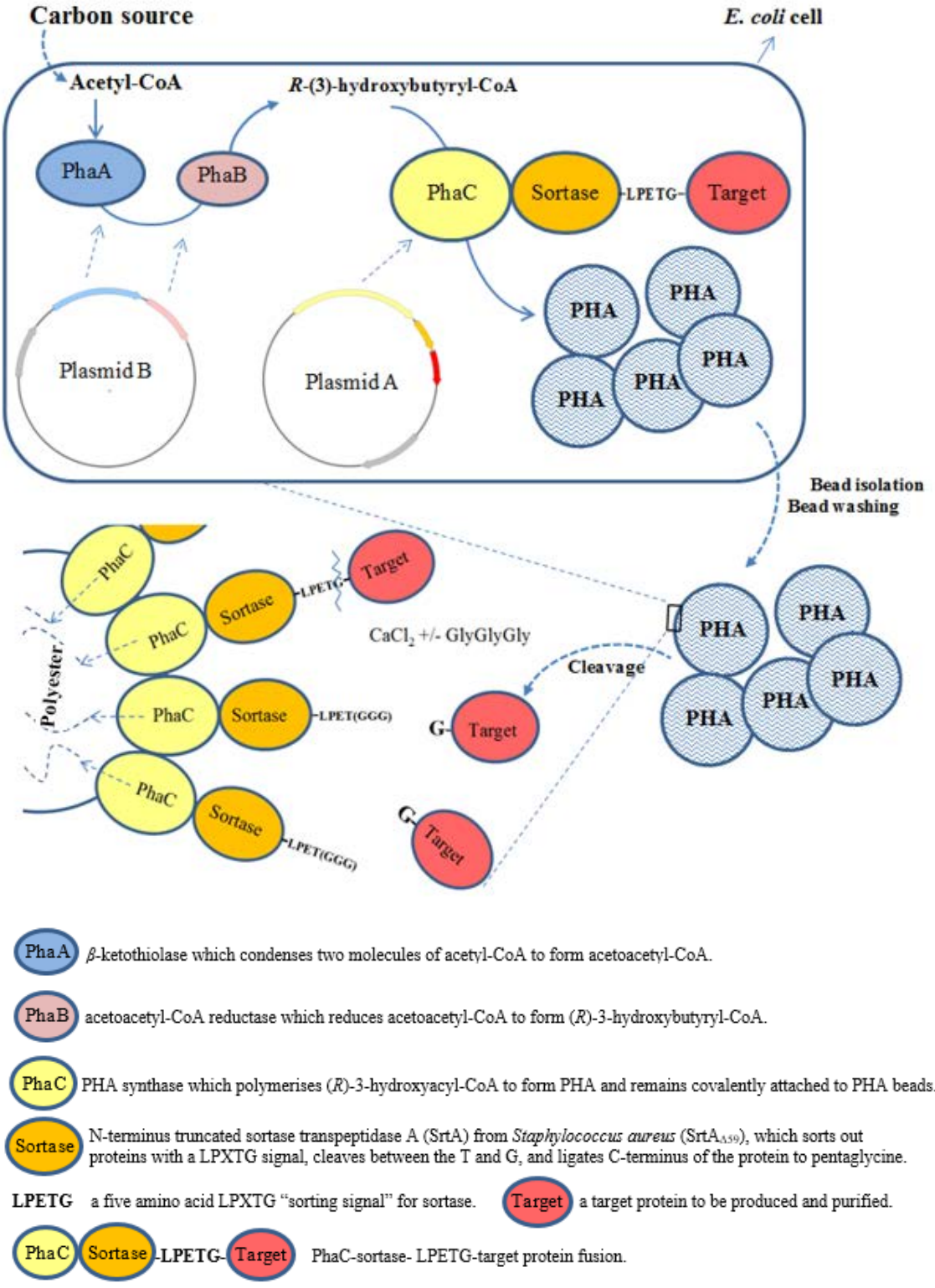



**Figure 3.1 Schematic representation of PHA beads as self-cleavable protein production resins mediated via PhaC-sortase-LPETG-target protein fusion.** Adapted from Hay *et al.* (2015a), Appendix 7.2.

To overcome the potential problems of the requirement for $Ca^{2+}$ +/- triglycine for sortase activation, as well as possible unwanted impacts of the G scars on target proteins in the PhaC-sortase-LPETG-target protein fusion strategy, the second objective of this PhD study involved use of a self-cleaving intein tag as an alternative to sortase. The self-cleaving tag *Ssp* DnaB intein derived from the commercial pTwin1 vector (NEB) was displayed on the surface of PHA beads which serve as purification resins. As depicted in Figure 3.2, it was hypothesised that by introducing a PhaC-intein-target protein fusion expressed by plasmid A, along with PhaA and PhaB proteins expressed by helper plasmid B, *E. coli* cells could be engineered to produce PHA beads displaying PhaC-intein-target protein fusions. Therefore, the target protein could be first produced and sequestered on the natural PHA resins, then separated from contaminating host proteins via simple PHA bead isolation / washing steps, and finally cleaved taglessly by specific release into the soluble fraction via intein activation triggered by a simple pH drop.

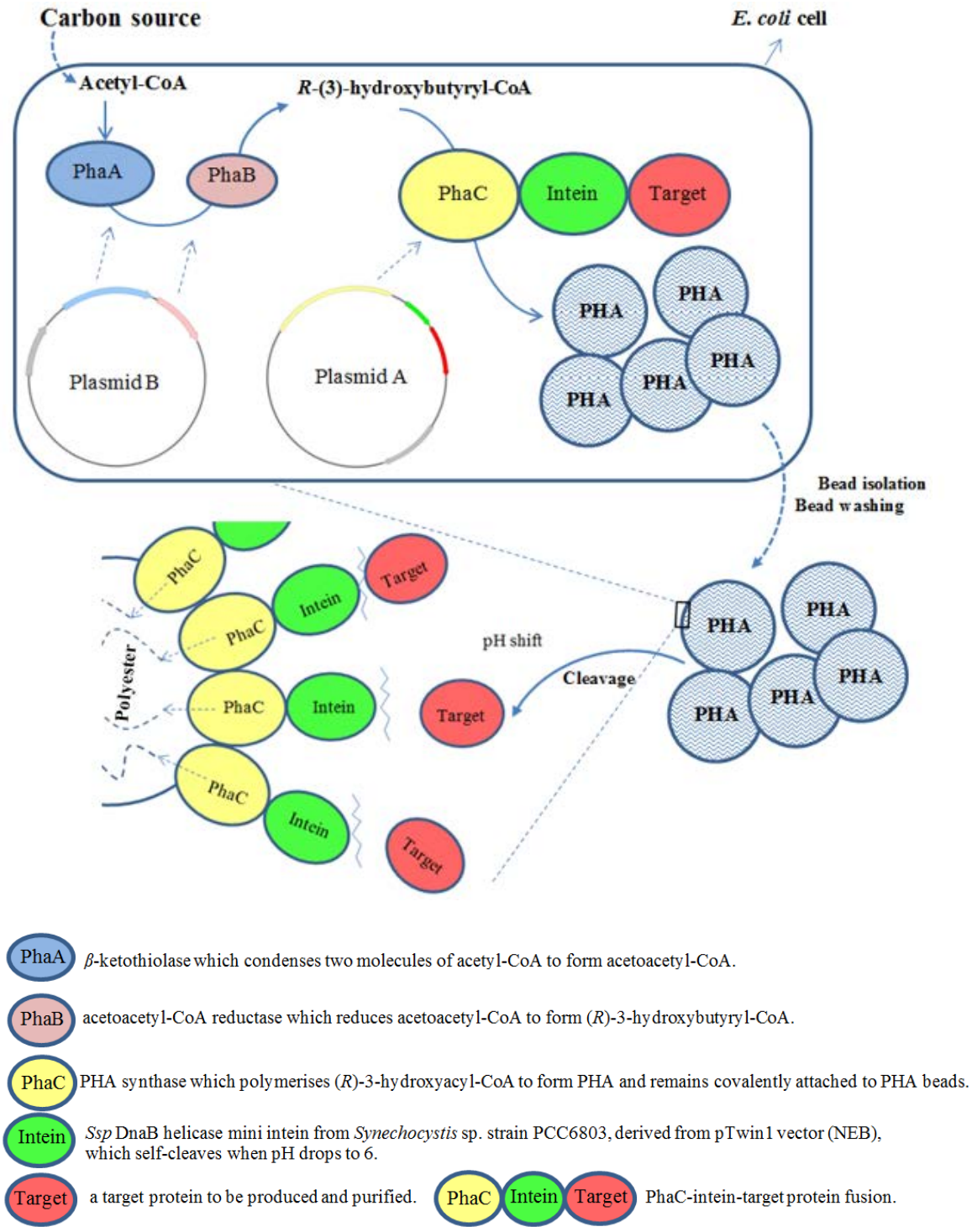



**Figure 3.2 Schematic representation of PHA beads as self-cleavable protein production resins mediated via PhaC-intein-target protein fusion.** Adapted from Du & Rehm (2017b) (Appendix 7.2).

## 3.1.1 PHA beads as self-cleavable protein purification resins mediated via PhaC-sortase-LPETG-target protein fusion

### 3.1.1.1 Sortase from *S. aureus* can be functionally displayed on PHA beads

To first assess whether the N-terminus truncated sortase transpeptidase A (SrtA) from *S. aureus* ($SrtA_{\Delta 59}$) could be functionally displayed on the surface of PHA beads, a PhaC-SrtA fusion expressing plasmid pET14-PhaC-SrtA (Table 2.2) was prepared according to section 2.4.8. The plasmid DNA sequence was confirmed (2.4.7) and its map is shown in Figure 7.1, Appendix 7.1. Plasmid transformation, cell cultivation for production of beads displaying PhaC-sortase, standard bead isolation and SDS-PAGE were performed according to sections 2.4.6, 2.2.2.3, 2.5.2 and 2.6.2.1, respectively.

It was established previously that wild type PHA beads (PhaC beads) could be produced using BL21 (DE3) *E. coli* strain (Table 2.1) harbouring plasmids A and B (namely plasmid pETC encoding PhaC the PHA synthase and plasmid pMCS69 encoding PhaA and PhaB) (Table 2.2) (Peters *et al.* 2007). Similarly, in the current work, when plasmid pETC was replaced by pET14-PhaC-SrtA (with SrtA fused in frame to the C-terminus of PhaC as compared to pETC), a dominant band at the size corresponding to PhaC-SrtA was found as shown in Figure 3.3, suggesting that PHA beads displaying PhaC-SrtA were produced successfully.

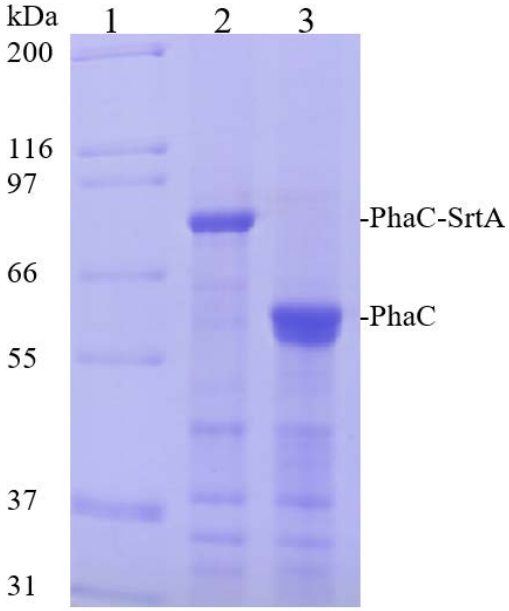



**Figure 3.3 Protein profiles of isolated PHA beads.** 10% SDS-PAGE was performed to examine whether PHA beads were produced with the correct protein profiles. Lane 1, Molecular weight marker (Mark12TM Unstained Standard (Novex®, InvitrogenTM, Thermo Fisher Scientific, Waltham, MA, USA)); Lane 2, PhaC-SrtA beads (~ 83 kDa); Lane 3, PhaC beads (~ 64 kDa).

To assess whether the sortase was functional in the PhaC-SrtA beads, a sortase assay was performed by using DABCYL-LPETG-EDANS, a synthetic fluorescently self-quenched peptide FRET (Förster resonance energy transfer) substrate (2.6.4.2). Specific cleavage of the LPETG by SrtA would separate fluorophore EDANS from quencher DABCYL thus give a fluorescence signal. PhaC or PhaC-SrtA beads were incubated with 5 µM of the FRET substrate in the presence of 5 mM $CaCl_2$ in a black non-binding flat bottom microplate (Greiner Bio-One 655900, Frickenhausen, Germany) in the dark at 37°C. As it was reported that triglycine was not essential for SrtA cleavage activity on the TG peptide bond in the LPETG signal (Ton-That *et al.* 2000), the sortase assay was performed in the absence of triglycine. PhaC-SrtA beads inactivated before assay (by denaturing at 95°C for 15 min) were also included as a bead blank control. Listed in Table 3.1 is a summary of assay reagents added.

**Table 3.1** *Sortase assay to assess cleavage activity of SrtA displayed at the bead surface*

| Column | Test material | FRET substrate* | 5 mM $CaCl_2$* |
|---|---|---|---|
| 1 | PhaC beads | ✓ | ✓ |
| 2 | PhaC-SrtA beads | ✓ | ✓ |
| 3 | PhaC-SrtA beads | ✓ | |
| 4 | PhaC-SrtA beads inactivated at 95°C before assay | ✓ | ✓ |
| 5 | PhaC-SrtA beads | | ✓ |
| 6 | TBS blank | ✓ | ✓ |

* Addition of these reagents is indicated with a tick.

No significant activity could be detected from the PhaC beads as compared to the PhaC-SrtA beads, whereas significant activity could be detected from the PhaC-SrtA beads (Figure 3.4). The activity was dependent on the presence of $CaCl_2$ and could be removed by inactivating the PhaC-SrtA beads at 95°C for 15 min before conducting the assay, indicating that this activity was the result of the SrtA displayed at the bead surface.

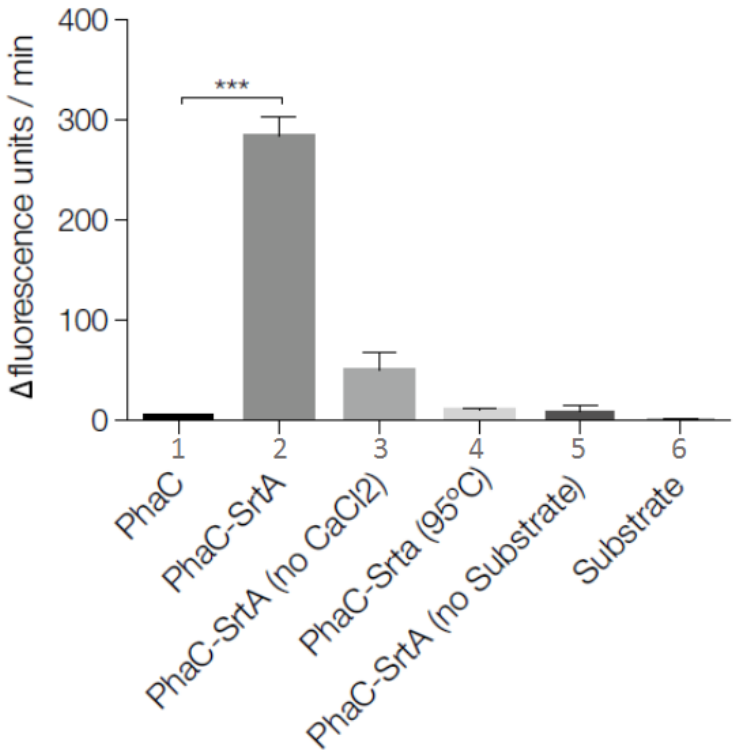



**Figure 3.4 Activity of the PhaC-SrtA beads measured via cleavage of the synthetic DABCYL-LPETG-EDANS substrate.** All reactions were performed in duplicate, and the error bars represent standard deviations. Fluorescent units are arbitrary. ***p = 0.0006. Published in Hay *et al.* (2015a), Appendix 7.2.

### 3.1.1.2 PhaC-sortase-LPETG-target protein fusion facilitated purification of target proteins

To assess whether the sortase displayed at the bead surface could be used for therapeutic protein production and purification, human tumour necrosis factor alpha (TNFα) and human interferon alpha-2b (IFNα2b) were tested. Overexpression of heterologous eukaryotic genes in *E. coli* tends to result in proteins as insoluble inclusion bodies (IBs). For example, one recent study reported that TNFα production in *E. coli* resulted in an IB format product of up to 50% (Zhang *et al.* 2014). Similarly IFNα2b production in *E. coli* was frequently reported to be in the form of IB aggregates (Rabhi-Essafi *et al.* 2007). It was thus also an additional aim to avoid IB formation of these proteins by immobilising

them on PHA beads *in vivo*, as it is widely accepted that immobilisation improves protein stability / solubility (Rehm *et al.* 2016; Rehm *et al.* 2017).

Plasmids encoding PhaC-sortase-LPETG-TNFα or PhaC-sortase-LPETG-IFNα2b (Table 2.2) were prepared according to section 2.4.8. The plasmid DNA sequences were confirmed (2.4.7) and their maps are shown in Figures 7.2-7.3, Appendix 7.1. Plasmid transformation, cell cultivation for production of beads displaying PhaC-sortase-LPETG-target protein fusions, isolation of beads displaying PhaC-sortase-LPETG-target protein fusions, cleavage of target proteins from beads displaying PhaC-sortase-LPETG-target protein fusions, SDS-PAGE, densitometry analysis and mass spectrometry were performed according to methods sections 2.4.6, 2.2.2.3, 2.5.2, 2.6.1, 2.6.2.1, 2.6.3.2 and 2.6.2.3, respectively.

Here, *E. coli* SHuffle® T7 express (Table 2.1) was chosen as the production strain, as TNFα contains one disulphide bond while IFNα2b contains two required for stability / functionality. The SHuffle® T7 express strain has an oxidizing cytosol due to *trxB/gor* mutations and also constitutively produces a cytosolic form of disulphide isomerase (DsbC) that acts as a chaperone. Also pMCS69E was used as helper plasmid B which encodes sulfhydryl oxidase (Erv1p) that can facilitate disulphide bond formation as well as containing the *phaA* and *phaB* genes (Table 2.2).

A dominant band at the size corresponding to PhaC-sortase-TNFα or PhaC-sortase-IFNα2b was found as shown in Figure 3.5 A & B, suggesting that PHA beads displaying corresponding proteins were produced successfully. The isolated PhaC-Sortase-TNFα and PhaC-Sortase-IFNα2b beads each were subjected to a cleavage reaction according to section 2.6.1. As shown in Figure 3.5 A & B, soluble TNFα and IFNα2b could be released

upon activation of the sortase with $CaCl_2$ and triglycine. This amounted to about 7.0 µg TNFα per gram of wet beads or 0.5 µg per gram of wet cell biomass, and about 1.8 µg or 0.1 µg respectively for IFNα2b (Table 3.2). In both cases, the therapeutic test proteins were the predominant proteins in the soluble fractions that resulted from cleavage reactions of the respective isolated PHA beads, accounting for about 80-90% of the soluble proteins (Figure 3.5 A & B and Table 3.2).

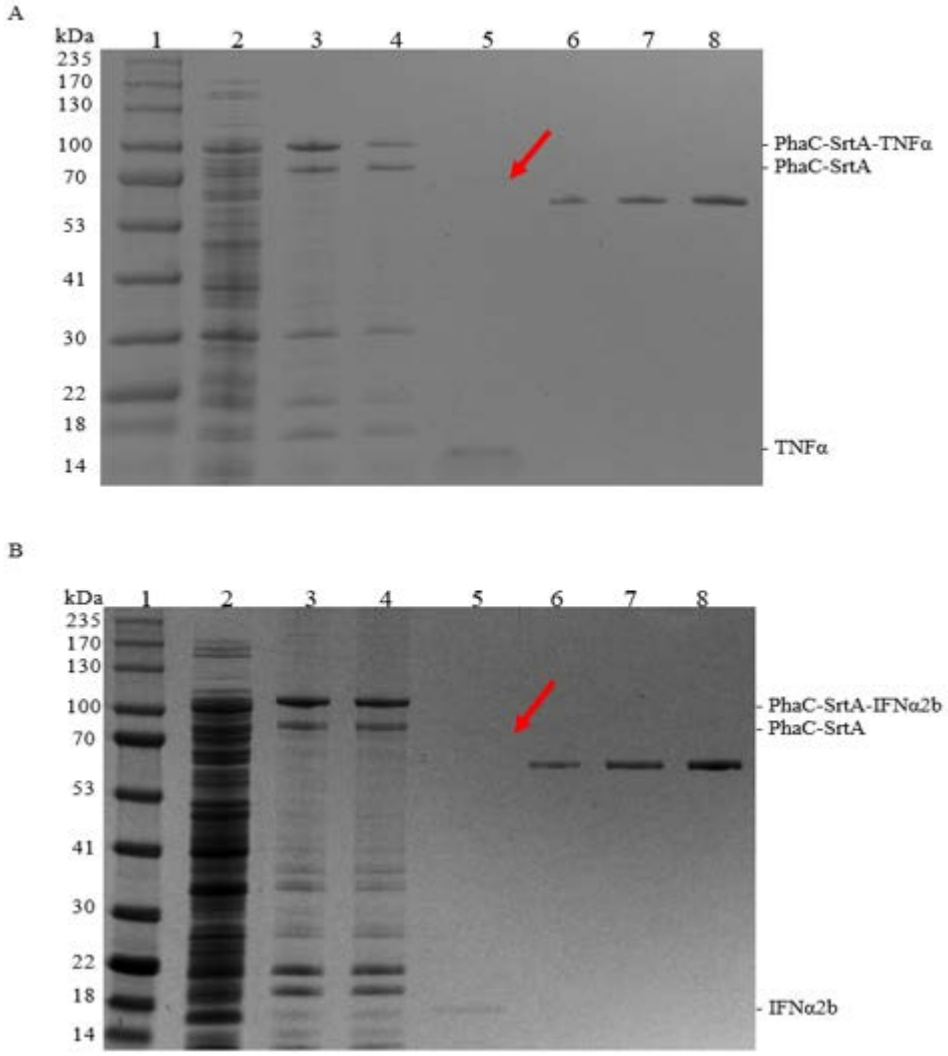



**Figure 3.5 Protein profiles of PHA-bead-producing whole cell lysate, PHA beads isolated and post-cleavage, and the resulting soluble fractions.** 10% SDS-PAGE was performed to examine protein profiles for (A) TNFα: Lane 1, Molecular weight marker (GangNam-STAIN$^{TM}$ Prestained Protein Ladder (iNtRON Biotechnology, Sungnam, Korea)); Lane 2, PhaC-SrtA-TNFα bead producing whole cell lysate; Lane 3, Isolated PhaC-SrtA-TNFα beads (pre-cleavage); Lane 4, PhaC-SrtA-TNFα beads post-cleavage; Lane 5, Cleaved TNFα in the soluble fraction, corresponding to the lowest bottom band (~ 17.4kDa); Lane 6-8, 50, 100 and 200 ng BSA; and (B) IFNα2b: Lane 1, Molecular weight marker (GangNam-STAIN$^{TM}$ Prestained Protein Ladder); Lane 2, PhaC-SrtA-IFNα2b bead producing whole cell lysate; Lane 3, Isolated PhaC-SrtA-IFNα2b beads (pre-cleavage); Lane 4, PhaC-SrtA-IFNα2b beads post-cleavage; Lane 5, Cleaved IFNα2b in the soluble fraction, corresponding to the lowest bottom band (~ 19.3kDa); Lane 6-8, 25, 50 and 100 ng BSA. *Arrow in lane 5 indicates a minor co-purifying protein. *Image contrast was adjusted for visibility. Published in Du & Rehm (2017a) (Appendix 7.2).

**Table 3.2** *Quantification of target proteins obtained from beads displaying PhaC-sortase-LPETG-target protein fusions*[*]

| Target protein | µg/L | µg per gram wet beads | µg per gram wet cell mass | % Cleavage ratio[1] | % purity based on SDS-PAGE[2] | % purity based on ELISA with conformation-specific antibody[3,4] |
|---|---|---|---|---|---|---|
| TNFα | 7.7 | 7.0 | 0.5 | 10.7 | 90.1 | 72.2 |
| IFNα2b | 2.1 | 1.8 | 0.1 | 4.3 | 80.7 | 39.7 |

* Published in Du & Rehm (2017a) (Appendix 7.2).

[1] Estimated cleavage ratio, calculated as the % reduction in PhaC-sortase-LPETG-target protein band levels post-cleavage as compared to pre-cleavage, based on densitometry analysis of the SDS-PAGE images shown in Figure 3.5 A & B;

[2] Estimated purity, calculated as the % of the cleaved target protein band content in the soluble fractions resulting from the cleavage reaction, based on densitometry analysis of the SDS-PAGE images shown in Figure 3.5 A & B;

[3] Estimated purity, calculated as the % of the mean absorbance value for the cleaved target proteins as compared to that for the respective commercial standard based on the ELISA assay results shown in Figure 3.6 A & B;

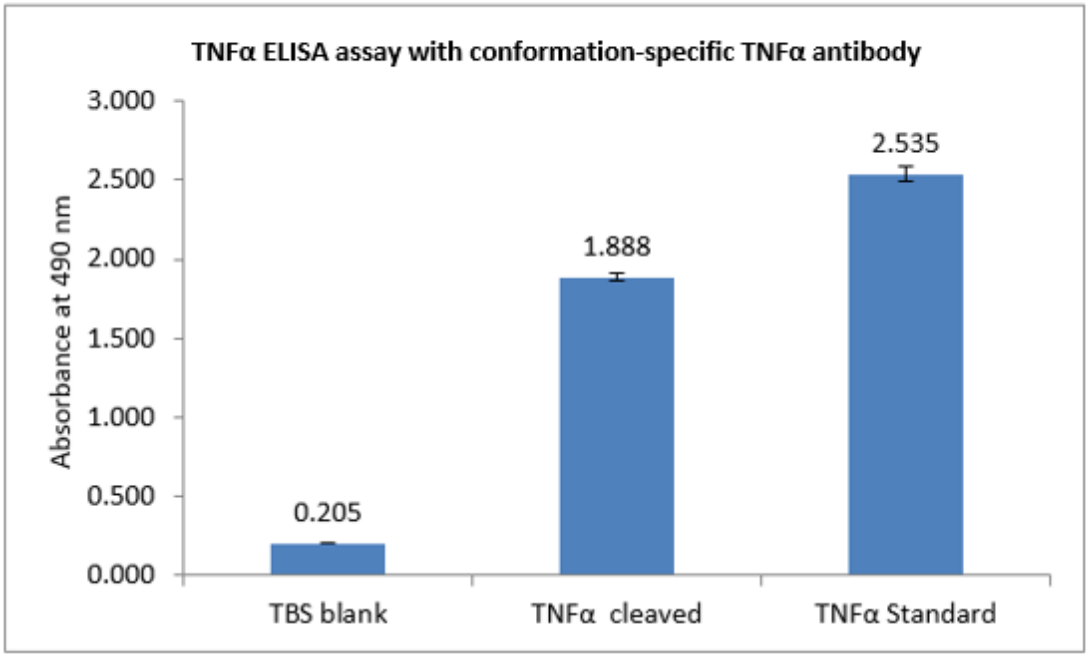
[4] Conformation-specific antibodies as used here were commercial antibodies that had been thoroughly tested to recognise only conformationally folded epitopes presented by correctly folded protein antigens in ELISA assays.

To confirm the identity of both purified target proteins, the bands corresponding to these proteins were excised and subjected to an in-gel trypsin digestion (section 2.6.2.3) followed by LC-MS/MS analysis. The peptide coverage ratios for TNFα and IFNα2b were 90% and 36% respectively (Figure 7.21 A & B, Appendix 7.1). The co-purifying host cell protein was similarly processed with in-gel trypsin digestion (section 2.6.2.3); and the subsequent LC-MS/MS analysis suggested a match to the *E. coli* chaperone protein DnaK with a peptide coverage ratio of 72% (Figure 7.21 C, Appendix 7.1).

### 3.1.1.3 Assessment of proper folding / specific antibody recognition of target proteins cleaved from PHA beads displaying PhaC-sortase-LPETG-target protein

In order to assess the proper folding of the TNFα and IFNα2b cleaved from respective PHA beads (section 3.1.1.2), both TNFα and IFNα2b were analysed by ELISA assays according to section 2.6.4.1 with corresponding conformation-specific antibodies (Sino Biological Inc., Beijing, China) that had been thoroughly tested to recognise only conformationally folded epitopes presented by correctly folded protein antigens in ELISA assays. The ELISA results showed that, at the end of the assay, readings for both cleaved proteins were significant higher than that for TBS blank control, but relatively lower than that for their respective standard(Figure 3.6 A & B). This verified their specific binding with respective conformation-specific antibodies, and indicated their proper folding.

A




B

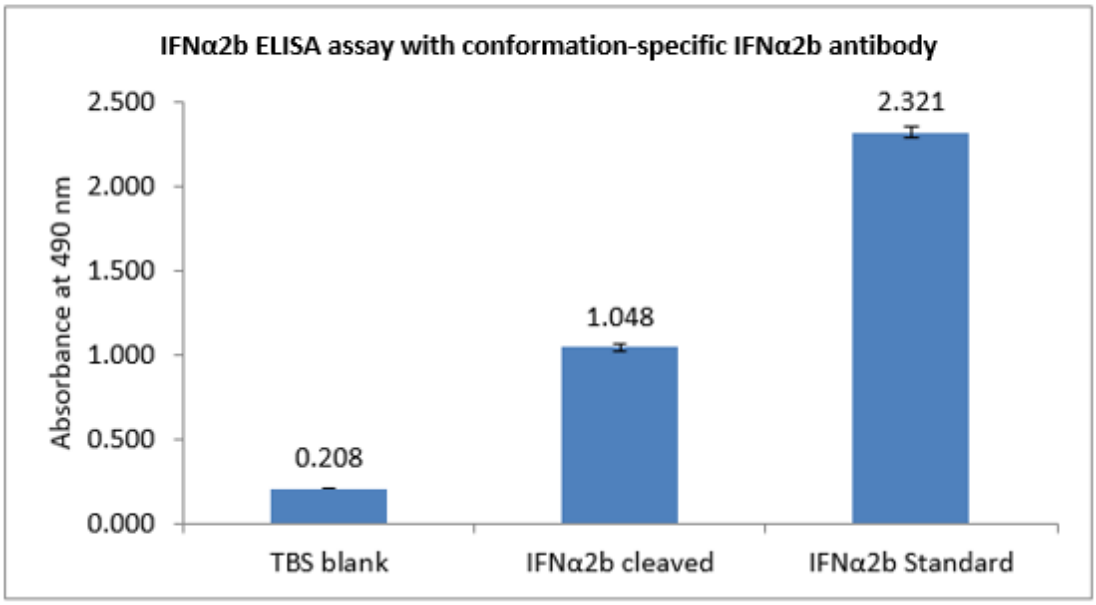



**Figure 3.6 ELISA assay result showing specific antibody recognition of (A) TNFα and (B) IFNα2b.** Protein amount used was 160 ng for cleaved proteins, standards as well as primary conformation-specific antibodies; secondary goat anti-rabbit IgG HRP-conjugate (Abcam ab6721) was used at 1:3000 dilution from commercial stock; TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7.8) was used as blank. All assays were performed in duplicate, and the error bars represent standard deviations. Published in Du & Rehm (2017a) (Appendix 7.2).

### 3.1.2 PHA beads as self-cleavable protein purification resins mediated via PhaC-intein-target protein fusion

The next objective involved testing the self-cleaving intein instead of sortase in the same protein production system. It was expected that in the PhaC-intein-target protein fusion strategy, by exploiting the pH inducible self-cleavage of the *Ssp* DnaB intein, a simpler cleavage of the target protein could be achieved. Namely, in contrast to the addition of $Ca^{2+}$ +/- triglycine required to trigger sortase self-cleavage, and the residual G-scars left on target proteins by sortase self-cleavage, there would be no additional requirement of chemical reagents for intein self-cleavage apart from a pH drop, nor any concern of residual scars left after intein self-cleavage.

#### 3.1.2.1 Engineering of PhaC-intein-target protein fusion enabled PHA bead production and facilitated purification of three model proteins

To assess whether beads displaying PhaC-intein-target protein fusions could be used for protein production and purification, three model proteins were first tested, including *Aequorea victoria* green fluorescent protein (GFP), *Mycobacterium tuberculosis* vaccine candidate Rv1626 and the synthetic immunoglobulin G (IgG) binding ZZ domain of protein A derived from *Staphylococcus aureus*. Corresponding plasmids are listed in Table 2.2 (as grouped for expression of PhaC-intein-target protein fusions) and were prepared according to methods section 2.4.8. The plasmid DNA sequences were confirmed (2.4.7) and their maps are shown in Figures 7.4-7.6, Appendix 7.1. Plasmid transformation, cell cultivation for production of beads displaying PhaC-intein-target protein fusions, isolation of beads displaying PhaC-intein-target protein fusions, cleavage

of target proteins from beads displaying PhaC-intein-target protein fusions, SDS-PAGE, densitometry analysis, and mass spectrometry were performed according to methods sections 2.4.6, 2.2.2.3, 2.5.2, 2.6.1, 2.6.2.1, 2.6.3.2 and 2.6.2.3, respectively.

PHA beads could be isolated from all strains (*E. coli* BL21 (DE3) containing pMCS69 as the helper plasmid B) producing the respective fusion proteins and, in each case, a dominant protein band corresponding to the PhaC-intein-target protein was detected by using SDS-PAGE (Figure 3.7). Each of the isolated beads displaying PhaC-intein-target protein fusions were subjected to a cleavage reaction according to section 2.6.1. In all cases, lowering the pH to 6 induced self-cleavage of intein and only the pure target protein became soluble, i.e. was released into the supernatant without any detectable contaminating proteins (Figure 3.7). Gel densitometry indicated that after 16 h incubation under cleavage conditions, about 7% of the bead immobilised target protein was converted into a soluble form. The purified target proteins amounted to around 100 µg per g of wet beads or slightly over 10 µg per g of wet cell biomass (Table 3.3).

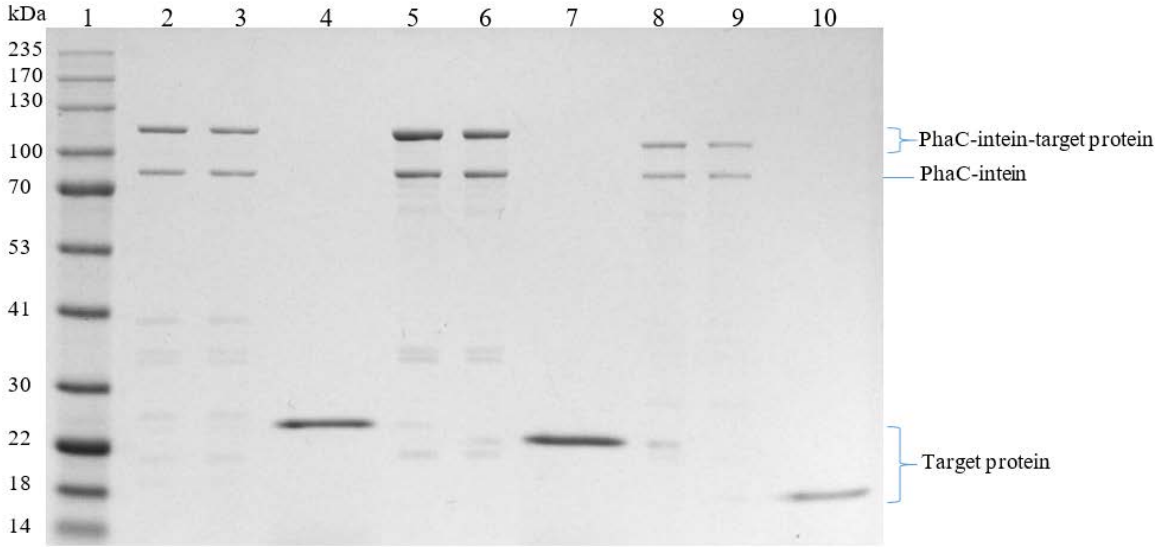



**Figure 3.7 Protein profiles of PHA beads isolated and post-cleavage, and the resulting soluble fractions.** 10% SDS-PAGE was performed to examine protein profiles. Lane 1 Molecular weight marker (GangNam-STAIN$^{TM}$ Prestained Protein Ladder); Lane 2 Isolated PhaC-Intein-GFP beads (pre-cleavage); Lane 3 PhaC-Intein-GFP beads post-cleavage; Lane 4 Cleaved GFP in the soluble fraction (~ 28 kDa); Lane 5 Isolated PhaC-Intein-Rv1626 beads (pre-cleavage); Lane 6 PhaC-Intein-Rv1626 beads post-cleavage; Lane 7 Cleaved Rv1626 in the soluble fraction (~ 24 kDa); Lane 8 Isolated PhaC-Intein-ZZ beads (pre-cleavage); Lane 9 PhaC-Intein-ZZ beads post-cleavage; Lane 10 Cleaved ZZ in the soluble fraction (~ 18 kDa). Published in Du & Rehm (2017b) (Appendix 7.2).

**Table 3.3** *Quantification of model proteins obtained from beads displaying PhaC-intein-target protein fusions**

| Target protein | µg/L | µg per gram wet beads | µg per gram wet cell mass | % Cleavage ratio[1] |
|---|---|---|---|---|
| GFP | 225.3 | 104.9 | 11.6 | 6.6 |
| Rv1626 | 278.1 | 136.7 | 14.5 | 6.5 |
| ZZ | 258.5 | 95.7 | 14.6 | 7.6 |

* Published in Du & Rehm (2017b) (Appendix 7.2).

[1] Estimated cleavage ratio, calculated as the % reduction in PhaC-intein-target protein band levels post-cleavage as compared to pre-cleavage, based on densitometry analysis of the SDS-PAGE image shown in Figure 3.7.

All of the bands corresponding to the three target proteins were separately excised and subjected to an in-gel trypsin digestion (section 2.6.2.3) followed by LC-MS/MS analysis, which confirmed their identity (Figure 7.22, Appendix 7.1).

In order to study whether the target protein was increasingly released over continuous incubation time with the low pH buffer, the cleavage of PhaC-Intein-GFP was examined over a 16 h time course. As expected, an increase in soluble GFP concentration and a shift of the PhaC-Intein-GFP band to the PhaC-Intein band was observed over time (Figure 3.8).

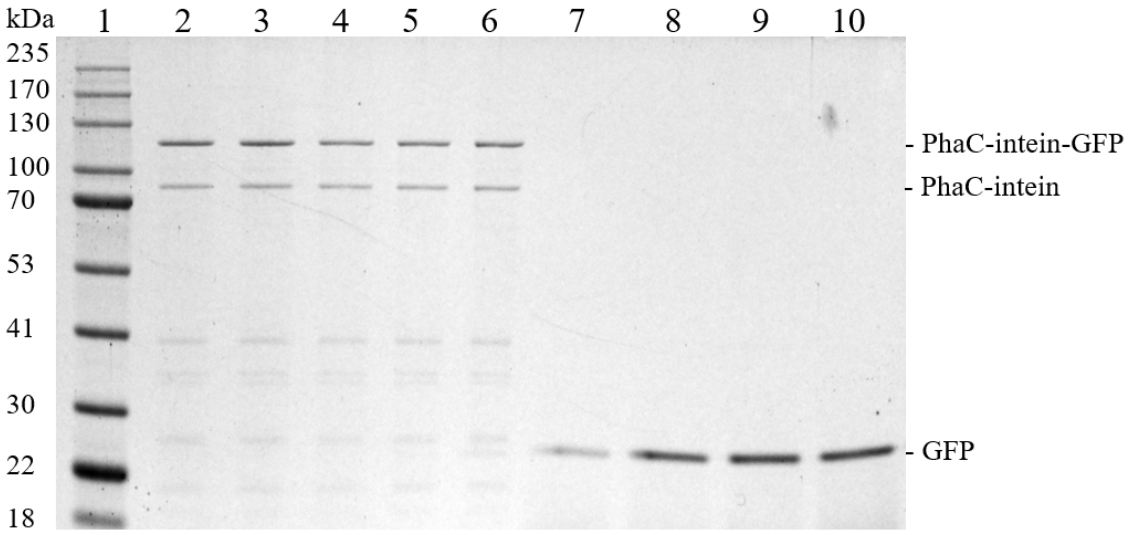



**Figure 3.8 Time course GFP cleavage off PhaC-Intein-GFP beads as revealed by 10% SDS-PAGE.** Lane 1, Molecular weight marker (GangNam-STAIN$^{TM}$ Prestained Protein Ladder); Lane 2, Isolated beads (pre-cleavage, 0 h); Lane 3, Beads after 16 h incubation at pH 8.6 as control; Lane 4-6, Beads after 4 h (4), 8 h (5) and 16 h (6) incubation at pH 6; Lane 7, Soluble fraction after 16 h incubation at pH 8.6 as control; Lane 8-10, GFP cleaved in the soluble fraction after 4 h (8), 8 h (9) and 16 h (10) incubation at pH 6. Published in Du & Rehm (2017b) (Appendix 7.2).

A densitometry analysis based on the SDS-PAGE gel image shown in Figure 3.8 was performed according to section 2.6.3.2, which provided a relative quantification of the protein band conversion from PhaC-intein-GFP into PhaC-intein plus GFP over the 16 h cleavage incubation at pH 6 as outlined in Table 3.4.

**Table 3.4** *Quantification of cleavage reaction for PhaC-intein-GFP beads*

| Protein band | Incubation time at pH 6 (h) | | | | |
|---|---|---|---|---|---|
| | 0 | 4 | 8 | 16 | 2nd 16* |
| % PhaC-intein-GFP band[1] | 68.7 | 62.7 | 62.3 | 62.0 | 54.9 |
| % PhaC-intein band[1] | 31.3 | 37.3 | 37.7 | 38.0 | 45.1 |
| % Cleavage ratio[2a] | - | 6 | 6.4 | 6.7 | 13.8# |
| Intensity of GFP band[3] | - | 1,546,293 | 1,651,727 | 1,737,535 | 1,817,167 |

[1] Estimated value, calculated as the % relative ratio between PhaC-intein-GFP and PhaC-intein band in the same lane.

[2] Estimated cleavage ratio, calculated as the % reduction in PhaC-intein-GFP band levels post-cleavage as compared to pre-cleavage starting material at 0 h.

[3] Estimated pixel counts.

* A second 16h cleavage cycle.

# Estimated total cleavage ratio after two 16h cleavage cycles.

As a significant amount of non-cleaved PhaC-intein-GFP still remained on the PHA beads even after 16 h of incubation as indicated in Lane 6 in Figure 3.8, a further attempt was made to see whether an additional round of incubation with the low pH buffer could lead to further cleavage of GFP. Specifically, remaining post-cleavage PhaC-intein-GFP bead samples after the 16 h cleavage reaction shown in Figure 3.8 were washed and re-suspended to a 40% slurry in fresh cleavage buffer and subjected to a second round of 16 h cleavage (section 2.6.1). Interestingly, GFP could again be cleaved off in a further round of 16 h cleavage reaction at a similar level to the first round of 16 h cleavage reaction (Figure 3.9). Densitometry quantification based on the SDS-PAGE image shown in Figure 3.9 was listed in the last column of Table 3.4.

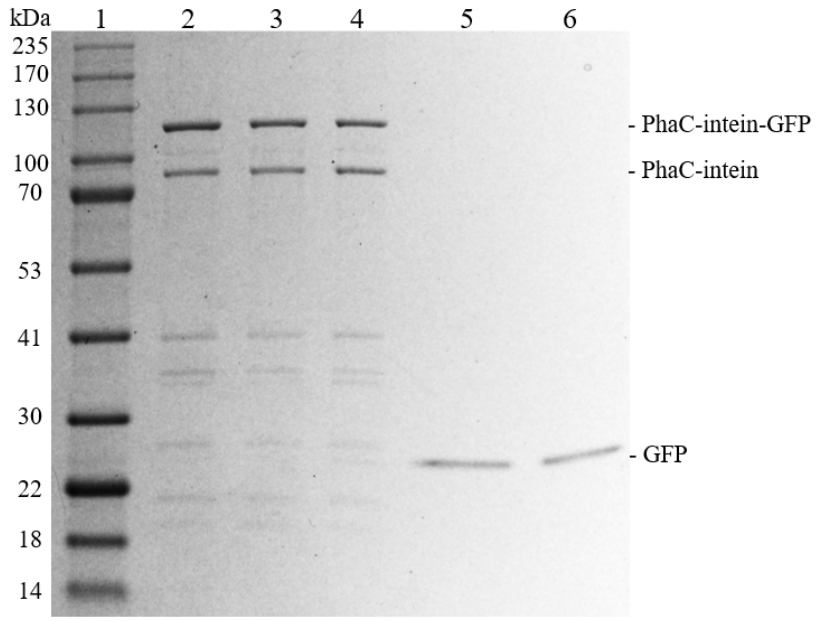



**Figure 3.9 GFP cleavage off PhaC-intein-GFP beads from two consecutive rounds of 16 h cleavage reaction as revealed by 10% SDS-PAGE.** Lane 1, Molecular weight marker (GangNam-STAIN$^{TM}$ Prestained Protein Ladder); Lane 2, Isolated PhaC-intein-GFP beads (pre-cleavage); Lane 3, Beads after the 1$^{st}$ cleavage reaction (16 h, pH 6); Lane 4, Beads after the 2$^{nd}$ cleavage reaction (2$^{nd}$ 16 h, pH 6); Lane 5, Cleaved GFP in the soluble fraction from the 1$^{st}$ cleavage reaction (16 h, pH 6); Lane 6, Cleaved GFP in the soluble fraction from the 2$^{nd}$ cleavage reaction (2 $^{nd}$ 16 h, pH 6). Published in Du & Rehm (2017b) (Appendix 7.2).

### 3.1.2.2 Functionality assessment of the model proteins cleaved from beads displaying PhaC-intein-target protein fusion

In order to assess the functionality of the GFP cleaved from PHA-Intein-GFP beads (section 3.1.2.1), fluorescence of the soluble fractions resulting from a 16 h cleavage reaction (corresponding to lanes 8-10 in Figure 3.8) was measured using a FLUOstar Omega (BMG labtech, Offenburg, Germany) microplate reader according to section 2.6.4.3. Results depicted in Figure 3.10 showed an increase of GFP fluorescence intensity

over the 16 h cleavage cycle (Figure 3.10). Fluorescence of GFP is indicative of functional folding (Ormö *et al.* 1996) throughout this production / purification process.

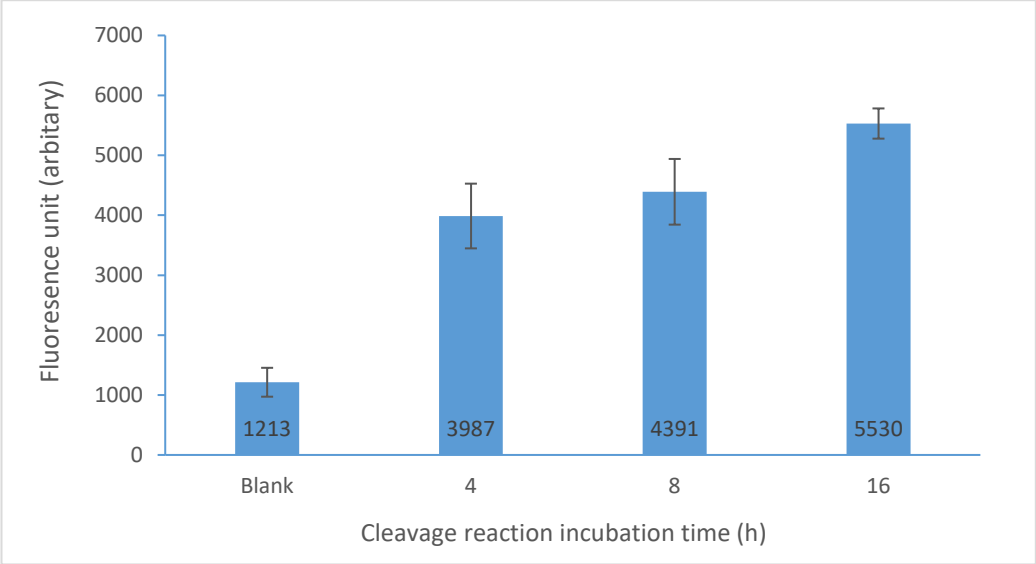



**Figure 3.10 GFP Fluorescence measurement of the soluble fractions resulting from a 16 h cleavage reaction.** All assays were performed in triplicate, and the error bars represent standard deviations. Published in Du & Rehm (2017b) (Appendix 7.2).

In the case of the RV1626 protein cleaved from PHA-Intein-Rv1626 beads (section 3.1.2.1), a western blotting assay (methods section 2.6.2.2) using a primary mouse polyclonal anti-Rv1626 antibody (Rubio Reyes *et al.* 2016) (1:10000) in combination with a secondary goat anti-mouse IgG HRP-conjugate (1:10000) (Abcam ab6789, UK), specifically revealed the cleaved Rv1626 protein, as well as the fusion protein PhaC-intein-Rv1626 both in the pre-cleavage bead sample and in the post-cleavage bead sample, respectively (Figure 3.11). PhaC-intein-GFP beads or cleaved GFP did not react with the antibody and served as negative controls (Figure 3.11).

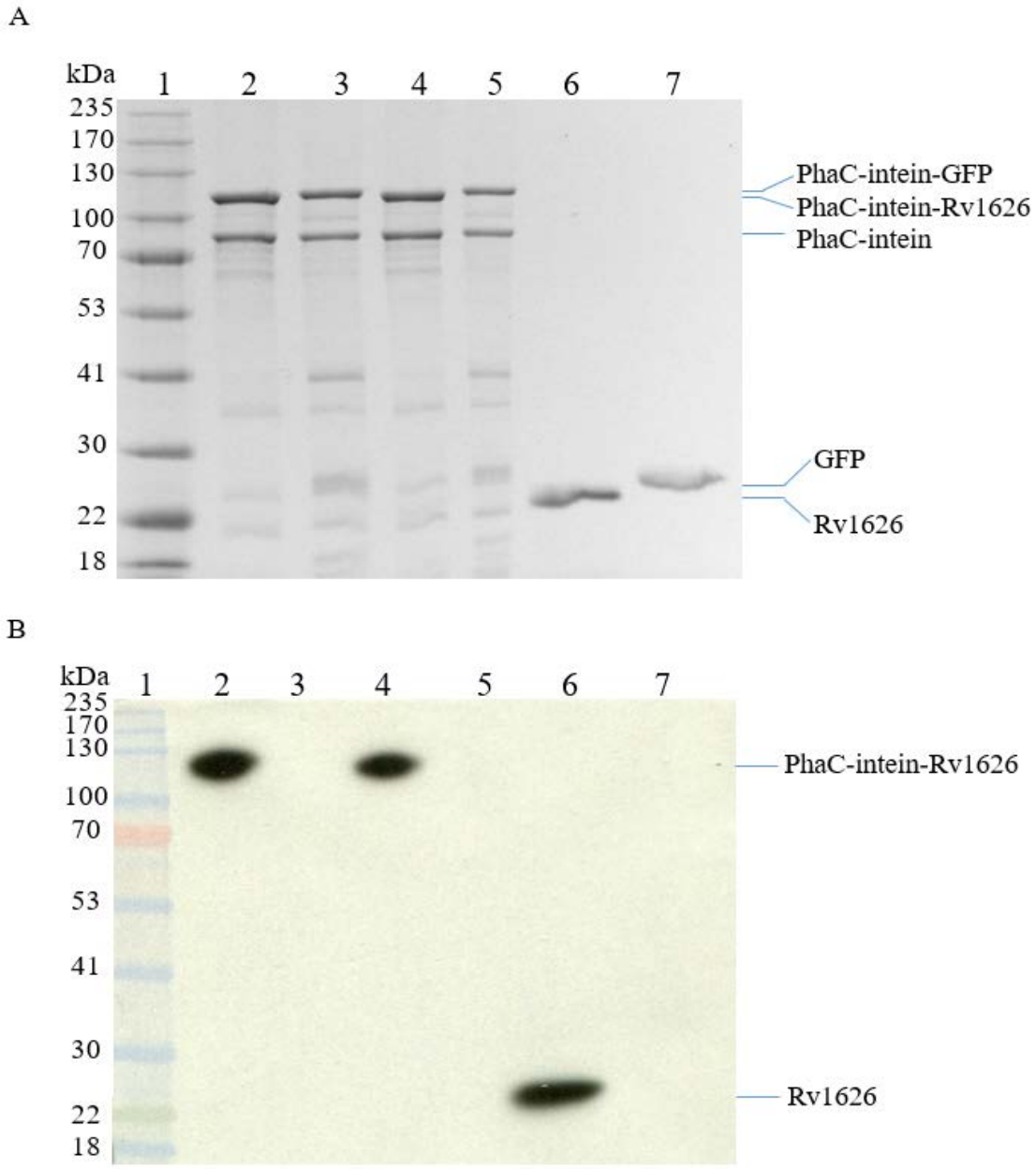



**Figure 3.11 Specific recognition of the Rv1626 antigen by a polyclonal anti-Rv1626 antibody from mice immunized with beads displaying Rv1626. A, SDS-PAGE; B, immunoblot.** Lane 1, Molecular weight marker (GangNam-STAIN$^{TM}$ Prestained Protein Ladder); Lane 2, Isolated PhaC-Intein-Rv1626 beads (pre-cleavage); Lane 3, Isolated PhaC-Intein-GFP beads (pre-cleavage) as negative control; Lane 4, PhaC-Intein-Rv1626 beads post-cleavage; Lane 5, PhaC-Intein-GFP beads post-cleavage as negative control; Lane 6, Rv1626 in soluble fraction (~ 24 kDa); Lane 7, GFP in soluble fraction as negative control (~ 28 kDa). Published in Du & Rehm (2017b) (Appendix 7.2).

Because the previous immunoblot was done with denatured Rv1626 antigen, in order to further investigate the specific binding between native Rv1626 and its antibody, as well as the IgG binding function of the synthetic ZZ domain of protein A derived from *S. aureus*, an ELISA assay was performed according to methods section 2.6.4.1 using the Rv1626 and ZZ proteins cleaved from respective beads (section 3.1.2.1). Briefly, ELISA plates were incubated overnight with 50 ng of Rv1626 or ZZ protein. Wild type PhaC beads or ZZ-PhaC beads (Brockelbank *et al.* 2006) were included as controls. Addition of the test material, the specific mouse polyclonal anti-Rv1626 antibody (anti-Rv1626 PAb) or a non-specific mouse polyclonal anti-DDA antibody (non-specific anti-DDA PAb) (Rubio Reyes *et al.* 2016), as well as the secondary goat anti-mouse IgG HRP-conjugate antibody (2nd goat anti-mouse IgG) is summarised in Table 3.5.

**Table 3.5** *ELISA assay to assess specific antibody binding of Rv1626 and IgG binding of ZZ domain*[*]

| Column | Test material | Non-specific anti-DDA PAb[#] | anti-Rv1626 PAb[#] | 2nd goat anti-mouse IgG[#] | Absorbance[1] value ± SD[2] |
|---|---|---|---|---|---|
| 1 | PBS Blank | | ✓ | ✓ | 0.050 ± 0.001 |
| 2 | Rv1626 cleaved | | ✓ | ✓ | 0.413 ± 0.007 |
| 3 | Rv1626 cleaved | ✓ | | ✓ | 0.054 ± 0.001 |
| 4 | Rv1626 cleaved | | | ✓ | 0.047 ± 0.002 |
| 5 | ZZ cleaved | | | ✓ | 0.767 ± 0.015 |
| 6 | ZZ-PhaC beads | | | ✓ | 0.802 ± 0.025 |
| 7 | WT PhaC beads | | | ✓ | 0.050 ± 0.003 |

[*] Published in Du & Rehm (2017b) (Appendix 7.2).
[#] Addition of these reagents is indicated with a tick.
1 Absorbance at 490 nm, which is proportional to the amount of IgG bound to the well.
2 All assays were performed in triplicate.

The ELISA assay result as summarised in Table 3.5 confirmed the specific binding between native Rv1626 and its antibody (Table 3.5, column 2), while in the absence of

the primary anti-Rv1626 antibody, only background absorbance similar to PBS was measured (Table 3.5). Both results suggested Rv1626 was produced and purified in a functional form.

In contrast to the background absorbance similar to PBS as measured for wild type PhaC beads, the purified ZZ domain was able to bind IgG as shown in the ELISA analysis, at a similar level as compared with the positive bead control, ZZ-PhaC beads, containing the same amount of ZZ protein (Table 3.5, column 5 & 6), which indicated the successful retention of its IgG binding capacity when produced and purified based on the PhaC-intein-target protein fusion approach.

### 3.1.2.3 Validation of the PhaC-intein-target protein fusion strategy to produce therapeutic proteins

To further assess the applicability of PhaC-intein-target protein fusion strategy to produce and purify high-value therapeutic proteins, three first-line anti-cancer therapeutic proteins were targeted including human tumour necrosis factor alpha (TNFα), human interferon alpha 2b (IFNα2b) and human granulocyte colony-stimulating factor (G-CSF). G-CSF production in *E. coli* tends to aggregate and forms inclusion bodies (IBs) (Do *et al.* 2014), like TNFα and IFNα2b mentioned in section 3.1.1.2. It was thus also an additional aim to avoid IB formation of these proteins by *in vivo* immobilising them on PHA beads, as it is widely accepted that immobilisation improves protein stability / solubility (Rehm *et al.* 2016; Rehm *et al.* 2017).

Corresponding plasmids are listed in Table 2.2 as grouped for expression of PhaC-intein-target protein fusions) and were prepared according to methods section 2.4.8. The plasmid

DNA sequences were confirmed (2.4.7) and their maps are shown in Figures 7.7-7.9, Appendix 7.1. Plasmid transformation, cell cultivation for production of beads displaying PhaC-intein-target protein fusions, isolation of beads displaying PhaC-intein-target protein fusions, cleavage of target proteins from beads displaying PhaC-intein-target protein fusions, SDS-PAGE, densitometry analysis, and mass spectrometry were performed according to methods sections 2.4.6, 2.2.2.3, 2.5.2, 2.6.1, 2.6.2.1, 2.6.3.2 and 2.6.2.3, respectively.

Here, *E. coli* SHuffle® T7 express (Table 2.1) was chosen as the production strain, and pMCS69E was used as helper plasmid B, as TNFα contains one disulphide bond while IFNα2b and G-CSF both contain two required for stability / functionality. As mentioned earlier under section 3.1.1.2, SHuffle® T7 express strain has an oxidizing cytosol due to *trxB/gor* mutations and also constitutively produces a cytosolic form of disulphide isomerase (DsbC) that acts as a chaperone. Furthermore pMCS69E encodes sulfhydryl oxidase (Erv1p) helpful for disulphide bond formation, in addition to containing the *phaA* and *phaB* genes (Table 2.2).

PHA beads could be isolated from the respective recombinant bacteria and the PhaC-intein-target protein could be detected as the dominant protein on the beads (Figure 3.12). Levels of premature cleavage (PhaC-intein) were similar to that of the GFP displaying PHA beads. Gel densitometry indicated that activation of the beads with a pH shift to 6 released about 10 to 20 µg therapeutic proteins per g of wet beads or about 1 to 1.6 µg per g of wet cell biomass (Table 3.6).

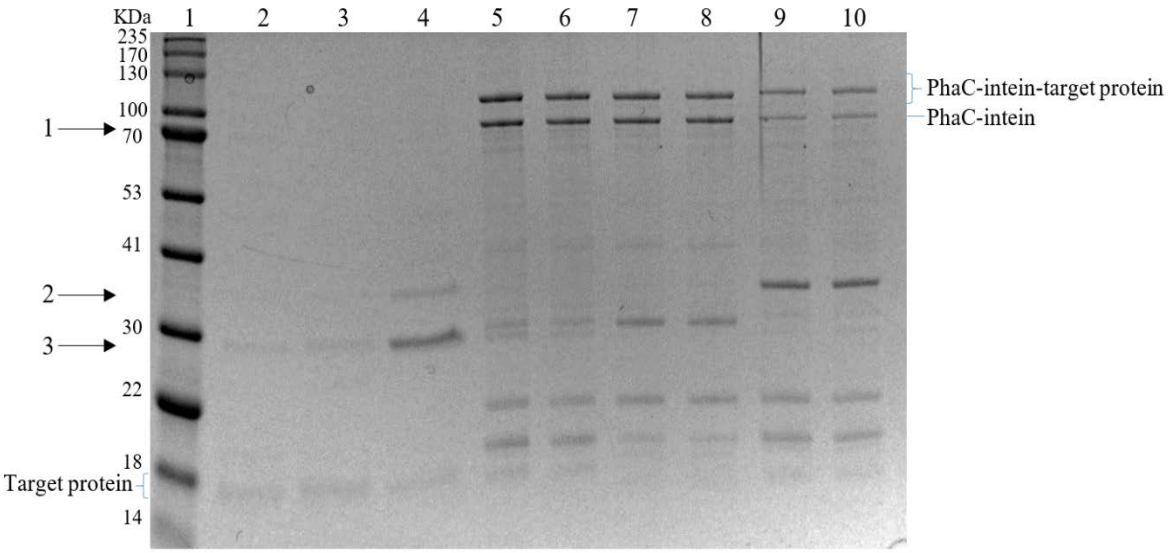



**Figure 3.12 Protein profiles of the resulting soluble fractions upon a 16 h cleavage incubation as well as the PHA beads isolated and post-cleavage.** 15% SDS-PAGE was performed to examine protein profiles. Lane 1, Molecular weight marker (GangNam-STAIN$^{TM}$ Prestained Protein Ladder); Lane 2, Cleaved TNFα in the soluble fraction, corresponding to the lowest bottom band (~ 18.6 kDa); Lane 3, Cleaved G-CSF in the soluble fraction, corresponding to the lowest bottom band (~ 19.9 kDa); Lane 4, Cleaved IFNα2b in the soluble fraction, corresponding to the lowest bottom band (~ 20.5 kDa); Lane 5, Isolated PhaC-Intein-TNFα beads (pre-cleavage); Lane 6, PhaC-Intein-TNFα beads post-cleavage; Lane 7; Isolated PhaC-Intein-G-CSF beads (pre-cleavage); Lane 8, PhaC-Intein-G-CSF beads post-cleavage; Lane 9, Isolated PhaC-Intein-IFNα2b beads (pre-cleavage); Lane 10, PhaC-Intein-IFNα2b beads post-cleavage. *Note that in lanes 2-4 there are two to three co-purified carry-over proteins which are respectively numbered and indicated with arrows on the left side. Published in Du & Rehm (2017b) (Appendix 7.2).

**Table 3.6** *Quantification of therapeutic proteins obtained from beads displaying PhaC-intein-target protein fusions*

| Target protein | µg/L | µg per gram wet beads | µg per gram wet cell mass | % Cleavage ratio[1] |
|---|---|---|---|---|
| TNFα | 20.6 | 12.2 | 1.5 | 2.7 |
| G-CSF | 23.5 | 10.8 | 1.6 | 2.9 |
| IFNα2b | 16.2 | 18.2 | 1.0 | 1.1 |

[*] Published in Du & Rehm (2017b) (Appendix 7.2).

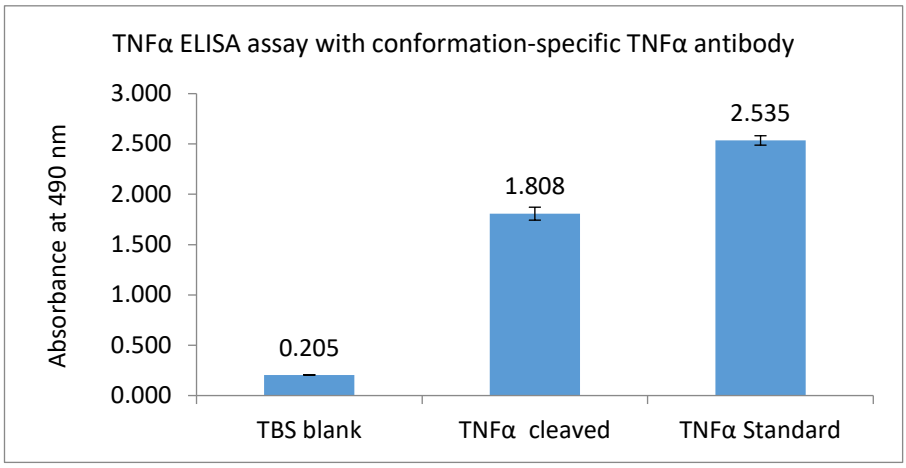
[1] Estimated cleavage ratio, calculated as the % reduction in PhaC-intein-target protein band levels post-cleavage as compared to pre-cleavage, based on densitometry analysis of the SDS-PAGE image shown in Figure 3.12.

TNFα, G-CSF and IFNα2b were visable in the soluble fractions that resulted from cleavage reactions of the respective isolated PHA beads (Figure 3.12). Minor impurities could likely be avoided by optimisation of bead isolation / washing steps or after cleavage by applying further purification steps, as would be required to achieve biopharmaceutical grade purity.

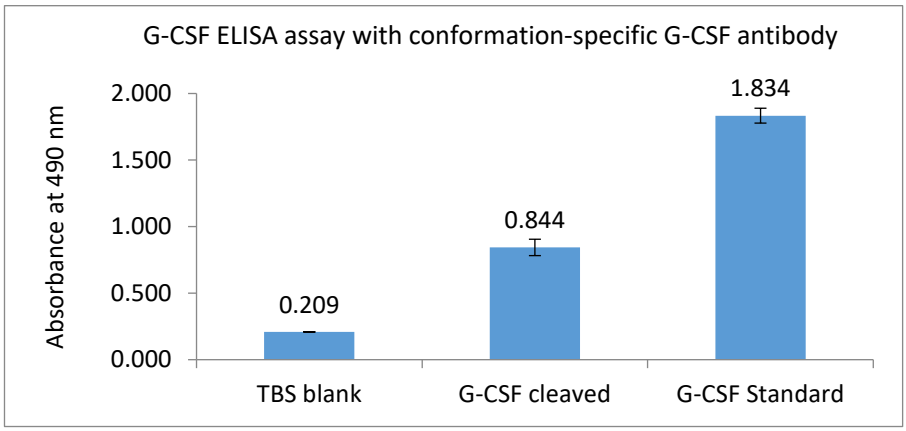
The bands corresponding to the therapeutic proteins were separately excised and subjected to an in-gel trypsin digestion (section 2.6.2.3) followed by LC-MS/MS analysis, which confirmed their identity (Figure 7.23, Appendix 7.1). The co-purifying host cell proteins were similarly identified and the results suggested them to be the *E. coli* chaperone protein DnaK, and both full length and truncated outer membrane protein A, respectively (Figure 7.24, Appendix 7.1).

In order to assess the proper folding of these therapeutic proteins cleaved from respective PHA beads, they were analysed by ELISA assays according to section 2.6.4.1 with corresponding conformation-specific antibodies (Sino Biological Inc.). The ELISA results as shown in Figure 3.13 A, B & C verified their specific binding with respective conformation-specific antibodies, and suggested that they were properly folded.

A




B




C

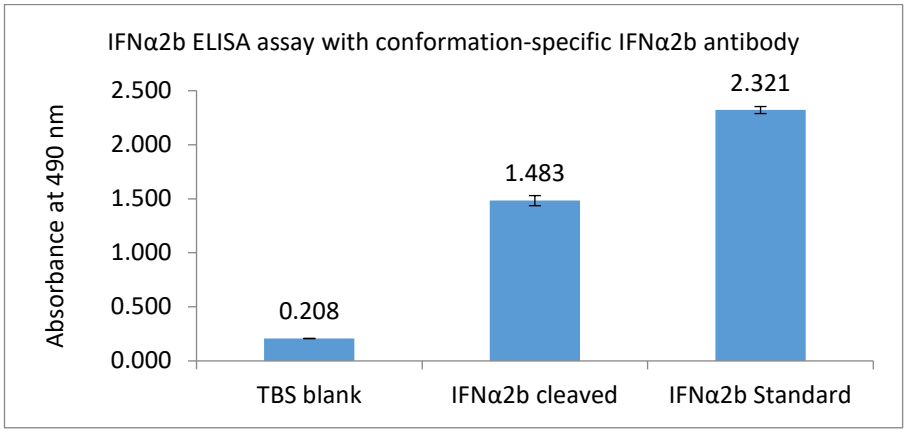



**Figure 3.13 ELISA assay results with respective conformation-specific antibody for (A) TNFα, (B) G-CSF and (C) IFNα2b showing the specific antibody recognition.** Protein amount used was 160 ng for cleaved proteins and standards as well as primary conformation-specific antibodies; secondary goat anti-rabbit IgG HRP-conjugate (Abcam ab6721) was used at 1:3000 dilution from commercial stock; TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7.8) was used as blank. All assays were performed in duplicate, and the error bars represent standard deviations. Published in Du & Rehm (2017b) (Appendix 7.2).

A brief comparison of this PhaC based protein purification method versus previous methods based on PhaP / PhaR tag is summarized in Table 3.7.

**Table 3.7** *Comparison of the PhaC based protein purification method with previous methods based on PhaP / PhaR tag*[*]

| PHA bead tag | PhaC-intein-target<br>This thesis | PhaP-intein-target<br>(Banki *et al.* 2005) | PhaP-intein-target<br>(Zhou *et al.* 2011) | PhaP-intein-target<br>(Barnard *et al.* 2005) | PhaP-intein-target<br>(Wang *et al.* 2008) | PhaR-intein-target<br>(Zhang *et al.* 2010) |
|---|---|---|---|---|---|---|
| Association with beads | Covalently attached *in vivo* in *E. coli* | Non-covalently attached *in vivo* in *E. coli* | Non-covalently attached *in vivo* in *E. coli* | Non-covalently attached *in vivo* in *R. eutropha* | *In vitro* non-covalent association | *In vitro* non-covalent association |
| Intein / activate condition | *Ssp* DnaB intein / low pH | ΔI-CM engineered from *Mtu* recA intein / low pH | ΔI-CM engineered from *Mtu* recA intein / low pH | *Mxe* GyrA intein/ thiols | *Ssp* DnaB intein / low pH | *Ssp* DnaB intein / low pH |
| Efforts on extra protein / polymer production | No concerns | Extra triple or dual PhaP protein production | Extra triple PhaP protein production | Extra PhaP production | Extra separate chemical synthesis of PHB polymer | Extra separate chemical synthesis of PHB polymer |
| Tailoring efforts on washing / elution conditions to prevent PhaP / PhaR leaching | No concerns | Tailoring efforts needed for an intermediate salt content of 50-150 mM in washing / elution | No discussion but necessary to prevent PhaP leaching while only allow target elution | No discussion but necessary to prevent PhaP leaching while only allow target elution | No discussion but necessary to prevent PhaP leaching while only allow target elution | No discussion but necessary to prevent PhaR leaching while only allow target elution |
| Purification for a potentially bead associated target | No concerns (except for PhaC itself) | With limitations, e.g. β-lactamase unsuitable | No discussion but with limitations in order to maintain non-covalent PhaP : bead binding | No discussion but with limitations in order to maintain non-covalent PhaP : bead binding | With limitations, e.g. PhaC unsuitable | No discussion but with limitations in order to maintain non-covalent PhaR : bead binding |
| Potentially multiple short cleavage cycles | Exemplified with PhaC-intein-GFP beads | Unexplored | Unexplored | Unexplored | Unexplored | Unexplored |
| Purification for therapeutic proteins containing S-S bond | Exemplified with human TNFα, IFNα2b and G-CSF | Unexplored | Exemplified with porcine IFNα | Unsuitable due to use of thiols | Unexplored | Unexplored |

[*] Adapted from Du & Rehm (2017b) (Appendix 7.2).

## 3.2 Design of PHA beads as affinity resins for molecular recognition by immobilising OBody ligands on bead surface

Affinity resins based on the biorecognition (specific noncovalent interaction (McCammon 1998)) between a biomolecule and its ligand (or binding partner) are so far the most efficient techniques, and are widely used to separate or detect biologically important molecules from natural sources.

However, manufacture of a typical affinity resin (with an antibody or antibody fragment immobilised on a support matrix) (Tozzi *et al.* 2003; Crivianu-Gaita & Thompson 2016) generally requires the expensive recombinant production of the antibody (or antibody fragment) (Dias & Roque 2017), the separate preparation of the support matrix (Vařilová *et al.* 2006), and complicated chemical cross-linking of the antibody (or antibody fragment) to the support matrix (Shen *et al.* 2017).

Therefore, developing alternative non-antibody affinity ligands or resins that enable simplified ligand immobilisation are necessary, such as in the cases of commercial lysozyme separation from hen egg white in the food and pharmaceutical industry (Shahmohammadi 2017), and progesterone (P4) detection in the dairy industry (Jang *et al.* 2017).

OBody is one of the emerging engineered non-antibody protein scaffolds serving as alternative elements for molecular recognition (Skerra 2007; Zhao *et al.* 2013; Škrlec *et al.* 2015; Dias & Roque 2017). OBody is engineered from OB-fold domain (see section 1.2.2 for detail) of aspartyl-tRNA synthetase (aspRS) from *Pyrobactulum aerophilum* by using phage display technology (Steemson 2011; Steemson *et al.* 2014). For example, an

OBody with a 3 nM affinity for hen egg-white lysozyme has been demonstrated (Steemson *et al.* 2014). It would be of great interest to immobilise such engineered biomolecular recognising OBodies on a proper support matrix for the purpose of affinity separation or detection.

Polyhydroxyalkanoate (PHA) beads could potentially serve as support matrices for OBody proteins. As mentioned in section 3.1, PHA beads have been produced in recombinant *E. coli* cells harbouring two plasmids encoding three key enzymes (PhaC the PHA synthase, either as a wild type protein or as a fusion with a protein of interest, encoded by plasmid A, as well as PhaA the *β*-ketothiolase and PhaB the acetoacetyl-CoA reductase encoded by helper plasmid B) (Jahns & Rehm 2009). Also, PHA beads have been previously functionalised to display various affinity domains and used as affinity resins, including a ZZ domain of protein A derived from *S. aureus* for IgG binding / separation (Brockelbank *et al.* 2006; Lewis & Rehm 2009), a streptavidin for biotin binding (Peters & Rehm 2008), and a single-chain variable fragment antibody (scFv) for *β*-galactosidase antigen binding (Grage & Rehm 2008).

Hence, it was envisaged that by translationally fusing an engineered OBody to PhaC, PHA beads with immobilised OBody could be produced in a single step, and the resulting "OBody beads" upon isolation as a whole would be useful as affinity resins for binding of an external target biomolecule recognised by said OBody ligand. This potential application of PHA bead fusions is described in this section, in contrast to the previous section 3.1 where a target protein was internally produced as immobilised on beads and subsequently cleaved in a free soluble form.

Two types of previously engineered OBody ligands based on the same OB-fold domain from *P. aerophilum* aspRS were examined, one recognising hen egg-white lysozyme as disclosed in a PhD thesis (Steemson 2011), the other recognising progesterone (P4) as generously shared by Vickery Arcus and the company "OBodies limited" (http://www.obodies.com/). These two types of engineered OBody ligands were chosen because both hen egg-white lysozyme separation and bovine milk P4 detection processes are of industrial significance as mentioned above.

In order to develop a new type of affinity resin enabling efficient and streamlined lysozyme separation, the third objective of this PhD study was to immobilise on PHA beads the lysozyme recognising OBody ligand L200EP-06 (Steemson 2011) (L200EP-06 was a clone code given for this particular OBody ligand during library screening (Steemson 2011) and abbreviated as O6 hereinafter), and to test the performance of these beads for external lysozyme separation. It was expected that either a PhaC-O6 or O6-PhaC fusion could mediate formation of PHA beads, and the resulting PhaC-O6 or O6-PhaC beads upon isolation could retain the O6 affinity for lysozyme and thus could be used as affinity resins for lysozyme separation.

In addition, in order to develop a new type of economic bovine milk progesterone (P4) detection method involving non-antibody ligands bound to a matrix, the fourth objective of this PhD study was to immobilise on PHA beads the P4 recognising OBody ligand P4013-D7 (abbreviated as D7 hereinafter), and test the performance of these beads for P4 detection. It was expected that different fusions between PhaC and D7 could mediate formation of PHA beads, and the resulting PHA beads upon isolation could retain the D7 affinity for progesterone and thus be used as affinity resins for P4 detection. It should be noted that, when later a 2$^{nd}$ generation of P4 recognising OBody ligand B7 became

available, it was similarly immobilised on PHA beads by fusing to the N-terminus of PhaC (based on results available for the 1st generation D7), and respective beads upon isolation were assessed for P4 detection as well.

## 3.2.1 Design of PHA beads as affinity resins for lysozyme separation

### 3.2.1.1 Production of lysozyme binding OBody beads

To test whether a lysozyme-binding OBody ligand could be functionally immobilised on PHA beads, OBody L200EP-06 (O6) was examined. The O6 ligand is a synthetic peptide that was engineered based on the OB-fold domain of aspartyl-tRNA synthetase (aspRS) from *P. aerophilum* and binds specifically to hen white egg lysozyme with an affinity of 612.8 nM (Steemson 2011). It is 113 amino acids in length (with amino acid sequence shown in Figure 7.10, Appendix 7.1), and contains 4 amino acid insertions plus 19 substitutions as compared to the wild type OB-fold (residues 1-109, GenBank ID NP_558783.1) (see sequence alignment in Figure 7.25, Appendix 7.1).

The O6 peptide was fused either N- or C- terminally to PhaC, the PHA synthase. Corresponding plasmids are listed in Table 2.2 (as grouped for expression of fusion proteins between PhaC and lysozyme binding OBody) and were prepared according to methods section 2.4.8. The plasmid DNA sequences were confirmed (2.4.7) and their maps are shown in Figures 7.11-7.12, Appendix 7.1. Plasmid transformation, cell cultivation for standard bead production, bead isolation, SDS-PAGE, Bradford assay and densitometry analysis were performed according to methods sections 2.4.6, 2.2.2.3, 2.5.2, 2.6.2.1, 2.6.3.1 and 2.6.3.2, respectively.

It was established previously that wild type PHA beads (PhaC beads) could be produced using BL21 (DE3) *E. coli* strain (Table 2.1) harbouring plasmids A and B (namely plasmid pETC encoding PhaC the PHA synthase and plasmid pMCS69 encoding PhaA and PhaB) (Table 2.2) (Peters *et al.* 2007). Similarly, in the current work, PHA beads were produced by respective strains (*E. coli* BL21 (DE3) containing either pET14b-O6-PhaC or pET14b-O6-PhaC as plasmid A and pMCS69 as the helper plasmid B), as shown by the dominant fusion protein bands in Figure 3.14.

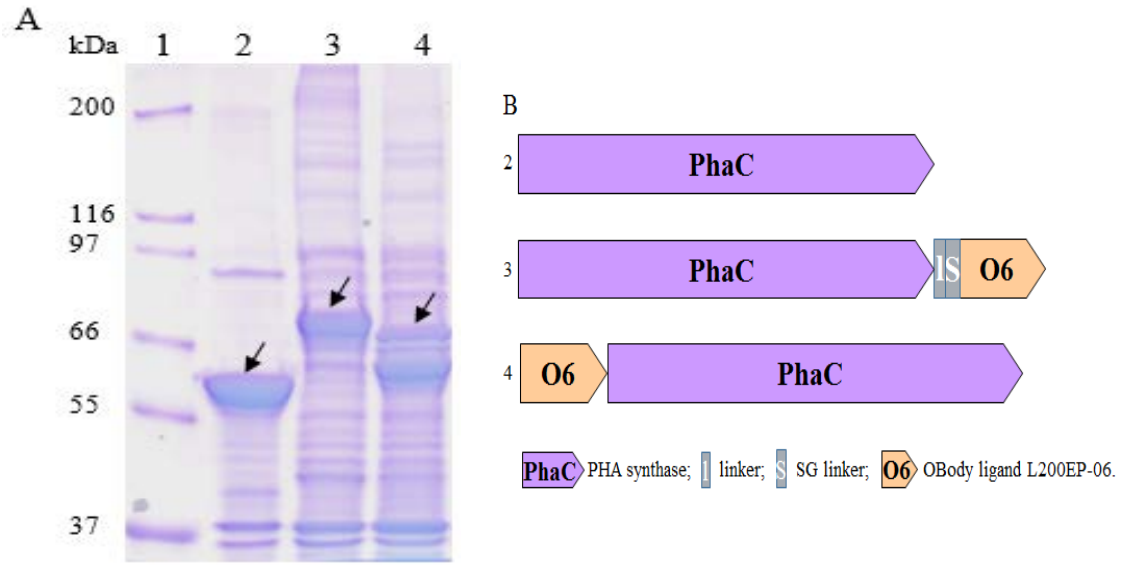



**Figure 3.14 Protein profiles of isolated PHA beads and schematic representation of relevant protein components.** (A) 10% SDS-PAGE was performed to examine whether PHA beads were produced with the correct protein profiles. Lane 1, Molecular weight marker (Mark12$^{TM}$); Lane 2, WT PhaC beads (~ 64 kDa); Lane 3, PhaC-O6 beads (~ 80 kDa); Lane 4, O6-PhaC beads (~ 77 kDa). Arrows indicate the PhaC protein or respective fusion proteins. (B) Schematic representation of relevant protein components. Number in front of each linear diagram corresponds to lane number in A.

As expected, fusion protein PhaC-O6 was slightly bigger than O6-PhaC, which was reflected by the size difference between corresponding protein bands on the SDS-PAGE gel (Figure 3.14, Lanes 3 & 4). This was caused by the two different starting plasmids

used for corresponding cloning work (pET14b-phaC-linker-SG linker-GFP containing an extra "*linker-SG linker*" region at 3' end of *phaC* gene, and plasmid pET14b-GFP-phaC without that region, section 2.4.8). Briefly, during construct preparation, the respective *gfp* gene in the starting plasmids was simply swapped with the corresponding O6 coding region; therefore the extra "*linker-SG linker*" region was kept in the final pET14b-PhaC-O6 construct but absent in the final pET14b-O6-PhaC construct, contributing to the observed size difference between the final fusion protein products. The "*linker-SG linker*" region was designed in a previous study (Jahns & Rehm 2009) on PhaC tolerance for C-terminal fusions. Upon translation, the "linker" component could allow flexible folding of a peptide or protein fused to the C-terminus of PhaC, while the "SG linker" component (a triplicate SGGGG peptide) could maintain the hydrophobic environment at C-terminus of PhaC which is essential for its synthase functionality (Jahns & Rehm 2009)

#### 3.2.1.2 Lysozyme binding function assessment

To evaluate whether these OBody beads could be used for lysozyme separation, a lysozyme binding assay was carried out with corresponding OBody beads (as well as WT PhaC beads as negative controls) and using a mixture solution of skim milk, BSA and lysozyme as the complex substrate, according to Methods section 2.6.4.4.

SDS-PAGE analysis results for the elution fractions as shown in Figure 3.15 indicated that both PhaC-O6 and O6-PhaC beads could be used to purify lysozyme from the mixture solution, whereas the negative control WT PhaC beads did not. These PhaC-O6 and O6-PhaC beads showed low levels of nonspecific binding (Figure 3.15). Based on densitometry, the purity of the eluted products was higher and the amount of lysozyme eluted was higher for the O6-PhaC beads than for the PhaC-O6 beads (about 95% vs. 72%

in terms of purity, and 465 vs. 360 nmol of lysozyme per g of beads in terms of lysozyme binding capacity), whereas the binding of the negative control WT PhaC beads was negligible (Figure 3.15 and Table 3.8).

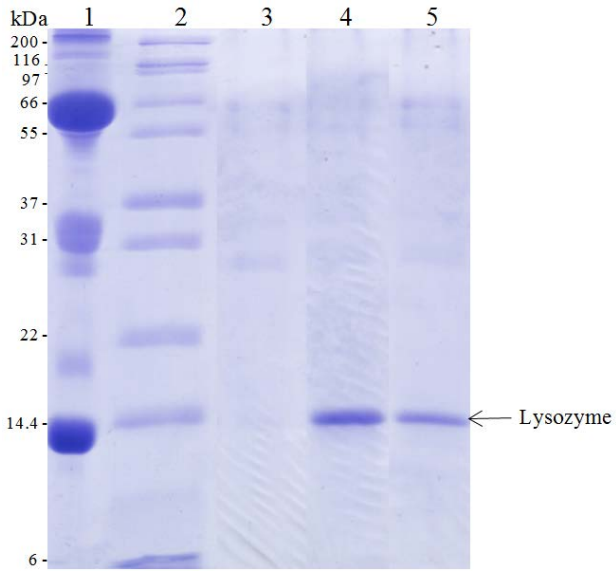



**Figure 3.15 Use of OBody beads for lysozyme separation from a solution of skim milk, BSA and lysozyme.** 15% SDS-PAGE was performed for resolving proteins in respective samples. Lane 1, mixture solution of skim milk, BSA and lysozyme; Lane 2, Molecular weight marker (Mark12$^{TM}$); Lane 3, elution fraction from WT PhaC beads; Lane 4, elution fraction from O6-PhaC beads; Lane 4, elution fraction from PhaC-O6 beads. Published in Hay *et al.* (2015b), Appendix 7.2.

**Table 3.8** *Lysozyme binding capacity and purification power of the OBody beads*[*]

| Bead prototype | Amount of target protein eluted (mg/g beads ± SD)[1] | Amount of target protein (nmol/g beads) | % purity (based on SDS-PAGE)[2] |
|---|---|---|---|
| WT PhaC | 0.32 ± 0.11 | 22 ± 8 | NA |
| O6-PhaC | 6.65 ± 0.17[a,b] | 465 ± 12[a,b] | 95 |
| PhaC-O6 | 5.13 ± 0.20[a] | 358 ± 14[a] | 72 |

[*] Published in Hay *et al.* (2015b), Appendix 7.2.

[1] Value obtained from Bradford assay (2.6.3.1), n = 3;

[2] Estimated purity, calculated as the % of the lysozyme band content in the elution fraction resulted from lysozyme binding assay (2.6.4.4) based on densitometry analysis of the SDS-PAGE image as shown in Figure 3.15.

[a] p = 0.00001 vs. WT PhaC control beads.

[b] p = 0.00008 vs. PhaC-O6 beads.

### 3.2.2 Design of PHA beads as affinity resins for progesterone detection

Given the encouraging lysozyme recognition power demonstrated by PHA beads displaying the OBody ligand L200EP-06 (O6) (section 3.2.1.2), attempts were made to immobilise a progesterone-recognising OBody ligand on PHA beads for the purpose of progesterone detection.

#### 3.2.2.1 Production of progesterone binding OBody beads

In this section, a different OBody ligand recognising a different molecule, progesterone (P4), would be similarly immobilised on PHA beads and tested for P4 detection usage. To test whether a P4 binding OBody ligand could be functionally immobilised on PHA beads, OBody P4013-D7 (D7) was tested. The D7 ligand is also a synthetic peptide that was engineered based on the OB-fold domain of aspartyl-tRNA synthetase (aspRS) from *P. aerophilum* and binds specifically to P4 with an affinity of about 300 - 400 nM (information generously shared by Vickery Arcus and the company "OBodies limited" (http://www.obodies.com/), and P4013-D7 was a clone code given by them for this particular OBody ligand during library screening). It is 106 amino acids in length (with amino acid sequence shown in Figure 7.13, Appendix 7.1), and contains 3 amino acid deletions plus 23 substitutions as compared to the wild type OB-fold (residues 1-109, GenBank ID NP_558783.1) (see sequence alignment in Figure 7.26, Appendix 7.1).

This D7 peptide was fused N- and / or C- terminally to PhaC (PHA synthase). Corresponding plasmids (pET14b-PhaC-D7, pET14b-3xD7-PhaC and pET14b-3xD7-PhaC-D7) are listed in Table 2.2 (as grouped for expression of fusion proteins between

PhaC and progesterone binding OBody) and prepared according to methods section 2.4.8. As noted in methods section 2.4.8, during preparation of the N-terminal fusion construct by using a single enzyme insertion site (*Spe*I), a triplicate N-terminal insertion (pET14b-3xD7-PhaC) was obtained instead. This construct pET14b-3xD7-PhaC was mistaken as a normal single insertion due to a false positive sequencing result (covering restriction site and further 200 bp or so from each direction), accordingly further cloning work was continued based on it, resulting in the plasmid pET14b-3xD7-PhaC-D7 as the dual terminal fusions. The correct plasmid DNA sequences were reconfirmed later (2.4.7) and their maps are shown in Figures 7.14-7.16, Appendix 7.1. Plasmid transformation, cell cultivation for standard bead production, bead isolation, SDS-PAGE and densitometry analysis were performed according to methods sections 2.4.6, 2.2.2.3, 2.5.2, 2.6.2.1 and 2.6.3.2, respectively.

As mentioned in sections 3.2.1.1, it was established previously that wild type PHA beads (PhaC beads) could be produced using BL21 (DE3) *E. coli* strain (Table 2.1) harbouring plasmids A and B (namely plasmid pETC encoding PhaC the PHA synthase and plasmid pMCS69 encoding PhaA and PhaB) (Table 2.2) (Peters *et al.* 2007). Similarly, in the current work, PHA beads were produced by respective strains (*E. coli* BL21 (DE3) containing either pET14b-PhaC-D7, pET14b-3xD7-PhaC or pET14b-3xD7-PhaC-D7 as plasmid A and pMCS69 as the helper plasmid B), as shown by the dominant fusion protein bands in Figure 3.16.

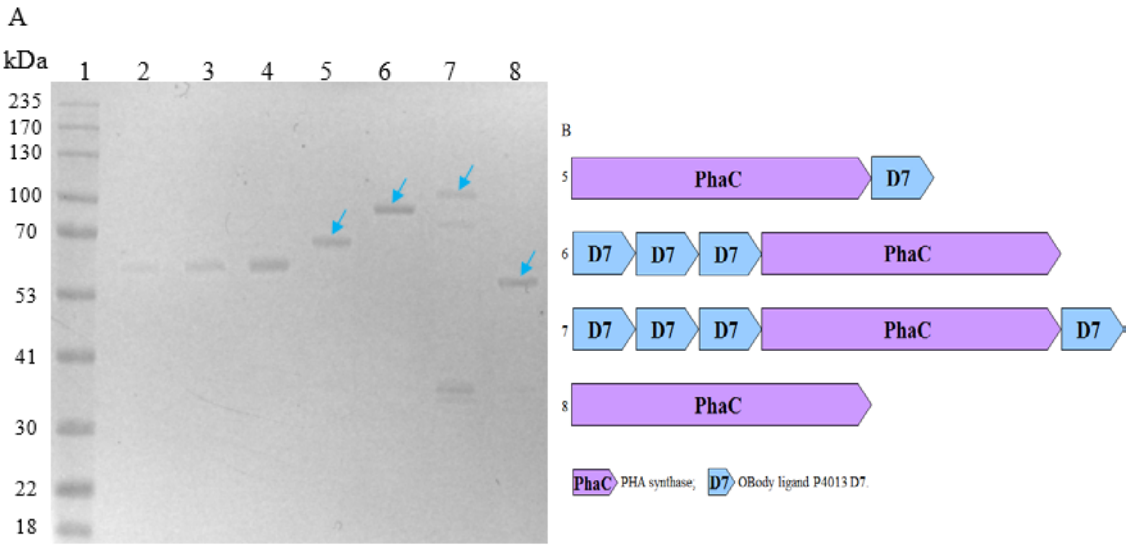



**Figure 3.16 Protein profiles of isolated PHA beads and schematic representation of relevant protein components.** (A) 10% SDS-PAGE was performed to examine whether beads were produced with the correct protein profiles. Lane 1, Molecular weight marker (GangNam-STAIN$^{TM}$ Prestained Protein Ladder); Lanes 2-4, 100, 200 and 400 ng BSA; Lane 5, PhaC-D7 beads (~ 78 kDa); Lane 6, 3xD7-PhaC beads (~ 106 kDa); Lane 7, 3xD7-PhaC-D7 beads (~ 117 kDa); Lane 8, WT PhaC beads (~ 64 kDa). Arrows indicate the PhaC protein or respective fusion proteins. (B) Schematic representation of relevant protein components. Number in front of each linear diagram corresponds to lane number in A.

The triplicate insertion problem was finally realised at this stage when a significant size difference was noticed between the C-terminal and N-terminal fusion protein bands as shown in lanes 5 and 6 in Figure 3.16. A repeated round of sequencing (2.4.7) for concerned plasmids confirmed the existence of 3xD7 at the N-terminus of PhaC.

As a remedy, a new construct coding for a single copy of D7 at the N-terminus of PhaC was under preparation. However, as it would take several weeks from plasmid cloning work to final bead product, and P4 assay kit and reagents had been purchased for a while and about to expire, further progesterone binding tests proceeded with just the beads available at that time, ie. without single copy N-terminal insertion beads.

### 3.2.2.2 Assessment of the progesterone binding capacity

In order to assess progesterone (P4) binding capacity of corresponding isolated beads (3.2.2.1), beads were incubated with a P4 solution, and a progesterone binding capacity test was performed by using a commercial Progesterone EIA Kit (Cayman Chemical, Ann Arbor, USA) according to Methods section 2.6.4.5. The EIA kit assay is based on competitive antibody binding between free P4 in a liquid sample and progesterone-acetylcholinesterase (P4 tracer) included with the kit, thus colour intensity upon addition of Ellman's reagent is proportional to the amount of P4 tracer bound to the well, which is inversely proportional to the amount of free P4 present in the well during the incubation. Therefore, the EIA kit cannot directly measure the amount of P4 bound by beads, but works indirectly by detecting P4 levels in the solution before and after bead incubation. Thus the amount of P4 reduced after bead incubation could be deemed as that bound by beads (either through the specific biorecognition of P4 by the OBody ligand D7, or through non-specific attachment of P4 to PHA beads). Therefore, to consider only the specific binding between P4 and OBody ligand D7, a simple wash step was introduced for beads after P4 incubation to wash off any non-specifically attached P4, which could be detected by the EIA kit, and should be subtracted as well in the calculation of bead P4 binding capacity.

A starting P4 solution (400 µg/ml dissolved in PBS buffer containing 20% ethanol as detailed in section 2.6.4.5) was used for bead binding capacity assay, and is defined as a feed fraction for later quantification of total P4 added. Following a pilot test, ~ 20 mg of WT PhaC beads or OBody beads outlined in the previous section were each mixed and incubated with about 75 Volumes (1.5 ml) of the feed fraction at 37ºC for 30 min, after

which beads were sedimented by centrifugation. The supernatant was collected as an unbound fraction for later quantification of the remaining soluble P4. Then beads were washed once with about 75 Volumes (1.5 ml) of PBS and the wash fraction was collected as well for later quantification of P4 that was non-specifically attached to beads and washed off. The feed, unbound and wash fractions were diluted as appropriate to be within the concentration ranges of progesterone standard provided by the kit and respectively assayed for P4 content. The P4 binding capacity of beads was calculated as:

$$\frac{Feed\ concentration \times Feed\ volume - Unbound\ concentration \times Unbound\ volume - Wash\ concentration \times Wash\ volume}{grams\ of\ beads\ used}$$

As shown in Table 3.9, non-specific P4 binding of WT PhaC control beads was observed, which was possibly caused by the hydrophobic interactions between the lipophilic P4 and PHA polymers, a process of physical absorption reported recently (Schäfer *et al.* 2011). Besides, the PBS buffer used and / or the single wash step might be insufficient to remove the non-specifically adsorbed P4 off the beads. Nevertheless, as compared to the WT PhaC beads, all OBody beads demonstrated a significantly higher binding capacity for progesterone, though no significant difference was found between any two types of OBody beads.

As fusion protein production level varied among different types of beads (Figure 3.16), P4 binding capacity based on bead biomass might not accurately reflect the role of OBody ligand D7 in P4 binding. Data conversion based on fusion protein amount again revealed similarly high P4 binding capacity for all OBody beads, and suggested that the recognition between bead fusion protein and P4 was not based on a 1:1 molar ratio (Table 3.9). It was unclear whether this was a reflection of the P4 recognition nature of D7, or

an indication of high background non-specific binding noises caused by PHA beads, as the exact recognition mechanism between D7 itself and P4 is still unknown so far.

Although the P4 EIA Kit assay provided a measurement of how much P4 could be bound by OBody beads, there were disadvantages in working with the Kit. As mentioned above, it could not be used for a direct quantification of P4 bound by beads, which involved a relatively more complicated procedure to separately collect feed, unbound and washing fractions, as well as related P4 level measuring efforts, thus it was inevitably error-prone. Besides, it could not reveal the binding strength (or binding affinity) between P4 and OBody beads, which would be an important factor in assessing whether OBody beads could be used for detection of bovine milk P4 usually at a low biological level of 1-10 ng/ml (Daems *et al.* 2017). Therefore, an alternative assay leading to a simplified direct quantification of P4 bound on beads and a measurement of binding affinity between P4 and OBody beads was desirable.

**Table 3.9** *Progesterone binding capacity of OBody beads*

| Bead prototype | Amount of P4 bound (value ± SD) | | MW | Fusion protein[a] | Beads loaded | Fusion protein concentration[a] | Amount of P4 bound (value ± SD) | |
|---|---|---|---|---|---|---|---|---|
| | mg/g beads[1] | µmol/g beads | kg/mol | ng[2] (pmole) | µg | pmol/µg =µmol/g beads | mg/µmol fusion protein[a] | µmol/µmol fusion protein[a] |
| PhaC-D7 | 23.2 ± 3.1* | 73.8 ± 9.9* | 78 | 236.2 (3.0) | 32 | 0.093 | 249.5 ± 33.3* | 793.5 ± 106.5* |
| 3xD7-PhaC | 29.5 ± 2.3* | 93.8 ± 7.3* | 106 | 382.2 (3.6) | 32 | 0.113 | 261.1 ± 20.4* | 830.1 ± 64.6* |
| 3xD7-PhaC-D7 | 29.7 ± 2.7* | 94.4 ± 8.6* | 117 | 211.3 (1.8) | 16 | 0.113 | 262.8 ± 23.9* | 835.4 ± 76.1* |
| WT PhaC | 3.8 ± 1.3 | 12.0 ± 4.1 | 64 | 258.1 (4.0)[a] | 24 | 0.16 [a] | 22.8 ± 7.8[a] | 71.9 ± 24.6[a] |

[1] Value obtained from the commercial Progesterone EIA Kit assay (2.6.4.5), n = 2.
[2] Estimated value based on densitometry analysis of the SDS-PAGE image as shown in Figure 3.16.
[a] Fusion protein for OBody beads or PhaC protein in the case of WT PhaC beads.
* $p < 0.005$ vs. WT PhaC controls (with a specific p value of 0.00473, 0.00164 and 0.00159 for PhaC-D7, 3xD7-PhaC and 3xD7-PhaC-D7, respectively).

**Table 3.10** *Calculation of bead dilution and theoretical amount of P4 needed*

| Bead prototype | MW | Fusion protein[1] | Beads Loaded[2] | Fusion protein concentration | | Assay at 0.5 µM | Assay at 0.1 µM | Assay at 0.01 µM | Assay at 0.01 µM |
|---|---|---|---|---|---|---|---|---|---|
| | kg/mol | ng | µl | ng/µl | nmol/ml = µM | bead dilution times (µg bead used per 50 µl assay)[3] | | | µM P4 for saturation[4] |
| PhaC-D7 | 78 | 236.2 | 0.16 | 1476.3 | 18.9 | 37.8 (264.6) | 189 (52.9) | 1890 (5.3) | 7.9 ± 1.1 |
| 3xD7-PhaC | 106 | 382.2 | 0.16 | 2388.8 | 22.5 | 45.0 (222.2) | 225 (44.4) | 2250 (4.4) | 8.3 ± 0.6 |
| 3xD7-PhaC-D7 | 117 | 211.3 | 0.08 | 2641.3 | 22.6 | 45.2 (221.2) | 226 (44.2) | 2260 (4.4) | 8.4 ± 0.8 |
| WT PhaC | 64 | 258.1 | 0.12 | 2150.8 | 33.6 | 67.2 (148.8) | 336 (29.8) | 3360 (3.0) | 0.7 ± 0.2 |

[1] Estimated value based on densitometry analysis of the SDS-PAGE image as shown in Figure 3.16.
[2] Value calculated as $\frac{\text{µl of diluted bead sample loaded on gel}}{\text{bead sample dilution times}}$.
[3] Value calculated based on the 20% (200 mg/ml) (w/v) bead stocks, the dilution listed, and the assay volume of 50 µl.
[4] Value calculated based on P4 binding capacity shown in Table 3.9.

### 3.2.2.3 Assessment of the equilibrium dissociation constant ($K_D$)

Before entering into descriptions about the alternative assay, it is necessary to summarize the concept of binding affinity. Binding affinity is commonly quantitatively represented by the equilibrium dissociation constant ($K_D$). When a target molecule A and its recognising molecule B are associated to form a complex (AB) in a solution, the corresponding equilibrium dissociation constant is defined as: $K_D = \frac{[A][B]}{[AB]}$, where [A], [B] and [AB] represent the concentrations of A, B and bound complex, respectively. The dissociation constant $K_D$ has units of molarity (M), and corresponds to the concentration of target molecule [A] at which the binding site on the recognising molecule B is half occupied, i.e. when the concentration of B with target bound [AB] equals the concentration of B with no target bound [B]. Therefore, a $K_D$ value is traditionally deduced from a typical saturated binding curve obtained through an ELISA format binding assay, and is approximately the concentration of target molecule [A] at which the binding reaction reading corresponds to half of the saturation reading. The smaller the $K_D$ value is, the more tightly or strongly bound the two molecules are, which contributes to significant binding even at very low concentrations of A and B.

In order to determine an equilibrium dissociation constant ($K_D$), a progesterone binding assay with a fixed amount of beads but varying amount of a biotin labelled P4 (progesterone 3-PEG11-biotin (Cayman Chemical), gifted by Vickery Arcus) was performed according to Methods section 2.6.4.6. WT PhaC beads were included as negative controls. A monoclonal anti-progesterone antibody produced in rat (1:2000) (P1922, Sigma) was used as a positive control. As compared to the above-mentioned progesterone EIA kit, the use of a biotin labelled P4 in combination with HRP-Conjugated

Streptavidin (Thermo Fisher Scientific, gifted by Vickery Arcus) would enable a sensitive direct quantification of P4 bound on beads. Here an OPD (*o*-Phenylenediamine dihydrochloride) tablet (P9187, Sigma) was used as HRP substrate, the reaction was stopped by addition of $H_2SO_4$ and colour intensity was monitored at 490 nm.

First, as a pilot experiment, all bead samples were tested in a volume of 50 µl at a concentration of 0.5 µM protein (in terms of respective target fusion protein for OBody beads or PhaC itself for WT PhaC beads), based on densitometry analysis of the SDS-PAGE image as shown in Figure 3.16 and calculations shown in Table 3.10. The biotin labelled P4 added was 50 µl of serial dilutions with a concentration range from 0.02 to 43.84 µM. Results shown in Figure 3.17 indicated the lack of saturation for all samples with 0.5 µM bead fusion protein at the P4 concentrations tested, and non-specific binding occurred as well in this biotin labelled P4 binding assay for the negative controls WT PhaC beads, which bound P4 at a similar level to positive control rat anti-progesterone antibody (Figure 3.17). Nevertheless, OBody beads did show relatively high levels of P4 binding, particularly the 3xD7-PhaC PHA beads (Figure 3.17).

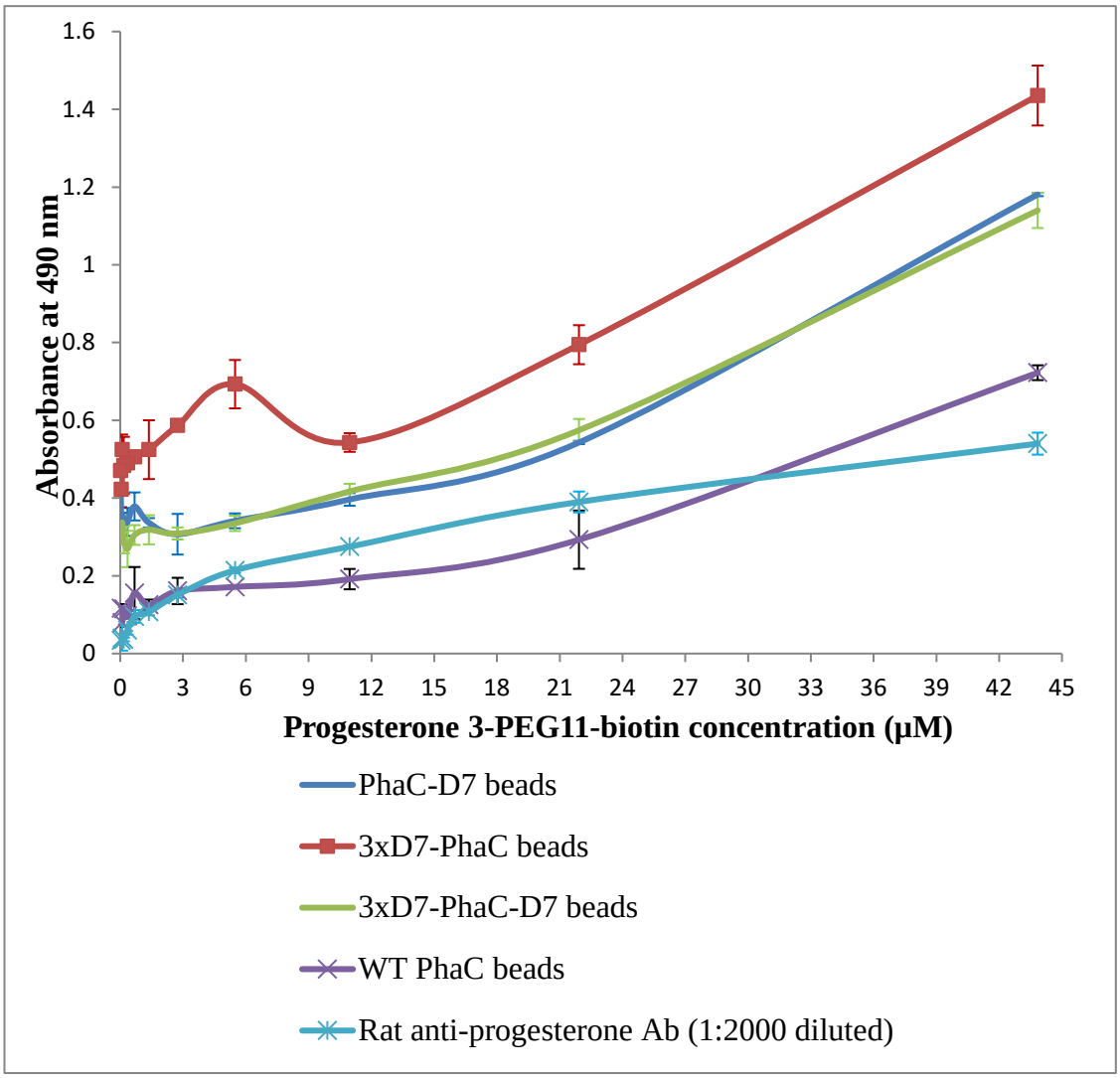



**Figure 3.17 P4 binding of PHA beads at 0.5 µM bead fusion protein.** All reactions were performed in duplicate, and the error bars represented standard deviations.

Too much bead fusion protein and to too little biotin labelled P4 used in the experiment could have both caused the lack of saturation. Either reducing the fusion protein amount or increasing the P4 amount might be helpful to solve this problem. However, lab budget was limited and the biotin labelled P4 was expensive and preciously gifted by the collaborator (Vickery Arcus and the "OBodies Limited"), therefore further tests were carried out with only OBody beads, at concentrations lower than 0.5 µM bead fusion protein. PBS was included as blank, and wells coated with OBody beads only (without

addition of P4) were included as bead background control. The bead background reading was subtracted for each bead sample reading before graphing.

In an attempt to get a saturated binding curve, OBody beads were diluted further down to 0.1 µM bead fusion protein. But results still indicated the lack of saturation as shown in Figure 3.18.

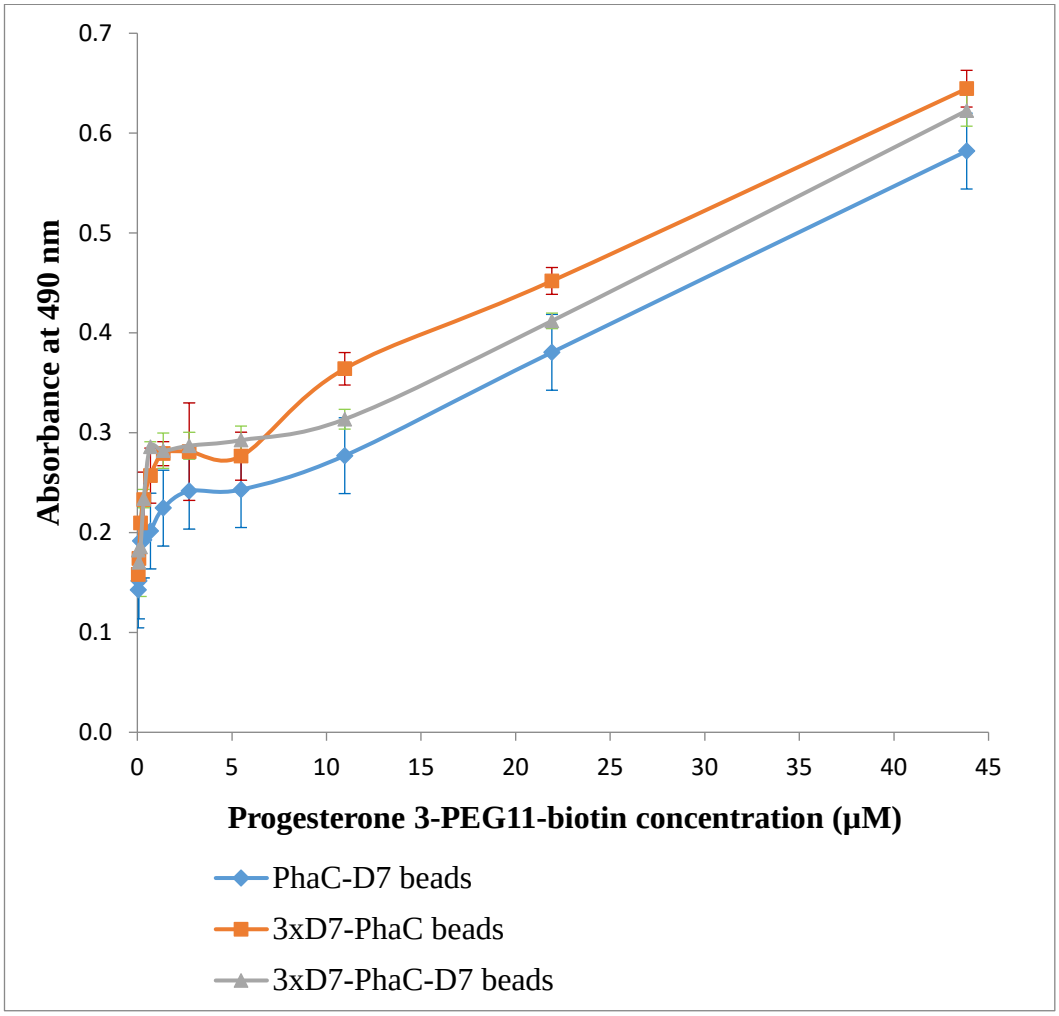



**Figure 3.18 P4 binding of OBody beads at 0.1 µM bead fusion protein.** All reactions were performed in duplicate, and the error bars represented standard deviations.

In another attempt to get a saturated binding curve, OBody bead samples were diluted further down to 0.01 µM bead fusion protein. Note that at this assay concentration of 0.01 µM, theoretical concentration of P4 needed to saturate all the binding sites on beads (as calculated in Table 3.10) was within the tested P4 concentration range (0.02 - 43.84 µM). But unfortunately, results still indicated the lack of saturation (Figure 3.19), which meant no $K_D$ value could be deduced from the experiment. The lack of saturation in this biotin labelled P4 binding assay implied a binding capacity even higher than that was indicated by the P4 EIA kit assay, which might be attributed to the sensitivity of the biotin labelled P4 assay. Therefore, the lack of saturation might not be a bad thing, especially considering the intended usage in detecting bovine milk P4 generally at a low level of 1-10 ng/ml (Daems *et al.* 2017).

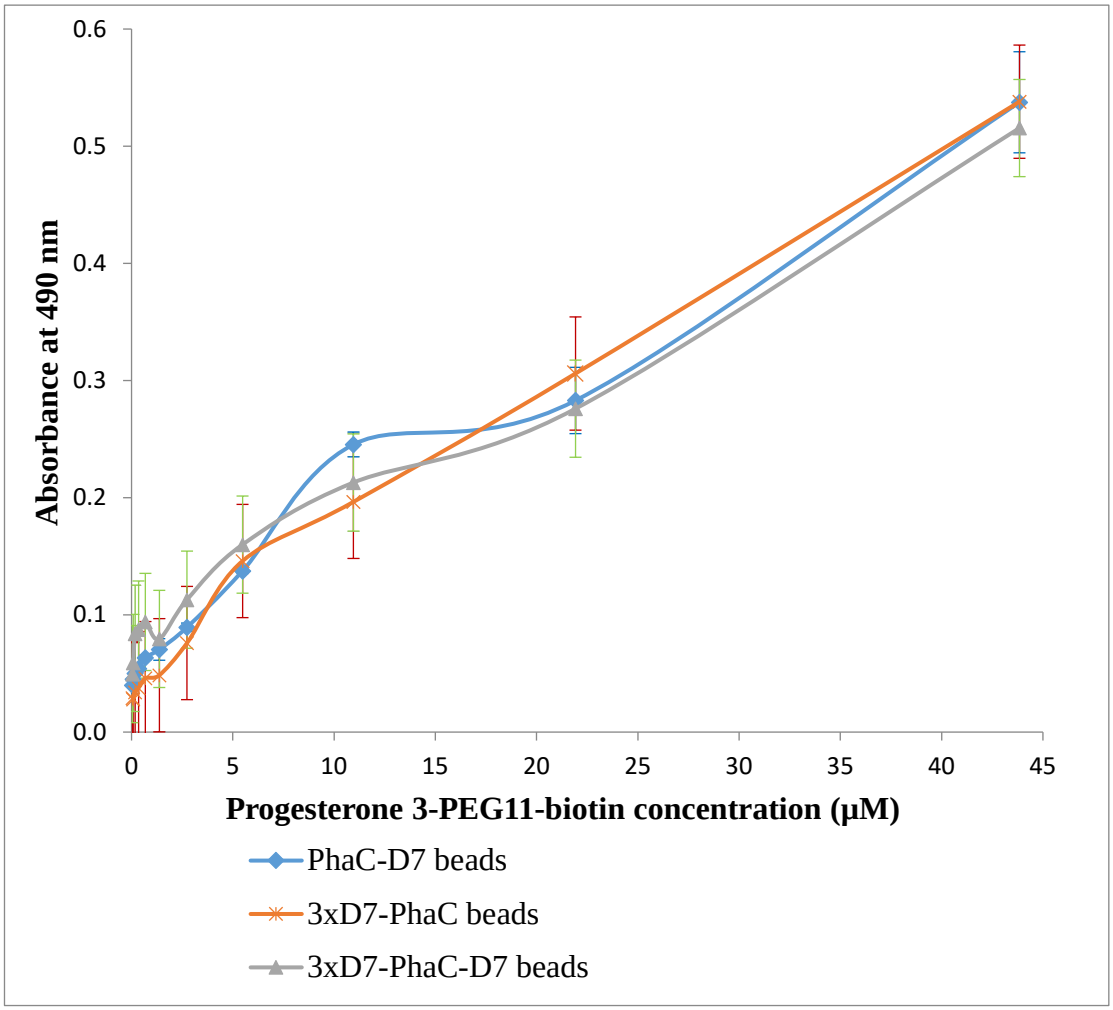



**Figure 3.19 P4 binding of OBody beads at 0.01 µM bead fusion protein.** All reactions were performed in duplicate, and the error bars represented standard deviations.

No further dilution was assayed due to the following two condensations. Firstly, as shown in Table 3.10, to obtain an assay dilution of 0.01 µM bead fusion protein from respective stock bead samples (200 mg/ml (w/v) bead slurry indicated in section 2.5.2), a $10^{-3}$ order of dilution was already necessary which resulted in an extremely clear diluted suspension. Considering the small volume (50 µl) of beads added in each plate well (corresponding to 3 - 5.3 µg of bead biomass shown in Table 3.10), further bead dilution was undesirable for an ideally even bead sampling to typically represent the real binding events. Secondly, a new generation of P4 binding OBody B7 became available from the collaborator (Vickery Arcus and the "OBodies Limited") with an improved affinity as compared to

D7 (below 300 nM vs. 300 - 400 nM). It would be interesting to immobilise B7 on PHA beads and assess the P4 binding functionality in comparison with the 1st generation D7 counterparts, which is detailed in the next section.

### 3.2.2.4 Production and function assessment of the 2nd generation of progesterone binding OBody beads

In this section, the 2nd generation OBody B7 with improved affinity for P4 was similarly immobilised on PHA beads and tested for P4 detection usage. The B7 ligand is also a synthetic peptide that was engineered based on OB-fold domain of aspartyl-tRNA synthetase (aspRS) from *P. aerophilum*, and developed by the collaborator (Vickery Arcus and "OBodies Limited") for a stronger binding towards P4 (with an affinity below 300 nM as compared to 300 - 400 nM for the 1st generation OBody D7). B7 was a clone code given by them for this particular OBody ligand during library screening and was adopted accordingly in this study. It is 106 amino acids in length (with amino acid sequence shown in Figure 7.18, Appendix 7.1), and contains 3 amino acid deletions plus 24 substitutions as compared to the wild type OB-fold (residues 1-109, GenBank ID NP_558783.1) (see sequence alignment in Figure 7.27, Appendix 7.1). In addition, as compared to the 1st generation D7 (with amino acid sequence shown in Figure 7.13, Appendix 7.1), B7 (with amino acid sequence shown in Figure 7.18, Appendix 7.1) contains 8 amino acid substitutions (see sequence alignment in Figure 7.28, Appendix 7.1).

This B7 peptide was fused only to the N-terminus of PhaC (PHA synthase) based on the following two considerations. Firstly, for OBody beads displaying the 1st generation D7 (abbreviation for its clone code P4013-D7), no significant statistical difference was found

among the three types beads of PhaC-D7, 3xD7-PhaC and 3xD7-PhaC-D7 in P4 binding (Table 3.9, section 3.2.2.2), regardless of the fusion copy and/or orientation of D7. It was thus anticipated that whether a N- or a C-terminal fusion of B7 would not make much difference either. Secondly, the construct bearing a single copy N-terminal fusion of D7 (pET14b-D7-PhaC as shown in Table 2.2 which was prepared according to methods section 2.4.) was ready for assessment by then. It was thus decided to prepare counterpart N-terminal B7 fusion constructs (pET14b-B7-PhaC and pET14b-3xB7-PhaC as shown in Table 2.2) according to methods section 2.4.8 for affinity comparison purpose with D7.

The plasmid DNA sequences were confirmed (2.4.7) and their maps are shown in Figures 7.17, 7.19 and 7.20, Appendix 7.1. Plasmid transformation, cell cultivation for standard bead production, bead isolation, SDS-PAGE and densitometry analysis were performed according to methods sections 2.4.6, 2.2.2.3, 2.5.2, 2.6.2.1 and 2.6.3.2, respectively.

As mentioned in sections 3.2.1.1, it was established previously that wild type PHA beads (PhaC beads) could be produced using BL21 (DE3) *E. coli* strain (Table 2.1) harbouring plasmids A and B (namely plasmid pETC encoding PhaC the PHA synthase and plasmid pMCS69 encoding PhaA and PhaB) (Table 2.2) (Peters *et al.* 2007). Similarly, in the current work, PHA beads were produced by respective strains (*E. coli* BL21 (DE3) containing either pET14b-D7-PhaC, pET14b-B7-PhaC, pET14b-3xD7-PhaC or pET14b-3xB7-PhaC as plasmid A and pMCS69 as the helper plasmid B), as shown by the dominant fusion protein bands in Figure 3.20.

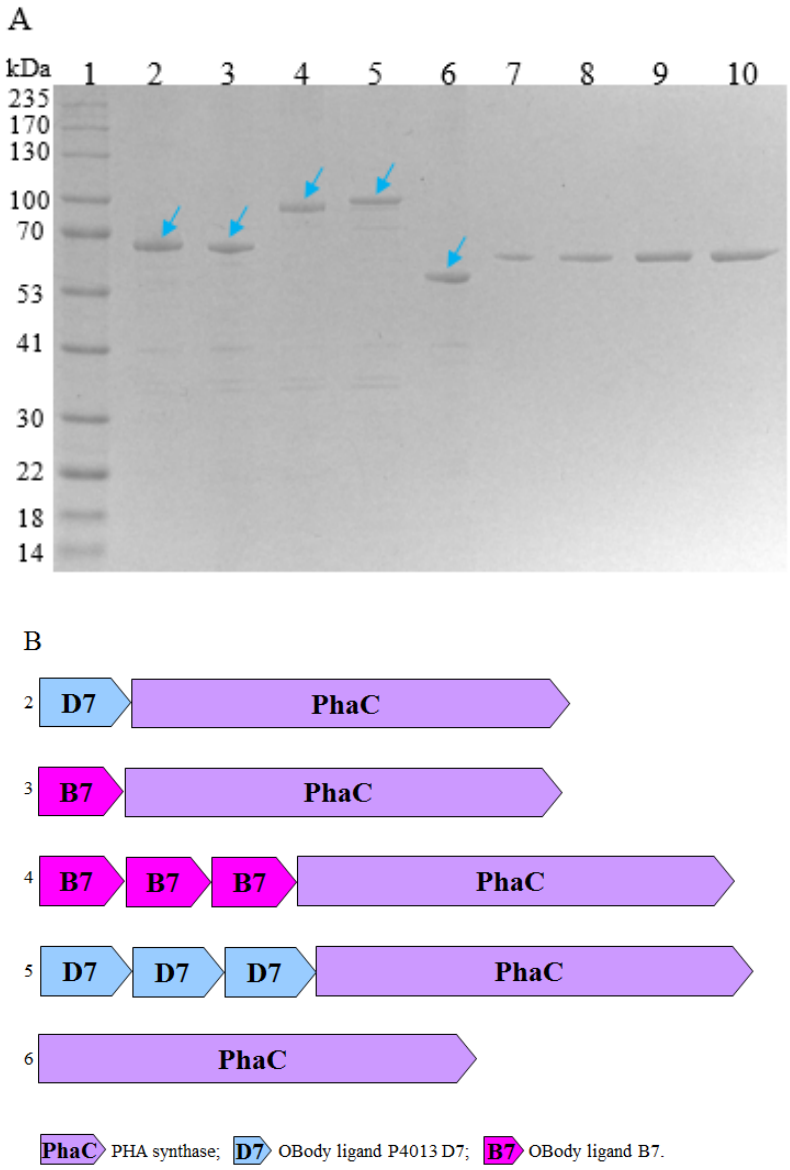



**Figure 3.20 Protein profiles of isolated PHA beads and schematic representation of relevant protein components.** (A) 10% SDS-PAGE was performed to examine whether PHA beads were produced with the correct protein profiles. Lane 1, Molecular weight marker (GangNam-STAIN$^{TM}$ Prestained Protein Ladder); Lane 2, D7-PhaC beads (~ 78 kDa); Lane 3, B7-PhaC beads (~ 77 kDa); Lane 4, 3xB7-PhaC beads (~ 103 kDa); Lane 5, 3xD7-PhaC beads (~ 106 kDa); Lane 6, WT PhaC beads (~ 64 kDa); Lane 7-10, 100, 200, 400 and 500 ng BSA. Arrows indicate the PhaC protein or respective fusion proteins. (B) Schematic representation of relevant protein components. Number in front of each linear diagram corresponds to lane number in A.

A biotin labelled P4 based progesterone binding assay was performed according to Methods section 2.6.4.6 using these OBody beads at 0.01 µM bead fusion protein (see calculations in Table 3.11). Soluble D7 and B7 were also assayed at 0.5 µM (as recommended by Vickery Arcus), which were gifted by the collaborator upon request as it would be interesting to see what difference could be brought about by immobilising respective OBody ligands on the surface of PHA beads. According to the collaborator, the soluble D7 has an affinity of 300 - 400 nM for P4, while B7 is below 300 nM when analysed with a similar biotin labelled P4 based progesterone binding assay. Thus the P4 concentration tested (0.02 to 43.84 µM) in this study theoretically was sufficient to saturate all the binding sites on D7 or B7. WT PhaC beads were included as negative controls, and rat anti-progesterone antibody (1:4000) was used as a positive control. PBS was included as blank, and wells coated with beads only (without addition of P4) were included as bead background control. Before graphing, the PBS blank reading was subtracted from soluble OBody sample reading, and the bead background reading was subtracted from each bead sample reading.

**Table 3.11** *Dilution calculation of bead immobilised and soluble OBodies*

| Bead prototype | MW | ng fusion | Beads Loaded[3] | Fusion protein concentration[a] | | Assay at 0.01 or 0.5 µM | |
|---|---|---|---|---|---|---|---|
| | kg/mol | protein[a] | µl | ng/µl | nmol/ml = µM | bead dilution (µg bead used per assay)[4] | |
| D7-PhaC | 78 | 355.7[1] | 0.15 | 2371.3 | 30.4 | 3040 (3.3) | - |
| B7-PhaC | 77 | 242.3[1] | 0.15 | 1615.3 | 21.0 | 2100 (4.8) | 42 (238.0) |
| 3xB7-PhaC | 103 | 200.6[1] | 0.15 | 1337.3 | 13.0 | 1300 (7.7) | - |
| 3xD7-PhaC | 106 | 189.4[1] | 0.15 | 1262.7 | 11.9 | 1190 (8.4) | - |
| WT PhaC | 64 | 226.6[1] | 0.75 | 302.1 | 4.7 | 470 (21.2) | - |
| WT PhaC | 64 | 221.5[2] | 0.15 | 1476.7 | 23.1 | - | 46.2 (216.5) |
| Soluble OBody | MW | | | Soluble OBody concentration | | | Assay at 0.5 µM |
| | kg/mol | | | µg/ml[b] | µmol/l = µM | | OBody dilution times |
| D7 | 13 | | | 690 | 54.3 | | 108.6 |
| B7 | 13 | | | 2230 | 175.6 | | 351.2 |

[a] Fusion protein for OBody beads or PhaC protein in the case of WT PhaC beads.
[1] Estimated value based on densitometry analysis of the SDS-PAGE image as shown in Figure 3.20.
[2] Estimated value based on densitometry analysis of a batch of WT PhaC beads.
[3] Value calculated as $\frac{\text{µl of diluted bead sample loaded on gel}}{\text{bead sample dilution times}}$.
[4] Value calculated based on the 20% (200 mg/ml) (w/v) bead stocks, the dilution listed, and the assay volume of 50 µl.
[b] Recorded value with the gifted soluble OBody from the collaborator.

The results shown in Figure 3.21 again indicated a lack of saturation for all samples with 0.01 µM bead fusion protein at the P4 concentrations (0.02 - 43.84 µM) tested. Negative control WT PhaC beads showed an abnormally high level of P4 binding (Figure 3.21), although this might have been due to the fact that over 2-6 times more of WT PhaC bead biomass was used (as compared to OBody beads) to get a consistent PhaC protein concentration of 0.01 µM (see Table 3.11), which naturally means more chances of non-specific binding.

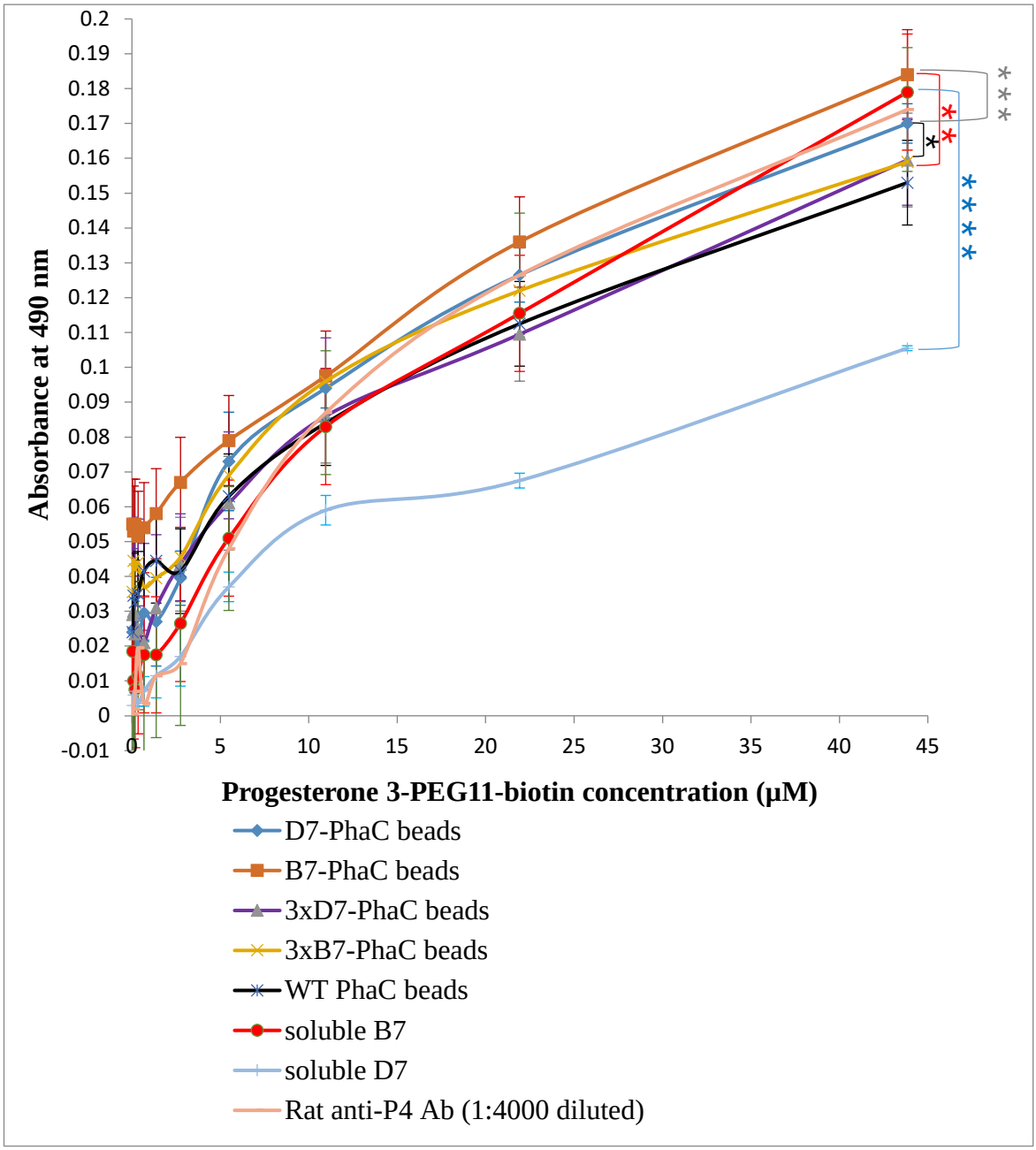



**Figure 3.21 P4 binding at 0.01 µM bead fusion protein for immobilised OBodies and 0.5 µM protein for soluble OBodies.** All reactions were performed in duplicate, and the error bars represented standard deviations. * p = 0.0056 at the maximum absorbance. ** p = 0.0016 at the maximum absorbance. *** p = 0.0050 at the maximum absorbance. **** p = 0.0001 at the maximum absorbance.

However, in general, the 2nd generation of OBody showed a relatively higher degree of P4 binding than its 1st generation counterpart, namely, soluble B7 higher than soluble D7,

and immobilised B7 (B7-PhaC beads) higher than immobilised D7 (D7-PhaC beads) (Figure 3.21). Interestingly, for both generations of immobilised OBodies, the single copy fusion performed better than the triplicate fusion, namely D7-PhaC beads bound more P4 than 3xD7-PhaC beads did, and B7-PhaC beads more than 3xB7-PhaC beads (Figure 3.21). Promisingly, the best performer, immobilised B7 (B7-PhaC beads) at 0.01 µM bead fusion protein performed comparably to the soluble B7 at 0.5 µM, and considering the 50-fold difference in their assay concentrations, bead immobilised B7 (B7-PhaC beads) might be useful in detecting P4 in bovine milk.

Unexpectedly, the P4 concentration (0.02 - 43.84 µM) tested did not even saturate the binding sites of soluble D7 or B7, which was in contradiction to observations from the collaborator. This led to questions on the analysis system used in this study. A careful examination of the experimental protocol used in this study and that from the collaborator revealed that both protocols were essentially the same, except for a house made PBS in this study vs. a commercial DPBS buffer (Thermo Fisher Scientific), an OPD (*o*-Phenylenediamine dihydrochloride) substrate (which requires plate reading at 490 nm, Sigma #P9187) in this study vs. a TMB (3,3',5,5' tetramethylbenzidine) substrate (which requires plate reading at 450 nm, Thermo Fisher scientific # 34028), as well as a high-binding microtitre plate (Greiner Bio-One #655061) in this study vs. a low-binding one (Greiner Bio-One #655101). Buffers and substrates (and respective reading wavelength as required) were unlikely to be responsible for differences seen between my results and those of our collaborator. However plates made of high-binding and low-binding polystyrene could have possibly caused a significant difference in non-specific physical absorption of P4 through hydrophobic interactions with the lipophilic P4 (Longman & Buehring 1986; Schäfer *et al.* 2011).

Because both PHA beads and the high-binding ELISA plate used in this study could have possibly caused the lack of saturation for either immobilized or soluble OBodies (as well as the abnormally high P4 binding of the negative control WT PhaC beads) due to non-specific physical absorption processes (Figure 3.21), then using an approximately equivalent bead biomass for each bead sample and adding blocking agents throughout the assay process might help to counteract these effects. In order to test this, a modified progesterone binding assay was done still using the high-binding ELISA plate but in the presence of both 1% BSA and 5% skim milk. Again, due to budget limit, only the same batch of best performing B7-PhaC beads (Figure 3.21) and a freshly isolated batch of control WT PhaC beads were tested, such that to get consistent protein concentration of 0.5 µM, about equivalent WT PhaC bead biomass was used as compared to B7-PhaC beads (Table 3.11).

Preliminary results showed that with an approximately equivalent bead biomass and the presence of blocking agents, a counteracting effect on the non-specific binding noise for both negative control WT PhaC beads and B7-PhaC beads was obvious (Figure 3.22). Particularly, within the tested range of progesterone concentration (0.086-21.92 µM), WT PhaC beads tended to give relatively flat low level of absorbance, whereas B7-PhaC beads showed a typical binding curve which saturated at around 5.5 µM (Figure 3.22). The half-maximum binding capacity corresponded to an absorbance of about 0.028, which was close to the value when about 0.7 µM progesterone 3-PEG11-biotin were used, therefore the deduced apparent equilibrium dissociation constant ($K_D$) between B7-PhaC beads and P4 was about 0.7 µM or $7\times10^{-7}$ M (Figure 3.22). This $K_D$ value was not ideal as compared to that for the soluble B7 (< 0.3 µM as shared by the collaborator), pointing again to the issue of non-specific P4 binding possibly caused by PHA beads and / or the high-binding ELISA plate, which led to a less tight overall binding between B7-PhaC beads and P4.

Nevertheless, the result was really encouraging considering the fact that 5.5 µM of P4 could saturate B7-PhaC beads assayed at 0.5 µM in the presence of blocking agents (Figure 3.22), whereas up to 43.84 µM of P4 could not when the same batch of B7-PhaC beads was assayed at a concentration as low as 0.01 µM without blocking agents (Figure 3.21).

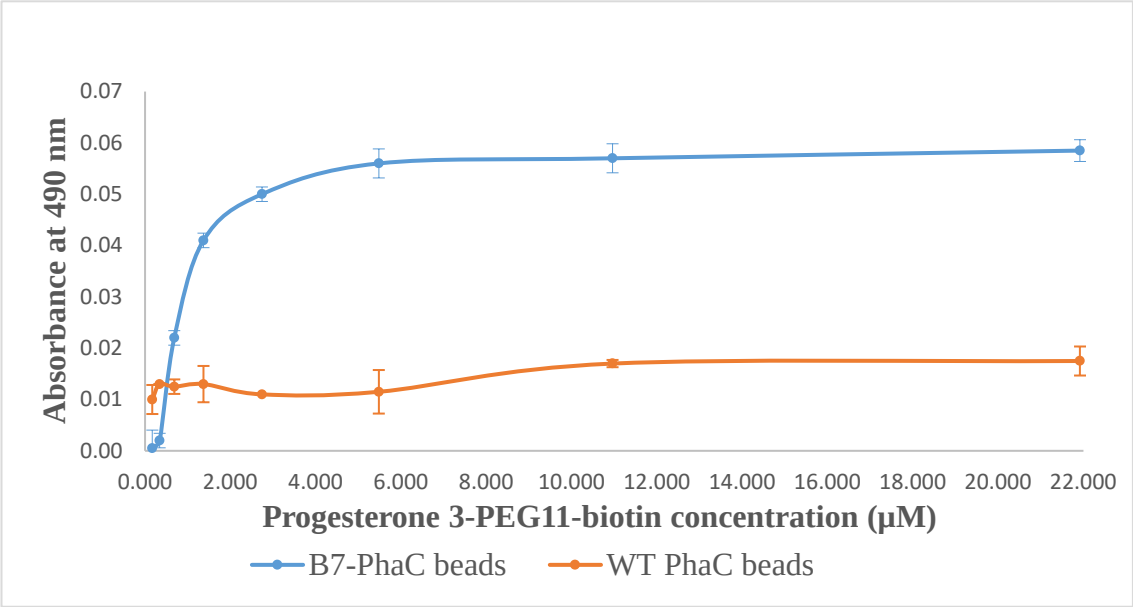



**Figure 3.22 P4 binding at 0.5 µM bead fusion protein.** Both 1% BSA and 5% skim milk were used as blocking agent. All reactions were performed in duplicates, and the error bars represented standard deviations.

Further study is necessary to examine whether the remaining non-specific binding noise shown in Figure 3.22 could be further eliminated by using a low-binding ELISA plate and / or choosing an optimum blocking agent, which will be elaborated upon in the Discussion / Future work sections.

# Chapter 4: Discussion

## 4.1 Design of PHA beads as self-cleavable protein purification resins

Recombinant protein production and purification from *Escherichia coli* is usually complicated and costly, especially for therapeutic proteins. For example, in terms of production and purification for medically important proteins such as human tumour necrosis factor alpha (TNFα), human granulocyte colony-stimulating factor (G-CSF) and interferon alpha-2b (IFNα2b), several recently published studies focus on optimising protein refolding from inclusion bodies as well as optimising chromatography operations (Vemula *et al.* 2015; Wang *et al.* 2015; Romanov *et al.* 2017). Many are trying to avoid tedious refolding by adopting solubility-enhancing / affinity tags in combination with specialized affinity resins plus additional protease treatment efforts in tag removal (Rabhi-Essafi *et al.* 2007; Do *et al.* 2014; Alizadeh *et al.* 2015). Hence, there is a need for a streamlined process with less complicated steps toward purification of recombinant proteins.

Production of recombinant target proteins as part of PhaC (PHA synthase) fusion was recently conceived as means to bind the target to PHA inclusions *in vivo* for facilitated purification. The target protein was covalently anchored to PHA beads via PhaC gene fusion, enabling efficient enrichment of the target protein (Grage *et al.* 2011). An enterokinase cleavage site was inserted between PhaC and the target protein, such that enterokinase treatment of respective PHA beads enabled release of the target protein (Grage *et al.* 2011). This platform had the advantages of being convenient and cost-

effective. Nevertheless, this process used additional expensive protease to initiate the release of the target protein. Further downstream processing might be necessary to remove the enterokinase, especially when purity is of critical importance, such as for biopharmaceuticals or protein crystallography.

In the current study, purification resins based on PHA beads were developed by inserting a self-cleavable tag (sortase or intein) between PhaC and the target protein, thereby the target protein was produced and immobilised on PHA beads during cell growth. During PHA bead isolation / washing step the target protein was separated from other cell components, and finally during inducible self-cleavage of the sortase or intein tag the protein was released off the isolated beads as a soluble fraction. Therefore, the common need for complicated chromatography or specialised affinity resin or costly protease treatment could be eliminated, which is particularly advantageous over previous studies.

Furthermore, the current platform is beneficial for the stability and solubility of target proteins due to their *in vivo* immobilisation on PHA beads. It is widely accepted that immobilisation improves protein stability and solubility (Rehm *et al.* 2016; Rehm *et al.* 2017), therefore the *in vivo* immobilisation of target proteins on the surface of PHA beads may aid in the functional folding of hard-to-express proteins (i.e., those prone to inclusion body (IB) formation). Indeed, all of the IB-prone therapeutic proteins exemplified in the current study could be cleaved off isolated PHA beads in a soluble form without any refolding effort.

Conformation-specific antibodies suggested a high fraction of the target protein population is properly folded. Making the assumption that commercial protein standards (Sino Biological Inc.) tested in the ELISA with respective conformation-specific

antibodies were 100% correctly folded (Figure 3.6, section 3.1.1.3 and Figure 3.13, section 3.1.2.3), a large proportion of the exemplified therapeutic proteins were suggested to be properly folded; for example 72.2% of TNFα and 39.7% of IFNα2b appeared to fold properly when the sortase tag was used, and 68.8% of TNFα, 39.1% of G-CSF and 60.4% of IFNα2b when the intein tag was used (Figure 3.6, section 3.1.1.3 and Figure 3.13, section 3.1.2.3, respectively).

Nevertheless, the observed fraction of misfolded proteins might be due to the complexity of disulphide bond formation within these proteins. TNFα only contains one S-S bond, while both G-CSF and IFNα2b need two S-S bonds for correct folding. This is undoubtedly the limitation when producing disulphide-bonded eukaryotic proteins in *E. coli* cytoplasm (Lobstein *et al.* 2012). In *E. coli*, disulphide bond formation / isomerization is confined to the periplasm through the Dsb (disulphide bond formation) family of oxidoreductases (Denoncin & Collet 2013). Thus eukaryotic proteins requiring disulphide bond for folding / stability are generally misfolded and inactive in the reducing environment of *E. coli* cytoplasm (Lobstein *et al.* 2012). Some *E. coli* cells were engineered and commercialised to provide an oxidizing cytosol to ease disulphide bond formation of recombinant proteins. For example, SHuffle® T7 express contains mutated *trxB/gor* genes (coding for thioredoxin reductase and glutaredoxin reductase) as well as an extra chaperone *dsbC* gene without a signal sequence that ensures the disulphide isomerase stays in the cytoplasm, therefore allowing disulphide bond formation in the cytoplasm. The Shuffle® T7 express strain was used in this study in combination with a sulfhydryl oxidase (Erv1p) expressing plasmid pMCS69E to aid in disulphide bond formation of such proteins as TNFα, G-CSF and IFNα2b. However, all these efforts seemed not so efficient to ensure a 100% correct folding.

One or more co-purifying host cell proteins were observed along with each cleaved therapeutic target protein by SDS-PAGE analysis (indicated by an arrow in Figure 3.5 A & B, section 3.1.1.2, or numbered as 1 to 3 in Figure 3.12, section 3.1.2.3), and were suggested via mass spectrometry to be the *E. coli* chaperone protein DnaK and both full length and truncated outer membrane protein A, respectively (Figure 7.21C and 7.24, Appendix 7.1), which are known as common impurities in *E. coli* inclusion bodies (Carrió & Villaverde 2005; Jürgen *et al.* 2010). Further purity improvement will likely be achievable via process optimisation during bead isolation and washing steps or by applying further purification steps after cleavage. As mentioned in the Methods section 2.5.2, a lysis buffer containing 0.04% (w/v) SDS with or without 150 mM NaCl at a pH of either 8.8 or 9.0 was adopted in this study for PHA bead isolation from *E. coli*. Increasing salt concentration and / or pH as well as addition of 0.05% (v/v) nonionic surfactant Tween 20 has been recommend for cleaning PHA beads displaying proteins of interest (Hay *et al.* 2014). It has also been previously reported that mechanical disruption of cell biomass suspended in a high concentration of SDS could improve both the purity of final PHA products and cell disruption efficiency. For example PHA of 95% purity could be obtained from *Methylobacterium* sp V49 cells suspended in 5% (w/v) SDS (no pH information) (Ghatnekar *et al.* 2002), and a cell disruption efficiency >99.99% could be reached for *Cupriavidus necator* cells suspended in 1% (w/v) SDS (pH 12) (Koller *et al.* 2013). However, these studies aimed at only PHA material as the final target, rather than any valuable protein coated on PHA beads as is the case of this study, therefore the bead treatment of using 1% to 5% SDS might be too harsh for proteins. Nevertheless, using a lysis buffer containing SDS > 0.04% (w/v) and/or at a pH > 9, along with more washing steps, might yield purer PHA beads, although optimisation would be required to avoid protein inactivation due to high pH or high SDS levels. Furthermore, use of the nonionic surfactant Triton X-100 in the lysis buffer might also help to remove the non-

specifically attached chaperone protein DnaK and membrane protein A, because Triton X-100 has long been used in membrane protein solubilisation (Schnaitman 1971), and has been proven to inhibit nonspecific adhesion of proteins onto a hydrophobic surface (Numata *et al.* 2006). Additionally, further chromatograph purification (or simple rebuffering / dialysis) after cleavage might be desirable to remove the $CaCl_2$ and triglycine introduced during cleavage, in the case of the sortase tag, to achieve biopharmaceutical grade purity.

### 4.1.1 Design of PHA beads as self-cleavable protein purification resins mediated via PhaC-sortase-LPETG-target protein fusion

Here it was demonstrated that PHA beads based purification resins mediated via PhaC-sortase-LPETG-target protein fusions could be used for the production and purification of therapeutic proteins. Target therapeutic proteins were produced with a single G scar on the N-terminus. Nevertheless, G's properties (uncharged and the smallest possible amino acid) should make it a relatively innocuous addition to most proteins.

Premature cleavage of sortase before subjecting PHA bead fusions to the cleavage buffer that contained 5 mM $CaCl_2$ (section 2.6.1) amounted to about 30-35% as measured via densitometry (Figure 3.5 A & B, section 3.1.1.2), despite the fact that *E. coli* only maintains cytosolic free $Ca^{2+}$ homeostasis at about 100-300 nM (Dominguez 2004). This was despite efforts to reduce premature cleavage by adopting a relatively short bead production period (about 24 h vs. normal 48 h) as well as using high EDTA content during bead isolation and washing steps to chelate $Ca^{2+}$. This suggested a less than stringent $Ca^{2+}$ control of the sortase activation, which was consistent with previous studies (Mao 2004; Bellucci *et al.* 2013), and inevitably impacted the final yield of target proteins. Future

studies investigating molecular mechanisms on how to improve control of sortase activation might be necessary in this regard.

### 4.1.2 Design of PHA beads as self-cleavable protein purification resins mediated via PhaC-intein-target protein fusion

Previously, two different PHA granule associated proteins (namely, PHA structure protein phasin (PhaP), and PHA regulatory protein (PhaR)) were utilized for protein purification by serving as PHA bead anchors fused via a self-cleavable intein to a target protein (Banki *et al.* 2005; Barnard *et al.* 2005; Wang *et al.* 2008; Zhang *et al.* 2010; Zhou *et al.* 2011). However, the intrinsically less stable hydrophobic interaction between the PhaP / PhaR and the PHA beads imposed constraints on its application. For example, during PHA bead isolation and even during target the protein elution process, the salt concentration had to be kept within an intermediate range of 50-150 mM in order to reduce association of non-PhaP proteins with the bead surface whilst maintaining binding of PhaP fusions with PHA beads (Banki *et al.* 2005), otherwise impurities of non-PhaP proteins, PhaP-intein-target or PhaP-intein would occur in the final soluble fraction of target protein. Moreover, this approach required that target proteins do not interact with PHA beads either before or after cleavage, otherwise they would associate with the preformed beads rather than staying in the soluble fraction under the mild salt elution conditions (Banki *et al.* 2005; Wang *et al.* 2008).

These problems were solved here by using PhaC that, upon polymerisation of PHA, was covalently bound to beads (see the summary in Table 3.7, section 3.1.2.3). PhaC maintained the attachment of PhaC-intein-target to the beads regardless of salt concentrations, therefore eliminating any concerns about non-elution of a bead associated

target protein, or detachment of PhaC-intein-target or PhaC-intein under various washing / elution conditions that could potentially contaminate the target protein. Moreover, as exemplified by PhaC-intein-GFP beads, multiple shorter cleavage reactions (2 rounds of cleavage reaction at pH 6, 16 h each) were feasible in this study (Figure 3.9, section 3.1.2.1) to favour completion of intein cleavage while minimizing the risk of degradation or activity loss of the target protein, which was advantageous as compared to one extended cleavage reaction over 24 h or even 36 h (Zhang *et al.* 2010; Zhou *et al.* 2011).

It is known that protein secondary structures tend to be disrupted at an acidic pH, though certain proteins remain stable at pH 6, as was the case of the proteins tested in the current study. For example, wild type GFP is known to be fluorescently stable over a broad pH range from 6 to 10 (Patterson *et al.* 1997). Also, for the solution structure of Rv1626, a putative transcriptional antiterminator from *Mycobacterium tuberculosis* and a potential vaccine candidate against this bacterium (Morth *et al.* 2004; Rubio Reyes *et al.* 2016), no change in X-ray scattering curves has been observed even when the pH dropped from 8 to 4, suggesting a structure stability within this pH range (Morth *et al.* 2004). Moreover, a backbone amide hydrogen/deuterium exchange rate study on the Z domain of staphylococcal protein A revealed that amide protons of all its three helices are protected from rapid exchange at both pH 6.5 and 4.4 demonstrating the intact structure of Z domain at acidic pH (Tashiro *et al.* 1997). Recently researchers have developed a strategy of cation exchange chromatography at pH 6.0 for successful purification of recombinant human tumour necrosis factor alpha (TNFα), and they found via circular dichroism (CD) analysis that the secondary structure of TNFα was perturbed only when pretreated below pH 5.0 (Zhang *et al.* 2014). Furthermore, a study on pH dependence of structural stability of human granulocyte colony-stimulating factor (G-CSF) disclosed that G-CSF displayed similar secondary helical content across a pH range of 4 through 7 via CD analysis (Ricci

*et al.* 2003). In addition, there is evidence that the highly helical secondary structure of human interferon alpha-2b (IFNα2b) is very conserved over a broad pH range from 2 to 10 (Beldarraín *et al.* 2001). In agreement with these findings, in this study, even though all the proteins tested remained in an environment of pH 6 for 16 h, they appeared to be quite stable in terms of functionality / proper folding.

Premature cleavage was really a challenge when working with the pH inducible *Ssp* DnaB intein, as was obvious with previous studies in which this intein system was used (Wood 2003), and this inevitably impacted the final yield of target proteins. *Ssp* DnaB intein is activated at pH 6-7 at 25°C (New England Biolabs), while the *E. coli* cytosolic pH is documented to drop below 7 due to media acidification and acetate production during shake-flask cultivation (Losen *et al.* 2004). In order to mitigate this pH drop i.e. minimize the intracellular premature cleavage of the target protein, a pH buffered Terrific broth (pH 8.6 with 25mM HEPES) was established as PHA bead production media using a 24 h incubation time. Besides, buffers with high pH (pH 8.8) were used during cell disruption and bead isolation in order to reduce extracellular premature cleavage of the target proteins. In addition, bead production after IPTG induction was carried out at 22°C rather than 25°C described elsewhere (Jahns & Rehm 2009; Hay *et al.* 2014; Rubio Reyes *et al.* 2016) to avoid activation of the *Ssp* DnaB intein. This avoided undesirable premature cleavage to a level within about 30-40% as measured via densitometry analysis (Figure 3.7, section 3.1.2.1, and Figure 3.12, section 3.1.2.3). To our knowledge, this is the first detailed effort in controlling the premature intein cleavage at shake-flask level. Bioreactors enabling accurate real-time pH and temperature control might further reduce intracellular premature cleavage.

However, on the other hand, it has to be admitted that both the pH buffered alkaline medium and the relatively shorter growth period were not ideal for the growth of *E. coli* cells or accumulation of PHA beads, which might have contributed to the relatively low final protein yield, with microgram levels per litre of cell culture (Table 3.3, section 3.1.2.1, and Table 3.6, section 3.1.2.3), as compared to a milligram/L level of protein yield commonly achieved with intein tags without medium buffering (Shi *et al.* 2013).

## 4.2 Design of PHA beads as affinity resins for molecular recognition by immobilising OBody ligands on bead surface

There are often industrial needs to separate or detect biologically important molecules from natural sources, like for instance commercial lysozyme separation from hen egg white in the food and pharmaceutical industry (Shahmohammadi 2017), and progesterone (P4) detection in the dairy industry (Jang et al. 2017). So far affinity separation and affinity detection techniques exploiting the biorecognition between a biomolecule and its ligand (or binding partner) are commonly regarded as the most efficient techniques. These technologies are mostly exemplified by affinity resins with immobilised antibodies that are capable of recognising antigenic epitopes of the molecule of interest (Tozzi *et al.* 2003; Crivianu-Gaita & Thompson 2016). Manufacture of this type of affinity resin generally requires three main processes: (1) the preparation of a support matrix (either through chemical treatment of existing natural resources such as agarose, dextrose and cellulose, or through *de novo* chemical synthesis such as polyacrylamide and polymethacrylate derivatives) (Vařilová *et al.* 2006), (2) the recombinant production and purification of an antibody or antibody fragment which is usually costly (Dias & Roque 2017), and (3) chemical cross-linking of the antibody or antibody fragment to the support matrix to avoid random orientation or denaturation of the antibody or antibody fragment,

which can be a complex process (Shen *et al.* 2017). Therefore, developing alternative non-antibody affinity ligands or resins that enable simplified ligand immobilisation are necessary.

This study addressed these problems by generating affinity resins *in vivo* in a single step through recombinant production of OBody ligands covalently attached to a PHA support matrix. The OBody ligands were produced simultaneously with the PHA bead support, and immobilised on the bead surface in functional orientations, thus eliminating any extra effort for chemical cross-linking. Furthermore, OBody ligands examined in this study are small (~13 kDa) with non-antibody protein scaffolds that do not require any S-S bonds for self-stability or affinity activity. These features, in contrast to antibodies or antibody fragments, contribute to a simpler production process and lower overall production cost.

### 4.2.1 Design of PHA beads as affinity resins for lysozyme separation

Here it was demonstrated that OBody beads made recombinantly in which the OBody ligand L200EP-06 (O6) was immobilised on PHA beads, retained lysozyme recognition functionality and thus could be used as affinity resins for lysozyme separation. The resulting OBody beads (both O6-PhaC and PhaC-O6 beads) showed encouraging purification power from a complex substrate consisting of BSA, skimmed milk and lysozyme (Figure 3.15, section 3.2.1.2), with significantly higher lysozyme binding capacities as compared to control WT PHA beads (Table 3.8, section 3.2.1.2).

It was also observed that O6-PhaC beads performed significantly better than PhaC-O6 beads in terms of both purity and yield for the lysozyme purified from the complex substrate (Figure 3.15, Table 3.8, section 3.2.1.2). This was not uncommon when using

PhaC fusion strategy for protein immobilisation, as some proteins do tolerate fusion to one terminus better than the other as found previously (Hay *et al.* 2014). It is well known that in designing a protein fusion strategy, the orientation and distance of the fusion partners as well as linker region choice can all impact on the performance of the final fusion protein (Yu *et al.* 2015). Further structural analysis of O6, O6-PhaC and PhaC-O6 might shed light on the molecular mechanism behind this.

### 4.2.2 Design of PHA beads as affinity resins for progesterone detection

Here it was demonstrated that OBody beads made by immobilisation of either the 1st generation of the OBody progesterone-binding ligand P4013-D7 (D7), or the 2nd generation ligand B7 on PHA beads, retained their respective progesterone (P4) recognition functionality (Table 3.9, section 3.2.2.2; Figure 3.17-19, section 3.2.2.3; and Figure 3.21-22, section 3.2.2.4). Of the six prototypes of OBody beads (PhaC-D7, 3xD7-PhaC, 3xD7-PhaC-D7, D7-PhaC, 3xB7-PhaC and B7-PhaC beads), the most promising B7-PhaC beads under the current assay conditions, without much optimisation, had a P4 binding affinity of about 0.7 μM (Figure 3.21, section 3.2.2.4), which is close to the soluble B7 counterpart that has an affinity of less than 0.3 μM when analysed with a similar biotin labelled P4 based binding assay, according to personal communication with the collaborator.

It was noticed that three types of the 1st generation D7 beads (PhaC-D7, 3xD7-PhaC and 3xD7-PhaC-D7 beads) performed similarly in P4 binding, regardless of the fusion copy and/or orientation of D7 (Table 3.9, section 3.2.2.2). Further P4 binding analysis with both generations of beads with a single or triplicate copy N-terminal orientation of D7 or B7 ligand (namely D7-PhaC, B7-PhaC, 3xD7-PhaC and 3xB7-PhaC beads) showed that

the single copy N-terminal orientation was preferred, in terms of P4 binding level, over triplicate copies (Figure 3.21, section 3.2.2.4). Further structural analysis of these ligands and the respective fusion proteins might help to elucidate why a single copy N-terminal orientation showed better P4 binding. As the N-terminal orientation preference was similar to what was observed for the lysozyme-biding OBody ligand L200EP-06 (O6) discussed in section 4.2.1, further structural studies might also help to explain whether it is merely a coincidence or a universal prerequisite for optimum functionality of all OBody ligands.

Previously it has been reported that lipophilic P4 tends to attach (or absorb) to polymer materials such as polystyrene plasticware and polyester membranes via non-specific hydrophobic interactions (Longman & Buehring 1986; Schäfer *et al.* 2011). Consistent with those observations, high background P4 binding noise brought by the PHA polymer materials and/or high-binding ELISA plates used in this study was obvious (Table 3.9, section 3.2.2.2; Figure 3.17, section 3.2.2.3; and Figure 3.21, section 3.2.2.4). Preliminary efforts that involved using approximately equivalent PHA bead biomasses for the different samples, as well as both 1% BSA and 5% skim milk as blocking agents throughout the assay, showed a promising counteracting effect on the non-specific P4 absorption noise (Figure 3.22, section 3.2.2.4).

In future, changing to the low-binding ELISA plates (Greiner Bio-One #655101, Frickenhausen, Germany) and / or optimising blocking conditions could be examined to try to eliminate non-specific P4 absorption. A study in 2015 suggested that since BSA only binds weakly to different ELISA plates tested (including polypropylene, polystyrene and polycarbonates) and tends to be washed away easily, BSA blocking is not necessary for ELISA assays as long as PBST (PBS containing 1% Tween 20) washing is performed

(Ahirwar *et al.* 2015). The blocking condition used in that study was 2% BSA in a neutral 10 mM PBS buffer (0.85% NaCl, pH 7.2) (Ahirwar *et al.* 2015). In a more recent paper, it was demonstrated that by adjusting the pH of blocking buffer to the isoelectric point of BSA (pH 4.6), a 2% BSA solution in a 0.1 M sodium citrate/citric acid buffer showed the best surface blocking performance on an optical-fiber biosenseor (Wang *et al.* 2017). Therefore it would be interesting to test these new methods in the current P4 binding ELISA assay to improve non-specific blocking performance. In addition, it would be worthwhile to test other regular protein blockers (such as casein and ovalbumin proteins) along with nonionic detergent blockers (such as Tween 20 and Triton X-100) using different combinations and concentrations.

## 4.3 General Conclusions

### 4.3.1 PHA beads as self-cleavable protein purification resins

New self-cleavable protein purification resins based on PHA beads were developed in this study. It was shown that a target protein fused to PhaC via a self-cleavable linker tag mediates *in vivo* production of PHA beads displaying the target protein. Functional target protein could be obtained at high purity from isolated PHA beads by incubation with $CaCl_2$ and triglycine (in the case of the self-cleavable sortase tag) or by a pH shift to 6 (in the case of the self-cleavable intein tag). Here the target protein was firstly produced as immobilized to the surface of PHA beads *in vivo*, then separated from contaminating host proteins via simple bead isolation / washing steps and finally purified by specific release into the soluble fraction.

This process requires neither expensive protein purification resins nor toxic chemicals or additional costly enzymes, and promises to serve as an economic and simplified platform for protein production and purification. The current platform was utilized for production and purification of the medically important proteins TNFα, IFNα2b and G-CSF, hence provides a promising approach to lower production costs of therapeutic proteins. PHA production itself has already been established as a commercially scalable process (Rehm 2010). There have been extensive studies devoted to bioprocessing strategies for large-scale PHA production (Kaur & Roy 2015), and globally there were 24 companies commercialising PHA products (mainly PHB (poly-*β*-hydroxybutyrate)) in the year 2009 (Chen 2009). Therefore this study provided the foundation for scalable and industrial PHA bead-based protein production.

### 4.3.2 PHA beads as affinity resins

Overall, in this study it was demonstrated that *in vivo* one-step production of PHA affinity resins is enabled by genetically fusing PhaC (PHA synthase) to differently customised OBody ligands. Furthermore, resulting recombinant OBody beads with appropriate ligands were used to achieve lysozyme separation from a complex substrate, or progesterone (P4) binding. Further optimisation of the P4 binding assay is necessary before the OBody bead system can be used for P4 detection in bovine milk. However, recombinant immobilisation of OBody ligands on the surface of PHA beads expands not only the attractiveness of these emerging OBody scaffolds, but also the utility scope of PHA beads as affinity resins.

# Chapter 5: Future work

## 5.1 PHA beads as self-cleavable protein purification resins

As discussed in sections 4.1.1 and 4.1.2, the main challenge in our PHA bead based self-cleavable protein production / purification platforms was the unwanted premature self-cleavage of either sortase or intein, as previously reported by others (Wood 2003; Mao 2004; Bellucci *et al.* 2013).

### 5.1.1 Future measures to control premature cleavage of sortase

Several measures could be taken to control premature cleavage of sortase, such as modification of the LPXTG signal sequence, adjustment of sortase position in the tripartite fusions (PhaC, sortase, target protein), optimisation of linker region design, and production of target protein independently of sortase, as described below.

Firstly, an alternative LPXTG (X represents any amino acid) signal sequence recognised by sortase could be conceived. LPETG was chosen in the current study merely because it exists in native sortase A (SrtA) substrates (such as *S. aureus* protein A) (Kruger *et al.* 2004) and is commonly used in sortase mediated recombinant protein purification (Mao 2004; Matsunaga *et al.* 2010; Bellucci *et al.* 2013). Previous studies of amino acid substitutions in the LPXTG signal showed that six alternatives (LPM/Y/L/F/Q/N/FTG) actually were preferred by SrtA over LPETG. Although amino acid substitutions at other positions in this signal generally lowered the SrtA reaction rate, some variants such as MPETG, LPEAG and IPKTG were quite well tolerated by SrtA (Kruger *et al.* 2004;

Piotukh *et al.* 2011). Further, LPGAG was found to significantly reduce premature self-cleavage of SrtA in a SrtA-ELP-LPGAG-target protein fusion effort for protein purification (Bellucci *et al.* 2013). Therefore, it would be worthwhile to try the LPGAG signal sequence or a differently tailored one in the current sortase mediated self-cleavable PHA beads based purification platform as a measure to control the premature cleavage of sortase.

Secondly, optimisation of sortase orientation in the tripartite fusions (PhaC, sortase, target protein) as well as the linker region design might be helpful in achieving more stringent calcium control. Previous NMR and crystal structure models revealed that the sortase A (SrtA) catalytic domain (residues 60-206) contains a closed eight-strand β-barrel fold with a disordered β6/β7 loop that is involved in the binding of both LPXTG and $Ca^{2+}$, but the LPXTG binding pocket and $Ca^{2+}$ binding pocket are located on opposing faces of SrtA (Ilangovan *et al.* 2001; Zong *et al.* 2004; Naik *et al.* 2006; Suree *et al.* 2009). Further studies demonstrated that $Ca^{2+}$ binding allosterically activates and stabilises SrtA by inducing partial closure and ordering of the β6/β7 loop, leading to preorganisation of the LPXTG binding pocket, which in turn upon contacting with the LPXTG substrate becomes fully organised; but not the other way aroud - LPXTG binding alone can not lead to conformation change or stabilisation (Naik *et al.* 2006; Kappel *et al.* 2012; Moritsugu *et al.* 2012; Pang & Zhou 2015). Therefore, it would be interesting to study whether the current fusion context of the PhaC-sortase-LPETG-target protein caused less dependency on $Ca^{2+}$ binding towards conformation transition to an active state, as compared to the stringent $Ca^{2+}$ dependency of a native sortase. Positioning sortase alternatively, such as in the form of sortase-PhaC-LPETG-target protein, and / or introducing differently designed linker regions between each of the fusion partners, might affect the SrtA conformation and reduce the premature cleavage.

Another way to reduce premature cleavage might be by producing the target protein and sortase on separate beads and separating the protein production and cleavage steps. In this way the concern of either intracellular or extracellular $Ca^{2+}$ concentration or premature cleavage of sortase would be completely abolished. In particular, a target protein could be designed as a PhaC-LPXTG-target protein fusion in the absence of sortase, which would lead to the production of PHA beads displaying PhaC-LPXTG-target protein. Then, the isolated PHA beads displaying PhaC-LPXTG-target protein could be incubated with the PhaC-sortase PHA beads described in 3.1.1.1 in the presence of $CaCl_2$ and triglycine so that only the G-tagged target protein is cleaved and released as a soluble fraction. An obvious disadvantage associated with this approach would be the two separate PHA bead production processes vs. one bead production process in the current study. Further, the optimum mixing ratio between the two beads (PhaC-LPXTG-target beads and PhaC-sortase beads) to give the best target protein yield might need to be worked out. These extra production / testing efforts in this approach might seem complicated and less economic, but might be worthwhile if significant yield improvements could be gained.

### 5.1.2 Future measures to control premature cleavage of intein

In terms of the pH-cleavable intein approach, in addition to use of an alternative production host other than *E. coli* that could tolerate a higher pH than *E. coli*, the use of split inteins or a different type of intein that did not rely on pH shifts for stringent control could be considered.

To overcome the limited growth of *E. coli* cells on a pH buffered medium, an alternative host might be considered, for example, the endotoxin free GRAS (Generally Recognized

As Safe) grade *Bacillus subtilis* (Vavrová *et al.* 2010) could probably adapt better to alkaline pH (Krulwich *et al.* 1994; Wiegert *et al.* 2001; Padan *et al.* 2005) and could be engineered to produce PHA (Law *et al.* 2003; Wang *et al.* 2006; Singh *et al.* 2009; Lin & Chen 2017). By using engineered *B. subtilis* in combination with an even higher pH buffered medium, it is likely that a simultaneous reduction in premature cleavage and yield increase in target proteins could be obtained.

Another option to be considered is split inteins. Split-inteins are inteins that are either naturally or artificially separated into two individually inactive segments which undergo *trans*-splicing only upon reassembly of the two segments (Li 2015). Split inteins can also be engineered to deliver cleavage only at their N- or C-termini. Therefore by separately producing the inactive segments (usually one fused with affinity tag, while the other fused with target protein), cleavage could be controlled and induced only upon reassembly of the two segments, thereby eliminating the problem of premature cleavage (Lahiry *et al.* 2017). So far, natural or artificial split inteins such as *Ssp* split-inteins based on *Synechocystis* sp. PCC6803 DnaE (*Ssp* DnaE) and *Mtu* ΔI-CM split inteins based on *Mycobacterium tuberculosis* recA (*Mtu* RecA) have been used for protein purification via either spontaneous *in vivo* / *in vitro* trans-splicing or *in vitro* pH-controllable N / C-terminal cleavage (Evans *et al.* 2000; Miao *et al.* 2005; Lu *et al.* 2011; Shi *et al.* 2013). A newly published patent application disclosed *Npu* split intein pairs based on *Nostoc punctiforme* DnaE (*Npu* DnaE) showing pH sensitive cleavage without premature cleavage (Wood & Shi 2016). Combining any of those split intein pairs with the current PHA beads based protein purification platform might prove promising in solving the premature cleavage issues.

An example of one pair of artificially engineered *Ssp* DnaE split inteins comprised a N-segment of 106 residues and a C-segment of 48 residues. Reassembly of Ala1-DnaE$^{N}$ (amino acid change at position 1 from original Cysteine to Alanine) and DnaE$^{C}$ leads to pH-controllable C-terminal cleavage, whereas reassembly of DnaE$^{N}$ and Ala154-DnaE$^{C}$ (amino acid change at position 154 from original Asparagine to Alanine) leads to pH-controllable N-terminal cleavage (Lu *et al.* 2011). These could be incorporated into the current platform, by designing respective fusions with PhaC and/or target protein in the format of PhaC-Ala1-intein$^{N}$ and PhaC-intein$^{C}$-target protein (or target protein-intein$^{N}$-PhaC and PhaC-Ala154intein$^{C}$), thus two different PHA beads displaying respective fusion proteins could be separately produced and isolated, and then, upon mixing of those beads, a pH-controllable C- or N-terminal cleavage mediated via the split intein pairs would potentially release the target protein into the soluble fraction.

Alternatively, the spontaneous *in vitro* trans-splicing function of the naturally occurring *Ssp* DnaE split intein pairs (DnaE(N) of amino acids 5-123 and DnaE(C) of the C-terminal 36 residues) published by Evans *et al.* (Evans *et al.* 2000) could also be exploited in the current platform. Namely, by splitting a target protein in two halves (target protein(N) and target protein(C)) and fusing each with a corresponding split intein segment along with PhaC in the format of target protein(N)-DnaE(N)-PhaC and PhaC-DnaE(C)-target protein(C), two different PHA beads displaying respective fusion proteins could be separately produced and isolated, and then, upon mixing of those beads, spontaneous *in vitro* trans-splicing medicated via the split intein pairs would potentially yield an intact target protein in the soluble fraction. This would be of particular advantage in production and purification of proteins toxic to host cells.

Another potential mechanism to eliminate unwanted premature cleavage is by replacing the *Ssp* DnaB intein with a differently non-pH stringently controlled one, such as the salt-dependent MCM2 intein derived from *Halorhabdus utahensis* (*Hut* MCM2 intein) whose cleavage occurs in 4 M NaCl in the absence of reducing agents (Ciragan *et al.* 2016). Moreover, this *Hut* MCM2 intein was artificially split into two highly soluble segments (*Hut* $MCM2_{\Delta C62}$ and *Hut* $MCM2_{C42}$) and exhibited high rates of trans-splicing upon reassembly in the presence of 4 M NaCl (Ciragan *et al.* 2016). Therefore, it might be feasible to introduce either the intact continuous *Hut* MCM2 intein or the two split segments into the current platform to produce proteins in a salt controllable manner.

In addition, the above mentioned *Hut* MCM2 intein showed higher cleavage efficiency in the presence of additional reducing agents (Ciragan *et al.* 2016), which was desirable for increasing the cleavage yield of a target protein containing no thiol sensitive residues. Along with this line, when a target protein is not sensitive to thiols, the *Ssp* DnaB intein in the current platform could also be substituted with the salt sensitive *Hsa* PolII intein derived from *Halobacterium salinarum* (*Hsa* PolII intein) that is cleavable at > 1.5 M NaCl in the presence of reducing agents (Reitter *et al.* 2016). Alternatively, a more stringently controlled thiol-inducible intein might be considered, such as *Mxe* GyrA or *Mth* RIR1 intein as provided in the commercial pTWIN1 or pTWIN 2 system (NEB).

## 5.2 PHA beads as affinity resins

As demonstrated in Figure 3.21, section 3.2.2.4, the best-performing progesterone (P4) recognising OBody beads, B7-PhaC beads, showed promising P4 binding affinity and capacity. As discussed in section 4.2.2, by using low-binding ELISA plates and / or choosing an optimum blocking condition, the high background non-specific P4 binding

noise brought by the PHA polymer materials and/or high-binding ELISA plates might be potentially solved.

In this section, some future work to practically use the B7-PhaC beads in bovine milk P4 detection is outlined.

## 5.2.1 Future bovine milk progesterone detection assay using OBody beads

### 5.2.1.1 Milk spike test

Briefly, a fixed amount of biotin-labelled progesterone (bio-P4) would be spiked into a real bovine milk sample before performing a bead P4 binding assay similarly to the protocol described in section 2.6.4.6. The progesterone (P4) naturally present in the milk sample would compete with the bio-P4 to bind the binding sites on OBody beads, and thus the colour intensity upon addition of OPD (*o*-Phenylenediamine dihydrochloride) reagent (section 2.6.4.6) would be proportional to the amount of bio-P4 bound to the beads, which is inversely proportional to the amount of natural P4 present in the milk sample.

To ensure detection sensitivity and accuracy, optimisation would be required to find the right amount of bio-P4 to be spiked, as well as the right bead biomass to be used for the assay.

#### 5.2.1.2 Strip test

Different to the ELISA format assay outlined in section 5.2.1.1, the strip test is a lateral flow format of analysis, but still based on competitive binding. It is envisaged that a fixed amount of BSA conjugated progesterone (BSA-P4) would be immobilised on the top "test line" of a nitrocellulose membrane, and the bottom of the resulting strip would be dipped in to a mixture suspension composed of bovine milk sample and a fixed amount of OBody beads (previously stained with Nile Red or recombinantly tagged with GFP protein). Thus natural P4 present in the milk sample would occupy the binding sites on OBody beads, leaving only limited empty binding sites on the OBody beads. Hence the OBody beads, upon migrating along to the top "test line" comprising BSA-P4, would only stay there (by binding to the BSA-P4) in amounts proportional to the empty binding sites left (and inversely proportionally to the amount of natural P4 present in the milk sample). Therefore, by quantification of the bead fluorescence on the test line, a measure of bovine milk P4 level could be revealed.

Again, to ensure bead migration along the strip as well as detection sensitivity and accuracy, optimisation would be required to find the right bead size or dilution condition, as well as the best Nile Red staining condition (or recombinant GFP tagging condition).

## 5.3 Overall summary

In summary, both aims of this study were achieved. On one hand, by fusing a target protein to PhaC via a self-cleavable linker tag (either sortase or intein), the target protein was firstly produced as immobilized to the surface of PHA beads *in vivo*, then separated from contaminating host proteins via simple bead isolation / washing steps and finally

purified by specific release into the soluble fraction, either by incubation with $CaCl_2$ ± triglycine, or by a pH shift to 6. In this way, the first aim was fulfilled, namely to provde a streamlined process with less complicated steps toward the production and purification of recombinant proteins, especially therapeutic proteins,. On the other hand, by fusing PhaC (PHA synthase) to differently customised OBody ligands, functional OBody beads could be obtained and used for lysozyme separation from a complex substrate, or for progesterone (P4) detection. Therefore, the second aim was fulfilled, namely to develop a simplified process for preparation of affinity resins with non-antibody ligands that could be used for separation and detection of industrially important biomolecules.

Nevertheless, future work to overcome the premature self-cleavage problem could include for example optimising the LPXTG signal sequence and sortase orientation in the tripartite fusions (PhaC, sortase, target protein) and testing different linkers between the two fusion partners in the sortase approach, or choosing non-pH stringently controlled inteins in the intein approach. An alternative option is to separate the protein production and cleavage steps, by producing the target protein and sortase on separate beads in the sortase approach, or producing the inactive segments of a split intein (one fused with the target protein while the other not) on separate beads in the intein approach, before mixing counterpart beads to induce a controlled self-cleavage.

With the knowledge gained in this study on production and purification of recombinant proteins (including medically important therapeutic proteins) and with plasmid contructs made as part of this study, it would be straightforward to develop further constructs that incorporate the ideas outlined above. Transformation of these constructs into *E. coli* or other appropriate hosts would allow for assessment of bead protein production and tag cleavage for the purpose of target protein purification.

In addition, the ‘ready to use’ B7-PhaC beads (the best-performing P4-recognising OBody beads), as well as knowledge gained in the biotin-labelled progesterone (bio-P4) based P4 binding assay, pave the way for future bovine milk progesterone (P4) detection methods that could easily be applied in industrial situations, such as by milk spike tests or strip tests.

# Chapter 6: References


Abe T, Kobayashi T & Saito T (2005) Properties of a novel intracellular poly(3-hydroxybutyrate) depolymerase with high specific activity (PhaZd) in *Wautersia eutropha* H16. *Journal of Bacteriology* **187**:6982-6990.

Abeyrathne EDNS, Lee HY & Ahn DU (2013) Egg white proteins and their potential use in food processing or as nutraceutical and pharmaceutical agents-A review. *Poultry Science* **92**:3292-3299.

Ahirwar R, Bariar S, Balakrishnan A & Nahar P (2015) BSA blocking in enzyme-linked immunosorbent assays is a non-mandatory step: a perspective study on mechanism of BSA blocking in common ELISA protocols. *RSC Advances* **5**:100077-100083.

Albuquerque PBS & Malafaia CB (2018) Perspectives on the production, structural characteristics and potential applications of bioplastics derived from polyhydroxyalkanoates. *International Journal of Biological Macromolecules* **107**:615-625.

Ali I & Jamil N (2016) Polyhydroxyalkanoates: current applications in the medical field. *Frontiers in biology* **11**:19-27.

Alizadeh AA, Hamzeh-Mivehroud M, Farajzadeh M, Moosavi-Movahedi AA & Dastmalchi S (2015) A simple and rapid method for expression and purification of functional TNF-α using GST fusion system. *Current Pharmaceutical Biotechnology* **16**:707-715.

Altintaş EB & Denizli A (2006) Monosize poly(glycidyl methacrylate) beads for dye-affinity purification of lysozyme. *International Journal of Biological Macromolecules* **38**:99-106.

Amara AA & Rehm BHA (2003) Replacement of the catalytic nucleophile cysteine-296 by serine in class II polyhydroxyalkanoate synthase from *Pseudomonas aeruginosa*-mediated synthesis of a new polyester: identification of catalytic residues. *Biochem J* **374**:413-421.

Arcus V (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Current Opinion in Structural Biology* **12**:794-801.

Arica MY, Yilmaz M, Yalçin E & Bayramoğlu G (2004) Affinity membrane chromatography: relationship of dye-ligand type to surface polarity and their effect on lysozyme separation and purification. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **805**:315-323.

Arikawa H, Sato S, Fujiki T & Matsumoto K (2016) A study on the relation between poly(3-hydroxybutyrate) depolymerases or oligomer hydrolases and molecular weight of polyhydroxyalkanoates accumulating in *Cupriavidus necator* H16. *Journal of Biotechnology* **227**:94-102.

Asenjo JA & Andrews BA (2009) Protein purification using chromatography: selection of type, modelling and optimization of operating conditions. *Journal of Molecular Recognition* **22**:65-76.

Atwood JA & Rehm BHA (2009) Protein engineering towards biotechnological production of bifunctional polyester beads. *Biotechnology Letters* **31**:131-137.

Bäckström BT, Brockelbank JA & Rehm BHA (2007) Recombinant *Escherichia coli* produces tailor-made biopolyester granules for applications in fluorescence activated cell sorting: functional display of the mouse interleukin-2 and myelin oligodendrocyte glycoprotein. *BMC Biotechnology* **7**

Balakrishna Pillai A & Kumarapillai HK (2017) Bacterial polyhydroxyalkanoates: recent trends in production and applications. In: Shukla P (ed) *Recent advances in Applied Microbiology*. Springer, Singapore, pp 19-53.

Banki MR, Gerngross TU & Wood DW (2005) Novel and economical purification of recombinant proteins: intein-mediated protein purification using *in vivo* polyhydroxybutyrate (PHB) matrix association. *Protein Science* **14**:1387-1395.

Barnard GC, McCool JD, Wood DW & Gerngross TU (2005) Integrated recombinant protein expression and purification platform based on *Ralstonia eutropha*. *Applied and Environmental Microbiology* **71**:5735-5742.

Baydemir G, Andaç M, Derazshamshir A, Uygun DA, Özçalişkan E, Akgöl S & Denizli A (2013) Synthesis and characterization of amino acid containing Cu(II) chelated nanoparticles for lysozyme adsorption. *Materials Science & Engineering C-Materials for Biological Applications* **33**:532-536.

Beeby M, Cho M, Stubbe J & Jensen GJ (2012) Growth and localization of polyhydroxybutyrate granules in *Ralstonia eutropha*. *Journal of Bacteriology* **194**:1092-1099.

Beldarraín A, Cruz Y, Cruz O, Navarro M & Gil M (2001) Purification and conformational properties of a human interferon α2b produced in *Escherichia coli*. *Biotechnology and Applied Biochemistry* **33**:173-182.

Bello-Gil D, Maestro B, Fonseca J, Dinjaski N, Prieto MA & Sanz JM (2017) Poly-3-hydroxybutyrate functionalization with BioF-tagged recombinant proteins. *Applied and Environmental Microbiology*

Bellucci JJ, Amiram M, Bhattacharyya J, McCafferty D & Chilkoti A (2013) Three-in-one chromatography-free purification, tag removal, and site-specific modification of recombinant fusion proteins using sortase A and elastin-like polypeptides. *Angewandte Chemie-International Edition* **52**:3703-3708.

Berry MJ, Davies J, Smith CG & Smith I (1991) Immobilization of Fv antibody fragments on porous silica and their utility in affinity chromatography. *Journal of Chromatography A* **587**:161-169.

Berry MJ & Pierce JJ (1993) Stability of immunoadsorbents comprising antibody fragments: comparison of Fv fragments and single-chain Fv fragments. *Journal of Chromatography A* **629**:161-168.

Bhubalan K, Rathi DN, Abe H, Iwata T & Sudesh K (2010) Improved synthesis of P(3HB-co-3HV-co-3HHx) terpolymers by mutant *Cupriavidus necator* using the PHA synthase gene of *Chromobacterium* sp USM2 with high affinity towards 3HV. *Polymer Degradation and Stability* **95**:1436-1442.

Blatchford PA, Scott C, French N & Rehm BHA (2012) Immobilization of organophosphohydrolase OpdA from *Agrobacterium radiobacter* by overproduction at the surface of polyester inclusions inside engineered *Escherichia coli*. *Biotechnology and Bioengineering* **109**:1101-1108.

Bohmert-Tatarev K, McAvoy S, Daughtry S, Peoples OP & Snell KD (2011) High levels of bioplastic are produced in fertile transplastomic tobacco plants engineered with a synthetic operon for the production of polyhydroxybutyrate. *Plant Physiology* **155**:1690-1708.

Bohmert K, Balbo I, Kopka J, Mittendorf V, Nawrath C, Poirier Y, Tischendorf G, Trethewey RN & Willmitzer L (2000) Transgenic *Arabidopsis* plants can accumulate polyhydroxybutyrate to up to 4% of their fresh weight. *Planta* **211**:841-845.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Analytical Biochemistry* **72**:248-254.Bresan S & Jendrossek D (2017) New insights into PhaM-PhaC-mediated localization of polyhydroxybutyrate granules in *Ralstonia eutropha* H16. *Applied and Environmental Microbiology* **83**:e00505-00517.

Bresan S, Sznajder A, Hauf W, Forchhammer K, Pfeiffer D & Jendrossek D (2016) Polyhydroxyalkanoate (PHA) granules have no phospholipids. *Scientific Reports* **6**:26612.

Breuer U, Terentiev Y, Kunze G & Babel W (2002) Yeasts as producers of polyhydroxyalkanoates: genetic engineering of *Saccharomyces cerevisiae*. *Macromolecular Bioscience* **2**:380-386.

Brockelbank JA, Peters V & Rehm BHA (2006) Recombinant *Escherichia coli* strain produces a ZZ domain displaying biopolyester granules suitable for immunoglobulin G purification. *Applied and Environmental Microbiology* **72**:7394-7397.

Byszewska-Szpocińska E & Markiewicz A (2006) New RIA kit for the determination of progesterone in cows' milk. *Journal of Immunoassay and Immunochemistry* **27**:279-288.

Carrió MM & Villaverde A (2005) Localization of chaperones DnaK and GroEL in bacterial inclusion bodies. *Journal of Bacteriology* **187**:3599-3601.

Cegielska-Radziejewska R, Lesnierowski G & Kijowski J (2008) Properties and application of egg white lysozyme and its modified preparations - a review. *Polish Journal of Food and Nutrition Sciences* **58**:5-10.

Chek MF, Kim SY, Mori T, Arsad H, Samian MR, Sudesh K & Hakoshima T (2017) Structure of polyhydroxyalkanoate (PHA) synthase PhaC from *Chromobacterium* sp USM2, producing biodegradable plastics. *Scientific Reports* **7**:5312.

Chen G-Q (2009) A microbial polyhydroxyalkanoates (PHA) based bio- and materials industry. *Chemical Society Reviews* **38**:2434-2446.

Chen G-Q, Hajnal I, Wu H, Lv L & Ye J (2015) Engineering biosynthesis mechanisms for diversifying polyhydroxyalkanoates. *Trends in Biotechnology* **33**:565-574.

Chen S-Y, Chien Y-W & Chao Y-P (2014) *In vivo* immobilization of D-hydantoinase in *Escherichia coli*. *Journal of Bioscience and Bioengineering* **118**:78-81.

Chen S, Parlane NA, Lee J, Wedlock DN, Buddle BM & Rehm BHA (2014) New skin test for detection of bovine tuberculosis on the basis of antigen displaying polyester inclusions produced by recombinant *Escherichia coli*. *Applied and Environmental Microbiology* **80**:2526-2535.

Cho M, Brigham CJ, Sinskey AJ & Stubbe J (2012) Purification of polyhydroxybutyrate synthase from its native organism, *Ralstonia eutropha*: implications for the initiation and elongation of polymer formation *in vivo*. *Biochemistry* **51**:2276-2288.

Ciragan A, Aranko AS, Tascon I & Iwaï H (2016) Salt-inducible protein splicing *in cis* and *trans* by inteins from extremely halophilic archaea as a novel protein-engineering tool. *Journal of Molecular Biology* **428**:4573-4588.

Clancy KW, Melvin JA & McCafferty DG (2010) Sortase transpeptidases: insights into mechanism, substrate specificity, and inhibition. *Biopolymers* **94**:385-396.

Crivianu-Gaita V & Thompson M (2016) Aptamers, antibody scFv, and antibody Fab' fragments: an overview and comparison of three of the most versatile biosensor biorecognition elements. *Biosensors and Bioelectronics* **85**:32-45.

Daems D, Lu J, Delport F, Mariën N, Orbie L, Aernouts B, Adriaens I, Huybrechts T, Saeys W, Spasic D & Lammertyn J (2017) Competitive inhibition assay for the detection of progesterone in dairy milk using a fiber optic SPR biosensor. *Analytica Chimica Acta* **950**:1-6.

Denoncin K & Collet J-F (2013) Disulfide bond formation in the bacterial periplasm: major achievements and challenges ahead. *Antioxidants & Redox Signaling* **19**:63-71.

Derazshamshir A, Ergün B, Peşint G & Odabaşi M (2008) Preparation of $Zn^{2+}$-chelated poly(HEMA-MAH) cryogel for affinity purification of chicken egg lysozyme. *Journal of Applied Polymer Science* **109**:2905-2913.

Dias AMGC & Roque ACA (2017) The future of protein scaffolds as affinity reagents for purification. *Biotechnology and Bioengineering* **114**:481-491.

Do BH, Ryu HB, Hoang P, Koo BK & Choe H (2014) Soluble prokaryotic overexpression and purification of bioactive human granulocyte colony-stimulating factor by maltose binding protein and protein disulfide isomerase. *PLoS One* **9**:e89906.

Dominguez DC (2004) Calcium signalling in bacteria. *Molecular Microbiology* **54**:291-297.

Draper J, Du J, Hooks DO, Lee JW, Parlane N & Rehm BHA (2013) Polyhydroxyalkanoate inclusions: polymer synthesis, self-assembly and display technology. In: Rehm BHA (ed) *Bionanotechnology biological self-assembly and its applications*. Caister Academic Press, Norfolk, pp 1-36.

Dunlop WF & Robards AW (1973) Ultrastructural study of poly-β-hydroxybutyrate granules from *Bacillus cereus*. *Journal of Bacteriology* **114**:1271-1280.

El Khoury G, Khogeer B, Chen C, Ng KT, Jacob SI & Lowe CR (2015) Bespoke affinity ligands for the purification of therapeutic proteins. *Pharmaceutical Bioprocessing* **3**:139-152.

Ergün B, Derazshamshir A & Odabaşı M (2007) Preparation of Fe(III)-chelated poly(HEMA-MAH) cryogel for lysozyme adsorption. *Hacettepe J Biol Chem* **35**:143–148.

Evans TC, Martin D, Kolly R, Panne D, Sun L, Ghosh I, Chen LX, Benner J, Liu XQ & Xu MQ (2000) Protein trans-splicing and cyclization by a naturally split intein from the *dnaE* gene of *Synechocystis species* PCC6803. *Journal of Biological Chemistry* **275**:9091-9094.

Fong RB, Ding ZL, Hoffman AS & Stayton PS (2002) Affinity separation using an Fv antibody fragment-"smart" polymer conjugate. *Biotechnology and Bioengineering* **79**:271-276.

Gao F, Zhang Q, Li X, Zhang Q, Mao T, Lu Y, Zhang W & Li H (2016) Comparison of standard addition and conventional isotope dilution mass spectrometry for the quantification of endogenous progesterone in milk. *Accreditation and Quality Assurance* **21**:395-401.

Geng Y, Wang S & Qi Q (2010) Expression of active recombinant human tissue-type plasminogen activator by using *in vivo* polyhydroxybutyrate granule display. *Applied and Environmental Microbiology* **76**:7226-7230.

Gerngross TU & Martin DP (1995) Enzyme-catalyzed synthesis of poly [(*R*)-(-)-3-hydroxybutyrate]: formation of macroscopic granules *in vitro*. *Proceedings of the National Academy of Sciences of the United States of America* **92**:6279-6283.

Gerngross TU, Reilly P, Stubbe J, Sinskey AJ & Peoples OP (1993) Immunocytochemical analysis of poly-beta-hydroxybutyrate (PHB) synthase in *Alcaligenes eutrophus* H16: localization of the synthase enzyme at the surface of PHB granules. *Journal of Bacteriology* **175**:5289-5293.

Gerngross TU, Snell KD, Peoples OP, Sinskey AJ, Csuhai E, Masamune S & Stubbe J (1994) Overexpression and purification of the soluble polyhydroxyalkanoate synthase from

*Alcaligenes eutrophus*: evidence for a required posttranslational modification for catalytic activity. *Biochemistry* **33**:9311-9320.

Ghatnekar MS, Pai JS & Ganesh M (2002) Production and recovery of poly-3-hydroxybutyrate from *Methylobacterium* sp V49. *Journal of Chemical Technology and Biotechnology* **77**:444-448.

Goyon A, Cai JZ, Kraehenbuehl K, Hartmann C, Shao B & Mottier P (2016) Determination of steroid hormones in bovine milk by LC-MS/MS and their levels in Swiss Holstein cow milk. *Food Additives and Contaminants Part A - Chemistry Analysis Control Exposure & Risk Assessment* **33**:804-816.

Grage K, Jahns AC, Parlane N, Palanisamy R, Rasiah IA, Atwood JA & Rehm BHA (2009) Bacterial polyhydroxyalkanoate granules: biogenesis, structure, and potential use as nano-/micro-beads in biotechnological and biomedical applications. *Biomacromolecules* **10**:660-669.

Grage K, Peters V & Rehm BHA (2011) Recombinant protein production by *in vivo* polymer inclusion display. *Applied and Environmental Microbiology* **77**:2706-2709.

Grage K & Rehm BHA (2008) *In vivo* production of scFv-Displaying biopolymer beads using a self-assembly-promoting fusion partner. *Bioconjugate chemistry* **19**:254-262.

Grönwall C & Ståhl S (2009) Engineered affinity proteins- generation and applications. *Journal of Biotechnology* **140**:254-269.

Guo W, Duan J, Geng W, Feng J, Wang S & Song C (2013) Comparison of medium-chain-length polyhydroxyalkanoates synthases from *Pseudomonas mendocina* NK-01 with the same substrate specificity. *Microbiological Research* **168**:231-237.

Hafizi A, Malboobi MA, Jalali-Javaran M, Maliga P & Alizadeh H (2017) Covalent-display of an active chimeric-recombinant tissue plasminogen activator on polyhydroxybutyrate granules surface. *Biotechnology Letters* **39**:1683-1688.

Hahn R, Podgornik A, Merhar M, Schallaun E & Jungbauer A (2001) Affinity monoliths generated by in situ polymerization of the ligand. *Analytical Chemistry* **73**:5126-5132.

Han J, Hou J, Liu H, Cai S, Feng B, Zhou J & Xiang H (2010) Wide distribution among halophilic archaea of a novel polyhydroxyalkanoate synthase subtype with homology to bacterial type III synthases. *Applied and Environmental Microbiology* **76**:7811-7819.

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* **166:**557-580.

Hay ID, Hooks DO & Rehm BHA (2014) Use of bacterial polyhydroxyalkanoates in protein display technologies. In: McGenity TJ (ed) *Hydrocarbon and Lipid Microbiology Protocols*. Springer, Berlin, pp 71-86.

Hoffmann N & Rehm BHA (2004) Regulation of polyhydroxyalkanoate biosynthesis in *Pseudomonas putida* and *Pseudomonas aeruginosa*. *Fems Microbiology Letters* **237**:1-7.

Hooks DO, Blatchford PA & Rehm BH (2013) Bioengineering of bacterial polymer inclusions catalyzing the synthesis of *N*-acetylneuraminic acid. *Applied and Environmental Microbiology* **79**:3116-3121.

Hooks DO, Venning-Slater M, Du J & Rehm BHA (2014) Polyhydroyxalkanoate synthase fusions as a strategy for oriented enzyme immobilisation. *Molecules* **19**:8629-8643.

Ilangovan U, Ton-That H, Iwahara J, Schneewind O & Clubb RT (2001) Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proceedings of the National Academy of Sciences* **98**:6056-6061.

Jaeger KE, Steinbüchel A & Jendrossek D (1995) Substrate specificities of bacterial polyhydroxyalkanoate depolymerases and lipases: bacterial lipases hydrolyze poly(ω-hydroxyalkanoates). *Applied and Environmental Microbiology* **61**:3113-3118.

Jahns AC, Haverkamp RG & Rehm BH (2008) Multifunctional inorganic-binding beads self-assembled inside engineered bacteria. *Bioconjugate chemistry* **19**:2072-2080.

Jahns AC, Maspolim Y, Chen S, Guthrie JM, Blackwell LF & Rehm BHA (2013) *In vivo* self-assembly of fluorescent protein microparticles displaying specific binding domains. *Bioconjugate chemistry* **24**:1314-1323.

Jahns AC & Rehm BH (2009) Tolerance of the Ralstonia eutropha class I polyhydroxyalkanoate synthase for translational fusions to its C terminus reveals a new mode of functional display. *Applied and Environmental Microbiology* **75**:5461-5466.

Jahns AC & Rehm BHA (2009) Tolerance of the *Ralstonia eutropha* class I polyhydroxyalkanoate synthase for translational fusions to its C terminus reveals a new mode of functional display. *Applied and Environmental Microbiology* **75**:5461-5466.

Jahns AC & Rehm BHA (2015) Immobilization of active lipase B from *Candida antarctica* on the surface of polyhydroxyalkanoate inclusions. *Biotechnology Letters* **37**:831-835.

Jang H, Ahmed SR & Neethirajan S (2017) GryphSens: a smartphone-based portable diagnostic reader for the rapid detection of progesterone in milk. *Sensors* **17**:1079.

Jendrossek D (2005) Fluorescence microscopical investigation of poly(3-hydroxybutyrate) granule formation in bacteria. *Biomacromolecules* **6**:598-603.

Jendrossek D & Handrick R (2002) Microbial degradation of polyhydroxyalkanoates. *Annual Review of Microbiology* **56**:403-432.

Jendrossek D & Pfeiffer D (2014) New insights in the formation of polyhydroxyalkanoate granules (carbonosomes) and novel functions of poly(3-hydroxybutyrate). *Environmental Microbiology* **16**:2357-2373.

Jendrossek D, Selchow O & Hoppert M (2007) Poly(3-hydroxybutyrate) granules at the early stages of formation are localized close to the cytoplasmic membrane in *Caryophanon latum*. *Applied and Environmental Microbiology* **73**:586-593.

Jensen TE & Sicko LM (1971) Fine structure of poly-β-hydroxybutyric acid granules in a blue-green alga, *Chlorogloea fritschii*. *Journal of Bacteriology* **106**:683-686.

Jia K, Cao R, Hua DH & Li P (2016) Study of class I and class III polyhydroxyalkanoate (PHA) synthases with substrates containing a modified side chain. *Biomacromolecules* **17**:1477-1485.

Jia Y, Kappock TJ, Frick T, Sinskey AJ & Stubbe J (2000) Lipases provide a new mechanistic model for polyhydroxybutyrate (PHB) synthases: characterization of the functional residues in *Chromatium vinosum* PHB synthase. *Biochemistry* **39**:3927-3936.

Jia Y, Yuan W, Wodzinska J, Park C, Sinskey AJ & Stubbe J (2001) Mechanistic studies on class I polyhydroxybutyrate (PHB) synthase from *Ralstonia eutropha*: Class I and III synthases share a similar catalytic mechanism. *Biochemistry* **40**:1011-1019.

Jozala AF, Geraldes DC, Tundisi LL, Feitosa VdA, Breyer CA, Cardoso SL, Mazzola PG, Oliveira-Nascimento L, Rangel-Yagui CO, Magalhães PO, Oliveira MA & Pessoa A, Jr. (2016) Biopharmaceuticals from microorganisms: from production to purification. *Brazilian Journal of Microbiology* **47**:51-63.

Jürgen B, Breitenstein A, Urlacher V, Büttner K, Lin H, Hecker M, Schweder T & Neubauer P (2010) Quality control of inclusion bodies in *Escherichia coli*. *Microbial Cell Factories* **9**:41.

Kappel K, Wereszczynski J, Clubb RT & McCammon JA (2012) The binding mechanism, multiple binding modes, and allosteric regulation of *Staphylococcus aureus* Sortase A probed by molecular dynamics simulations. *Protein Science* **21**:1858-1871.

Käppel ND, Pröll F & Gauglitz G (2007) Development of a TIRF-based biosensor for sensitive detection of progesterone in bovine milk. *Biosensors and Bioelectronics* **22**:2295-2300.

Kaur G & Roy I (2015) Strategies for large-scale production of polyhydroxyalkanoates. *Chemical and Biochemical Engineering Quarterly* **29**:157-172.

Ke Y, Zhang XY, Ramakrishna S, He LM & Wu G (2016) Synthetic routes to degradable copolymers deriving from the biosynthesized polyhydroxyalkanoates: a mini review. *Express Polymer Letters* **10**:36-53.

Khanna S & Srivastava AK (2005) Recent advances in microbial polyhydroxyalkanoates. *Process Biochemistry* **40**:607-619.

Kim J, Kim YJ, Choi SY, Lee SY & Kim KJ (2017) Crystal structure of *Ralstonia eutropha* polyhydroxyalkanoate synthase C-terminal domain and reaction mechanisms. *Biotechnology Journal* **12**:1600648.

Kim YJ, Choi SY, Kim J, Jin KS, Lee SY & Kim KJ (2017) Structure and function of the N-terminal domain of *Ralstonia eutropha* polyhydroxyalkanoate synthase, and the proposed structure and mechanisms of the whole enzyme. *Biotechnology Journal* **12**:1600649.

Kobayashi T, Uchino K, Abe T, Yamazaki Y & Saito T (2005) Novel intracellular 3-hydroxybutyrate-oligomer hydrolase in *Wautersia eutropha* H16. *Journal of Bacteriology* **187**:5129-5135.

Koller M, Niebelschüetz H & Braunegg G (2013) Strategies for recovery and purification of poly (R)-3-hydroxyalkanoates (PHA) biopolyesters from surrounding biomass. *Engineering in Life Sciences* **13**:549-562.

Koller M, Salerno A, Dias M, Reiterer A & Braunegg G (2010) Modern biotechnological polymer synthesis: a review. *Food Technology and Biotechnology* **48**:255-269.

Kruger RG, Otvos B, Frankel BA, Bentley M, Dostal P & McCafferty DG (2004) Analysis of the substrate specificity of the *Staphylococcus aureus* sortase transpeptidase SrtA. *Biochemistry* **43**:1541-1551.

Krulwich TA, Cheng JB & Guffanti AA (1994) The role of monovalent cation/proton antiporters in $Na^+$-resistance and pH homeostasis in *Bacillus*: an alkaliphile *versus* a neutralophile. *Journal of Experimental Biology* **196**:457-470.

Lahiry A, Fan Y, Stimple SD, Raith M & Wood DW (2017) Inteins as tools for tagless and traceless protein purification. *Journal of Chemical Technology & Biotechnology*:5415.

Law KH, Cheng YC, Leung YC, Lo WH, Chua H & Yu HF (2003) Construction of recombinant *Bacillus subtilis* strains for polyhydroxyalkanoates synthesis. *Biochemical Engineering Journal* **16**:203-208.

Lee J, Jung S-G, Park C-S, Kim H-Y, Batt CA & Kim Y-R (2011) Tumor-specific hybrid polyhydroxybutyrate nanoparticle: surface modification of nanoparticle by enzymatically synthesized functional block copolymer. *Bioorganic & Medicinal Chemistry Letters* **21**:2941-2944.

Lee JH, Nam DH, Lee SH, Park JH, Park SJ, Lee SH, Park CB & Jeong KJ (2014) New platform for cytochrome P450 reaction combining *in situ* immobilization on biopolymer. *Bioconjugate chemistry* **25**:2101-2104.

Lee SJ, Park JP, Park TJ, Lee SY, Lee S & Park JK (2005) Selective immobilization of fusion proteins on poly(hydroxyalkanoate) microbeads. *Analytical Chemistry* **77**:5755-5759.

Levine AC, Heberlig GW & Nomura CT (2016) Use of thiol-ene click chemistry to modify mechanical and thermal properties of polyhydroxyalkanoates (PHAs). *International Journal of Biological Macromolecules* **83**:358-365.

Lewis JG & Rehm BHA (2009) ZZ polyester beads: an efficient and simple method for purifying IgG from mouse hybridoma supernatants. *Journal of Immunological Methods* **346**:71-74.

Li J, Shang G, You M, Peng S, Wang Z, Wu H & Chen G-Q (2011) Endotoxin removing method based on lipopolysaccharide binding protein and polyhydroxyalkanoate binding protein PhaP. *Biomacromolecules* **12**:602-608.

Li R, Zhang H & Qi Q (2007) The production of polyhydroxyalkanoates in recombinant *Escherichia coli*. *Bioresource Technology* **98**:2313-2320.

Li Y (2015) Split-inteins and their bioapplications. *Biotechnology Letters* **37**:2121-2137.

Lightfoot EN & Moscariello JS (2004) Bioseparations. *Biotechnology and Bioengineering* **87**:259-273.

Lin Y-Y & Chen PT (2017) Development of polyhydroxybutyrate biosynthesis in *Bacillus subtilis* with combination of PHB-associated genes derived from *Ralstonia eutropha* and *Bacillus megaterium*. *Journal of the Taiwan Institute of Chemical Engineers* **79**:110-115.

Liu J-W, Yang T, Chen S, Chen X-W & Wang J-H (2013) Nickel chelating functionalization of graphene composite for metal affinity membrane isolation of lysozyme. *Journal of Materials Chemistry B* **1**:810-818.

Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P & Berkmen M (2012) SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial Cell Factories* **11**:56.

Longman SM & Buehring GC (1986) A method for measuring steroid adsorption to tissue culture plasticware. *Journal of tissue culture methods* **10**:253-255.

Losen M, Frölich B, Pohl M & Büchs J (2004) Effect of oxygen limitation and medium composition on *Escherichia coli* fermentation in shake-flask cultures. *Biotechnology progress* **20**:1062-1068.

Lu J, Tappel RC & Nomura CT (2009) Mini-review: biosynthesis of poly(hydroxyalkanoates). *Polymer Reviews* **49**:226-248.

Lu W, Sun Z, Tang Y, Chen J, Tang F, Zhang J & Liu JN (2011) Split intein facilitated tag affinity purification for recombinant proteins with controllable tag removal by inducible auto-cleavage. *Journal of Chromatography A* **1218**:2553-2560.

Lundgren DG, Pfister RM & Merrick JM (1964) Structure of poly-β-hydroxybutyric acid granules. *Journal of General Microbiology* **34**:441-446.

Madison LL & Huisman GW (1999) Metabolic engineering of poly(3-hydroxyalkanoates): from DNA to plastic. *Microbiology and Molecular Biology Reviews* **63**:21-53.

Maestro B & Sanz JM (2017) Polyhydroxyalkanoate-associated phasins as phylogenetically heterogeneous, multipurpose proteins. *Microbial Biotechnology* **10**:1323-1337.

Mao H (2004) A self-cleavable sortase fusion for one-step purification of free recombinant proteins. *Protein Expression and Purification* **37**:253-263.

Martinez-Donato G, Piniella B, Aguilar D, Olivera S, Pérez A, Castañedo Y, Alvarez-Lajonchere L, Dueñas-Carrera S, Lee JW, Burr N, Gonzalez-Miro M & Rehm BHA (2016) Protective T

cell and antibody immune responses against hepatitis C virus achieved using a biopolyester-bead-based vaccine delivery system. *Clinical and Vaccine Immunology* **23**:370-378.

Matsumoto Ki, Murata T, Nagao R, Nomura CT, Arai S, Arai Y, Takase K, Nakashita H, Taguchi S & Shimada H (2009) Production of short-chain-length/medium-chain-length polyhydroxyalkanoate (PHA) copolymer in the plastid of *Arabidopsis thaliana* using an engineered 3-Ketoacyl-acyl carrier protein synthase III. *Biomacromolecules* **10**:686-690.

Matsunaga S, Matsuoka K, Shimizu K, Endo Y & Sawasaki T (2010) Biotinylated-sortase self-cleavage purification (BISOP) method for cell-free produced proteins. *BMC Biotechnol* **10**:42.

Mayer F & Hoppert M (1997) Determination of the thickness of the boundary layer surrounding bacterial PHA inclusion bodies, and implications for models describing the molecular architecture of this layer. *Journal of Basic Microbiology* **37**:45-52.

McCammon JA (1998) Theory of biomolecular recognition. *Current Opinion in Structural Biology* **8**:245-249.

McQualter RB, Bellasio C, Gebbie LK, Petrasovits LA, Palfreyman RW, Hodson MP, Plan MR, Blackman DM, Brumbley SM & Nielsen LK (2016) Systems biology and metabolic modelling unveils limitations to polyhydroxybutyrate accumulation in sugarcane leaves; lessons for C-4 engineering. *Plant Biotechnology Journal* **14**:567-580.

Meng D-C, Shen R, Yao H, Chen J-C, Wu Q & Chen G-Q (2014) Engineering the diversity of polyesters. *Current Opinion in Biotechnology* **29**:24-33.

Miao J, Wu W, Spielmann T, Belfort M, Derbyshire V & Belfort G (2005) Single-step affinity purification of toxic and non-toxic proteins on a fluidics platform. *Lab on a Chip* **5**:248-253.

Michalak M, Kurcok P & Hakkarainen M (2017) Polyhydroxyalkanoate-based drug delivery systems. *Polymer International* **66**:617-622.

Mittendorf V, Robertson EJ, Leech RM, Krüger N, Steinbüchel A & Poirier Y (1998) Synthesis of medium-chain-length polyhydroxyalkanoates in *Arabidopsis thaliana* using intermediates of peroxisomal fatty acid β-oxidation. *Proceedings of the National Academy of Sciences of the United States of America* **95**:13397-13402.

Moldes C, Farinós GP, de Eugenio LI, García P, Garcia JL, Ortego F, Hernández-Crespo P, Castañera P & Prieto MA (2006) New tool for spreading proteins to the environment: Cry1Ab toxin immobilized to bioplastics. *Applied Microbiology and Biotechnology* **72**:88-93.

Moldes C, García P, García JL & Prieto MA (2004) *In vivo* immobilization of fusion proteins on bioplastics by the novel tag BioF. *Applied and Environmental Microbiology* **70**:3205-3212.

Moritsugu K, Terada T & Kidera A (2012) Disorder-to-order transition of an intrinsically disordered region of sortase revealed by multiscale enhanced sampling. *Journal of the American Chemical Society* **134**:7094-7101.

Morth JP, Feng V, Perry LJ, Svergun DI & Tucker PA (2004) The crystal and solution structure of a putative transcriptional antiterminator from *Mycobacterium tuberculosis*. *Structure* **12**:1595-1605.

Mumtaz T, Abd-Aziz S, Rahman NA, Yee PL, Wasoh H, Shirai Y & Hassan MA (2011) Visualization of core-shell PHBV granules of wild type *Comamonas* sp. EB172 *in vivo* under transmission electron microscope. *International Journal of Polymer Analysis and Characterization* **16**:228-238.

Murzin AG (1993) OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *The EMBO journal* **12**:861-867.

Naik MT, Suree N, Ilangovan U, Liew CK, Thieu W, Campbell DO, Clemens JJ, Jung ME & Clubb RT (2006) *Staphylococcus aureus* Sortase A transpeptidase - Calcium promotes sorting signal binding by altering the mobility and structure of an active site loop. *Journal of Biological Chemistry* **281**:1817-1826.

Neumann L, Spinozzi F, Sinibaldi R, Rustichelli F, Pöetter M & Steinbüechel A (2008) Binding of the major phasin, PhaP1, from *Ralstonia eutropha* H16 to poly(3-hydroxybutyrate) granules. *Journal of Bacteriology* **190**:2911-2919.

Nomura CT, Tanaka T, Gan Z, Kuwabara K, Abe H, Takase K, Taguchi K & Doi Y (2004) Effective enhancement of short-chain-length-medium-chain-length polyhydroxyalkanoate copolymer production by coexpression of genetically engineered 3-ketoacyl-acyl-carrier-protein synthase III (*fabH*) and polyhydroxyalkanoate synthesis genes. *Biomacromolecules* **5**:1457-1464.

Normi YM, Hiraishi T, Taguchi S, Abe H, Sudesh K, Najimudin N & Doi Y (2005) Characterization and properties of G4X mutants of *Ralstonia eutropha* PHA synthase for poly (3-hydroxybutyrate) biosynthesis in *Escherichia coli*. *Macromolecular Bioscience* **5**:197-206.

Numata K, Kikkawa Y, Tsuge T, Iwata T, Doi Y & Abe H (2006) Adsorption of biopolyester depolymerase on silicon wafer and poly (R)-3-hydroxybutyric acid single crystal revealed by real-time AFM. *Macromolecular Bioscience* **6**:41-50.

Obruca S, Sedlacek P, Koller M, Kucera D & Pernicova I (2017) Involvement of polyhydroxyalkanoates in stress resistance of microbial cells: biotechnological consequences and applications. *Biotechnology Advances* **12**:006.

Oliveira C & Domingues L (2018) Guidelines to reach high-quality purified recombinant proteins. *Applied Microbiology and Biotechnology* **102**:81-92.

Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J, Sussman JL, Verschueren KHG & Goldman A (1992) The α/β hydrolase fold. *Protein Engineering* **5**:197-211.

Ormö M, Cubitt AB, Kallio K, Gross LA, Tsien RY & Remington SJ (1996) Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**:1392-1395.

Padan E, Bibi E, Ito M & Krulwich TA (2005) Alkaline pH homeostasis in bacteria: new insights. *Biochimica Et Biophysica Acta-Biomembranes* **1717**:67-88.

Pang X & Zhou H-X (2015) Disorder-to-order transition of an active-site loop mediates the allosteric activation of sortase A. *Biophysical Journal* **109**:1706-1715.

Parlane NA, Grage K, Lee JW, Buddle BM, Denis M & Rehm BHA (2011) Production of a particulate hepatitis C vaccine candidate by an engineered *Lactococcus lactis* strain. *Applied and Environmental Microbiology* **77**:8516-8522.

Parlane NA, Gupta SK, Rubio-Reyes P, Chen S, Gonzalez-Miro M, Wedlock DN & Rehm BHA (2017) Self-assembled protein-coated polyhydroxyalkanoate beads: properties and biomedical applications. *Acs Biomaterials Science & Engineering* **3**:3043-3057.

Parlane NA, Wedlock DN, Buddle BM & Rehm BHA (2009) Bacterial polyester inclusions engineered to display vaccine candidate antigens for use as a novel class of safe and efficient vaccine delivery agents. *Applied and Environmental Microbiology* **75**:7739-7744.

Patterson GH, Knobel SM, Sharif WD, Kain SR & Piston DW (1997) Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophysical Journal* **73**:2782-2790.

Perler FB, Davis EO, Dean GE, Gimble FS, Jack WE, Neff N, Noren CJ, Thorner J & Belfort M (1994) Protein splicing elements: inteins and exteins - a definition of terms and recommended nomenclature. *Nucleic Acids Research* **22**:1125-1127.

Peters V, Becher D & Rehm BHA (2007) The inherent property of polyhydroxyalkanoate synthase to form spherical PHA granules at the cell poles: the core region is required for polar localization. *Journal of Biotechnology* **132**:238-245.

Peters V & Rehm BHA (2005) *In vivo* monitoring of PHA granule formation using GFP-labeled PHA synthases. *Fems Microbiology Letters* **248**:93-100.

Peters V & Rehm BHA (2006) *In vivo* enzyme immobilization by use of engineered polyhydroxyalkanoate synthase. *Applied and Environmental Microbiology* **72**:1777-1783.

Peters V & Rehm BHA (2008) Protein engineering of streptavidin for *in vivo* assembly of streptavidin beads. *Journal of Biotechnology* **134**:266-274.

Pfeiffer D & Jendrossek D (2012) Localization of poly(3-hydroxybutyrate) (PHB) granule-associated proteins during PHB granule formation and identification of two new phasins, PhaP6 and PhaP7, in *Ralstonia eutropha* H16. *Journal of Bacteriology* **194**:5909-5921.

Pfeiffer D, Wahl A & Jendrossek D (2011) Identification of a multifunctional protein, PhaM, that determines number, surface to volume ratio, subcellular localization and distribution to

daughter cells of poly(3-hydroxybutyrate), PHB, granules in *Ralstonia eutropha* H16. *Molecular Microbiology* **82**:936-951.

Pina AS, Lowe CR & Roque ACA (2014) Challenges and opportunities in the purification of recombinant tagged proteins. *Biotechnology Advances* **32**:366-381.

Piotukh K, Geltinger B, Heinrich N, Gerth F, Beyermann M, Freund C & Schwarzer D (2011) Directed evolution of sortase A mutants with altered substrate selectivity profiles. *Journal of the American Chemical Society* **133**:17536-17539.

Poirier Y, Erard N & MacDonald-Comber Petétot J (2002) Synthesis of polyhydroxyalkanoate in the peroxisome of *Pichia pastoris*. *Fems Microbiology Letters* **207**:97-102.

Poli A, Di Donato P, Abbamondi GR & Nicolaus B (2011) Synthesis, production, and biotechnological applications of exopolysaccharides and polyhydroxyalkanoates by archaea. *Archaea* **2011**:693253.

Posthuma-Trumpie GA, van Amerongen A, Korf J & van Berkel WJH (2009) Perspectives for on-site monitoring of progesterone. *Trends in Biotechnology* **27**:652-660.

Pötter M, Madkour MH, Mayer F & Steinbüchel A (2002) Regulation of phasin expression and polyhydroxyalkanoate (PHA) granule formation in *Ralstonia eutropha* H16. *Microbiology* **148**:2413-2426.

Pötter M & Steinbüchel A (2005) Poly(3-hydroxybutyrate) granule-associated proteins: impacts on poly(3-hydroxybutyrate) synthesis and degradation. *Biomacromolecules* **6**:552-560.

Priji P, Sajith S, Sreedevi S, Unni KN, Kumar S & Benjamin S (2016) *Candida tropicalis* BPU1 produces polyhydroxybutyrate on raw starchy substrates. *Starch-Starke* **68**:57-66.

Priji P, Unni KN, Sajith S & Benjamin S (2013) *Candida tropicalis* BPU1, a novel isolate from the rumen of the Malabari goat, is a dual producer of biosurfactant and polyhydroxybutyrate. *Yeast* **30**:103-110.

Qi Q, Rehm BHA & Steinbüchel A (1997) Synthesis of poly(3-hydroxyalkanoates) in *Escherichia coli* expressing the PHA synthase gene *phaC2* from *Pseudomonas aeruginosa*: comparison of PhaC1 and PhaC2. *Fems Microbiology Letters* **157**:155-162.

Rabhi-Essafi I, Sadok A, Khalaf N & Fathallah DM (2007) A strategy for high-level expression of soluble and functional human interferon α as a GST-fusion protein in *E. coli*. *Protein Engineering, Design & Selection* **20**:201-209.

Rai R, Keshavarz T, Roether JA, Boccaccini AR & Roy I (2011) Medium chain length polyhydroxyalkanoates, promising new biomedical materials for the future. *Materials Science & Engineering: R: Reports* **72**:29-47.

Ran G, Tan D, Dai W, Zhu X, Zhao J, Ma Q & Lu X (2017) Immobilization of alkaline polygalacturonate lyase from *Bacillus subtilis* on the surface of bacterial

polyhydroxyalkanoate nano-granules. *Applied Microbiology and Biotechnology* **101**:3247-3258.

Rasiah IA & Rehm BHA (2009) One-step production of immobilized α-amylase in recombinant *Escherichia coli*. *Applied and Environmental Microbiology* **75**:2012-2016.

Rehm BHA (2003) Polyester synthases: natural catalysts for plastics *Biochemical Journal* **376** 15-33.

Rehm BHA (2006) Genetics and biochemistry of polyhydroxyalkanoate granule self-assembly: the key role of polyester synthases. *Biotechnology Letters* **28**:207-213.

Rehm BHA (2007) Biogenesis of microbial polyhydroxyalkanoate granules: a platform technology for the production of tailor-made bioparticles. *Current Issues in Molecular Biology* **9**:41-62.

Rehm BHA (2010) Bacterial polymers: biosynthesis, modifications and applications. *Nature Reviews Microbiology* **8**:578.

Rehm BHA, Krüger N & Steinbüchel A (1998) A new metabolic link between fatty acid *de novo* synthesis and polyhydroxyalkanoic acid synthesis - the *phaG* gene from *Pseudomonas putida* KT2440 encodes a 3-hydroxyacyl-acyl carrier protein coenzyme A transferase. *Journal of Biological Chemistry* **273**:24044-24051.

Rehm FBH, Chen S & Rehm BHA (2016) Enzyme engineering for *in situ* immobilization. *Molecules* **21**:1370.

Rehm FBH, Chen S & Rehm BHA (2017) Bioengineering toward direct production of immobilized enzymes: a paradigm shift in biocatalyst design. *Bioengineered* **9**:6-11.

Reitter JN, Cousin CE, Nicastri MC, Jaramillo MV & Mills KV (2016) Salt-dependent conditional protein splicing of an intein from *Halobacterium salinarum*. *Biochemistry* **55**:1279-1282.

Ricci MS, Sarkar CA, Fallon EM, Lauffenburger DA & Brems DN (2003) pH Dependence of structural stability of interleukin-2 and granulocyte colony-stimulating factor. *Protein Science* **12**:1030-1038.

Robins KJ, Hooks DO, Rehm BHA & Ackerley DF (2013) *Escherichia coli* NemA is an efficient chromate reductase that can be biologically immobilized to provide a cell free system for remediation of hexavalent chromium. *PLoS One* **8**:e59200.

Romanov VP, Kostromina TI, Miroshnikov AI & Feofanov SA (2017) Preparative method for obtaining recombinant human interferon α2b from inclusion bodies of *Escherichia coli*. *Russian Journal of Bioorganic Chemistry* **42**:631-637.

Rubio Reyes P, Parlane NA, Wedlock DN & Rehm BHA (2016) Immunogencity of antigens from *Mycobacterium tuberculosis* self-assembled as particulate vaccines. *International Journal of Medical Microbiology* **306**:624-632.

Sambrook J, Fritsch EF & Maniatis T (1989) *Molecular cloning: A laboratory manual*. New York, NY: Cold Spring Harbor Laboratory Press.

Samsonova JV, Safronova VA & Osipov AP (2015) Pretreatment-free lateral flow enzyme immunoassay for progesterone detection in whole cows' milk. *Talanta* **132**:685-689.

Schäfer AI, Akanyeti I & Semião AJC (2011) Micropollutant sorption to membrane polymers: a review of mechanisms for estrogens. *Advances in Colloid and Interface Science* **164**:100-117.

Schnaitman CA (1971) Solubilization of the cytoplasmic membrane of *Escherichia coli* by Triton X-100. *Journal of Bacteriology* **108**:545-552.

Schubert P, Steinbuchel A & Schlegel HG (1988) Cloning of the *Alcaligenes eutrophus* genes for synthesis of poly-beta-hydroxybutyric acid (PHB) and synthesis of PHB in *Escherichia coli*. *Journal of Bacteriology* **170**:5837-5847.

Seo H-M, Kim J-H, Jeon J-M, Song H-S, Bhatia SK, Sathiyanarayanan G, Park K, Kim KJ, Lee SH, Kim HJ & Yang Y-H (2016) *In situ* immobilization of lysine decarboxylase on a biopolymer by fusion with phasin immobilization of CadA on intracellular PHA. *Process Biochemistry* **51**:1413-1419.

Shahmohammadi A (2017) Lysozyme separation from chicken egg white: a review. *European Food Research and Technology*:1-17.

Shen M, Rusling JF & Dixit CK (2017) Site-selective orientated immobilization of antibodies and conjugates for immunodiagnostics development. *Methods* **116**:95-111.

Shevchenko A, Tomas H, Havlis J, Olsen JV & Mann M (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature protocols* **1**:2856-2860.

Shi C, Meng Q & Wood DW (2013) A dual ELP-tagged split intein system for non-chromatographic recombinant protein purification. *Applied Microbiology and Biotechnology* **97**:829-835.

Shi C, Miskioglu EE, Meng Q & Wood DW (2013) Intein-based purification tags in recombinant protein production and new methods for controlling self-cleavage. *Pharmaceutical Bioprocessing* **1**:441-454.

Simersky R, Swaczynova J, Morris DA, Franek M & Strnad M (2007) Development of an ELISA-based kit for the on-farm determination of progesterone in milk. *Veterinarni Medicina* **52**:19-28.

Singh M, Patel SKS & Kalia VC (2009) *Bacillus subtilis* as potential producer for polyhydroxyalkanoates. *Microbial Cell Factories* **8**

Skerra A (2007) Alternative non-antibody scaffolds for molecular recognition. *Current Opinion in Biotechnology* **18**:295-304.

Škrlec K, Štrukelj B & Berlec A (2015) Non-immunoglobulin scaffolds: a focus on their targets. *Trends in Biotechnology* **33**:408-418.

Slater SC, Voige WH & Dennis DE (1988) Cloning and expression in *Escherichia coli* of the *Alcaligenes eutrophus* H16 poly-beta-hydroxybutyrate biosynthetic pathway. *Journal of Bacteriology* **170**:4431-4436.

Somleva MN, Peoples OP & Snell KD (2013) PHA bioplastics, biochemicals, and energy from crops. *Plant Biotechnology Journal* **11**:233-252.

Ståhl S, Kronqvist N, Jonsson A & Löfblom J (2013) Affinity proteins and their generation. *Journal of Chemical Technology and Biotechnology* **88**:25-38.

Steemson JD (2011) Directed evolution and structural analysis of an OB-fold domain towards a specifc binding reagent. Doctoral dissertation, University of Waikato

Steemson JD, Baake M, Rakonjac J, Arcus VL & Liddament MT (2014) Tracking molecular recognition at the atomic level with a new protein scaffold based on the OB-fold. *PLoS One* **9**:e86050.

Suree N, Liew CK, Villareal VA, Thieu W, Fadeev EA, Clemens JJ, Jung ME & Clubb RT (2009) The structure of the *Staphylococcus aureus* sortase-substrate complex reveals how the universally conserved LP*X*TG sorting signal is recognized. *Journal of Biological Chemistry* **284**:24465-24477.

Sznajder A & Jendrossek D (2014) To be or not to be a poly (3-hydroxybutyrate)(PHB) depolymerase: PhaZd1 (PhaZ6) and PhaZd2 (PhaZ7) of *Ralstonia eutroph*a, highly active PHB depolymerases with no detectable role in mobilization of accumulated PHB. *Applied and Environmental Microbiology* **80**:4936-4946.

Tan GYA, Chen CL, Li L, Ge L, Wang L, Razaad IMN, Li Y, Zhao L, Mo Y & Wang JY (2014) Start a research on biopolymer polyhydroxyalkanoate (PHA): a review. *Polymers* **6**:706-754.

Tashiro M, Tejero R, Zimmerman DE, Celda B, Nilsson B & Montelione GT (1997) High-resolution solution NMR structure of the Z domain of staphylococcal protein A. *Journal of Molecular Biology* **272**:573-590.

Thompson T, Rehm BHA, Herbert AB, Saravolac EG, McDermott PB & Draper JL (2013) Compositions for separation methods. United States Patent US 13/992,813, 19 Dec. 2013.

Thomson N, Summers D & Sivaniah E (2010) Synthesis, properties and uses of bacterial storage lipid granules as naturally occurring nanoparticles. *Soft Matter* **6**:4045-4057.

Tian JM, Sinskey AJ & Stubbe J (2005) Kinetic studies of polyhydroxybutyrate granule formation in *Wautersia eutropha* H16 by transmission electron microscopy. *Journal of Bacteriology* **187**:3814-3824.

Timm A & Steinbüchel A (1992) Cloning and molecular analysis of the poly (3-hydroxyalkanoic acid) gene locus of *Pseudomonas aeruginosa* PAO1. *European Journal of Biochemistry* **209**:15-30.

Ton-That H, Mazmanian SK, Faull KF & Schneewind O (2000) Anchoring of surface proteins to the cell wall of *Staphylococcus aureus* - sortase catalyzed in vitro transpeptidation reaction using LPXTG peptide and NH2-Gly(3) substrates. *Journal of Biological Chemistry* **275**:9876-9881.

Tozzi C, Anfossi L & Giraudi G (2003) Affinity chromatography techniques based on the immobilisation of peptides exhibiting specific binding activity. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **797**:289-304.

Trower MK & Elgar GS (1996) Cloning PCR Products Using T-Vectors. In: Harwood A.J. (ed) *Basic DNA and RNA Protocols*. Methods in Molecular Biology™, Humana Press, **58**:313-324.

Tsuge T, Hyakutake M & Mizuno K (2015) Class IV polyhydroxyalkanoate (PHA) synthases and PHA-producing *Bacillus*. *Applied Microbiology and Biotechnology* **99**:6231-6240.

Tsuge T, Taguchi K, Taguchi S & Doi Y (2003) Molecular characterization and properties of (*R*)-specific enoyl-CoA hydratases from *Pseudomonas aeruginosa*: metabolic tools for synthesis of polyhydroxyalkanoates via fatty acid β-oxidation. *International Journal of Biological Macromolecules* **31**:195-205.

Uchino K, Saito T, Gebauer B & Jendrossek D (2007) Isolated poly(3-hydroxybutyrate) (PHB) granules are complex bacterial organelles catalyzing formation of PHB from acetyl coenzyme A (CoA) and degradation of PHB to acetyl-CoA. *Journal of Bacteriology* **189**:8250-8256.

Uchino K, Saito T & Jendrossek D (2008) Poly(3-hydroxybutyrate) (PHB) depolymerase PhaZa1 is involved in mobilization of accumulated PHB in *Ralstonia eutropha* H16. *Applied and Environmental Microbiology* **74**:1058-1063.

Ushimaru K & Tsuge T (2016) Characterization of binding preference of polyhydroxyalkanoate biosynthesis-related multifunctional protein PhaM from *Ralstonia eutropha*. *Applied Microbiology and Biotechnology* **100**:4413-4421.

Vandamme P & Coenye T (2004) Taxonomy of the genus *Cupriavidus*: a tale of lost and found. *International Journal of Systematic and Evolutionary Microbiology* **54**:2285-2289.

Vařilová T, Maděra M, Pacáková V & Štulík K (2006) Separation media in affinity chromatography of proteins - A critical review. *Current Proteomics* **3**:55-79.

Vavrová Ľ, Muchová K & Barák I (2010) Comparison of different *Bacillus subtilis* expression systems. *Research in Microbiology* **161**:791-797.

Vemula S, Thunuguntla R, Dedaniya A, Kokkiligadda S, Palle C & Ronda SR (2015) Improved production and characterization of recombinant human granulocyte colony stimulating factor from *E. coli* under optimized downstream processes. *Protein Expression and Purification* **108**:62-72.

Wahl A, Schuth N, Pfeiffer D, Nussberger S & Jendrossek D (2012) PHB granules are attached to the nucleoid via PhaM in *Ralstonia eutropha*. *BMC Microbiology* **12**:262.

Wang R, Zhou X, Zhu X, Yang C, Liu L & Shi H (2017) Isoelectric bovine serum albumin: robust blocking agent for enhanced performance in optical-fiber based DNA sensing. *Acs Sensors* **2**:257-262.

Wang S, Chen W, Xiang H, Yang J, Zhou Z & Zhu M (2016) Modification and potential application of short-chain-length polyhydroxyalkanoate (scl-PHA). *Polymers* **8**:273.

Wang Y, Ren W, Gao D, Wang L, Yang Y & Bai Q (2015) One-step refolding and purification of recombinant human tumor necrosis factor-α (rhTNF-α) using ion-exchange chromatography. *Biomed Chromatogr* **29**:305-311.

Wang Y, Ruan L, Chua H & Yu PHF (2006) Cloning and expression of the PHA synthase genes *phaC1* and *phaC1AB* into *Bacillus subtilis*. *World Journal of Microbiology & Biotechnology* **22**:559-563.

Wang Z, Wu H, Chen J, Zhang J, Yao Y & Chen GQ (2008) A novel self-cleaving phasin tag for purification of recombinant proteins based on hydrophobic polyhydroxyalkanoate nanoparticles. *Lab on a Chip* **8**:1957-1962.

Welling GW, van Gorkum J, Damhof RA, Drijfhout JW, Bloemhoff W & Welling-Wester S (1991) A ten-residue fragment of an antibody (mini-antibody) directed against lysozyme as ligand in immunoaffinity chromatography. *Journal of Chromatography* **548**:235-242.

Wiegert T, Homuth G, Versteeg S & Schumann W (2001) Alkaline shock induces the *Bacillus subtilis* $\sigma^W$ regulon. *Molecular Microbiology* **41**:59-71.

Williams MD, Rahn JA & Sherman DH (1996) Production of a polyhydroxyalkanoate biopolymer in insect cells with a modified eucaryotic fatty acid synthase. *Applied and Environmental Microbiology* **62**:2540-2546.

Wittenborn EC, Jost M, Wei Y, Stubbe J & Drennan CL (2016) Structure of the catalytic domain of the class I polyhydroxybutyrate synthase from *Cupriavidus necator*. *Journal of Biological Chemistry* **291**:25264-25277.

Wodzinska J, Snell KD, Rhomberg A, Sinskey AJ, Biemann K & Stubbe J (1996) Polyhydroxybutyrate synthase: evidence for covalent catalysis. *Journal of the American Chemical Society* **118**:6319-6320.

Wolman FJ, Copello GJ, Mebert AM, Targovnik AM, Miranda MV, Navarro del Cañizo AA, Díaz LE & Cascone O (2010) Egg white lysozyme purification with a chitin-silica-based affinity chromatographic matrix. *European Food Research and Technology* **231**:181-188.

Wood DW (2003) Simplified protein purification using engineered self-cleaving affinity tags. *Journal of Chemical Technology and Biotechnology* **78**:103-110.

Wood DW & Shi C (2016) Protein production systems and methods thereof. US20160207965 A1, 21 Jul. 2016.

Wu L, Xu C, Xia C, Duan Y, Xu C, Zhang H & Bao J (2014) Development and application of an ELISA kit for the detection of milk progesterone in dairy cows. *Monoclonal antibodies in immunodiagnosis and immunotherapy* **33**:330-333.

Xiao N, Jiao N & Liu Y (2015) *In vivo* and *in vitro* observations of polyhydroxybutyrate granules formed by *Dinoroseobacter* sp JL 1447. *International Journal of Biological Macromolecules* **74**:467-475.

Yao YC, Zhan XY, Zhang J, Zou XH, Wang ZH, Xiong YC, Chen J & Chen G-Q (2008) A specific drug targeting system based on polyhydroxyalkanoate granule binding protein PhaP fused with targeted cell ligands. *Biomaterials* **29**:4823-4830.

York GM, Stubbe J & Sinskey AJ (2002) The *Ralstonia eutropha* PhaR protein couples synthesis of the PhaP phasin to the presence of polyhydroxybutyrate in cells and promotes polyhydroxybutyrate production. *Journal of Bacteriology* **184**:59-66.

Young CL, Britton ZT & Robinson AS (2012) Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. *Biotechnology Journal* **7**:620-634.

Yu K, Liu C, Kim BG & Lee DY (2015) Synthetic fusion protein design and applications. *Biotechnology Advances* **33**:155-164.

Yuan L, Yue Z, Chen H, Huang H & Zhao T (2009) Biomacromolecular affinity: Interactions between lysozyme and regioselectively sulfated chitosan. *Colloids and Surfaces B: Biointerfaces* **73**:346-350.

Zhang C, Liu Y, Zhao D, Li X, Yu R & Su Z (2014) Facile purification of *Escherichia coli* expressed tag-free recombinant human tumor necrosis factor alpha from supernatant. *Protein Expression and Purification* **95**:195-203.

Zhang HY, Du XY, Liu Q, Xia C & Sun LW (2013) Detection of progesterone in bovine milk using an electrochemical immunosensor. *International Journal of Dairy Technology* **66**:461-467.

Zhang S, Kolvek S, Lenz RW & Goodwin S (2003) Mechanism of the polymerization reaction initiated and catalyzed by the polyhydroxybutyrate synthase of *Ralstonia eutropha*. *Biomacromolecules* **4**:504-509.

Zhang S, Wang ZH & Chen GQ (2010) Microbial polyhydroxyalkanote synthesis repression protein PhaR as an affinity tag for recombinant protein purification. *Microbial Cell Factories* **9**:28.

Zhao M, Wu T, Xiao X, Liu Y & Su X (2013) New advances in molecular recognition based on biomolecular scaffolds. *Analytical and Bioanalytical Chemistry* **405**:5679-5685.

Zhou X, Song Z, Liu X, Jia F & Wang Y (2011) Production of recombinant porcine interferon alpha using PHB-intein-mediated protein purification strategy. *Applied Biochemistry and Biotechnology* **163**:981-993.

Zinn M & Hany R (2005) Tailored material properties of polyhydroxyalkanoates through biosynthesis and chemical modification. *Advanced Engineering Materials* **7**:408-411.

Zong Y, Bice TW, Ton-That H, Schneewind O & Narayana SVL (2004) Crystal structures of *Staphylococcus aureus* sortase A and its substrate complex. *Journal of Biological Chemistry* **279**:31383-31389.

Zou H, Shi M, Zhang T, Li L, Li L & Xian M (2017) Natural and engineered polyhydroxyalkanoate (PHA) synthase: key enzyme in biopolyester production. *Applied Microbiology and Biotechnology* **101**:7417-7426.

# Chapter 7: Appendices

## 7.1 Supplementary figures

Key for plasmid components in the supplementary figures:

| | |
|---|---|
| T7 promoter | Promoter for bacteriophage T7 RNA polymerase |
| T7 terminator | Transcription terminator for bacteriophage T7 RNA polymerase |
| AmpR | Ampicillin resistance conferred by β-Lactamase |
| ori | Origin of replication |
| PhaC | PHA synthase from *Ralstonia eutropha*, wild type |
| SrtA | Sortase A from *Staphylococcus aureus* minus the N-terminal membrane anchor region |
| LPETG signal | SrtA recognises this five amino acid signal and cleaves between the T and G in the presence of $Ca^{2+}$ +/- triglycine |
| TNFα | Human tumour necrosis factor alpha, soluble form |
| IFNα2b | Human interferon alpha 2b, without signal peptide |
| *Ssp* Intein | *Ssp* DnaB helicase mini intein from *Synechocystis* sp. strain PCC6803, derived from pTwin1 vector (NEB), which self-cleaves when pH drops to 6. |
| CBD | Chitin binding domain derived from pTwin1 vector (NEB) |
| GFP | Green fluorescent protein from *Aequorea victoria* |
| Rv1626 | Putative transcriptional antiterminator Rv1626 from *Mycobacterium tuberculosis* |
| ZZ | Two copies of a IgG binding Z domain of protein A derived from *Staphylococcus aureus* |
| G-CSF | Human granulocyte colony-stimulating factor, short isoform without signal peptide or VSE after the QEKL residue |
| Linker | A flexible DNA linker inserted between 3' end of *phaC* gene and 5' end of a target coding region to facilitate the folding of the target protein or peptide |
| SG-linker | A triplicate SGGGG linker introduced at C-terminus of PhaC to maintain its hydrophobic environment for a proper PHA synthase functionality |
| O6 | Lysozyme binding OBody L200EP-06, a synthetic peptide engineered based on the OB-fold domain of aspartyl-tRNA synthetase (aspRS) from *Pyrobaculum aerophilum* |
| D7 | 1st generation of progesterone (P4) binding OBody P4013-D7, a synthetic peptide engineered based on the OB-fold domain of aspartyl-tRNA synthetase (aspRS) from *P. aerophilum* |
| B7 | 2nd generation of P4 binding OBody B7, a synthetic peptide engineered based on the OB-fold domain of aspRS from *P. aerophilum* |


**Figure 7.1 Plasmid map for pET14b-PhaC-SrtA (Methods section 2.4.8).**

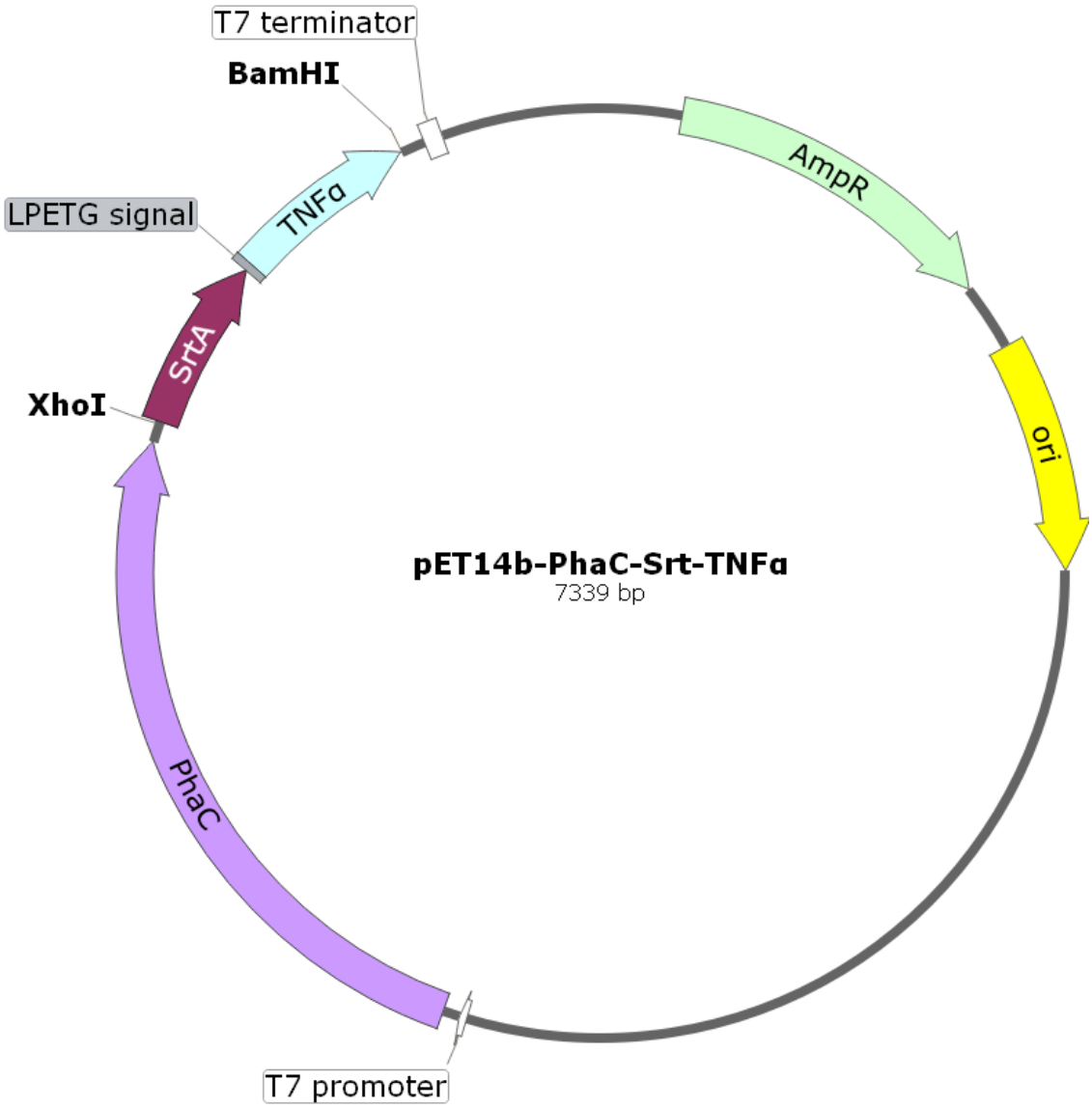



**Figure 7.2 Plasmid map for pET14b-PhaC-SrtA-TNFα (Methods section 2.4.8).**

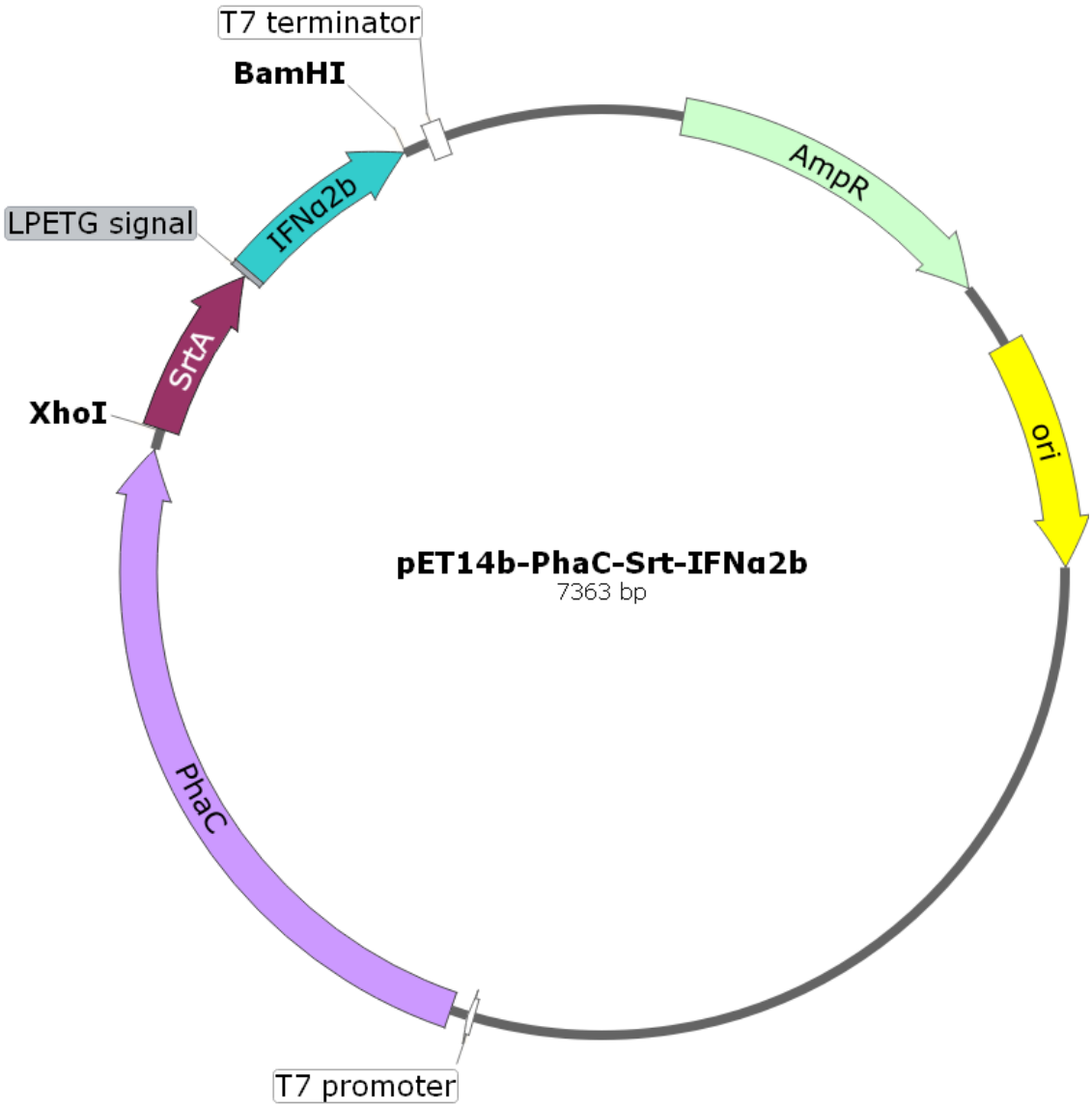



**Figure 7.3 Plasmid map for pET14b-PhaC-SrtA-IFNα2b (Methods section 2.4.8).**

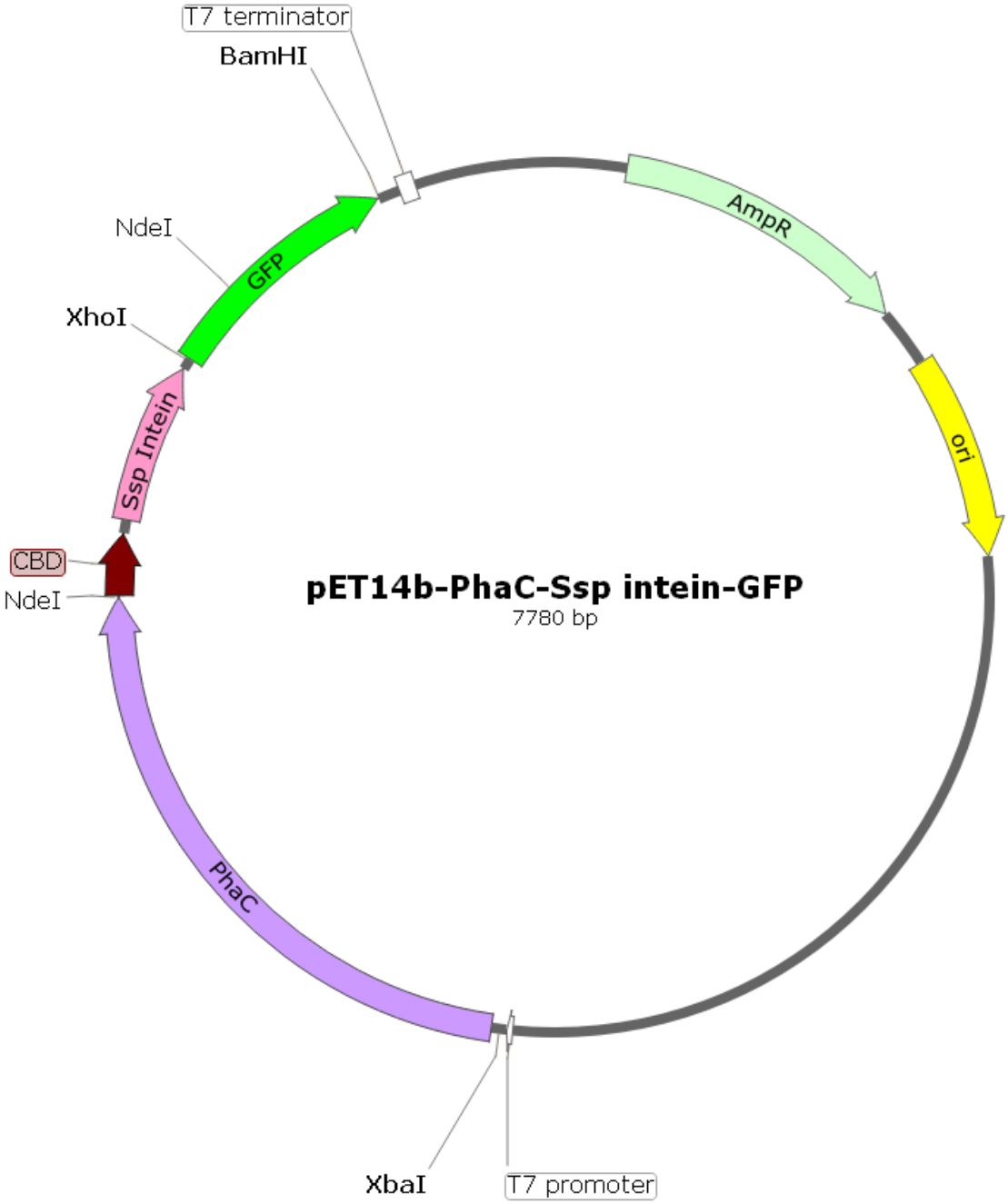



**Figure 7.4 Plasmid map for pET14b-PhaC-Intein-GFP (Methods section 2.4.8).**


**Figure 7.5 Plasmid map for pET14b-PhaC-Intein-Rv1626 (Methods section 2.4.8).**

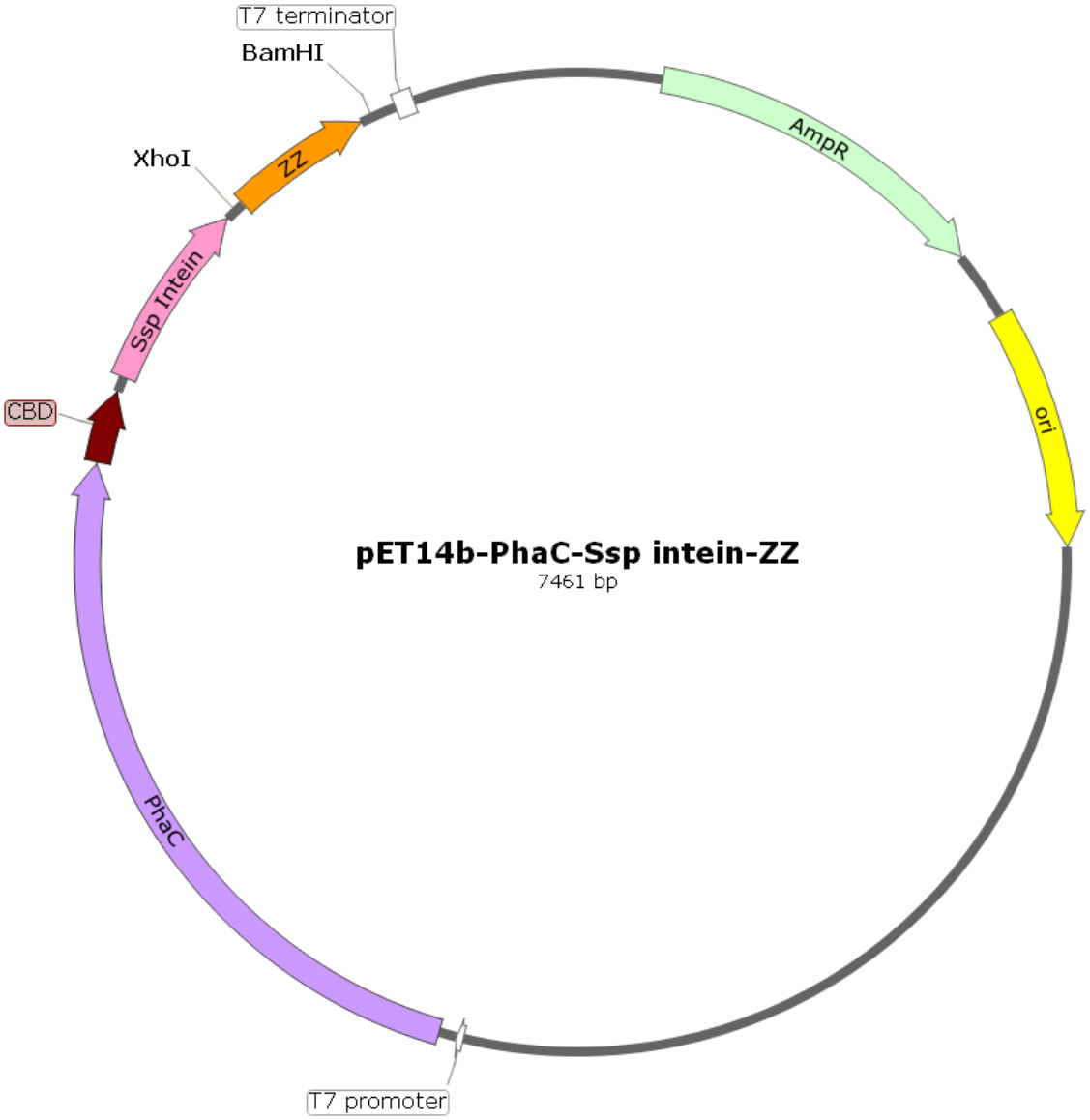



**Figure 7.6 Plasmid map for pET14b-PhaC-Intein-ZZ (Methods section 2.4.8).**

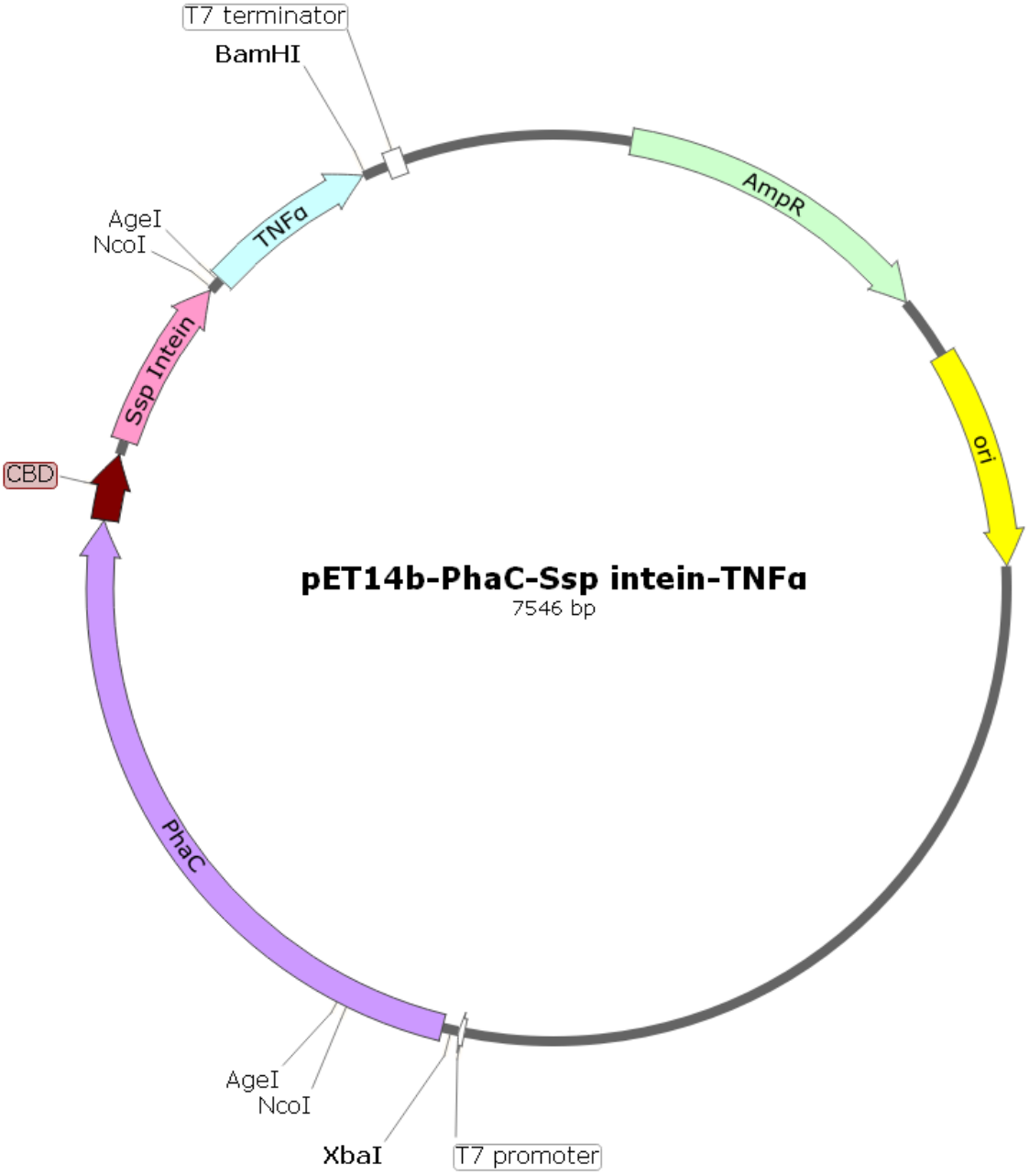



**Figure 7.7 Plasmid map for pET14b-PhaC-Intein-TNFα (Methods section 2.4.8).**

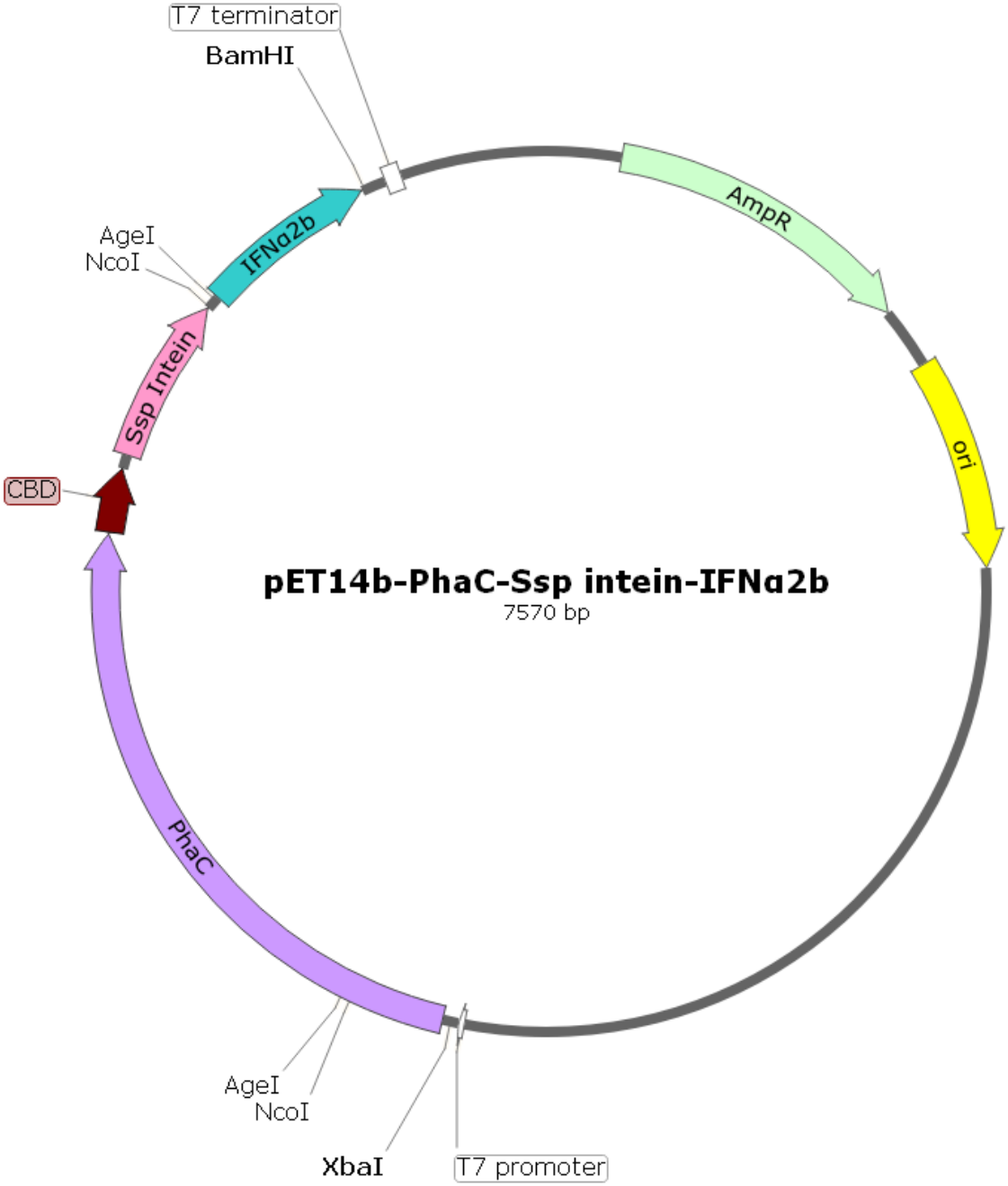



**Figure 7.8 Plasmid map for pET14b-PhaC-Intein-IFNα2b (Methods section 2.4.8).**

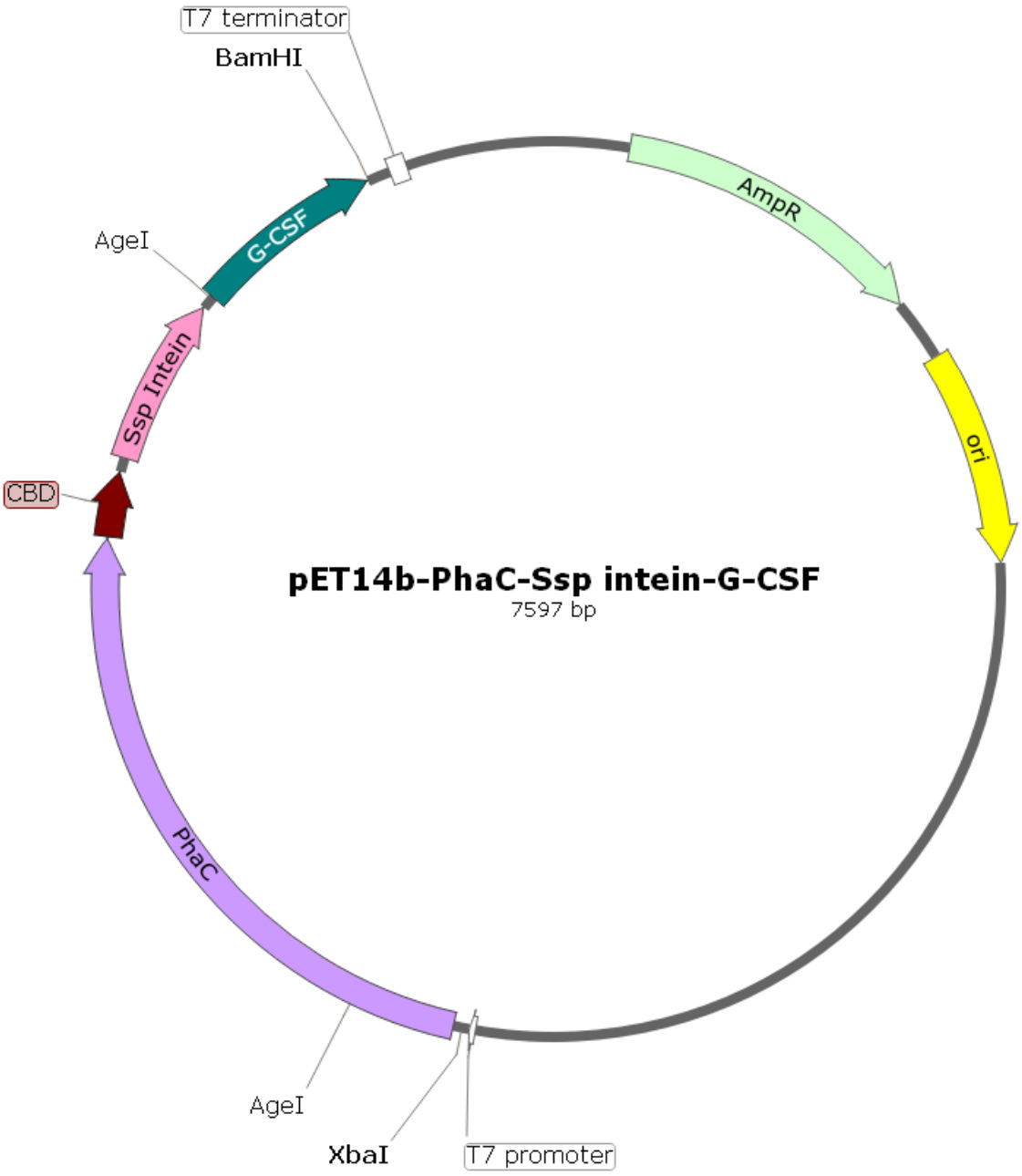



**Figure 7.9 Plasmid map for pET14b-PhaC-Intein- G-CSF (Methods section 2.4.8).**

```
  1 VYPKKTHWTA EITPNLHGSE VVVAGWVAHL GDYGRVKIVK VSDREGGAAV
 51 PVYLERGKTP DHLFKVFAEL SREDVVVIKG IVEASEQYGS DTGVEIFPSE
101 IWILNKAKPL PID
```

**Figure 7.10 Amino acid sequence of OBody L200EP-06 (O6) (Methods section 2.4.8).**

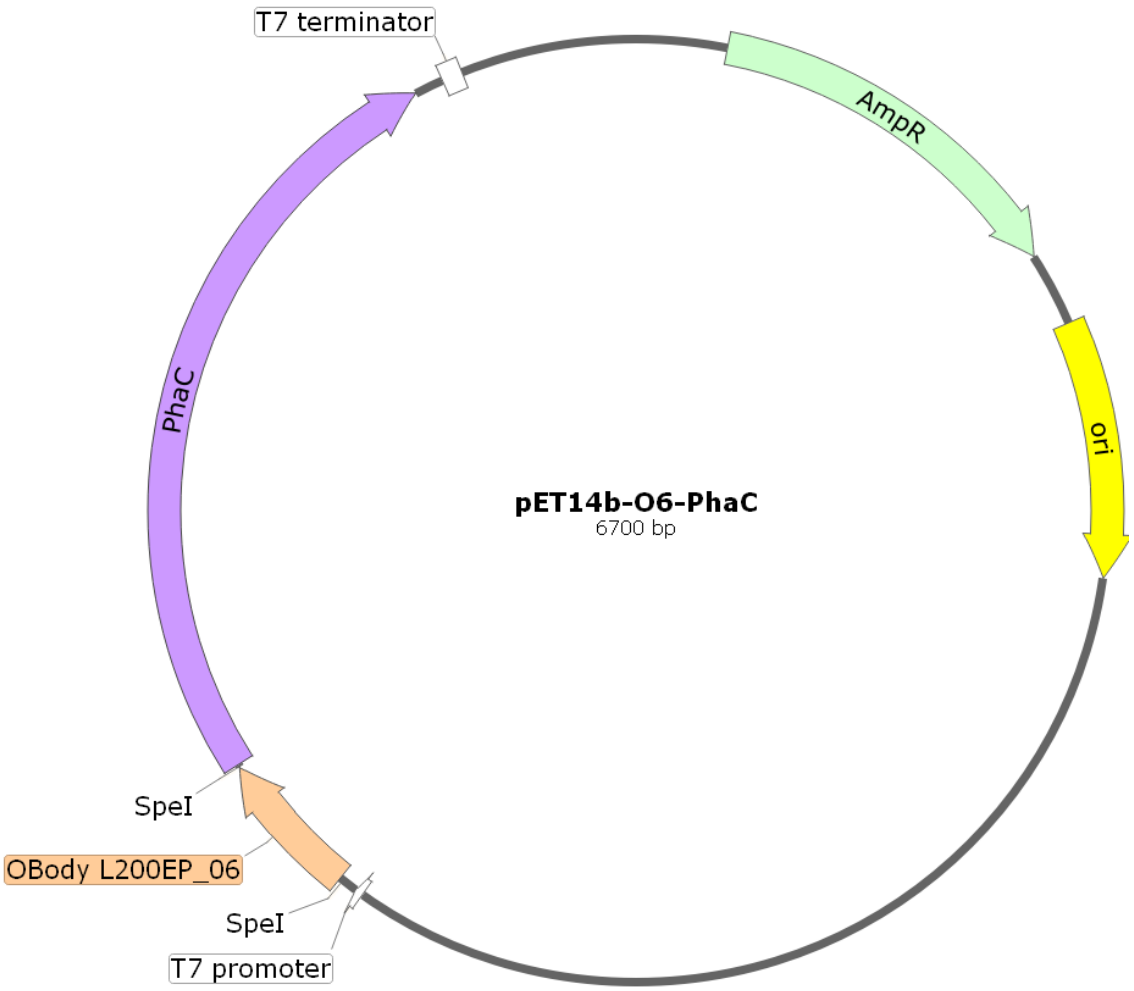



**Figure 7.11 Plasmid map for pET14b-O6-PhaC (Methods section 2.4.8).**

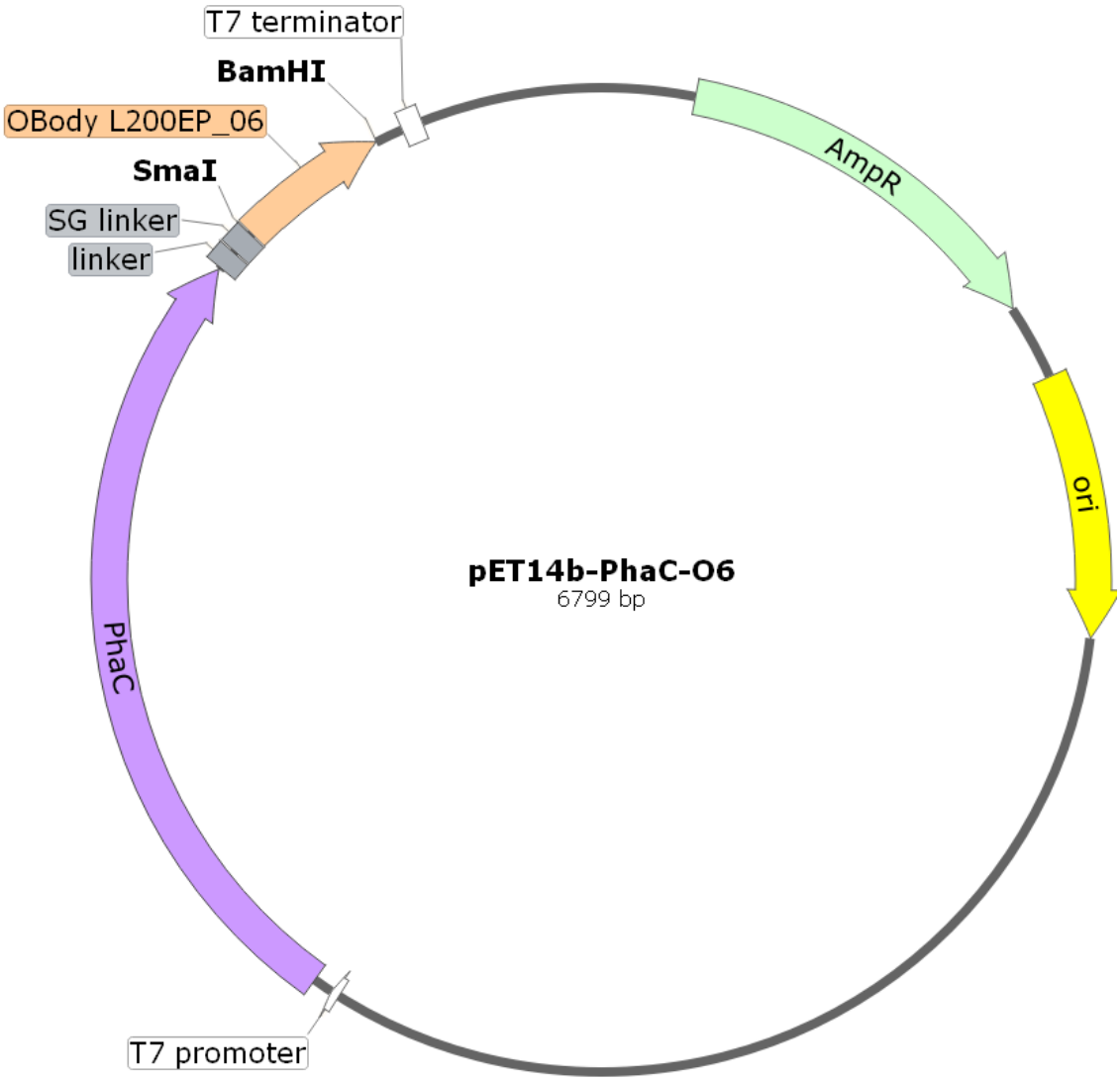



**Figure 7.12 Plasmid map for pET14b-PhaC-O6 (Methods section 2.4.8).**

```
  1 MATHWTAEIT PNLHGTEVVV AGWVSTLTDN GSYKCVGVSD HQGFVLVCLV
 51 AGSTPDHLFK VFAELSREDV VVIKGIVHAS CKHKCGVSIS PSEIWILNKA
101 KPLPID
```

**Figure 7.13 Amino acid sequence of OBody P4013-D7 (D7) (Methods section 2.4.8).**

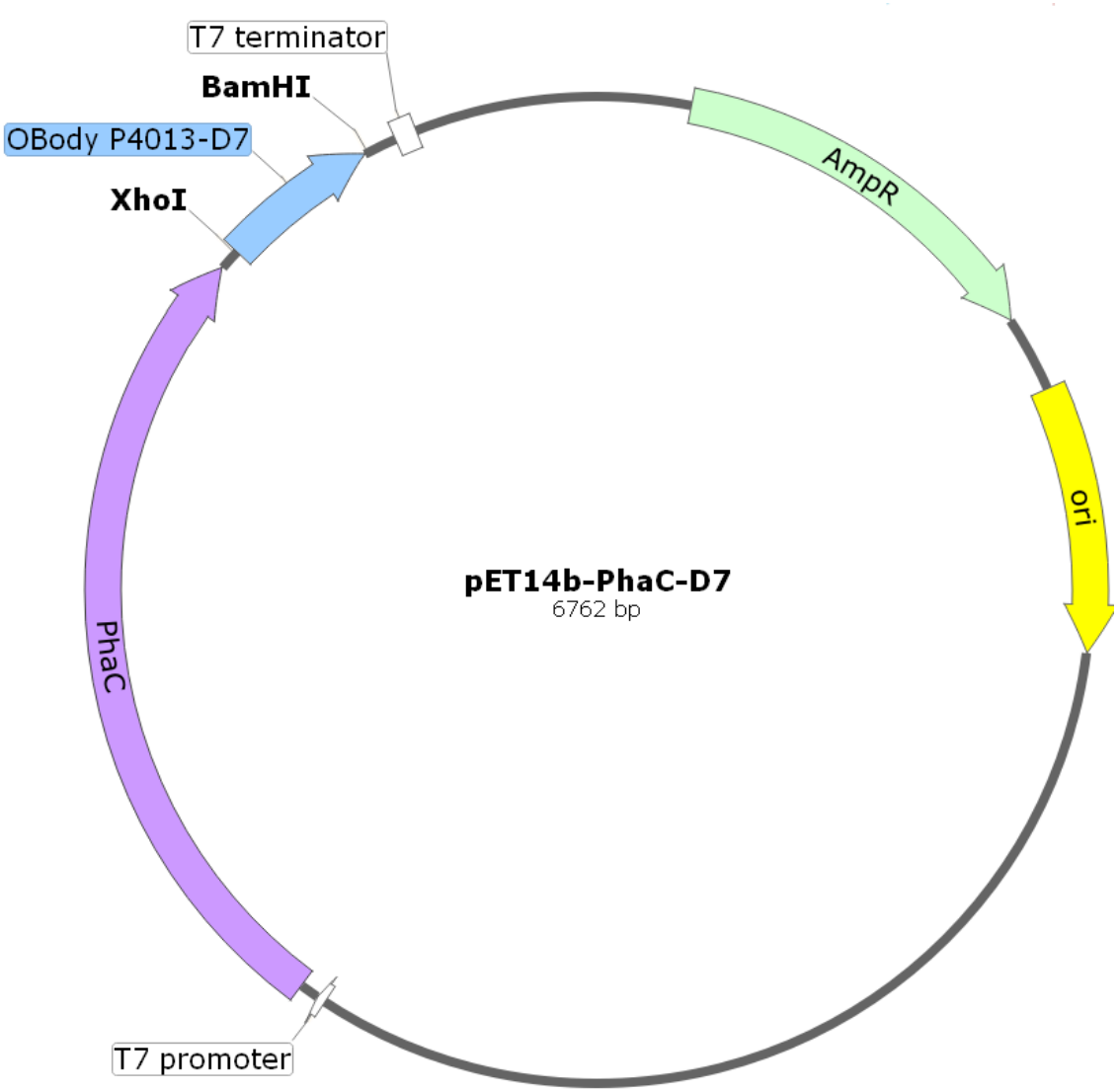



**Figure 7.14 Plasmid map for pET14b-PhaC-D7 (Methods section 2.4.8).**

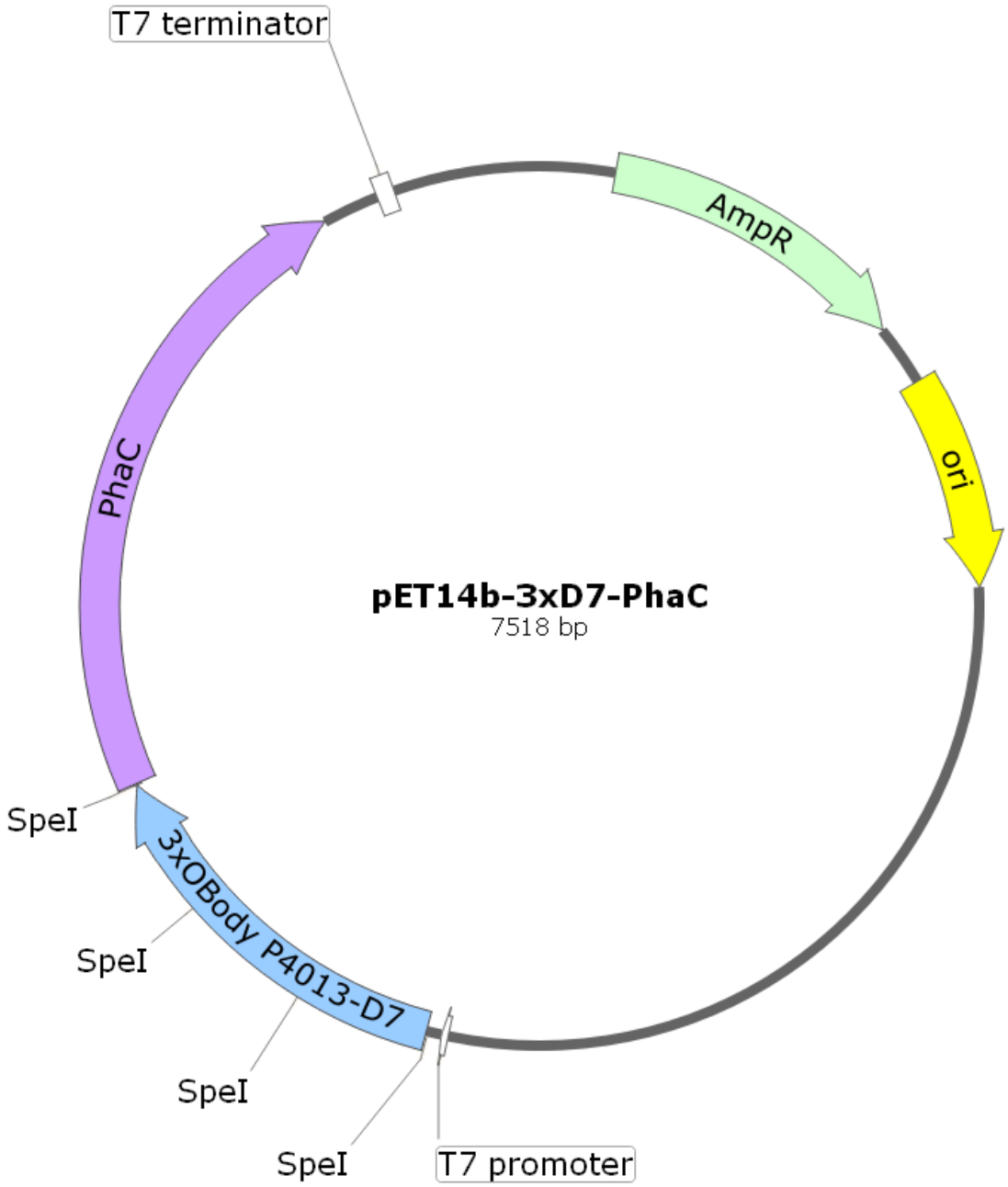



**Figure 7.15 Plasmid map for pET14b-3xD7-PhaC (Methods section 2.4.8).**

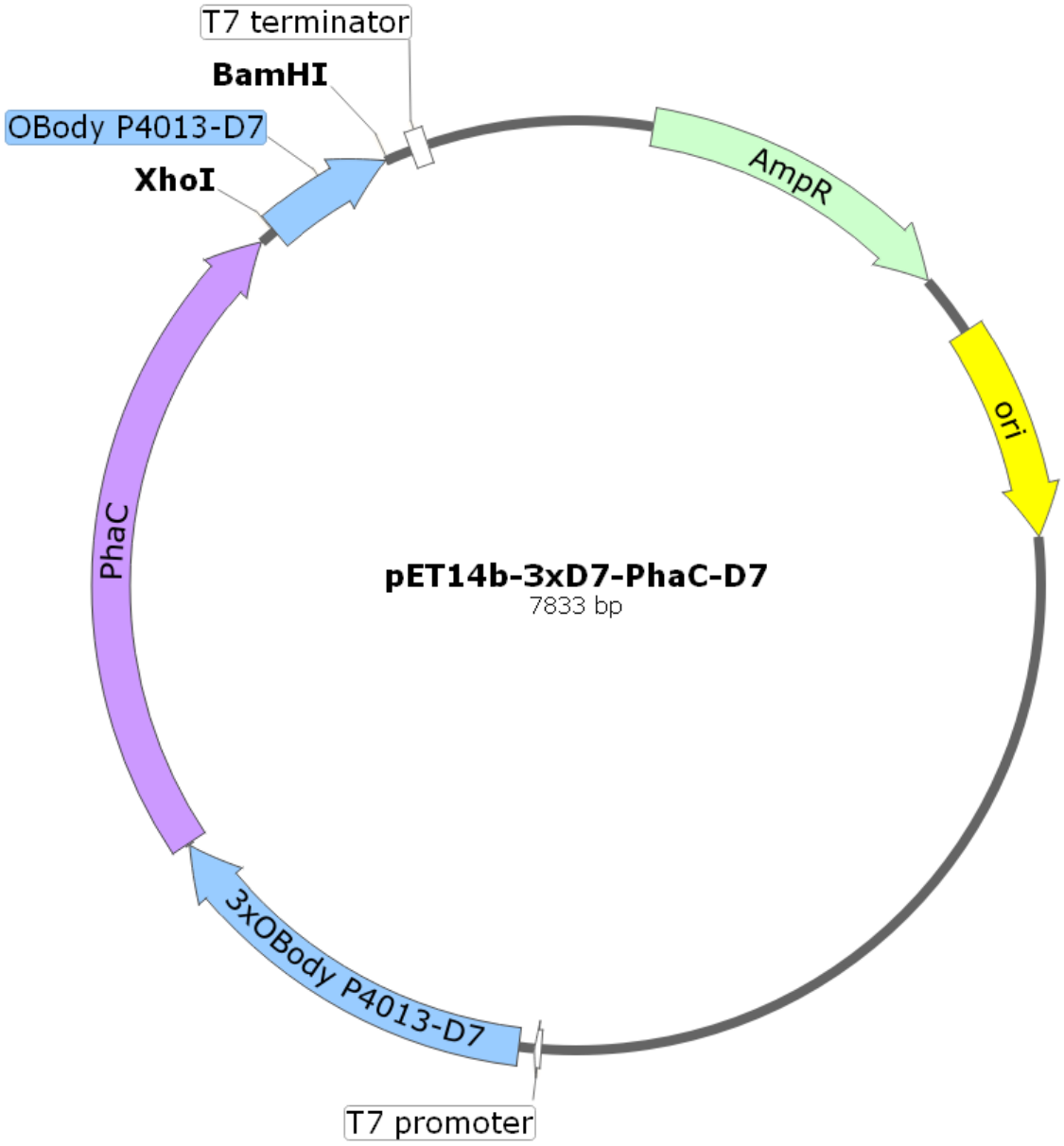



**Figure 7.16 Plasmid map for pET14b-3xD7-PhaC-D7 (Methods section 2.4.8).**

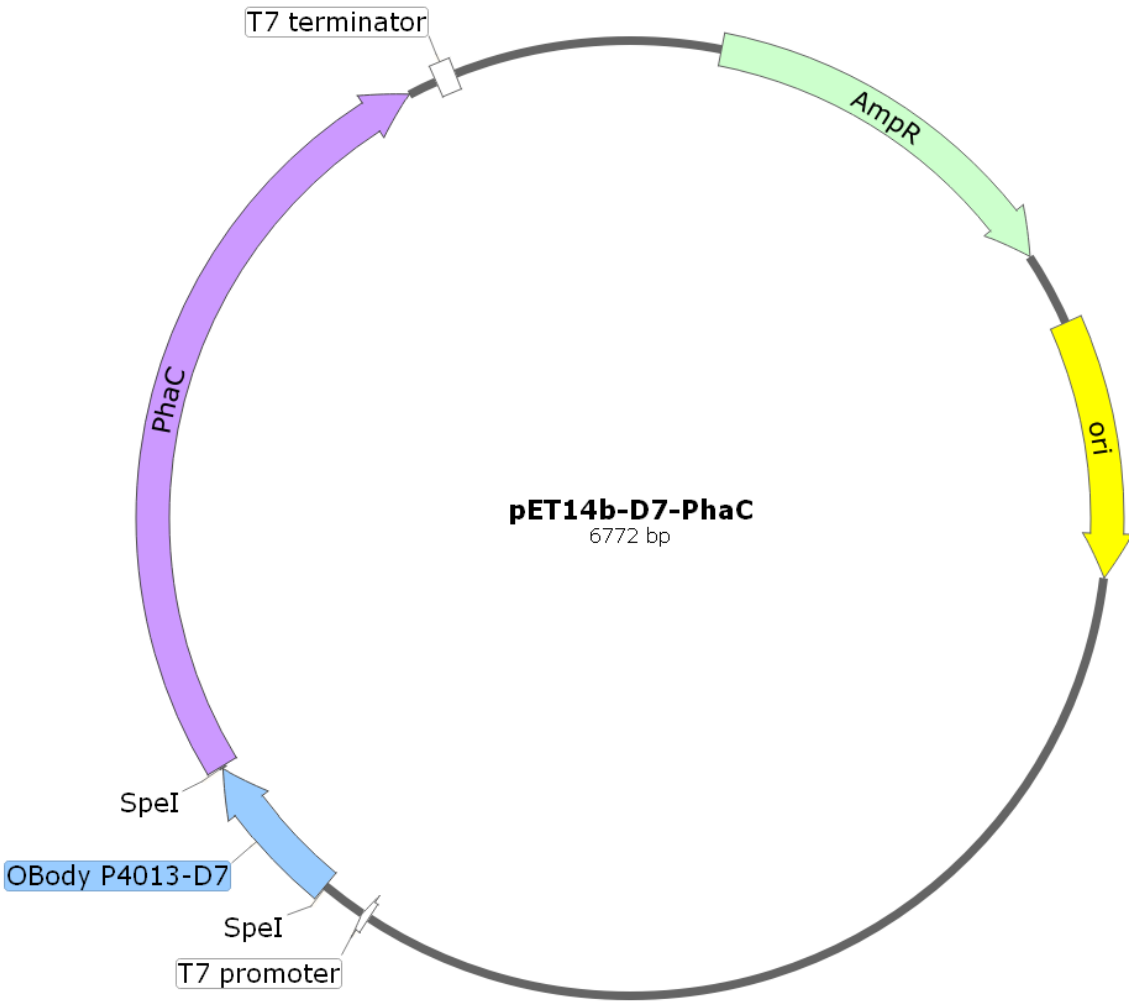



**Figure 7.17 Plasmid map for pET14b-D7-PhaC (Methods section 2.4.8).**

```
  1 MATHWTAEIT PNLHGTEVVV AGWVSTLTDN GSYKGVGVSD HQGFVLVSLV
 51 AGSTPDHLFK VFAELSREDV VVIKGIVHAH EPWDSGVSIS PSEIWILNKA
101 KPLPID
```

**Figure 7.18 Amino acid sequence of OBody B7 (B7) (Methods section 2.4.8).**

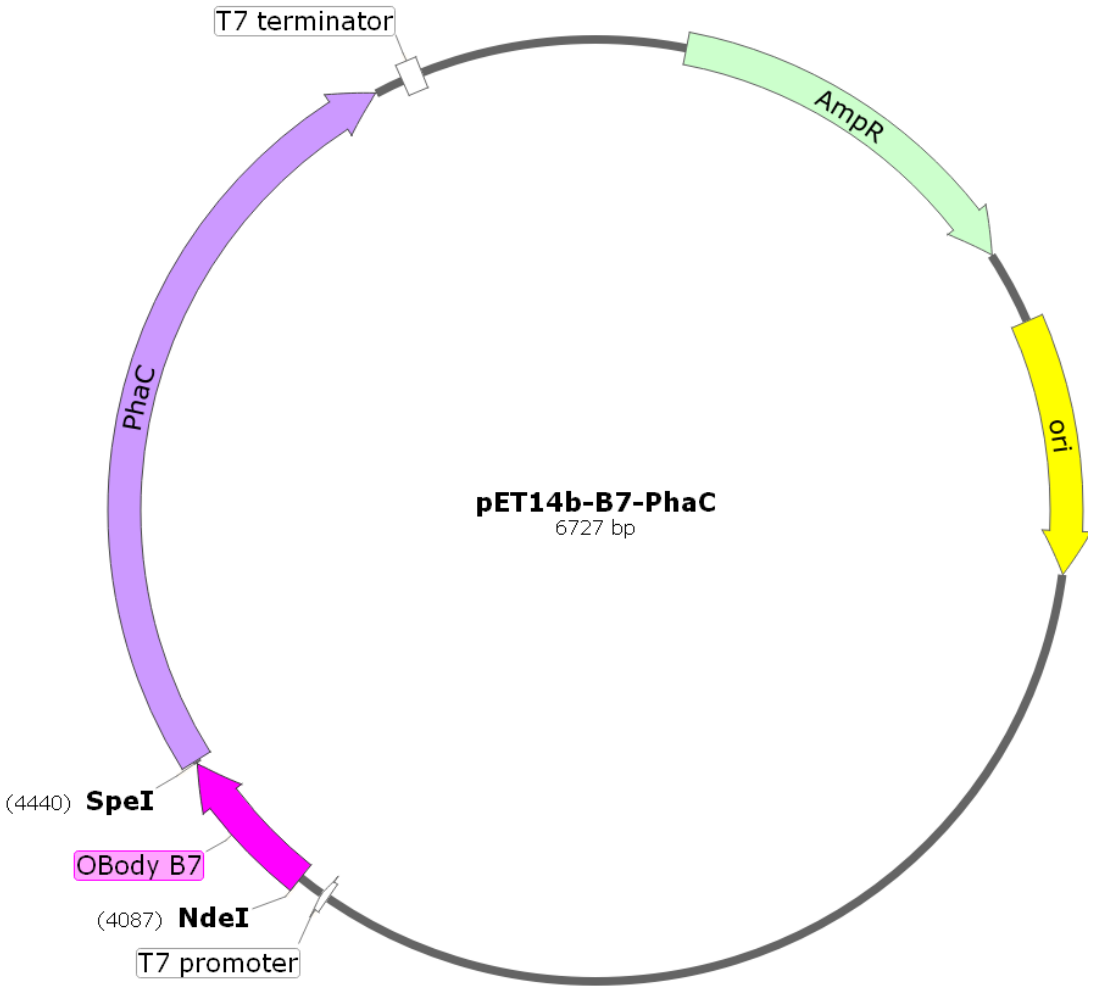



**Figure 7.19 Plasmid map for pET14b-B7-PhaC (Methods section 2.4.8).**

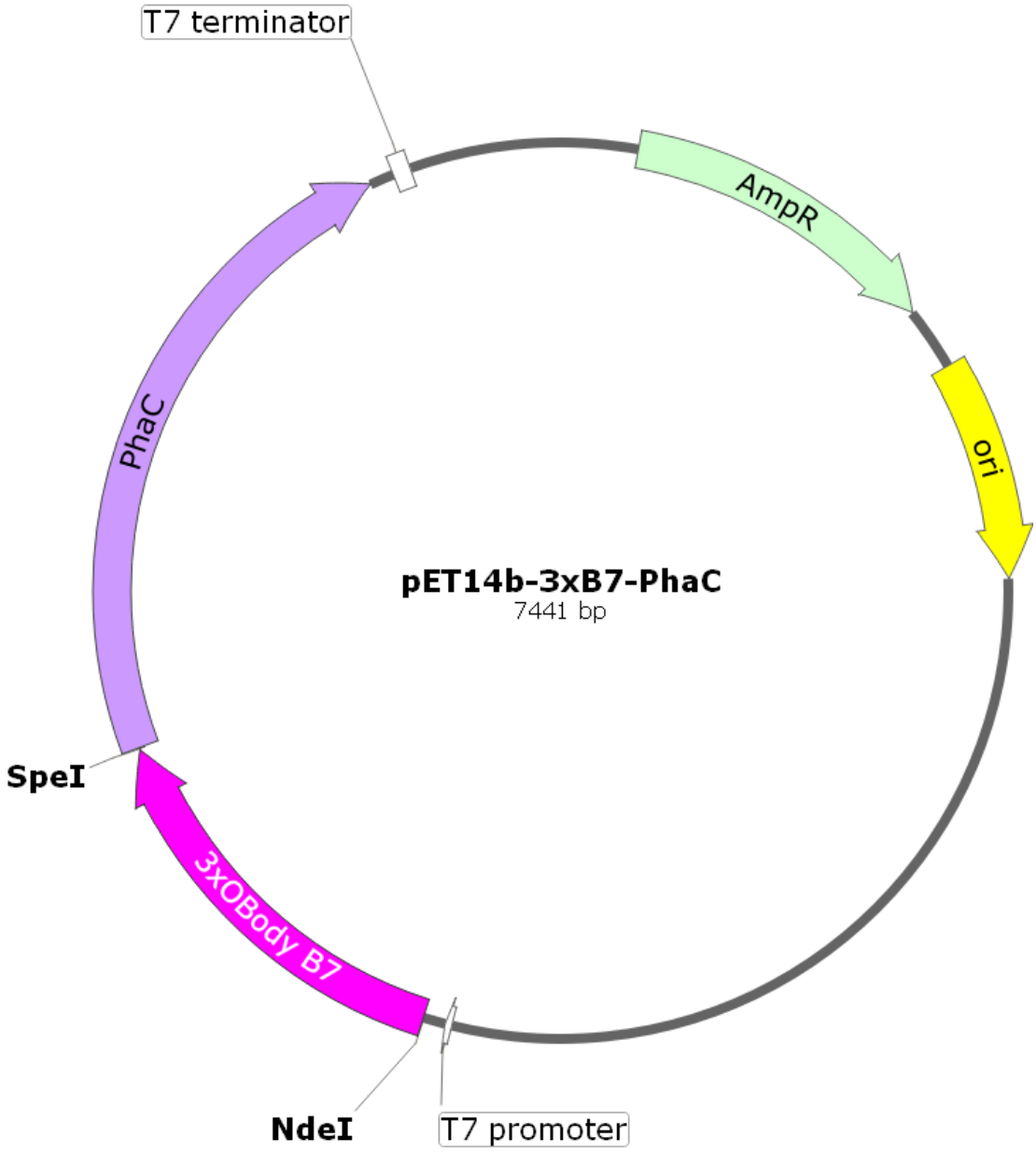



**Figure 7.20 Plasmid map for pET14b-3xB7-PhaC (Methods section 2.4.8).**

**A. Human tumour necrosis factor alpha (soluble form, 77-233 from original numbering)**

Protein sequence coverage: **90%** (142/157)

Matched peptides shown in **bold red**.

```
  1 VRSSSRTPSD KPVAHVVANP QAEGQLQWLN RRANALLANG VELRDNQLVV
 51 PSEGLYLIYS QVLFKGQGCP STHVLLTHTI SRIAVSYQTK VNLLSAIKSP
101 CQRETPEGAE AKPWYEPIYL GGVFQLEKGD RLSAEINRPD YLDFAESGQV
151 YFGIIAL
```

**B. Human Interferon alpha 2b (without SP, 24-188 from original numbering)**

Protein sequence coverage: **36%** (59/165)

Matched peptides shown in **bold red**.

```
  1 CDLPQTHSLG SRRTLMLLAQ MRRISLFSCL KDRHDFGFPQ EEFGNQFQKA
 51 ETIPVLHEMI QQIFNLFSTK DSSAAWDETL LDKFYTELYQ QLNDLEACVI
101 QGVGVTETPL MKEDSILAVR KYFQRITLYL KEKKYSPCAW EVVRAEIMRS
151 FSLSTNLQES LRSKE
```

**C. *Escherichia coli* chaperone protein DnaK**

Protein sequence coverage: **72%** (457/638)

Matched peptides shown in **bold red.**

```
  1 MGKIIGIDLG TTNSCVAIMD GTTPRVLENA EGDRTTPSII AYTQDGETLV
 51 GQPAKRQAVT NPQNTLFAIK RLIGRRFQDE EVQRDVSIMP FKIIAADNGD
101 AWVEVKGQKM APPQISAEVL KKMKKTAEDY LGEPVTEAVI TVPAYFNDAQ
151 RQATKDAGRI AGLEVKRIIN EPTAAALAYG LDKGIGNRTI AVYDLGGGTF
201 DISIIEIDEV DGEKTFEVLA TNGDTHLGGE DFDSRLINYL VEEFKKDQGI
251 DLRNDPLAMQ RLKEAAEKAK IELSSAQQTD VNLPYITADA TGPKHMNIKV
301 TRAKLESLVE DLVNRSIEPL KVALQDAGLS VSDIDDVILV GGQTRMPMVQ
351 KKVAEFFGKE PRKDVNPDEA VAIGAAVQGG VLTGDVKDVL LLDVTPLSLG
401 IETMGGVMTT LIAKNTTIPT KHSQVFSTAE DNQSAVTIHV LQGERKRAAD
451 NKSLGQFNLD GINPAPRGMP QIEVTFDIDA DGILHVSAKD KNSGKEQKIT
501 IKASSGLNED EIQKMVRDAE ANAEADRKFE ELVQTRNQGD HLLHSTRKQV
551 EEAGDKLPAD DKTAIESALT ALETALKGED KAAIEAKMQE LAQVSQKLME
601 IAQQQHAQQQ TAGADASANN AKDDDVVDAE FEEVKDKK
```

**Figure 7.21 Mass spectrometry analysis result for the therapeutic proteins as well as a minor co-purified carry-over protein obtained from beads displaying PhaC-sortase-LPETG-target protein fusions.** Published in Du & Rehm (2017a) (Appendix 7.2).

**A. *Aequorea victoria* Green fluorescent protein (without starting M)**

Protein sequence coverage: **70% (166/237)**

Matched peptides shown in **bold red**.

```
  1 SKGEELFTGV VPILVELDGD VNGHKFSVSG EGEGDATYGK LTLKFICTTG
 51 KLPVPWPTLV TTLTYGVQCF SRYPDHMKRH DFFKSAMPEG YVQERTIFFK
101 DDGNYKTRAE VKFEGDTLVN RIELKGIDFK EDGNILGHKL EYNYNSHNVY
151 IMADKQKNGI KVNFKIRHNI EDGSVQLADH YQQNTPIGDG PVLLPDNHYL
201 STQSALSKDP NEKRDHMVLL EFVTAAGITH GMDELYK
```

**B. *Mycobacterium tuberculosis* vaccine candidate Rv1626**

Protein sequence coverage: **79% (161/205)**

Matched peptides shown in **bold red**.

```
  1 MTGPTTDADA AVPRRVLIAE DEALIRMDLA EMLREEGYEI VGEAGDGQEA
 51 VELAELHKPD LVIMDVKMPR RDGIDAASEI ASKRIAPIVV LTAFSQRDLV
101 ERARDAGAMA YLVKPFSISD LIPAIELAVS RFREITALEG EVATLSERLE
151 TRKLVERAKG LLQTKHGMTE PDAFKWIQRA AMDRRTTMKR VAEVVLETLG
201 TPKDT
```

**C. The IgG binding ZZ domain of protein A derived from *Staphylococcus aureus***

Protein sequence coverage: **59% (68/116)**

Matched peptides shown in **bold red**.

```
  1 VDNKFNKEQQ NAFYEILHLP NLNEEQRNAF IQSLKDDPSQ SANLLAEAKK
 51 LNDAQAPKVD NKFNKEQQNA FYEILHLPNL NEEQRNAFIQ SLKDDPSQSA
101 NLLAEAKKLN DAQAPK
```

**Figure 7.22 Mass spectrometry analysis result for the model proteins obtained from beads displaying PhaC-intein-target protein fusions.** Published in Du & Rehm (2017b) (Appendix 7.2).

**A. Human tumor necrosis factor α** (soluble form, 77-233 from original numbering)

Protein sequence coverage: **69%** (108/157)

Matched peptides shown in **bold red**.

```
  1 VRSSSRTPSD KPVAHVVANP QAEGQLQWLN RRANALLANG VELRDNQLVV
 51 PSEGLYLIYS QVLFKGQGCP STHVLLTHTI SRIAVSYQTK VNLLSAIKSP
101 CQRETPEGAE AKPWYEPIYL GGVFQLEKGD RLSAEINRPD YLDFAESGQV
151 YFGIIAL
```

**B. Human Granulocyte colony-stimulating factor**

(short isoform without SP or VSE after the QEKL residue, 31-65, 69-207 from original numbering)

Protein sequence coverage: **18%** (32/174)

Matched peptides shown in **bold red**.

```
  1 TPLGPASSLP QSFLLKCLEQ VRKIQGDGAA LQEKLCATYK LCHPEELVLL
 51 GHSLGIPWAP LSSCPSQALQ LAGCLSQLHS GLFLYQGLLQ ALEGISPELG
101 PTLDTLQLDV ADFATTIWQQ MEELGMAPAL QPTQGAMPAF ASAFQRRAGG
151 VLVASHLQSF LEVSYRVLRH LAQP
```

**C. Human Interferon α2b** (without SP, 24-188 from original numbering)

Protein sequence coverage: **40%** (66/165)

Matched peptides shown in **bold red**.

```
  1 CDLPQTHSLG SRRTLMLLAQ MRRISLFSCL KDRHDFGFPQ EEFGNQFQKA
 51 ETIPVLHEMI QQIFNLFSTK DSSAAWDETL LDKFYTELYQ QLNDLEACVI
101 QGVGVTETPL MKEDSILAVR KYFQRITLYL KEKKYSPCAW EVVRAEIMRS
151 FSLSTNLQES LRSKE
```

**Figure 7.23 Mass spectrometry analysis result for the therapeutic proteins obtained from beads displaying PhaC-intein-target protein fusions.** Published in Du & Rehm (2017b) (Appendix 7.2).

**1. *Escherichia coli* chaperone protein DnaK**

Protein sequence coverage: **67%** (427/638)

Matched peptides shown in **bold red**.

```
  1 MGKIIGIDLG TTNSCVAIMD GTTPRVLENA EGDRTTPSII AYTQDGETLV
 51 GQPAKRQAVT NPQNTLFAIK RLIGRRFQDE EVQRDVSIMP FKIIAADNGD
101 AWVEVKGQKM APPQISAEVL KKMKKTAEDY LGEPVTEAVI TVPAYFNDAQ
151 RQATKDAGRI AGLEVKRIIN EPTAAALAYG LDKGTGNRTI AVYDLGGGTF
201 DISIIEIDEV DGEKTFEVLA TNGDTHLGGE DFDSRLINYL VEEFKKDQGI
251 DLRNDPLAMQ RLKEAAEKAK IELSSAQQTD VNLPYITADA TGPKHMNIKV
301 TRAKLESLVE DLVNRSIEPL KVALQDAGLS VSDIDDVILV GGQTRMPMVQ
351 KKVAEFFGKE PRKDVNPDEA VAIGAAVQGG VLTGDVKDVL LLDVTPLSLG
401 IETMGGVMTT LIAKNTTIPT KHSQVFSTAE DNQSAVTIHV LQGERKRAAD
451 NKSLGQFNLD GINPAPRGMP QIEVTFDIDA DGILHVSAKD KNSGKEQKIT
501 IKASSGLNED EIQKMVRDAE ANAEADRKFE ELVQTRNQGD HLLHSTRKQV
551 EEAGDKLPAD DKTAIESALT ALETALKGED KASIEAKMQE LAQVSQKLME
601 IAQQQHAQQQ TAGADASANN AKDDDVVDAE FEEVKDKK
```

**2. *Escherichia coli* Outer membrane protein A** (full length)

Protein sequence coverage: **82%** (283/346)

Matched peptides shown in **bold red**.

```
  1 MKKTAIAIAV ALAGFATVAQ AAPKDNTWYT GAKLGWSQYH DTGFINNNGP
 51 THENQLGAGA FGGYQVNPYV GFEMGYDWLG RMPYKGSVEN GAYKAQGVQL
101 TAKLGYPITD DLDIYTRLGG MVWRADTKSN VYGKNHDTGV SPVFAGGVEY
151 AITPEIATRL EYQWTNNIGD AHTIGTRPDN GMLSLGVSYR FGQGEAAPVV
201 APAPAPAPEV QTKHFTLKSD VLFNFNKATL KPEGQAALDQ LYSQLSNLDP
251 KDGSVVVLGY TDRIGSDAYN QGLSERRAQS VVDYLISKGI PADKISARGM
301 GESNPVTGNT CDNVKQRAAL IDCLAPDRRV EIEVKGIKDV VTQPQA
```

**3. *Escherichia coli* Outer membrane protein A** (without SP, 22-346 from original numbering)

Protein sequence coverage: **87%** (284/325)

Matched peptides shown in **bold red**.

```
  1 APKDNTWYTG AKLGWSQYHD TGFINNNGPT HENQLGAGAF GGYQVNPYVG
 51 FEMGYDWLGR MPYKGSVENG AYKAQGVQLT AKLGYPITDD LDIYTRLGGM
101 VWRADTKSNV YGKNHDTGVS PVFAGGVEYA ITPEIATRLE YQWTNNIGDA
151 HTIGTRPDNG MLSLGVSYRF GQGEAAPVVA PAPAPAPEVQ TKHFTLKSDV
201 LFNFNKATLK PEGQAALDQL YSQLSNLDPK DGSVVVLGYT DRIGSDAYNQ
251 GLSERRAQSV VDYLISKGIP ADKISARGMG ESNPVTGNTC DNVKQRAALI
301 DCLAPDRRVE IEVKGIKDVV TQPQA
```

**Figure 7.24 Mass spectrometry analysis result for the co-purified carry-over proteins obtained from beads displaying PhaC-intein-target protein fusions.** Published in Du & Rehm (2017b) (Appendix 7.2).

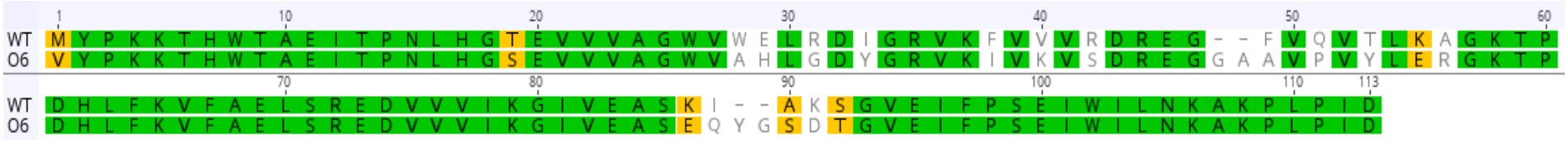



**Figure 7.25 Sequence alignment of wild type OB-fold from *P. aerophilum* aspRS (WT, residues 1-109, GenBank ID NP_558783.1) and the lysozyme binding OBody L200EP-06 (O6, sequence shown in Figure 7.10).** Identical residues are high lightened in green, similar ones in orange, and different ones are in grey without high lightening.

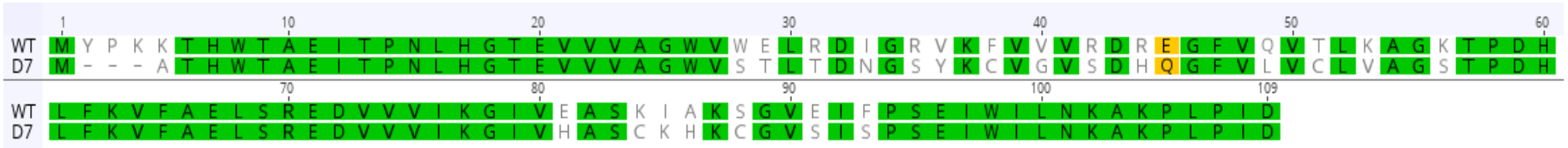



**Figure 7.26 Sequence alignment of wild type OB-fold from *P. aerophilum* aspRS (WT, residues 1-109, GenBank ID NP_558783.1) and the P4 binding OBody P4013-D7 (D7, sequence shown in Figure 7.13).** Identical residues are high lightened in green, similar ones in orange, and different ones are in grey without high lightening.

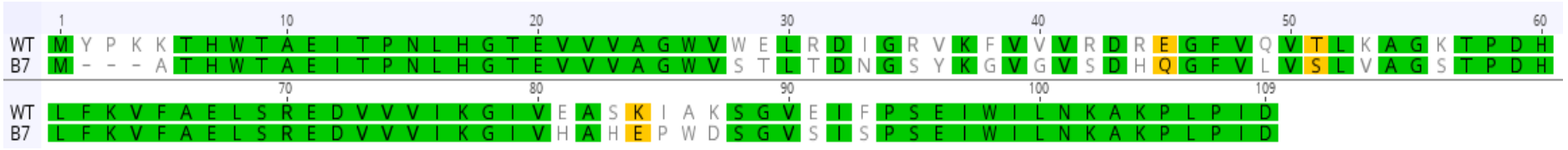



**Figure 7.27 Sequence alignment of wild type OB-fold from *P. aerophilum* aspRS (WT, residues 1-109, GenBank ID NP_558783.1) and the P4 binding OBody B7 (sequence shown in Figure 7.19).** Identical residues are high lightened in green, similar ones in orange, and different ones are in grey without high lightening.

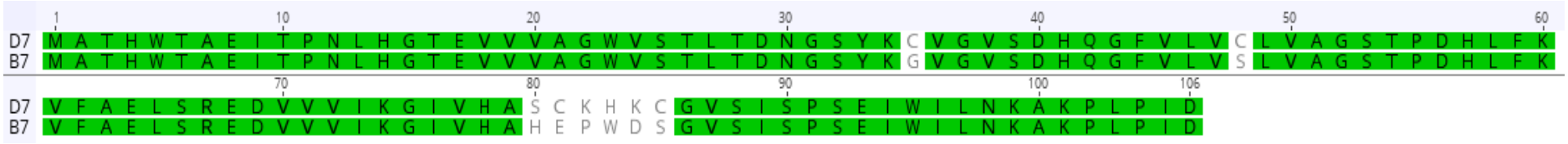



**Figure 7.28 Sequence alignment of the two generations of P4 binding D7 and B7 (sequences shown in Figures 7.13 and 7.19).** Identical residues are high lightened in green, and different ones are in grey without high lightening.

## 7.2 List of publications

The current study has contributed to the following publications:

1. Hay ID, **Du J**, Reyes PR & Rehm BHA: ***In vivo* polyester immobilized sortase for tagless protein purification.** *Microbial Cell Factories* 2015, **14:**190.
   * This paper was cited here as Hay *et al.* (2015a) in relavant figures and tables.
   *Contribution: construction of PhaC-SrtA plasmid, production and isolation of PhaC-SrtA beads, and sortase assay of PhaC-SrtA beads

2. **Du J** & Rehm BHA: **Purification of therapeutic proteins mediated by *in vivo* polyester immobilized sortase.** *Biotechnology Letters* 2017, 1-15.
   * This paper was cited here as Du & Rehm (2017a) in relavant figures and tables.
   *Contribution: designed and conducted all the experiments

3. **Du J** & Rehm BHA: **Purification of target proteins from intracellular inclusions mediated by intein cleavable polyhydroxyalkanoate synthase fusions.** *Microbial Cell Factories* 2017, **16:**184.
   * This paper was cited here as Du & Rehm (2017b) in relavant figures and tables.
   *Contribution: designed and conducted all the experiments

4. Hay ID, **Du J**, Burr N & Rehm BHA: **Bioengineering of bacteria to assemble custom-made polyester affinity resins.** *Applied and Environmental Microbiology* 2015, **81:**282-291.
   * This paper was cited here as Hay *et al.* (2015b) in relavant figures and tables.
   *Contribution: construction of PhaC-OB and OB-phaC plasmids, production and isolation of the OBody beads, and the lysozyme binding assay

5. Draper J, **Du J**, Hooks DO, Lee JW, Parlane N & Rehm BHA: **Polyhydroxyalkanoate inclusions: polymer synthesis, self-assembly and display technology.** In: Rehm BHA (ed) *Bionanotechnology biological self-assembly and its applications*. Caister Academic Press, Norfolk, 2013, pp 1-36.
   *Contribution: bead self-assembly and application in protein production

6. Hooks DO, Venning-Slater M, **Du J** & Rehm BHA: **Polyhydroyxalkanoate Synthase Fusions as a Strategy for Oriented Enzyme Immobilisation.** *Molecules* 2014, **19:**8629-8643.
   *Contribution: polyhydroxyalkanoate biobeads and *in vivo* immoboliation