Towards Development of an Immunoassay Utilizing Circularly Permutated Proteins to Detect Environmental Contaminants

by

Sara Zunnoon Khan

A Thesis
Presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Environmental Biology

Guelph, Ontario, Canada

© Sara Zunnoon Khan, August, 2013
ABSTRACT

DEVELOPMENT OF AN IMMUNOASSAY UTILIZING CIRCULARLY PERMUTATED PROTEINS TO DETECT ENVIRONMENTAL CONTAMINANTS

Sara Zunnoon Khan
University of Guelph, 2013

Advisor:
Dr. J. Christopher Hall

A fusion protein composed of antibody fragments and β-lactamase was earlier created by Kojima et al. (2011), with antigen specificities against a bone disease marker and a pesticide. The enzyme was circularly permutated and fused to the variable heavy and light chain antibody fragments, thereby ensuring inactivity until binding of the target antigen triggered enzyme activation. Upon activation, the β-lactamase produced a colorimetric signal, which indicated antigen presence. In this work, a similar strategy was used to create two novel fusion proteins composed of circularly permuted β-lactamase and superfolder green fluorescent protein with anti-benzo[a]pyrene variable antibody fragments. The fusion proteins were designed and expressed in E. coli for the development of a single-step visual immunoassay. It was hypothesized that the cp reporter proteins would be activated once the binding of B[a]P to the variable antibody fragments occurred, and this interaction was expected to produce a detectable colorimetric or fluorescent signal. Although positive results were obtained in one instance, substantial supportive evidence in favour of the hypothesis could not be obtained.
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all those who helped me in reaching this goal. To my advisor J. Christopher Hall for giving me this opportunity to attend graduate school, and allowing me the freedom to make mistakes and learn from them while relaying utmost patience and invaluable guidance. I will always appreciate our candid conversations and your non-PC sense of humour. Thank you to my advisory committee members Drs. Roger Mackenzie and Jamshid Tanha for their direction, advice and expertise in the navigation of this challenging project.

To my mentors Dr. Roberto Michelutti and Dr. Yaima Arocha-Rosete for believing in me, convincing me to pursue a Masters degree and for being genuinely interested in my future. I would not have made it this far without either of you.

Thank you to members of the Hall lab for their help and support. I would like to especially thank Dr. Haifeng Wang for his invaluable teachings, non-human cloning speed and for being my Human Google—you helped me overcome many obstacles in the lab.

Thank you to my friends in Guelph who understood and accepted my workaholic nature, and dragged me away from work to keep me sane when I did not know how to stop. Special thanks to Erin Gilbert for being a gracious friend. Thanks also to my friends in T.O. for their endless support, selfless love and understanding--thanks for being my weekend getaways, and keeping me “long distance company” especially when I did my overnight experiments--even when I went off grid Nee Nee, you always found a way to reach me.

Thank you to my family for their invaluable emotional and financial support. To my siblings Zara and Haris for always being just a phone call away--your zest for life is infectious. To
Sumera and my brother-in-law Arslan for always being the first ones to visit regardless of where I move (my next address is the moon, get your space suits ready!). Thank you sis for taking my calls at the most inconvenient hours and willing to talk me through those bouts of irrationality.

I would like to thank Mr. Anwer Zunnoon Khan and Mrs. Shaheena Anwer for their immeasurable sacrifices, endless support, encouragement, patience and understanding. You have been my strongest supporters through everything, believed in me more than I believe in myself, and resolutely stood by me while I treads on non-traditional paths in life. I feel blessed and privileged to call myself your daughter.

Last but not least, I would like to thank my chota sa jaanu Rayya for being a true embodiment of her name, and reminding me why life is beautiful every time I lost perspective.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................................................................... iii
LIST OF ABBREVIATIONS ............................................................................................................... vi
LIST OF FIGURES .......................................................................................................................... ix
LIST OF TABLES ............................................................................................................................ xi

i) GENERAL OVERVIEW
  1.1: Introduction ................................................................................................................................. 1
  1.2: Hypotheses & Research Objectives .............................................................................................. 2

ii) LITERATURE REVIEW
  2.1: Circular permutation ....................................................................................................................... 4
  2.2: Reporter Proteins
      2.2.1: Beta-lactamase ....................................................................................................................... 11
      2.2.1.1: Circular permutation of beta-lactamase ............................................................................. 12
      2.2.2: Green Fluorescent Protein ..................................................................................................... 14
      2.2.2.1: Circular permutation of GFP .............................................................................................. 17
  2.3: Restatement of research objectives .............................................................................................. 23
  2.4: Overview of Benzo[a]pyrene ....................................................................................................... 24

iii) MATERIALS AND METHODS .................................................................................................... 27

iv) ANTI-B[a]P cp-BLA FUSION PROTEIN
  4.1: Results ........................................................................................................................................ 35
  4.2: Discussion .................................................................................................................................... 38
  4.3: Conclusion ................................................................................................................................... 41

v) ANTI-B[a]P cp-GFP FUSION PROTEIN
  5.1: Results ........................................................................................................................................ 42
  5.2: Discussion .................................................................................................................................... 49
  5.3: Conclusion ................................................................................................................................... 59

vi) FUTURE RESEARCH ..................................................................................................................... 60

vii) LITERATURE CITED .................................................................................................................... 63

viii) APPENDIX
     Appendix I Anti-B[a]P cp-BLA ........................................................................................................... 69
     Appendix II Anti-B[a]P cp-GFP ......................................................................................................... 70
     Appendix III cp-GFP sequence ......................................................................................................... 71
     Appendix IV sf-GFP sequence ........................................................................................................... 72
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Victoria</td>
<td><em>Aequorea Victoria</em></td>
</tr>
<tr>
<td>ALAS</td>
<td>5-aminolevulinate synthase</td>
</tr>
<tr>
<td>anti-B[a]P cp-BLA</td>
<td>fusion protein comprising anti-B[a]P V&lt;sub&gt;H&lt;/sub&gt; and V&lt;sub&gt;L&lt;/sub&gt; domains fused to circularly permutated BLA</td>
</tr>
<tr>
<td>anti-B[a]P cp-GFP</td>
<td>fusion protein comprising anti-B[a]P V&lt;sub&gt;H&lt;/sub&gt; and V&lt;sub&gt;L&lt;/sub&gt; domains fused to circularly permutated sf-GFP</td>
</tr>
<tr>
<td>B[a]P</td>
<td>benzo[a]pyrene</td>
</tr>
<tr>
<td>Btk</td>
<td>Bruton’s tyrosine kinase</td>
</tr>
<tr>
<td>Btk-cpGFP</td>
<td>fusion protein with cp-GFP inserted into Btk</td>
</tr>
<tr>
<td>BPTI</td>
<td>bovine pancreatic trypsin inhibitor</td>
</tr>
<tr>
<td>BLA</td>
<td>β-lactamase</td>
</tr>
<tr>
<td>Ca&lt;sup&gt;2+&lt;/sup&gt;</td>
<td>calcium</td>
</tr>
<tr>
<td>CaM</td>
<td>calmodulin</td>
</tr>
<tr>
<td>CFP</td>
<td>cyan fluorescent protein</td>
</tr>
<tr>
<td>Con A</td>
<td>concanavalin A</td>
</tr>
<tr>
<td>Cp</td>
<td>circular permutation</td>
</tr>
<tr>
<td>cp-BLA</td>
<td>circularly permutated β-lactamase</td>
</tr>
<tr>
<td>cp-GFP</td>
<td>circularly permutated superfolder green fluorescent protein</td>
</tr>
<tr>
<td>DHFR</td>
<td>dihydrofolate reductase</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>EGFP</td>
<td>enhanced GFP</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>ePRAI</td>
<td>phosphoribosyl anthranilate isomerase</td>
</tr>
<tr>
<td>G-CaMP</td>
<td>calcium probe based on GFP</td>
</tr>
<tr>
<td>GC-MS</td>
<td>gas chromatography–mass spectrometry</td>
</tr>
<tr>
<td>HPLC</td>
<td>high-performance liquid chromatography</td>
</tr>
<tr>
<td>IARC</td>
<td>International Agency for Research on Cancer</td>
</tr>
<tr>
<td>IC₅₀</td>
<td>half-maximal inhibitory concentration</td>
</tr>
<tr>
<td>IMAC</td>
<td>immobilized metal-affinity chromatography</td>
</tr>
<tr>
<td>Ins(1,3,4,5)P₄</td>
<td>D-myoinositol-1,3,4,5-tetrakisphosphate</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>M13</td>
<td>fragment of myosin light chain kinase</td>
</tr>
<tr>
<td>MBP</td>
<td>maltose binding protein</td>
</tr>
<tr>
<td>MTX</td>
<td>methotrexate</td>
</tr>
<tr>
<td>NADPH</td>
<td>nicotinamide adenine dinucleotide phosphate</td>
</tr>
<tr>
<td>OD₆₀₀</td>
<td>optical density at 600 nm</td>
</tr>
<tr>
<td>PAH</td>
<td>polycyclic aromatic hydrocarbon</td>
</tr>
<tr>
<td>pelB</td>
<td>pectate lyase B from <em>Erwinia caratovora</em></td>
</tr>
<tr>
<td>PH</td>
<td>pleckstrin homology</td>
</tr>
<tr>
<td><em>Pseudomonas</em> exotoxin</td>
<td>exotoxin produced by <em>Pseudomonas aeruginosa</em></td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
</tbody>
</table>
rTHS recombinant single subunit thermosome from *Methanocaldococcus jannaschii*

scFv single-chain variable fragment antibody

SDS-PAGE sodium dodecyl sulfate polyacrylamide gel electrophoresis

sf-GFP superfolder GFP

$V_H$ variable domain of heavy chain of antibody

$V_L$ variable domain of light chain of antibody

wtGFP wild-type GFP

YFP yellow fluorescent protein

$Zn^{2+}$ zinc
## LIST OF FIGURES

| FIGURE 1.1: | Schematic of the hypothesis, illustrating functioning of the proposed fusion proteins consisting of (A) cp-GFP and (B) cp-BLA | 3 |
| FIGURE 2.1: | Schematic representation of circular permutation | 5 |
| FIGURE 2.2: | Ribbon diagram of TEM-1 BLA structure | 11 |
| FIGURE 2.3: | Schematic representation of the sequence of RG13 including linkers | 13 |
| FIGURE 2.4: | GFP structure | 15 |
| FIGURE 2.5: | Schematic representation of the sf-GFP | 17 |
| FIGURE 2.6: | Schematic representation of the fusion protein created with GFP and antiparallel leucine zippers joined via linkers | 18 |
| FIGURE 2.7: | Schematic representation of the sf-GFP insertion construct | 20 |
| FIGURE 2.8: | Schematic representation of Btk-cpGFP construct | 20 |
| FIGURE 2.9: | Schematic representation of the calcium probe with cp-GFP variant | 22 |
| FIGURE 2.10: | Schematic representation of the expression plasmid for anti-osteocalcin / imidacloprid cp-BLA | 23 |
| FIGURE 2.11: | Chemical structure of benzo[a]pyrene | 24 |
| FIGURE 3.1: | Schematic representation of the expression vector pIT2-SZ1 | 27 |
| FIGURE 3.2: | Schematic representation of the expression vectors pITFK-cpGFP and pIT2-SZ2 | 28 |
| FIGURE 4.1: | Expression of soluble anti-B[a]P cp-BLA protein (A) 12% SDS PAGE gel (B) Immunoblot | 35 |
FIGURE 4.2: Immunoblot visualizing purified anti-B[a]P cp-BLA protein
FIGURE 4.3: Preliminary assay performed with both fusion proteins and 1 µg/ml B[a]P
FIGURE 4.4: A simple schematic of a typical β-lactam interacting with class A TEM-1 β-lactamase
FIGURE 5.1: Immunoblot visualizing purified control cp-GFP protein
FIGURE 5.2: Immunoblot visualizing purified anti-B[a]P cp-GFP fusion protein
FIGURE 5.3: In vivo fluorescence assay
FIGURE 5.4: Mean fluorescence of crude anti-B[a]P cp-GFP and cp-GFP protein in sonication lysate
FIGURE 5.5: In vitro assay utilizing E1 anti-B[a]P cp-GFP protein A) mean fluorescence and B) time-course graph of fluorescence
FIGURE 5.6: In vitro assay utilizing E2 anti-B[a]P cp-GFP protein A) mean fluorescence and B) time-course graph of fluorescence
FIGURE 5.7: B[a]P concentration-response curve
FIGURE 5.8: A schematic representation of the intermolecular self-assembly of anti-B[a]P cp-GFP fusion protein
FIGURE 5.9: Indirect competitive ELISA of B[a]P-specific scFvs
LIST OF TABLES

TABLE 1: Preparation of B[a]P standards 34

TABLE 2: Results from the nitrocefin cross-reactivity assay 38
1. GENERAL OVERVIEW

1.1. Introduction

Hazardous contaminants can persist long-term in various environmental matrices like air, water, and soil, with the ability to migrate from one matrix to another (El-Shahawi et al. 2010). These contaminants include chemical pollutants like the polyaromatic hydrocarbons (PAHs), which are simultaneously detrimental to human health and the environment. Consequently, the concern of human exposure to these contaminants has led to the development of detection methods to identify their presence before reaching a harmful level of accumulation. Some existing methods include the enzyme-linked immunosorbent assay (ELISA), high performance liquid chromatography, and mass spectroscopy among others. Each is applicable for detection depending on the nature of the contaminant, and each has its associated shortcomings such as high cost of equipment and/or labour, time for completion of analysis, and complexity. Thus, there is a search for improved methods.

The aim of this thesis was to develop a detection method to detect an environmental contaminant utilizing fusion proteins, composed of variable heavy and light chain domains of an antibody and a circularly permutated reporter protein, in a single-step visual immunoassay. The PAH benzo[a]pyrene (B[a]P) was selected as the antigen, and circularly permutated forms of the bacterial enzyme β-lactamase along with a variant of GFP, the superfolder green fluorescent protein, were used as reporters.
1.2. Hypotheses and Research Objectives

Hypotheses:

i) The interaction of an antigen and variable heavy and light chain domains of an antibody, will activate the fused circularly permutated proteins.

ii) This interaction should subsequently result in a chromogenic or fluorescent signal, thereby indicating the presence of the target antigen (Figure 1A, 1B).

Research Objectives:

i) To design fusion proteins using DNA sequences of circularly permutated β-lactamase and superfolder GFP, with the variable heavy and light chain domains of an anti-B[a]P scFv.

ii) To express the anti-B[a]P cp-GFP and anti-B[a]P cp-BLA fusion proteins in *Escherichia coli*.

iii) To utilize the fusion proteins in a single-step visual immunoassay.
Figure 1.1: Schematic of the hypothesis, illustrating functioning of the proposed fusion proteins.

The addition of benzo[a]pyrene will bring $V_H$ and $V_L$ in closer proximity, which should activate the cp fragments of reporter proteins, causing a stabilization and conformational change in the fusion protein.

A) When the reporter protein is cp-GFP, excitement with UV light at 490 nm will cause emission at 510 nm.

B) When the reporter protein is cp-BLA, addition of its substrate nitrocefin will change the solution colour from yellow to red.
2. LITERATURE REVIEW

2.1. Circular Permutation

Circular permutation (cp) is the rearrangement of the amino acid sequence of a protein, resulting in a change in the sequence order, without a change in the number and type of amino acids. In other words it is an intramolecular relocation of the protein’s N- and C- termini. Cp can exist naturally by the occurrence of gene duplication, or can be introduced artificially by protein engineering techniques where the amino and carboxyl termini are linked via a short covalent linker to circularize the protein, and a different location is cleaved generating a protein with new N- and C- termini with a different order of sequence (Figure 2.1; Yu and Lutz 2011). It is common to select the location of the new termini to be at the surface loops of the protein, to ensure that the secondary and tertiary structures of the new protein are not dramatically affected and that a functional protein is still produced (Pan and Uhlenbeck 1993). It is also important to consider the proximity of the original termini in the native protein, which should be close to each other for a successful cp to take place (Topell and Glockshuber 2002).

Construction of a cp protein with these two requirements ensures that the cp protein not only folds into a similar conformation as the native protein but as well as retain its (native protein) activity (Goldenberg 1989). Proteomic studies (structure, folding, and conformational stability) comprising of at least 20 different circularly permutated proteins have reported a substantial number of successful circular permutations, and implicit in these successfully engineered proteins is the idea that most protein termini do not contain information for tertiary structures (Heinemann and Hahn 1995; Rojas et al. 1999; Topell and Glockshuber 2002).
The first example of a naturally occurring circularly permutated protein was discovered in 1979 by Cunningham *et al.* through comparing the sequence arrangement of lectins, favin and concanavalin A (Con A), and concluding the two proteins to be similar. Since then several examples of naturally occurring and artificially created circular permutations of proteins have been observed. Although Goldenberg and Creighton (1983) first reported an artificially created circularly permutated protein, i.e., the bovine pancreatic trypsin inhibitor (BPTI), it was not genetically engineered. Luger *et al.* (1989) reported the first genetically engineered artificial circular permutation in a single-domain βα barrel enzyme from *E. coli*. Two variants of the gene for this enzyme phosphoribosyl anthranilate isomerase (ePRAI), were created according to the criteria for constructing a functional cp protein, i.e., the N- and C- termini of the original protein were in close proximity (so they could be easily covalently linked), and the new site of cleavage

---

**Figure 2.1: Schematic representation of circular permutation.** To transform native protein into a permutated form, the amino and carboxyl termini are covalently linked via a short linker, followed by cleavage at a new position within the sequence, thus creating a protein with a new sequence order. Reprinted/adapted from Trends in Biotechnology, 29, Yu and Lutz, Circular permutation: a different way to engineer enzyme structure and function, 18-25, © 2011, with permission from Elsevier.
was in the surface loops. Both cp versions of ePRAI were enzymatically active, although one had activity equivalent to wild-type, while the second had lower activity due to the intentional disruption of the active site during the cp process. This study revealed information about protein folding as it was shown that the three-dimensional structure of the original protein could be attained through a different folding pathway, one in which the native amino and carboxyl termini or surface loops were not necessary.

In addition to studying the obvious effects of circular permutation on proteins, such as the effect of secondary structures on tertiary structure (Osuna, Pérez-Blancas, and Soberón 2002), function of specific residues (whether they comprise the active site) (Luger et al. 1989; Mathieu, Fastrez, and Soumillion 2010), or overall folding kinetics (more than one folding pathway, more or less stable pathways, unusual folding pathway intermediates) (Bulaj, Koehn, and Goldenberg 2004), several other interesting cases have been reported, including the use of circular permutation to reduce the susceptibility of a protein to proteolysis (Whitehead, Bergeron, and Clark 2009). A few such examples displaying the utilization of circular permutation in proteins, along with their conclusions and implications are discussed below.

Although the implicit goal of using circular permutation to study a protein is to potentially add to it an enhancement (structural, activity, folding), in reality such an attempt can either impart a negative or a positive effect. Bulaj et al. (2004) studied the folding efficiency of the circularly permutated bovine pancreatic trypsin inhibitor (cp-BPTI), through the alteration of the disulfide-coupled folding pathway of BPTI. Four variants of cp-BPTI were created and the kinetics of the folding and unfolding of each cp protein was compared. All four of the cp-BPTI were found to be less stable than the wild-type due to an enhanced rate of
protein unfolding, since an alternate folding pathway was favoured over the wild-type pathway. Even though a BPTI with enhanced activity could not be designed, knowledge about the importance of the specific interactions of residues and intermediates of the disulfide-coupled folding pathway was attained.

In contrast, Cheltsov et al. (2001) decided to not only study a protein’s folding and structure, but also the role of the polypeptide chain in function. They circularly permutated the 5-aminolevulinate synthase (ALAS), an enzyme that regulates the first reaction in the heme biosynthetic pathway of non-plant eukaryotes and some bacteria. Several circularly permuted ALAS variants were constructed, including some with disrupted residues normally involved in the forming of important secondary structures. These disruptions were expected to affect the assembly of the binding site of ALAS. Yet these cp-ALAS variants retained catalytic activity and showed no detrimental effects on the folding and assembly of subunits of the active site. Cheltsov et al. concluded that although the disruptions altered the sequence of formation of secondary structure elements, this alteration did not prevent the formation of the same overall tertiary structure as the wild-type ALAS. Therefore, this example of cp-ALAS showed that in some proteins, the native arrangement of a polypeptide chain is considerably resistant to significant modifications, and that the structure of a functional protein could be achieved through more than one folding pathway.

In a more interesting example involving receptor and ligand binding, Smith and Matthews (2001) circularly permutated dihydrofolate reductase (DHFR) of E. coli at position 86, a monomeric protein containing a binding domain for adenosine. The cp-DHFR was found to have compromised enzymatic activity, but was equally stable as wild-type, albeit with an
altered secondary structure. Upon addition of substrate (dihydrofolate) and cofactor (NADPH),
the enzymatic activity was restored to 50% of the original DHFR. This was due to a change in
the secondary structure of the binding domain induced upon the addition of substrate and
cofactor. It was hypothesized that although the cp-DHFR was quite different than the wild-type
with respect to secondary structure, the enzymatic activity was restored through a
conformational change which resembled the wild-type structure. The addition of a tight-
binding, inhibitory ligand methotrexate (MTX) confirmed the hypothesis, as the secondary and
three dimensional structures that formed upon MTX addition restored cp-DHFR to a native-like
structure. This work emphasized how the ligand-protein interaction of a modified protein can
stabilize a complex and restore the 3-D structure of the wild-type.

In contrast to the structural and functional studies discussed above, Whitehead et al.
(2009) used circular permutation to prevent proteolysis of a full-length recombinant protein.
The molecular chaperone thermosome (rTHS) from the archaeon Methanocaldoccoccius
jannaschii is highly prone to proteolysis when expressed in soluble form in E. coli, especially at
the 70 residues at the N-terminus which are essential for protein refolding. Upon cleavage of
these residues, a change in conformation results which makes the protein more sensitive to
proteolysis. Whitehead et al. circularly permutated rTHS by splitting the sequence between
residues 146 and 147, and expressing it in three different E. coli strains. In comparison to the
wild-type, the active form of the circularly permutated thermosome had increased expression
as well as decreased proteolysis. Circular permutation led to the occurrence of a different
folding pathway with the formation of conformational states that were not seen in wild-type,
and made cp-rTHS less susceptible to degradation.
In addition to basic functional and structural studies, circular permutation has the potential to be used in more unique applications. For instance, the creation of fusion proteins is usually done via attachment at the N- or C- terminus, which in some cases leads to decreased activity if the native proteins require their termini to be free of steric hindrances for full activity. Therefore, not only the affinity but also the folding of either of the fusion partners can be affected when using this method. Optimization is based on switching the order of proteins, mutating one of the proteins to decrease the detrimental effects caused by fusion, or modifying the linkers between the fusion partners (Edwards et al. 1989; Brinkmann, Buchner, and Pastan 1992; Kreitman, Puri, and Pastan 1994). In more recent endeavours, circular permutation has been used as an alternate protein engineering technique to create fusion proteins that can function as biomolecular protein switches.

A protein switch is a protein that responds to a stimulus (e.g., ligand concentration, temperature or pH change, covalent modification) by switching between at least two conformations and displaying a signal (ligand affinity, enzyme activity, fluorescence). Although protein switches exist naturally (e.g., allosteric regulators, transcription factors), cp can be used to generate switches that can be used as molecular and biological sensors (Guntas, Mitchell, and Ostermeier 2004). Such a protein switch can be created either by inserting the protein of interest within the linker region of the circularly permutated protein, or by inserting the circularly permutated protein into the linker region of the protein of interest. For optimization of the resulting fusion protein, further aspects can be modified such as the linker (length, sequence) and site selected for circular permutation in the permuted protein, or the site of
insertion within the cp or protein of interest, as well as the linkers (length and specific residues) fusing the cp and protein of interest together (Stratton and Loh 2011).

The circular permutation of bacteriophage T4 lysozyme between positions 36 and 37 joined via six-residue linker was the first example of cp being applied to a two domain protein. The permutation did not cause a dramatic change in protein structure or activity, thus it folded efficiently and retained cleavage specificity equivalent to wild-type. In addition, similar to previous studies it was seen that the termini are not essential in protein folding and structure formation (Zhang et al. 1993). A year later, Kreitman et al. (1994) designed a fusion protein with a circularly permutated interleukin 4 (IL4; growth factor / immune-response protein) and Pseudomonas exotoxin, which possessed higher antitumor activity than the control protein, i.e., IL4 with a C-terminal exotoxin fusion. Previous studies (Le et al. 1991; Ramanathan et al. 1993) had shown that the C-terminus of IL4 played an important role in binding. Thus the circular permutation of IL4 in the fusion was such that it allowed the residues that play a role in binding (carboxyl terminus of native protein) to be free of steric hindrance and remain functional, while the exotoxin was fused to the new C-terminus. Due to this decrease in steric interference from the exotoxin, the antitumor activity of the fusion was ten-fold higher as the IL4 bound more efficiently to its receptor.

For a protein switch to function as a biosensor, the choice and use of appropriate reporter proteins is essential. Two reporter proteins that have been used extensively in several in vitro and in vivo studies are β-lactamase and the green fluorescent protein.
2.2. Reporter Proteins

2.2.1. Beta-lactamase

β-lactamases (BLA) are bacterial enzymes present in Gram-negative and Gram-positive bacteria allowing them to be resistant to antibiotics such as penicillins, cephalasporins and similar analogs with a β-lactam ring in their structure. This resistance is due to the BLA catalyzing a hydrolysis reaction of the β-lactam ring that is found within these compounds. The TEM-1 BLA (29 kDa; Figure 2.2) is encoded by the Gram-negative bacteria *E. coli*, and the genetic sequence is commonly found in plasmids as a selectable marker that confers resistance against ampicillin (Fonzé et al. 1995).

![Ribbon diagram of TEM-1 BLA structure](image)

**Figure 2.2: Ribbon diagram of TEM-1 BLA structure.** Serine at position 70 is a crucial nucleophile in the active site, required for catalysis. (PDB: 1zg4). From Acta Cryst, D61, Stec et al., Structure of the wild-type TEM-1β-lactamase at 1.55 Å and the mutant enzyme Ser70 Ala at 2.1 Å suggest the mode of non-covalent catalysis for the mutant enzyme, 1072-1079, Reproduced with permission of the International Union of Crystallography, © 2005.

β-lactamases are classified into four molecular classes ranging from A to D depending on their amino acid sequence, substrate specificity and response to inhibitors. TEM-1 BLA belongs
to one of the most common molecular classes of β-lactamases, class A, that in addition to C and D include enzymes possessing a serine residue at the active site (Jiang, Xing, and Rao 2008). The nucleophilic serine is involved in the formation of a covalent intermediate (acyl-enzyme) upon addition of BLA substrate such as antibiotics or synthetic compounds (Stec et al. 2005).

There are various fluorogenic and chromogenic substrates for BLA which enable visualization of the hydrolysis reaction by fluorescence or a colour change, respectively. Nitrocefin is one of the most commonly used substrates that results in a colour change from yellow to red, thus visualizing the hydrolysis reaction (Jiang, Xing, and Rao 2008). Several properties of this enzyme, such as its small size, non-toxicity to cells, and sensitivity have made it an ideal candidate to be used as a biological reporter (Philippon et al. 1998) and these qualities have also made it readily applicable for the creation of fusion proteins.

2.2.1.1. **Circular permutation of β-lactamase**

Although BLA has been circularly permutated by itself e.g., TEM-1 β-lactamase was circularly permutated and the two domains were linked by a hinge region that allowed allosteric regulation upon binding of zinc and nickel ions (Mathieu et al. 2010), more complex studies have used cp-BLA with other proteins of interest to create a fusion protein that could serve as a biosensor / protein switch. An interesting example is the RG13 protein.

Guntas et al. (2004) reported the creation of a molecular switch, RG13, in which a TEM-1 cp-BLA was inserted into the maltose binding protein (MBP) of *E. coli* (Figure 2.3). In the absence of maltose, cp-BLA activity was negatively affected while in the presence of maltose hydrolysis activity increased 25-fold. RG13 expressed in *E. coli* also conferred a unique
phenotype of maltose-dependent resistance to ampicillin. MBP contains two domains linked by a hinge region, and at the interface of the two domains is the maltose binding site. This causes MBP to exist in an open form in the absence of maltose, and in a closed form in the presence of maltose (Sharff et al. 1992). This natural conformational change of MBP was used in RG13, as the structural change caused by maltose binding was used to induce a conformational change in the active site of cp-BLA. Once cp-BLA was fully functional it hydrolyzed the β-lactam ring in ampicillin making the antibiotic ineffective against *E. coli*. In further work, Ke et al. (2012) reported that although RG13 was positively regulated by maltose, it was negatively regulated by zinc (Zn$^{2+}$). Through structural analyses it was revealed that the zinc ion bridged the two linked domains of MBP and TEM-1 cp-BLA, which were joined by two linkers. The positioning of zinc disrupted one of the linkers which is a part of the BLA active site, and displaced it from between the MBP and cp-BLA domains, thus inactivating the BLA active site. The results of this study indicated that the site of linker attachment on the protein of interest, the residues comprising the linker, and the linker length play a critical role in creating a fusion protein that

**Figure 2.3: Schematic representation of the sequence of RG13 including linkers.**

MBP = Maltose binding protein; BLA = beta-lactamase. Numbers in parentheses indicate the number of amino acid in the sequence. Reprinted/adapted from Chemistry & Biology, 11, Guntas, Mitchell, and Ostermeier, A Molecular Switch Created by In Vitro Recombination of Nonhomologous Genes, 1483–1487, © 2004, with permission from Elsevier.
retains the native properties of both fusion partners; this is due to the fact that the structural changes / steric hindrances / and proximity of the fused domains ultimately determine the function of the new protein.

Although BLA is a suitable reporter in many cases, especially where sensitivity is essential (i.e., due to catalytic activity), one of its drawbacks is that it needs a substrate to visualize the reaction. This makes it more complicated to use, especially when experiments are performed in vivo. Therefore, an appropriate alternative is the green fluorescent protein, which requires no substrate.

2.2.2. Green Fluorescent Protein

The green fluorescent protein (GFP) is another common reporter used to study gene expression, protein localization, and visualization of whole cells or cellular compartments (Müller-Taubenberger and Anderson 2007). It was first discovered in 1962, in the extract of jellyfish Aequorea Victoria by Shimomura et al. The wild-type protein has two fluorescence excitation peaks, a major peak at 395 nm, and a minor one at 475 nm (Ormö et al. 1996), while the emission peak is at 508 nm (Morin and Hastings 1971). The native GFP (Figure 2.4) folds into an 11-stranded β-barrel (β-can) where each β-strand contributes to the structural integrity of the barrel (Topell and Glockshuber 2002), which surrounds the internal fluorophore formed by residues Ser65-Tyr66-Gly67 (Baird, Zacharias, and Tsien 1999). The fluorophore is also surrounded by two cysteine residues, C49 and C71 (Aronson, Costantini, and Snapp 2011).

One major characteristic of wild-type GFP (wtGFP) that makes it ideal for proteomic studies is that it does not need an enzymatic substrate or cofactor to be fully functional, since
Figure 2.4: GFP structure. Left: Ribbon diagram of the folded β-barrel structure (chromophore on the inside). Right: Topographical depiction of the GFP structure displaying arrangement of aa and N- and C- termini. 


the fluorophore forms autocatalytically (Topell and Glockshuber 2002). On the other hand there are also important disadvantages, arising especially when wtGFP is used as a marker in heterologous expression hosts. The protein is known to express poorly in some hosts, and is also highly sensitive to misfolding when fused to an insoluble or poorly folding partner. In addition, since A. victoria is found in the cold natural habitat of the Pacific Northwest, the mature GFP therefore folds efficiently at temperatures below 37 °C (optimum temperature ≤ 25°C), making it prone to folding inefficiently when expressed at ≥ 37 °C (the temperature at which one of the most common expression hosts, E. coli, is incubated; Jiang, Xing, and Rao 2008; Stepanenko et al. 2008). To circumvent some of these issues, many variants of GFP have been created such as the enhanced GFP (EGFP) for brighter fluorescence and better folding at
37°C (Cormack, Valdivia, and Falkow 1996), S65T-GFP for increased fluorescence and longer excitation (489 nm) wavelength (Heim and Tsien 1996), and other colour variants like the blue-shifted mutant P4 (Y66H) (Heim, Prasher, and Tsien 1994) Cyan Fluorescent Protein (CFP) and Yellow Fluorescent Protein (YFP) / Venus (Müller-Taubenberger and Anderson 2007; Nagai et al. 2002). Recently, Pédelacq et al. (2006) engineered a superfolder variant of GFP (sf-GFP) that folds more robustly, has higher emission and excitation spectra (excitation 488 nm; emission 510 nm), and most importantly is more tolerant of circular permutation. It is also one of the fastest folding GFPS reported to date. The sf-GFP (Figure 2.5) contains eleven mutations in total, in comparison to the wild-type GFP; F99S, M153T, V163A, F64L, S65T, S30R, Y39N, N105T, Y145F, I171V and A206V. Six of the above mutations, i.e., S30R, Y39N, N105T, Y145F, I171V and A206V are unique to sf-GFP (others have already been reported to exist in the ‘folding reporter’ variant) and participate in improving protein folding and resistance against urea denaturation. The improved qualities of sf-GFP have readily made it available for application in studies previously thwarted by limitations of earlier GFP variants. For instance, due to its robust nature, sf-GFP has been used to study protein folding in vivo at the extreme temperature of 70°C in the extreme thermophilic eubacterium, Thermus thermophilus (Cava et al. 2008). It has also been used as a fusion tag with a single-strand DNA binding protein to facilitate immobilization of target proteins (Zhong, Fang, and Wei 2010). In lieu of other variants of GFP, sf-GFP has also been used in circularly permutated protein studies.
2.2.2.1. Circular permutation of GFP

In comparison to the cp-BLA fusion proteins discussed earlier, many of the cp-GFP fusion proteins have been engineered in an opposite orientation. In several studies where a circularly permutated GFP has been used, the second strategy of creating a protein switch was followed, i.e., protein of interest was inserted into the circularly permutated GFP protein, and fused via linkers. For example, Ghosh et al. (2000) used the sg100 GFP variant possessing mutations that displayed only one excitation (475 nm) and one emission (505 nm) wavelength in contrast to wtGFP, and circularly permutated it between positions 157 and 158. Sequences for helices that

---

**Figure 2.5: Schematic representation of the sf-GFP.** Open circles represent mutations unique to sf-GFP. Black circles represent mutations reported previously. Circled 11 represents the site of cp between aa 172 and 173. Reprinted/adapted by permission from Macmillan Publishers Ltd: Nature Biotechnol, 24, Pédelacq et al., Engineering and characterization of a superfolder green fluorescent protein, 79-88, © 2006.
were predicted to form antiparallel leucine zippers were fused to the two GFP fragments via linkers (Figure 2.6). Since the three residues involved in the autocatalytic formation of the chromophore were intact in one of the GFP fragments, the fragmented GFP was reconstituted and fluorescence restored upon reassembly of the attached leucine zippers that brought the GFP fragments in closer proximity (Ghosh, Hamilton, and Regan 2000). It was concluded that the function of a protein could potentially be controlled through its three-dimensional structure, more specifically through non-covalent reconnection of the two termini of a fragmented protein via antiparallel leucine zippers.

**Figure 2.6: Schematic representation of the fusion protein created with GFP and antiparallel leucine zippers joined via linkers.** Numbers represent relevant amino acid positions. NZ = leucine zipper at the N-terminus, CZ = leucine zipper at the C-terminus. Adapted with permission from J. Am. Chem. Soc., 122, Ghosh et al., Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein, 5658-5659, © 2000 American Chemical Society.
In another study, Baird et al. (1999) constructed a fusion protein with the enhanced yellow fluorescent protein (another variant of GFP) by inserting calmodulin (CaM) or the zinc finger domain at position 145. It was observed that there was a significant increase in fluorescence, once binding of Ca\(^{2+}\) to CaM or of Zn\(^{2+}\) to the zinc finger domain occurred. A seven to eight-fold increase in fluorescence (increase in amplitude at emission 490 nm) was observed once Ca\(^{2+}\) bound, and 1.7-fold upon Zn\(^{2+}\) binding. The resulting fusion protein was intended to serve as an example of a physiological indicator. This study showed that it is possible to tightly control the GFP fluorescence if the attached receptor is more sensitive to its ligand, and also goes under a conformational change upon ligand binding.

Similar to the structural and functional studies discussed earlier, the superfolder GFP was used as a reporter in an assay to predict protein solubility and folding (Cabantous et al. 2008). A *Mycobacterium tuberculosis* protein Rv0113, was inserted into the cp sf-GFP (cleaved between positions 172-173) scaffold to increase solubility of Rv0113 (Figure 2.7). Previously, when Rv0113 was expressed in *E. coli*, the protein ended up in inclusion bodies. The aim of this study was to test protein folding and solubility, through the cp sf-GFP insertion fusions. The insertion fusions performed better than N-terminal fusions, as indicators of properly folded soluble protein.

In a contrasting study by Sakaguchi et al. (2009) a fusion protein was constructed with a different orientation than the experiments discussed above in which cp-GFP was inserted into the protein of interest. The cp-GFP was inserted into the split pleckstrin homology (PH) domain of a tyrosine kinase—Bruton’s tyrosine kinase (Btk), and this fusion protein was predicted to bind its cognate ligand D-myo-inositol-1,3,4,5-tetrakisphosphate (Ins(1,3,4,5)P\(_4\)). The split
PH domain had been previously shown to reconstitute into a fully functional protein once the two split halves were brought closer together. Therefore, the fusion protein Btk-cpGFP (Figure 2.8) was intended to serve as a fluorescent sensor for an intracellular signal messenger once substrate was added. Upon binding of Ins(1,3,4,5)P$_4$ to the PH domain, a conformational change in the fused cp-GFP occurred as observed by a change in the emission and excitation spectra. Specifically, when excited at 470 nm there was a decrease of 508 nm emission band, and an increase in the 508 nm emission band when excited at 396 nm. This corresponded to the changed state of GFP, i.e., an increase in protonated (excitation 396 nm) and decrease in deprotonated (excitation 470 nm) state of the chromophore. This study showed that proximity

---

**Figure 2.8: Schematic representation of Btk-cpGFP construct.** Numbers represent amino acid sequences. Reprinted/adapted from Bioorg. Med. Chem., 17, Sakaguchi et al., A single circularly permuted GFP sensor for inositol-1,3,4,5-tetrakisphosphate based on a split PH domain, 7381–7386, © (2009), with permission from Elsevier.
of the reporter to the binding domain of the protein of interest is essential when designing a fully functional fusion protein, since a conformational change in the latter (protein of interest) has to be transduced to the former (reporter) upon substrate binding. In the case of Btk-cpGFP, this had to be achieved without disrupting the residues of the split PH domain directly involved in substrate interaction, so that selectivity and affinity were not lost. The Btk-cpGFP was also designed according to the recommended cp protein creation criteria; hence the chosen insertion point in the PH domain for introducing the cp-GFP was the group of residues within the loop region. This work is also a good example of how biosensors can be designed where the protein of interest (target site for ligand binding) does not necessarily undergo a large conformational change, yet can still transduce a conformational change in the attached cp reporter protein. The fusion RG13 protein comprising the MBP and cp-BLA domains discussed earlier (Guntas et al. 2004) also represents such an example. The wild-type MBP used in the RG13 is comprised of two domains linked by a hinge region, and it is known to undergo a slight structural change upon maltose binding. Specifically, the hinge region bends 35° when maltose binds, which results in the closed conformation of MBP (Sharff et al. 1992). When maltose bound to MBP in the fusion protein, the closed conformation of MBP resulting from the slight change upon binding was strong enough to induce a change in the attached cp-BLA. The split PH in Btk-cpGFP also experienced a small conformational change resulting in GFP fluorescence, thus providing evidence that the end point of a fusion protein experiment can be controlled tightly with respect to the cp reporter protein through small structural changes.

In another study, Nakai et al. (2001) constructed a calcium probe based on GFP (G-CaMP) with a high signal-to-noise ratio. The EGFP variant was circularly permutated at 144 and
149, then inserted between a fragment from myosin light chain kinase (M13) at the N-terminus, and calmodulin at the C-terminus, fused via linkers (Figure 2.9 A, B). The M13 fragment was a target peptide for CaM which in turn was a binding target for calcium (Ca$^{2+}$). The resulting fusion protein G-CaMP, was intended to serve as an indicator of Ca$^{2+}$ in the cellular environment. Upon binding of calcium to CaM, a conformational change in the Ca$^{2+}$-CaM-M13 complex occurred which consequently resulted in a conformational change in the circularly permutated EGFP. Interestingly, this change was indicated as an alteration in the intensity of fluorescence, which indicated the presence of Ca$^{2+}$. Although there was a background signal owing to the affinity of CaM for M13, the fluorescent signal upon the addition of calcium was higher. As was the case for other cp studies, the orientation of the cp-GFP with respect to the

---

**Figure 2.9: Schematic representation of the calcium probe with cp-GFP variant.**
protein of interest was shown to be highly important. Specifically, when the orientation of G-CaMP was altered such that CaM was at the N-terminus and M13 at the C-terminus of cpEGFP, the resulting probe did not respond as well as the first orientation to the presence of calcium.

2.3. Restatement of research objectives

All of the studies discussed above used enzymes and other ligand binding proteins for circular permutations and resultant fusion proteins, but none used an antibody with a cp protein to construct a fusion protein. The first report of the construction of such a fusion protein was by Kojima et al. (2011) where the variable antibody domains against a bone disease marker (osteocalcin) and a pesticide (imidacloprid) were fused to a circularly permuted TEM-1 BLA via short linkers (Figure 2.10). The resulting fusion protein had a high background signal in vitro, but was able to confer antigen dependent ampicillin resistance to E. coli grown in the presence of ampicillin and the target antigen of the variable antibody domains.

![Figure 2.10: Schematic representation of the expression plasmid for anti-osteocalcin / imidacloprid cp-BLA.](image-url)
The aim of this thesis was to develop a detection method utilizing fusion proteins constructed using a similar design as Kojima et al. (2011). The circularly permuted superfolder GFP was added as a second reporter in addition to the cp-BLA, and these fusion proteins were to be used in the development of a single-step visual immunoassay. The hydrophobic polycyclic aromatic hydrocarbon benzo[a]pyrene was selected as the antigen.

2.4. Overview of Benzo[a]pyrene

B[a]P belongs to the class of polyaromatic hydrocarbons and exists in the environment from anthropogenic and natural sources (Figure 2.11). Polyaromatic hydrocarbons are soluble in lipids, but insoluble in aqueous solutions. However, the solubility can be increased by dissolving in organic solvents such as DMSO, acetone, methanol and ethanol (Kontir et al. 1986). Due to the hydrophobicity and the consequent insolubility in water, PAHs adsorb to particulate and oily matter found in contaminated water, sediment and soil. Although B[a]P is not commercially produced or used, it is widely distributed in the environment, wherever incomplete combustion of organic matter occurs (Osborne and Crosby 1987). It is also

![Chemical structure of benzo[a]pyrene](image)

**Figure 2.11: Chemical structure of benzo[a]pyrene.** $C_{20}H_{12}$, MW 252.31. Solubility in water: $1.61 \times 10^{-3}$ mg/L. CAS No.: 50-32-8.
a potent animal carcinogen and although considered previously as a Group 2A probable
carcinogen ("Limited evidence in humans and sufficient evidence in animals") in 1987 by the
International Agency for Research on Cancer, B[a]P has recently been elevated to a group I
human carcinogen ("Sufficient evidence in humans or sufficient evidence in animals and strong
mechanistic data in humans") in 2012 (IARC 1987; IARC 2012). The primary route of entry of
PAHs in humans is through ingestion and inhalation and due to the hydrophobicity they can
readily enter cells. However, excretion is facilitated once metabolization increases solubility
(Osborne and Crosby 1987).

Benzo[a]pyrene is not toxic in its natural state but once it is metabolized by phase-I and
phase-II enzymes it becomes bioactive. The mechanism of carcinogenesis in experimental
animals is explained through two pathways; through formation of (i) diolepoxides and (ii)
radical cations (Osborne and Crosby 1987). The metabolism of B[a]P yields such metabolites as
epoxide intermediates, dihydrodiols, phenols, and quinines (ATSDR 1995; Parkinson 1996). One
important metabolite is benzo[a]pyrene-7,8-diol resulting from the conversion of
benzo[a]pyrene-7,8-epoxide by epoxide hydrolase (phase-I reaction) through addition of water.
Metabolism of benzo[a]pyrene-7,8-diol by the mixed function oxidase cytochrome P-450 results
in the formation of a carcinogenic metabolite benzo[a]pyrene-7,8-diol-9,10-epoxide, which is
the ultimate toxicant that binds covalently to DNA and forms adducts (Rogers and Kavlock,
1996; Parkinson 1996).

B[a]P has been shown to be tumorigenic in animals by causing lung and skin tumours in
murine species. Occupational exposure to B[a]P in humans has been observed to cause cancers
as well, e.g., skin cancer during coal tar distillation, lung and oesophagus cancer by exposure to
soot, and lung/lip/larynx cancer from tobacco smoking. Therefore, based on the experimental studies performed with B[a]P to gather evidence for its toxicity, the IARC declared B[a]P to be a Group 1 carcinogen to humans (IARC 2012).

The more common methods of B[a]P detection in the environment are solid phase extraction or liquid/liquid partition of the sample followed by concentration determination of extracts using GC-MS or HPLC. However, they are time-consuming and labour intensive. Therefore, the development of an ELISA for B[a]P detection has been a desired goal. Previously, the type of anti-B[a]P antibodies that existed were not very specific and were mostly against metabolites of B[a]P or DNA adducts formed with the metabolites (Santella et al. 1985). Scharnweber et al. (2001) first reported the production of a monoclonal anti-B[a]P antibody with a broad specificity for PAHs that could be used in a competitive indirect ELISA. Four years later Matschulat et al. (2005) reported the development of the most sensitive indirect competitive ELISA for the detection of B[a]P also using an anti-B[a]P monoclonal antibody (lower limit of detection of 24 ng/L). More recently, in an effort to develop a traditional ELISA Karsunke et al. (2012) developed several anti-B[a]P scFvs using the previously generated monoclonal by Matschulat et al. (2005) and Karsunke et al. (2011). One of the scFvs generated in this binding study (L2H2) was used in the construction of the two fusion proteins (anti-B[a]P cp-BLA; anti-B[a]P cp-GFP) for this thesis.
3. MATERIALS & METHODS

3.1. Construction of Plasmids

All of the designed constructs mentioned below were synthesized by Genewiz (New Jersey, USA), unless otherwise stated.

3.1.1. Construction of anti-\(\text{B[a]P}\) cp-BLA (\text{pIT2-SZ1 plasmid})

The gene sequence for TEM-1 β-lactamase (BLA; E.C: 3.5.2.6) was obtained from Protein Data Bank (PDB ID: 1ZG4), circularly permutated according to literature (Kojima et al. 2011), and fused via short flexible linkers to variable heavy and light chain antibody fragments of an anti-B[α]P scFv (Karsunke et al. 2012). A His\(_6\) affinity tag was added to the carboxyl-terminal of the fusion protein to aid in purification. The designed construct was cloned into the phagemid vector pIT2 bearing a selectable marker for ampicillin resistance, and a pelB leader sequence directing nascent protein to the periplasm (Figure 3.1; Appendix I).

![Figure 3.1: Schematic representation of the expression vector pIT2-SZ1. The circularly permutated BLA sequence (N-terminal amino acids 168-286 linked via DKS linker to C-terminal amino acids 24-170) was attached via linkers to anti-B[a]P scFv \(V_H\) and \(V_L\) gene sequences, and cloned into the pIT2 vector bearing the selectable marker for ampicillin resistance.](image)

3.1.2. Construction of anti-B[α]P cp-GFP and cp-GFP (pIT2-SZ2 and pITFK-cpGFP plasmids)

The circularly permutated superfolder GFP gene sequence was obtained from literature (Pédelacq et al. 2006; Waldo and Cabantous 2013) and cloned into a modified pIT2 vector.
Figure 3.2: Schematic representation of the expression vectors pITFK-cpGFP and pIT2-SZ2. **A)** The circularly permutated superfolder GFP gene sequence (cp-GFP; N-terminal amino acids 173-238 linked to C-terminal amino acids 1-172) was cloned into pITKanM with kanamycin as a selectable marker. **B)** Anti-B[a]P scFv V\text{H} and V\text{L} gene sequences were joined via linkers to cp-GFP, and cloned into the original pIT2 vector to create the pIT2-SZ2 plasmid (Figure 3.2 B; Appendix II).

3.2. **Protein Expression and Extraction**

The plasmids pIT2-SZ1 and pIT2-SZ2 were used to transform *Escherichia coli* strain HB2151 by electroporation. The same strain was also transformed with an empty pIT2 plasmid to be used as a negative control, and with pITFK-cpGFP to be used as a positive control. The transformation mixes were plated on 2xYT (Yeast-Tryptone) agar plates supplemented with...
carbenicillin (or kanamycin for pITFK-cpGFP), and incubated overnight at 37°C. A single colony from the streak cultures was used to inoculate starter cultures (15 ml) of 2xYT media supplemented with 50 µg/ml of carbenicillin or kanamycin and 1% glucose, incubated overnight (16 h) at 200 rpm and 37 °C. The next day, glycerol stocks (15%) were prepared from the starter cultures, and remaining cultures were used in a 1:100 dilution to inoculate fresh 1L and 100 ml of 2xYT media for fusion proteins and cp-GFP, respectively, which were supplemented as described above. The suspension cultures were incubated at 37 °C and allowed to reach an O.D.₆₀₀ of 0.8-0.9 before centrifugation for 20 min at 3000 rpm, at room temperature. To induce protein expression, the cell pellets were resuspended in fresh 2xYT media (50 ml) supplemented with 50 µg/ml carbenicillin or kanamycin, 0.1% glucose, 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and incubated for 24 h with vigorous shaking at room temperature. The suspensions were incubated on ice for 0.5 h, followed by centrifugation for 10 min at 7500 rpm, and 4°C. Cell pellets were washed twice by resuspension in 10 ml of sterile phosphate buffered saline (PBS) solution and centrifuged for 15 minutes at 5000 rpm, and 4°C. After washing, the pellets were resuspended in 10 ml of PBS and stored at -20°C. The frozen suspensions were incubated on ice to thaw, and subsequent steps were all performed with the suspensions stored on ice. Cells were lysed by sonication (550 Sonic Dismembrator, Fisher Scientific) using short 30 second pulses (total time 3 min) with intermittent 30 second pauses for cooling. The sonication lysates were transferred to 2-ml microcentrifuge tubes and centrifuged for 20 min at 12000 rpm, 4°C to pellet cell debris and obtain soluble protein.
3.3. Protein Purification

The supernatant from the sonication lysate was applied to His SpinTrap™ columns (GE Healthcare) packed with Ni Sepharose™ High Performance affinity media (GE Healthcare) for IMAC purification of proteins with exposed histidine groups. Columns were equilibrated with Binding buffer (20 mM sodium phosphate and imidazole, 500 mM sodium chloride, pH 7.4), followed by sample application (600 µl max.) and washed with Binding buffer. Purified protein was eluted using 2 x 200 µl of Elution buffer (20 mM sodium phosphate, 500 mM imidazole and sodium chloride, pH 7.4) in 1.5-ml microcentrifuge tubes.

3.4. Protein analyses

Protein parameters such as theoretical molecular weight, amino acid composition and molar extinction coefficient were obtained using the ProtParam tool on the ExPASy server (Gasteiger et al. 2005). Quantification of protein in crude sonication lysate was determined using the Bio-Rad Protein (Bradford) Assay (Mississauga, Canada) with bovine serum albumin (Thermo Scientific) to derive a standard curve. Preliminary analysis was performed by resolving total soluble protein (30 µg/well) on 12% non-reducing SDS-PAGE gels stained with Coomassie blue dye. This procedure was followed by transferring resolved protein bands (i.e., from non-stained gels) to a polyvinylidene difluoride (PVDF) membrane (Bio-Rad) for further examination by Western blot analysis. The membrane was probed using the anti-6xHIS tag® antibody [HIS-1] (Abcam ab49746; 1/2000 dilution) conjugated to alkaline phosphatase.

Ten-fold dilutions of eluted fractions collected from protein purification were prepared and absorbance was measured at 280 nm using a spectrophotometer. The Beer-Lambert
equation for spectrophotometry was used to calculate protein concentration (Grimsley and Pace 2004):

\[
c = \frac{(A_{280} \times MW)}{(\varepsilon \times l)}
\]

where \(c\) is the protein concentration in mg/ml, \(A_{280}\) is the absorbance of sample at 280 nm, \(MW\) is the protein molecular weight, \(\varepsilon\) is the molar extinction coefficient (M\(^{-1}\) cm\(^{-1}\)) and \(l\) is the cell pathlength (cm). The calculated \(c\) was multiplied by the dilution factor to achieve the final protein concentration.

To verify protein purification, 200 ng of protein from eluted fractions were also separated on additional 12% non-reducing SDS-PAGE gels, and stained with Coomassie dye. Western blot analysis was also conducted on non-stained gels, and probed with anti-hexahistidine antibody as described above.

3.5. **Immunoassays**

3.5.1. **Anti-\([B[a]P]\) cp-BLA fusion protein assay**

Clear 96-well polysterene microplates (non-binding; Corning 3641) were used for the cp-BLA activity assays. Nitrocefin (Calbiochem) was prepared as a 2 ml, 1 mM solution using 1 mg of reagent in 5% DMSO (v/v), and 95% PBS (v/v; pH 7.4). Preliminary assays were done using both fusion proteins (crude (dialyzed) and purified; 50 µl), nitrocefin (50 µl), and 1 µg/ml \([B[a]P]\) or PBS (100 µl), and absorbance measured at 485 nm using a Wallac Envision 2100 microplate reader over 24 hours.

A nitrocefin cross reactivity assay was performed to determine cross reactivity between the reagents comprising the elution buffer and the substrate solution. Elution buffer and each of its reagents; 20 mM sodium phosphate, 500 mM sodium chloride and imidazole, were added
(50 µl) to wells followed by addition of either 150 µl of 1 mM nitrocefin solution, or only the DMSO:PBS solution without nitrocefin. Water was used as a negative control and elution buffer as a positive control. Absorbance was measured as stated above.

3.5.2. **Anti-B[a]P cp-GFP fusion protein assays**

Black, 96-well non-binding (Corning 3650) polystyrene microplates were used for the fluorescent assays, and fluorescence was measured using the Wallac Envision 2100 microplate reader (excitation at 490 nm and emission at 510 nm).

3.5.2.1. **In vivo assay**

Fluorescence of *E. coli* cultures expressing the cp-GFP or anti-B[a]P cp-GFP protein was measured with and without addition of B[a]P. Cultures were grown in 25 ml of 2xYT media supplemented with 1% glucose and 50 µg/ml carbenicillin or kanamycin, and incubated at 180 rpm and 37°C, until O.D._600_ = 0.8 - 1.0. Pellets were resuspended in fresh media supplemented with 0.1% glucose and 50 µg/ml carbenicillin / kanamycin, with the addition of 1 mM IPTG for induction. Following induction, cultures were incubated at 180 rpm and 26 °C for 2 h and subsequently split into two 12.5-ml fractions. B[a]P (11 µl of 1000 µg /ml in acetone) was added to only one fraction to achieve a final concentration of 0.9 µg/ml. PBS was used as a negative control. After overnight incubation at room temperature on a shaker, pellets were resuspended in 6.3 ml PBS. Fluorescence of 100 µl of pellet suspension was measured over 21 hours.

3.5.2.2. **In vitro assays**

**Assay 1:** Fusion protein (crude extract) was added to each well (150 µl), followed by 50 µl of 1 µg/ml B[a]P or PBS. Plates were covered with plastic wrap, and sealed with another microplate
to prevent sample evaporation. Incubation was performed at room temperature on a shaker, and fluorescent measurements were made over 21 hours.

**Assay 2:** Fusion protein (crude extract; two eluted fractions (E1, E2) from purification) and cp-GFP (purified) protein were added (80 µl/well) with the addition of 20 µl of 1 µg/ml B[a]P or PBS. Plates were incubated as described above, and read over 25 hours.

**Assay 3:** Fusion protein (two eluted fractions (E1, E2) from purification) were diluted eight fold and three fold respectively, while purified cp-GFP was diluted sixteen fold in elution buffer. Proteins were added (80 µl/well) along with one of three different concentrations (200 µl/well) of B[a]P (1 µg/ml, 10 ng/ml, 100 pg/ml) or PBS. Plates were incubated at 37°C on a shaker, sealed and placed in a plastic bag with moistened paper towels to prevent sample evaporation. Fluorescent measurements were taken over a period of 39 hours.

3.6. **B[a]P standards**

B[a]P standards of 100 µg/ml, 10 µg/ml, 1 µg/ml, 100 ng/ml, 10 ng/ml, 1 ng/ml, 100 pg/ml, 10 pg/ml, 1 pg/ml were prepared from the standard solution of 1 mg/ml B[a]P in acetone (Sigma-Aldrich; Supelco, PA, USA). Methanol was used as an organic solvent to enhance benzo[a]pyrene solubility in the liquid medium (Table 1). Standards were stored in sealed and covered glass volumetric flasks at 4°C to decrease adsorption to container walls as well as prevent exposure to light.

3.7. **Statistical Analysis**

Student’s t-test was used to determine a significant difference between the two groups, i.e., protein with B[a]P and without B[a]P. Statistics representing a value of p < 0.05 were considered significant. For *in vitro* assay 3 of anti-B[a]P cp-GFP fusion protein, a one-way
ANOVA was used followed by Bonferroni’s post-hoc test. All experiments were run in triplicates (n=3).

**Table 1: Preparation of B[a]P standards.** B[a]P standards of several concentrations were prepared using methanol as a B[a]P solubility enhancer.

<table>
<thead>
<tr>
<th>Solution Designation</th>
<th>Concentration of B[a]P standard(^a)</th>
<th>Dilutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100 µg/ml</td>
<td>1 ml stock solution to 9 ml methanol*</td>
</tr>
<tr>
<td>B</td>
<td>10 µg/ml</td>
<td>1 ml of A to 9 ml methanol</td>
</tr>
<tr>
<td>C</td>
<td>1 µg/ml</td>
<td>1 ml of B to 9 ml water</td>
</tr>
<tr>
<td>D</td>
<td>100 ng/ml</td>
<td>1 ml of C to 9 ml 10% methanol</td>
</tr>
<tr>
<td>E</td>
<td>10 ng/ml</td>
<td>1 ml of D to 9 ml 10% methanol</td>
</tr>
<tr>
<td>F</td>
<td>1 ng/ml</td>
<td>1 ml of E to 9 ml 10% methanol</td>
</tr>
<tr>
<td>G</td>
<td>100 pg/ml</td>
<td>1 ml of F to 9 ml 10% methanol</td>
</tr>
<tr>
<td>H</td>
<td>10 pg/ml</td>
<td>1 ml of G to 9 ml 10% methanol</td>
</tr>
<tr>
<td>I</td>
<td>1 pg/ml</td>
<td>1 ml of H to 9 ml 10% methanol</td>
</tr>
</tbody>
</table>

* Stock solution was 1 mg/ml in acetone.

\(^a\) Solubility of B[a]P in water is \(1.6 \times 10^{-3}\) mg/L
4. ANTI-B[a]P cp-BLA FUSION PROTEIN

4.1. Results

The fusion protein comprising DNA sequences of anti-B[a]P scFv and circularly permutated TEM-1 β-lactamase was expressed in *E. coli*. Immunoblots confirmed protein expression and the recombinant protein was successfully purified by affinity chromatography.

Figure 4.1: Expression of soluble anti-B[a]P cp-BLA protein. A) A 12% SDS-PAGE gel ran under non-reducing conditions and stained with Coomassie dye. Protein loaded: 30 µg/well. Lane 1: protein ladder as a molecular weight marker; Lane 2: 6xHis protein ladder used as positive control; Lane 3: crude extract from empty pIT2 plasmid used as negative control; Lane 4: crude protein extract of the anti-B[a]P cp-BLA fusion protein (~57 kDa) from pIT2-SZ1 plasmid. B) Immunoblot visualizing anti-B[a]P cp-BLA protein. Crude protein extracts (30 µg/well) were resolved on 12% SDS-PAGE gel then transferred to PVDF membrane. The membrane was probed with anti-6xHis antibody. Lane 1: protein ladder as a molecular weight marker; Lane 2: 6xHis protein ladder used as positive control; Lane 3: crude extract from empty pIT2 plasmid used as negative control; Lane 4: crude protein extract of the anti-B[a]P cp-BLA fusion protein (~57 kDa) from pIT2-SZ1 plasmid.
(IMAC). The Coomassie stained SDS-PAGE gel and immunoblot results confirmed the expression of the soluble protein (527 amino acids), as indicated by a band (Figure 4.1 A, B) of the expected molecular weight—57.3 kDa (MW scFv: 25.80; MW cpBLA: 29.57; excluding start codon, restriction sites and His\textsubscript{6} affinity tag). The results of further analysis to determine successful purification by IMAC confirmed a band of similar size from both eluted fractions E1 and E2 (Figure 4.2).

**Figure 4.2: Immunoblot visualizing purified anti-B[a]P cp-BLA protein.** Two eluted fractions (200 ng/well) were resolved on 12% SDS-PAGE gel under non-reducing conditions, then transferred to PVDF membrane. The membrane was probed with anti-6xHis antibody.

Lane 1: protein ladder as a molecular weight marker; Lane 2: 6xHis protein ladder used as positive control Lane 3: negative control (PBS); Lane 4: crude protein extract from pIT2-SZ1 plasmid expressing the fusion protein; Lane 5: wash flow through from IMAC purification; Lanes 6 and 7 (indicated by the arrows): eluted fractions 1 and 2 from IMAC purification using His SpinTrap\textsuperscript{TM} columns.
The preliminary assay done using dialyzed crude protein from sonication lysate displayed negative results, i.e., all wells containing the fusion anti-B[a]P cp-BLA protein and negative control (anti-B[a]P cp-GFP) remained yellow over 24 hours. Once purified protein was obtained, it was assayed the same way. Immunoassay data revealed unanticipated results since a colorimetric reaction occurred (yellow to red) in the absence of B[a]P, in wells containing the chromogenic substrate nitrocefin and either anti-B[a]P cp-BLA or anti-B[a]P cp-GFP (negative control). Also unexpectedly, wells containing B[a]P in contrast did not display a colour change and remained yellow (Figure 4.3). These results indicated that the substrate nitrocefin was being converted non-enzymatically. Consequently, a nitrocefin cross reactivity assay was performed to determine the cause of this non-enzymatic colorimetric reaction.

By adding components comprising the elution buffer to wells, in addition to nitrocefin (Table 2), it was found that imidazole (500 mM) was responsible for the colour change from yellow to red. Subsequent efforts to remove the imidazole from solution using desalting columns were unsuccessful.

<table>
<thead>
<tr>
<th>Elution Buffer</th>
<th>20 mM Sodium phosphate</th>
<th>500 mM NaCl</th>
<th>500 mM Imidazole</th>
<th>Water</th>
<th>DMSO + PBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 mM Nitrocefin</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>DMSO + PBS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(+ water)</td>
</tr>
</tbody>
</table>

Table 2: Results from the nitrocefin cross-reactivity assay. Elution buffer was used as a positive control and water as a negative control. The (++) denote the occurrence of a colorimetric reaction where solution turned to red from yellow. The (-) denotes no occurrence of a colour change, wells remained yellow.

4.2. Discussion

The anti-B[a]P cp-BLA fusion protein was successfully expressed, extracted and purified as shown by Coomassie dyed SDS-PAGE gels and immunoblots. However, immunoassay results were unable to support the hypothesis as the presence of B[a]P did not lead to a colorimetric reaction when cp-BLA was fused to the variable heavy and light chain scFv domains. Additionally, the occurrence of non-specific substrate hydrolysis created the presence of a confounding variable.

According to the hypothesis, a colorimetric reaction should have been observed with the recombinant protein, i.e., anti-B[a]P cp-BLA in the presence of B[a]P and nitrocefin. The cp-BLA in the fusion protein would be activated once B[a]P was bound by both heavy and light
chains which would bring the two parts of cp-BLA together. However, when B[a]P was added to wells containing protein (crude or purified) and nitrocefin, no colour change was observed. This was expected when the fusion protein was anti-B[a]P cp-GFP as the required enzyme (BLA) was absent to interact with nitrocefin. However, this was not the case for the anti-B[a]P cp-BLA protein which unpredictably displayed a similar result, i.e., no colour change. It is possible that the concentration of 1 µg/ml B[a]P used was too high and may have impeded the conformational change required to make the cp-BLA in the fusion protein fully functional, thereby preventing hydrolysis of nitrocefin. It is also possible that B[a]P could be responsible for steric hindrance and blocked the enzyme’s access to the β-lactam ring within the substrate. In addition, B[a]P could have created a steric hindrance to the occurrence of conformational change in the fusion protein and prevented the activation of the fused cp-BLA. In the case of crude protein, hindrance from host proteins present in the crude mix could also be responsible for preventing interaction between B[a]P and variable domains of fusion protein thereby impeding the necessary conformational change that would activate the cp-BLA thus resulting in the hydrolysis of nitrocefin (colour change from yellow to red).

Of note was another outcome where a colour change was observed, unexpectedly, in the absence of B[a]P. When purified fusion proteins anti-B[a]P cp-BLA and anti-B[a]P cp-GFP were both incubated with nitrocefin in the absence of B[a]P, the solution in both wells turned red. Theoretically, a colour change from yellow to red is observed when the amide bond in the β-lactam ring of nitrocefin is hydrolyzed by β-lactamase (Figure 4.4). The wells containing anti-B[a]P cp-GFP protein did not contain the β-lactamase capable of hydrolyzing the β-lactam ring
in nitrocefin, yet displayed a colour change similar to the wells with anti-B[a]P cp-BLA. This unexpected colour change in the absence of both essential factors B[a]P and β-lactamase represented a non-specific hydrolysis of the chromogenic substrate nitrocefin. It was demonstrated that the elution buffer used for protein purification was responsible for this non-enzymatic reaction. Specifically, the cross-reactivity assay confirmed that imidazole was interacting with the substrate, as only wells with elution buffer and imidazole displayed a colour change in the presence of nitrocefin and the absence of its specific enzyme β-lactamase. Since attempts to remove the imidazole from solution by using desalting columns were unsuccessful, further experiments with the anti-B[a]P cp-BLA protein were not performed.

Inference from the mechanism of β-lactam hydrolysis by β-lactamase, suggests a probability that the imidazole is acting as a nucleophile attacking the amide bond in the β-lactam ring and aiding the formation of an acyl intermediate prior to hydrolysis by water. Additionally, non-specific hydrolysis of nitrocefin has been reported earlier with materials such as...
as albumin, egg white, milk, thiols and glutathione (O’Callaghan et al. 1972), although clear evidence of the detailed mechanism of interaction is not yet available. This specific occurrence of imidazole and nitrocefin cross-reaction has not so far been reported in literature, and thus appears to be a novel addition to the list of substances that non-specifically interact with nitrocefin.

4.3. Conclusion

Although anti-B[a]P cp-BLA fusion protein was successfully expressed, extracted and purified, an effective immunoassay to detect B[a]P could not be developed. This was due to the absence of a colour change upon B[a]P addition to the reaction, as well as the presence of a confounding variable, i.e., non-specific hydrolysis of the substrate nitrocefin in the presence of imidazole. After attempting unsuccessful removal of imidazole, it was confirmed that an obstacle existed where a colour change was inevitable without B[a]P presence. This rendered the assay ineffective at exclusively indicating the existence of the chosen antigen. Consequently, further experiments with the anti-B[a]P cp-BLA protein were not pursued.
5. ANTI-B[a]P cp-GFP FUSION PROTEIN

5.1. Results

The control cp-GFP protein and the fusion protein (anti-B[a]P scFv fused to circularly permutated superfolder GFP) were expressed in *E. coli*. Western blot analysis (Figure 5.1 lane 5, Figure 5.2 lane 4) indicated bands of the expected molecular weights; cp-GFP—27.32 kDa, anti-B[a]P cp-GFP—54.96 kDa (MW scFv: 25.80 kDa; MW cp-GFP: 26.81 kDa; excluding linkers, start codon, restriction sites and His$_6$ affinity tag) thereby confirming expression of the soluble.

![Figure 5.1: Immunoblot visualizing purified control cp-GFP protein. Lane 1: protein ladder as a molecular weight marker; Lane 2: negative control (PBS); Lane 3: positive control (protein with His$_6$ affinity tag); Lane 4: 30 µg of crude extract from empty pIT2 plasmid used as negative control; Lane 5: 30 µg of crude protein extract from pITFK-cpGFP plasmid expressing the control cp-GFP protein (~27 kDa); Lane 6 and 7 (indicated by the arrows): 200 ng of eluted fractions 1 and 2 from IMAC purification using His SpinTrap™ columns.](image)
proteins. The recombinant proteins were subsequently purified by affinity chromatography (IMAC) and the results of further analysis to determine successful purification indicated bands of similar sizes from both eluted fractions E1 and E2 (Figures 5.1, 5.2).

Figure 5.2: Immunoblot visualizing purified anti-B[a]P cp-GFP fusion protein. Lane 1: protein ladder as a molecular weight marker; Lane 2: negative control (PBS); Lane 3: positive control (protein with His6 affinity tag); Lane 4: 30 µg of crude extract from pIT2-SZ2 plasmid (~55 kDa); Lane 5 and 6 (indicated by the arrows): 200 ng of eluted fractions 1 and 2 from IMAC purification using His SpinTrap™ columns.

Protein concentration of the crude anti-B[a]P cp-GFP protein was determined to be 1.5 mg/ml by using Bradford assay, while for purified E1 and E2 fractions was 4.52 mg/ml and 1.68 mg/ml respectively, using spectrophotometry.
In vivo ASSAY

*E. coli* cultures expressing pITFK-cpGFP and pIT2-SZ2 plasmids displayed no significant difference in fluorescence when B[a]P was or was not added, and emission measured over time (Figure 5.3). Specifically, there was no significant difference in mean fluorescence among cultures expressing the control cp-GFP protein that were and were not treated with B[a]P, or among cultures expressing the fusion anti-B[a]P cp-GFP protein that were and were not treated with B[a]P.

![Graph](image)

**Figure 5.3: In vivo fluorescence assay.** Data plotted represents mean fluorescence (excitation 490 nm, emission 510 nm) readings of IPTG-induced *E. coli* HB2151 cultures expressing cp-GFP or anti-B[a]P cp-GFP protein that were either untreated or treated with B[a]P. Fluorescence was measured 15 hours after initial addition of B[a]P with subsequent measurements every 2 hours. B[a]P addition did not significantly increase (p > 0.05) fluorescence of cp-GFP or fusion cultures. Error bars represent SEM (n=3).
**In vitro ASSAYS**

**Crude sonication lysate**

There was also no significant difference in fluorescence among crude control cp-GFP protein wells that were and were not treated with B[a]P, as well as among crude anti-B[a]P cp-GFP fusion protein wells that were and were not treated with B[a]P (Figure 5.4).

![Graph showing fluorescence levels](image)

**Figure 5.4: Mean fluorescence of crude anti-B[a]P cp-GFP and cp-GFP protein in sonication lysate.** Data represents average of untreated triplicates (no B[a]P), or triplicates treated with 1 µg/ml B[a]P and measured over 21 hours. Error bars represent SEM (n=3). The addition of B[a]P did not significantly increase (p > 0.05) fluorescence of crude anti-B[a]P cp-GFP protein or control cp-GFP protein over time.

**Purified protein**

The *in vitro* immunoassay utilizing the first eluted (E1) fraction of anti-B[a]P cp-GFP protein, did not display a higher final mean fluorescence when wells that were treated with
B[a]P were compared with wells that were not treated with B[a]P (Figure 5.5 A). A time course graph (Figure 5.5 B) of transformed data with respect to time 0, that represented the mean fluorescence over 25 hours (all time points that were statistically different when the two groups were compared), indicated that fluorescence was no different between wells that did and did not contain B[a]P. Similarly, the *in vitro* immunoassay utilizing the second eluted (E2) fraction of purified anti-B[a]P cp-GFP protein, also did not display a significant difference in fluorescence between wells that were and were not treated with B[a]P (Figure 5.6 A, B).
Figure 5.5: In vitro assay utilizing E1 anti-B[a]P cp-GFP protein. A) Data represents mean fluorescence of triplicates that were untreated (no B[a]P) or were treated with B[a]P, and measured over 25 hours. B) Mean fluorescence at each hour was expressed as a percentage change in fluorescence with respect to time 0 and plotted against time. Only the time points that displayed a significant difference in fluorescence between the two groups are shown. Error bars represent SEM (n=3). The addition of B[a]P to fusion protein did not result in a significant increase (p > 0.05) in fluorescence over time when compared to wells containing the fusion protein and no B[a]P.
Figure 5.6: *In vitro* assay utilizing E2 anti-B[a]P cp-GFP protein. A) Data represents mean fluorescence of triplicates that were untreated (no B[a]P) or were treated with B[a]P, and measured over 25 hours. B) Mean fluorescence at each hour was expressed as a percentage change in fluorescence with respect to time 0 and plotted against time. Only the time points that displayed a significant difference in fluorescence between the two groups are shown. Error bars represent SEM (n=3). The addition of B[a]P to fusion protein did not result in a significant increase (p > 0.05) in fluorescence over time when compared to wells containing the fusion protein and no B[a]P.
Concentration-response curve

The concentration-response curves generated from assays utilizing both E1 and E2 fractions (0.82 nmol) of anti-B[a]P cp-GFP protein treated with three different B[a]P concentrations, did not display a significant correlative relationship between fluorescence and B[a]P concentration (Figure 5.7 A, B). Although there was no significant difference in mean fluorescence among wells treated with the three different B[a]P concentrations, wells treated with B[a]P did display higher mean fluorescence than wells not treated with B[a]P.

5.2. Discussion

Anti-B[a]P cp-GFP and cp-GFP (control) were successfully expressed, extracted and purified as shown by Western blot analysis. In vivo data and in vitro immunoassay results were, however, unable to provide sufficient evidence in support of the hypothesis that an antigen binding to the \( V_h \) and \( V_l \) domains would cause a conformational change activating the fused \( \text{cp} \) reporter protein, which in turn would result in a discernible signal increase.

Data from the in vivo assay indicated that fluorescence was detected in all \( E. \text{coli} \) cultures, i.e., those expressing the \( \text{cp} \)-GFP, and those expressing the fusion anti-B[a]P \( \text{cp} \)-GFP protein. Although fluorescence was detected in all, there was no significant difference in emission among treated and untreated \( E. \text{coli} \) cultures expressing the \( \text{cp} \)-GFP and anti-B[a]P \( \text{cp} \)-GFP. The presence of fluorescence in \( E. \text{coli} \) cultures expressing the control protein indicated that the \( \text{cp} \)-GFP protein was active and functional in the circularly permutated form. This was expected since the fused variable domains were not present in the control \( \text{cp} \)-GFP to induce the reporter to exist in an alternate conformation which was hypothesized to keep the fusion
Figure 5.7: B[a]P concentration-response curve. Anti-B[a]P cp-GFP protein (0.82 nmol), A) E1, B) E2, were treated with three different concentrations of B[a]P; 10², 10⁴, 10⁶ pg/ml. Mean fluorescence was measured over 39 hours, and data expressed as a ratio of change in fluorescence with respect to concentration 0, then plotted against antigen concentration. Error bars represent SEM (n=3). There was no significant difference (p > 0.05) when mean fluorescence of fusion protein treated with the three different B[a]P concentrations was compared. However, there was a significant difference between fluorescence from fusion protein that was untreated (no B[a]P) and fusion protein treated with each of the three concentrations of B[a]P.
protein inactive in the absence of B[a]P. Thus the cp reporter proved to be a satisfactory control by confirming that the sf-GFP was indeed correctly circularly permutated in a fashion that did not abolish its fluorescence, and that the fluorescence was unaffected by B[a]P since a change in emission was not observed upon addition of antigen. The fluorescence of cultures expressing the fusion protein was high and also not affected by the absence or presence of B[a]P. The high fluorescence observed for the fusion protein may be due to the intermolecular self-assembly (Figure 5.8) of the fusion protein. The non-covalent interaction of variable heavy and light chain domains with neighbouring $V_H$ and $V_L$ domains may have led to the conformational change in activated the cp-GFP. This would cause the fusion protein to fluoresce even in the absence of B[a]P resulting in a high background signal. A study by Nakai et al. (2001) reported the
development of a Ca\(^{2+}\) probe using the circularly permutated form of another variant of GFP, the enhanced GFP (EGFP). Constructed similar in design to the anti-B[a]P cp-GFP protein, this probe was developed by fusing M13 and calmodulin fragments to cp-EGFP via linkers. Just as the binding of B[a]P to V\(_H\) and V\(_L\) domains was expected to induce a conformational change in the cp-GFP, the Ca\(^{2+}\) probe constructed by Nakai et al. (2001) worked in a similar fashion. The binding of the target (Ca\(^{2+}\)) resulted in a conformational change in the cp-EGFP which shifted the fluorescence intensity of the probe, thus presenting a higher signal than the background noise. Based on this reported shift, it is possible that a similar shift in fluorescence intensity also occurred for the anti-B[a]P cp-GFP protein. The cp-GFP alone was active but the fluorescence intensity was low (background), however, once the fusion protein self-assembled and the scFv was stabilized it subsequently also stabilized the cp-GFP, thus activating it and causing a shift in fluorescence intensity. Consequently, the background signal of the fusion protein appeared to be higher than the control. Additionally, the observation of an apparent negligible effect of B[a]P on fluorescence may be a result of B[a]P not being taken up by E. coli.

Several factors can influence the uptake of B[a]P including the strain of E. coli. It was proposed (not yet identified) by Moore and Harrison (1965) that a structure mainly responsible for B[a]P binding is associated with the cell wall-membrane complex of E. coli and possesses an affinity for B[a]P. Additionally, the concentration of this structure was proposed by the authors to vary from strain to strain; they also suggested that the physiological state of the cell influences the affinity of this structure for B[a]P by increasing or altering the structure. The non-suppressor E. coli HB2151 strain used in the in vivo assay might not possess this proposed cell wall structure in abundance, consequently imposing a barrier to sufficient B[a]P uptake. Also,
incubation temperature significantly affects the physiological state of a cell and incubation of cultures at 26°C may not have been the optimum temperature for B[a]P uptake. Since the affinity for B[a]P varies in accordance to the physiological state, cells at a suboptimal state may have a lower affinity and uptake for B[a]P. Therefore, if due to the preceding reasons B[a]P was not efficiently absorbed and taken up by E. coli, the fluorescence of B[a]P treated cultures would behave as if B[a]P was absent from the reaction and would consequently display the same response as cultures that were not treated with B[a]P. It is also possible that the 15-hour delay preceding data collection was a significant factor in the failure to observe a difference between treated and untreated groups. Due to the known physico-chemical properties of B[a]P, i.e., hydrophobicity and long incubation time before dissolving into solution (Osborne and Crosby 1987), it was hypothesized that B[a]P uptake by E. coli would take several hours. However, incubation of E. coli cultures for 15 hours after adding B[a]P and before recording the first fluorescent measurement, resulted in data not being collected for earlier time points where fluorescence could have gradually increased prior to reaching a plateau at, or shortly before, 15 hours. Subsequent in vitro assay data later revealed that the time period for observing the occurrence of a significant difference in fluorescence between treated and untreated groups was between 0 to 12 hours after B[a]P addition. Although other factors such as B[a]P uptake by E. coli likely affected the outcome of the in vivo assay, it is possible that a different trend might have been observed and a significant increase in fluorescence of B[a]P treated wells would have been perceived if measurements had been taken every hour immediately following B[a]P addition to wells. To remedy these deficiencies, further experiments were performed in vitro and fluorescence was recorded at each hour.
Preliminary immunoassays with crude protein taken directly from the sonication lysate provided results similar to the \textit{in vivo} assay. Wells containing crude protein from sonication lysates of cp-GFP and anti-B[a]P cp-GFP did not display a difference in fluorescence among treated and non-treated groups, i.e., no significant difference in fluorescence was observed among wells containing cp-GFP where B[a]P was or was not added, or among anti-B[a]P cp-GFP protein wells where B[a]P was or was not added. The failure of B[a]P to affect fluorescence of anti-B[a]P cp-GFP could be due to interference from other compounds and host proteins present in the crude mix. This interference could inhibit dissolving of B[a]P into the reaction or prevent the interaction between fusion protein and B[a]P, which would in turn prevent the variable scFv domains from coming together and resulting in a conformational change to stabilize and activate the cp-GFP in the fusion protein. In addition, the fluorescence observed in both treated and untreated wells was likely due to the high background resulting from self-assembly, as discussed earlier. To eliminate potential hindrance from interfering substances and to decrease the background noise, protein purification was performed and the purified fractions E1 and E2 were used in an assay.

Imunoassay utilizing the E1 and E2 fractions of purified anti-B[a]P cp-GFP provided similar results, and no difference in fluorescence was observed between treated and untreated groups. Fluorescence was measured over 25 hours and only time points that displayed a significant difference in fluorescence were exclusively selected for further statistical analysis. A significant difference in fluorescence was observed at seven to eight time points between 0 – 12 hours, but the overall mean fluorescence of both E1 and E2 wells treated with B[a]P was not significantly higher than wells not treated with B[a]P (Figure 5.5 A, 5.6 A). When fluorescence
was compared at each time point, the time course graphs (5.5 B, 5.6 B) corroborated the final mean fluorescence results. In addition, it was observed that the trend for both treated and untreated control groups was similar, as fluorescence increased within an hour of B[a]P addition and kept increasing gradually until reaching a maximum at 7 (E1) and 10 (E2) hours, followed by a plateau. Both E1 and E2 fractions were expected to give similar results since the only difference between them was the concentration of protein. The concentration of protein in the E1 fraction was higher than E2, thus the overall mean and percentage change in fluorescence was consequently observed to be higher in the E1 assay. Of note was the fluorescence of the untreated purified protein group in comparison to the untreated crude protein group. The untreated purified protein displayed higher fluorescence than untreated crude protein. This observation supports the former supposition, that interfering substances were indeed present in the crude mixture hindering not only interaction between antigen and variable antibody fragments, but also between neighbouring variable heavy and light chain domains thus preventing self-assembly and generation of a background signal (Figure 5.8). Furthermore, data support the idea that the quality of protein used is a crucial factor in the development of a successful immunoassay. Purified protein alone, however, could not be the only critical factor, since assays utilizing the purified E1 and E2 fractions of anti-B[a]P cp-GFP protein did not provide supporting data in favour of the hypothesis. Since interference from substances assumed to be present in the crude protein mixture and preventing the B[a]P from dissolving into solution was not a concern in this assay, the specific reason for not seeing a difference between B[a]P treated and untreated groups was therefore attributed to the intermolecular self-assembly of the fusion protein. This generation of a high background signal
as a result of self-assembly prevented observation of a significant difference in fluorescence in the presence of B[a]P. Consequently, experiments were performed with lower and equivalent moles of E1 and E2, investigating the concentration-response relationship between the fusion anti-B[a]P cp-GFP protein and three B[a]P concentrations; 1 µg/ml, 10 ng/ml, 100 pg/ml.

A significant difference in fluorescence was observed between wells not treated with B[a]P and wells treated with the three B[a]P concentrations when the same number of moles of E1 and E2 purified protein were used in a concentration-response experiment. In contrast, no correlative relationship could be established between B[a]P concentration and anti-B[a]P cp-GFP fluorescence, i.e., there was no significant difference in fluorescence among wells treated with the different B[a]P concentrations. As expected, the trend for both fractions (Figure 5.7 A, B) was similar, and while in both cases there was a significant difference in fluorescence between treated and untreated wells, there was no difference in treated wells among the three B[a]P concentrations. Karsunke et al. (2012) calculated the sensitivity of the anti-B[a]P scFv (comprising the fusion protein in this study) by performing an indirect competitive ELISA using different B[a]P concentrations including 1 µg/ml, 10 ng/ml and 100 pg/ml to construct a calibration curve (Figure 5.9). The results of the concentration-response experiment did not conform to the prior indirect competitive ELISA study. Fluorescence of anti-B[a]P cp-GFP did not increase with increasing B[a]P concentration. This discrepancy could be due to several reasons, including the presence of 10% methanol used to prepare the B[a]P standards.

According to calculations the number of moles of B[a]P evidently decreased with respect to concentration (1 µg/ml—0.79 nmol, 10 ng/ml—7.9x10⁻³ nmol, 100 pg/ml—7.9x10⁻⁵ nmol), thus theoretically a varied response in fluorescence should have been observed.
Although these numbers were above the low solubility of B[a]P (6.34 x 10^{-6} nmol/µl), the presence of 10% methanol as a co-solvent ensured that the solubility was enhanced. However, the effect of methanol on the fusion protein was not quantitatively determined. Matrix effects have been known to affect the sensitivity of an ELISA. In a previous study it was reported that the concentration of methanol dramatically affected the affinity of the monoclonal anti-B[a]P antibody used to perform an assay, which consequently affected the sensitivity of the ELISA.

![Graph](image)

**Figure 5.9: Indirect competitive ELISA of B[a]P-specific scFvs.** “L2H2” was the anti-B[a]P scFv used to design the fusion protein. IC_{50} = 13.1µg/L. Error bars represent standard deviation (n=3).

(Scharnweber et al. 2001). Also, the fusion protein is considerably larger than the scFv so it is possible that it would be more sensitive to the presence of the solvent than the much smaller scFv even at a lower methanol concentration, due to the larger surface area available for interaction including the binding site. Hence the affinity would likely be affected more dramatically for the fusion protein than the scFv—not enough to completely retard binding to B[a]P, but enough to affect sensitivity. Therefore, it is possible that the affinity of the variable domains was weakened and the IC$_{50}$ of 13.1 µg/L increased due to the presence of methanol.

Additionally, the B[a]P standards were stored in glass containers. Significant sorption of PAHs including B[a]P to glass has been reported before (Qian, Posch, and Schmidt 2011). Therefore, the amount of B[a]P in all three treatments might in reality have been close to being equivalent instead of the expected differences between the concentrations 1 µg/ml, 10 ng/ml, and 100 pg/ml. This would explain the fixed response in fluorescence as a plateau with respect to varying B[a]P concentrations.

Lastly, several aspects in the design of the fusion protein construct itself could have negatively affected the scFv affinity. The composition and length of linkers, the site of attachment of linkers to the variable domains, and the proximity of cp reporter causing steric hindrance, are all probable factors that in previous circular permutation studies have been shown to affect protein stability and ligand affinity. These structural traits of the fusion protein could have prevented the development of the native conformation of scFv which was previously responsible for tightly binding to B[a]P and the resultant strong affinity.

Ultimately, regardless of the concentrations of B[a]P, the data does clearly show that at 0.82 nmol of fusion protein the background fluorescence was low enough that a statistically
significant increase in fluorescence was observed upon B[a]P addition. Furthermore, this increase in fluorescence was greater than that seen when the same number of moles of cp-GFP were or were not supplemented with the same concentrations of B[a]P.

5.3. Conclusion

There is encouraging evidence in support of the hypothesis demonstrating that the presence of B[a]P does indeed bring the variable domains of the anti-B[a]P scFv together thus causing a conformational change in the circularly permutated GFP, which results in an explicit high intensity fluorescent signal indicative of antigen presence. The quality (purified protein free from interfering contaminants) and a low amount of fusion protein are important factors in developing a successful and working single-step visual immunoassay to detect an antigen such as B[a]P. Although supportive data is limited, it is simultaneously highly informative and does not preclude the prospect of developing a working assay with this fusion protein.
6. FUTURE RESEARCH

The most essential step in the development of a single-step immunoassay is the construction of the fusion protein. Assuming that the circularly permutated sequence of the reporter protein is obtained from literature (since several studies exist that have already established functional permutations), emphasis should focus on the characteristics of the scFv with which it will be fused. The length and composition of the linkers (residues that allow flexibility), the proximity of the cp reporter (close enough for transduction of conformational change), and the site of linker attachment of the cp reporter to the scFv domains (without disruption of residues of the binding site) are all important factors to consider when designing the construct of a fusion protein that retains its affinity and reporter function. The use of a high affinity scFv is ideal as it would predictably lead to the creation of a high affinity fusion protein. On the other hand, if fusion of the cp reporter does affect the scFv then the high affinity may hypothetically only be weakened rather than abolished. It is also important to have sufficient knowledge of the scFv residues involved in antigen binding, epitope distance, and folding that led to three-dimensional structure of the wild-type scFv so that an appropriate site of attachment can be selected for the reporter that will not interrupt the affinity / sensitivity, or the formation of secondary structures leading to the development of the scFv’s tertiary structure. It is also important to consider the orientation of the circularly permutated protein to be fused to the variable domains (i.e., $\text{V}_\text{H}$-(N-terminal cp)-(cp C-terminal)-$\text{V}_\text{L}$ VS $\text{V}_\text{L}$-(N-terminal cp)-(cp C-terminal)-$\text{V}_\text{H}$). Differences in the activity of fusion proteins have been observed when altered orientations e.g., N-protein vs protein-C- terminal fusions were compared (Casey et al.
2000; Nakai, Ohkura, and Imoto 2001; Huang and Shusta 2006), as some orientations are more optimal than others.

Once synthesized, the affinity of the scFv in the fusion protein should also be verified to be as close to the original scFv as possible by performing an ELISA. If the affinity is negatively affected due to linkers or fusion to the cp reporter then the sensitivity of the assay will be compromised.

With respect to the anti-B[a]P cp-GFP protein, the length and composition of linkers fusing the circularly permutated reporter protein and variable antibody fragments can also be adjusted or modified completely to prevent intermolecular self-assembly and, hence, generation of a high background signal. The fusion protein was also synthesized in only one orientation (VH-(N-terminal cp)-(cp C-terminal)-VL) and it is possible that the opposite orientation, i.e., VL-(N-terminal cp)-(cp C-terminal)-VH may perform better. Therefore, fusion proteins of different orientations can be designed and compared before following up with in vitro experiments.

The affinity of the scFv in the anti-B[a]P cp reporter fusion should also be verified by an ELISA. It is also possible that another scFv with higher affinity for the target antigen could be used that would result in a higher intensity of fluorescence thereby creating a higher signal-to-noise ratio.

In addition, the effect of solvent concentration (10% methanol used to prepare B[a]P standards) on the activity of the fusion protein should be quantitatively established. Alternatively, since B[a]P has low water solubility and takes a long time to dissolve into solution before reaching a concentration equilibrium, an antigen with higher water solubility may prove
to be an easier target for the development of a correlative relationship derived from the concentration-response curve. In this case sorption to glass containers would also be minimized and enhancement of low solubility with a co-solvent would not be needed; thus factors that may negatively affect scFv affinity and sensitivity could be reduced.

In the case of the fusion protein with cp-BLA as a reporter, substrate selection can be modified so that a different β-lactamase substrate, e.g., fluorocillin (a fluorescent alternative to the chromogenic nitrocefin), could be used thus preventing non-specific interaction with imidazole in the elution buffer. If nitrocefin is used then a different method of protein purification must be used so that the elution buffer does not contain imidazole. Finally, a further attempt should be made to remove imidazole from the eluted fraction of purified protein by employing desalting columns packed with a different resin than the one used in this study, or by using buffer exchange to transfer protein into a solution without imidazole.

Ultimately, this thesis showed that the development of a single-step visual immunoassay utilizing circularly permutated reporter proteins is a practical idea and an attainable goal. The results of this research have elucidated the specific aspects that can be improved upon and add to the knowledge that is already established, thereby leading the way toward a practical cp-based, one-step immunoassay.
7. LITERATURE CITED


APPENDIX I: Anti-B[a]P cp-BLA amino acid sequence (527 aa; MW 57274.5 Da)

1 MAMAEVKLMESSGGGLVQPQGGLSKLCAASGFTFSNYGMAMWRQTPDKRLELVATIYSDGV

61 ITYYPDVKGRFTISRDNAKNTYLMGSLRSEDASVYCSRRGNGNYGDYWQGQTTLTV

121 SSSGGGGSGFNEAIPNDEDDTTMPAAPATTLRKLTTLLASRRQQLIDWMEADKVAGP

181 LLRSALPGWFIADKSGAGERGSGGIAALGPDGKPSRIVVIYTTGSAATMDERNRQIAE

241 IGASLKHWDKSHPEVTLVKDAEDQGARVGYIELDLNGLSKEFPEERFPMMSTFK

301 VLLCGAVLSRIDAQGQGRLRHIYSQNDLVEYSPVTEKHLTDGMTVRLECSAAITMSDNT

361 AANLTTTIGGPKELTAFHNMGDHVTRLRDEPELNEAKGSGGGGGSTDIVMTQSHKF

421 MSTPGDVSITCKASQDVSTAVAWYQQPGQSPKLLIIYWASTRHTGVPDRFTGSGSGTD

481 YTTLISSVQAEDLALYYCQQHYSTPFTFGSGTKLEIKRAAAHHHHHH

Underlined residues = Variable heavy and light chain domains of anti-B[a]P scFv

Residues in italics = linkers (in order: GFG, DKS, KGSG)

Residues in grey = His6 affinity tag

Residues in bold = cp-BLA (168-286 aa-linker-24-170 aa)
APPENDIX II: Anti-B[α]P cp-GFP amino acid sequence (505 aa; MW 54961.3 Da)

1  MAMAEVKLMESSGGGLVQPQGSLKLSCAASGFTSNYGMAMWIRQTPDKRLLEVATIYSDGV

61  ITYYPDSVKGRFTISRDNKNTLYQMGSLRSEDSAVVYCSRRGNNGYDYWGGQGTTLTV

121  SSSGGGGGSDGSVQPADHYQQNTPIDGPVLLPDNYLSTQSVLSKDPNEKRDMV

181  EFVTAGITHGMDELKGGS GG SMKGEELFTGVPVILVLDGDVNGHKFSVRGEGE

241  DATNGKLTLKFIETGKLPWPWTTLTYVQCFSPDPHMKRDFTDRSAMSAPGEYVQE

301  RTISKFDDGNYKTRAEVKFDGLTVNRIELKGTDFKEDGNLGHKLEYNFNSHNYTAD

361  KQKNGIKANFKIRHNVEGGSGGGSTDIVMTQSSHKFMSTPVGDRSITCKASQDVSTA

421  VAWYQQPKGQSPKLIYWASTRHTGVPDRFTGSGTDTYTLTSSVQAE DLALYYCQHY

481  STPTFGSGTKLEIKRAAHHHHHH

Underlined residues = Variable heavy and light chain domains of anti-B[α]P scFv

Residues in italics = (G₅S)₂ linker

Residues in grey = His₆ affinity tag

Residues in bold = cp-GFP (173-238 aa-linker-1-172 aa)
APPENDIX III: cp-GFP sequence (246 aa)

1 DGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSDKPNEKRHDHMLLEFVTAAAGITHG

61 MDELYKGGS5GGGSMKGEELFTGVVIPVAILGEDVNGHKFSVRGEDEGDATNGKLTG

121 ICTTGKLVPWPLVTTLYGVQCFSYPDHMKRHDFRSAMPEGYVQERTISFDDDNGY

181 KTRAEMKFEGDTLVRRIELGTDFKEGNILGHKLEYNFNSHNYTVADSKNGIKANFK

241 IRHNVE

Residues in italics = (G,S)2 linker
Underlined residues = 173-238 aa
Non-underlined residues = 1-172 aa
**APPENDIX IV: sf-GFP sequence (non-cp)**

1) Superfolder GFP sequence in the fusion protein (obtained from Waldo and Cabantous 2013):

```
1 MSKGEELFTG VVPILVE LDG DVNGHKFSVR GEGEGDATNG KTLKFICTT GKLVPWPWT L
61 VTTLT GYVQC FSRYPDHMKR HDFFRSAMPE GYVQERTISF KDDGNVKTRA EVKFEGDTLV
121 NRIELKGD TDF KEDGNILGHK LEYNFNSHNV YITADKQKNG IKANFKIRHN VEDGSVQLAD
181 HYQQNTPIGD GPVLLPDNHY LSTQSVLSKD PNEKRDMVL LEFVTAGIT HGMDELYK
```

Bolded residues = residue of the deposited sequence in PDB that did not match the reported amino acid mutation as stated in the publication, hence was changed accordingly.

Residues in bold and italics = amino acids that were shown to be different from the deposited sequence in PDB by sequence alignment but were not changed.

2) Superfolder GFP sequence from protein data bank (PDB ID: 2B3P)

```
1 MSKGEELFTG VVPILVE LDG DVNGHKFSVR GEGEGDATNG KTLKFICTT GKLVPWPWT L
61 VTTLG YVQC FSRYPDHMKR HDFFSSAMPE GYVQERTISF KDDGTYKTRA EVKFEGDTLV
121 NRIELKGD TDF KEDGNILGHK LEYNFNSHNV YITADKQKNG IKANFKIRHN VEDGSVQLAD
181 HYQQNTPIGD GPVLLPDNHY LSTQSVLSKD PNEKRDMVL LEFVTAGIT HGMDELYKGS
241 HHHHHH
```

Underlined residues = residues not included in the construction of the fusion protein (sequence alignment did not match the 238 aa long wtGFP).