University of Crete

Master's Program "Protein Biotechnology"

Departments of Biology and Chemistry

Laboratory of Natural Biomaterials – Department of Materials Science and Technology

Master's Thesis: Georgilis Evangelos

"Natural fibrous proteins as scaffolds for nanotechnology"

Supervisor: Prof. Michael Kokkinidis

Co-supervisor: Prof. Anna Mitraki

Heraklion 2016

Πανεπιστήμιο Κρήτης

Μεταπτυχιακό πρόγραμμα «Πρωτεϊνική Βιοτεχνολογία»

Τμήματα Βιολογίας και Χημείας

Εργαστήριο Φυσικών Βιοϋλικών – Τμήμα Επιστήμης και Τεχνολογίας Υλικών

Μεταπτυχιακή διατριβή: Γεωργιλής Ευάγγελος

«Φυσικές ινώδεις πρωτεΐνες ως ικριώματα νανοτεχνολογίας»

Επιβλέπων Καθηγητής: Καθ. Μιχαήλ Κοκκινίδης

Συνεπιβλέπων Καθηγητής: Καθ. Άννα Μητράκη

Ηράκλειο 2016

ACKNOWLEDGEMENTS

The completion of this thesis is the second milestone in my academic pathway. For this reason I am grateful to Prof. Anna Mitraki, who has been supportive since my undergraduate studies. During our 5-year collaboration, I learnt a lot of things by her side and became qualified enough to continue my studies as a PhD candidate. I do hope that our collaboration is going to continue. Along with Prof. Mitraki, I would like to thank the other two members of my three-member committee, Prof Michael Kokkinidis and Principal Researcher Kyriacos Petratos for their excellent collaboration and mentoring.

Furthermore, I am grateful to the staff members of the laboratory of Electron Microscopy: Ms Alexandra Siakouli, Ms Aleka Manousaki, Mr Stefanos Papadakis and Ms Sevasti Papadogiorgaki for their contribution in electron microscopy experiments and excellent collaboration. Also, I would like to thank Prof. Georges Chalepakis for his help in those experiments and for offering space in his laboratory for me. I would also like to thank PhDc Nikos Louros from Prof. Hamodrakas's laboratory, University of Athens, for helping with part of the electron microscopy experimentation.

Next, I would like to thank my colleagues: Dr Fani Tsitouroudi, Dr Paraskevas Lamprou, PhDc Ariadne Prigipaki, PhDc Chrysoula Kokotidou, PhDc Graziano Deidda, MSc Eirini Ornithopoulou and the undergraduate students Marita Vassila, Clyto Katsara and Apostolos Hatzoudis. Their help has been of paramount importance for the progress of my thesis and they created a nice working environment, which was equally important.

The contribution of my family to this work has been crucial throughout my studies in the University of Crete. I appreciate their loving support and I am grateful to them for their help. I would also like to thank my friends for the good times we had together and for their help at hard times.

This work is dedicated to the loving memory of my uncle Antonis Axiotis and my grandmother Stella Triantafyllou.

Last but not least, I would like to acknowledge the financial support from "PHOTOPEPMAT" Design, production and Laser PHOTO structuring of self-assembling PEPtides and proteins destined for bioMATerials applications" Greek Secretariat for Research and Technology, ARISTEIA II Excellence grant number 3941 which is part of the Action "Education and Lifelong Learning" and co-financed by the European Social Fund and National Funds.

Contents:

Chapter 1 – Abstract-Περίληψη

The present study was conducted in the laboratory of Natural Biomaterials, Dept. of Materials Science and Technology, University of Crete. The research was focused on the development of self-assembled protein materials that can be used as scaffolds in nanotechnology. The system used for this purpose was the β-helical domain of the P22 phage tailspike (Bhx). The sequence of Bhx is aggregation-prone and thermoresponsive. Moreover, it binds to oligosaccharides and has been shown to form irreversible, fibrous aggregates during refolding experiments. In this study, spherical and fibrous microstructures of Bhx were developed and observed via electron microscopy. These studies revealed a hierarchical formation of the microstructures from nano-filaments. Inspired by the β-helical virulence factor of *Bordetella pertussis* that mediates mammalian cell adhesion, an RGD-motif was inserted to the sequence of Bhx via site-directed mutagenesis for the development of biocompatible fibrous scaffolds that might promote cell adhesion and proliferation. Furthermore, an emerald green fluorescent protein (GFPem) and a biotinylation peptide sequence were introduced to the sequence of Bhx via the seamless cloning method. GFPem was fused to both the N- and C-terminus of Bhx, providing data on its degradation resistance and functionality of enzymes in inclusion bodies.

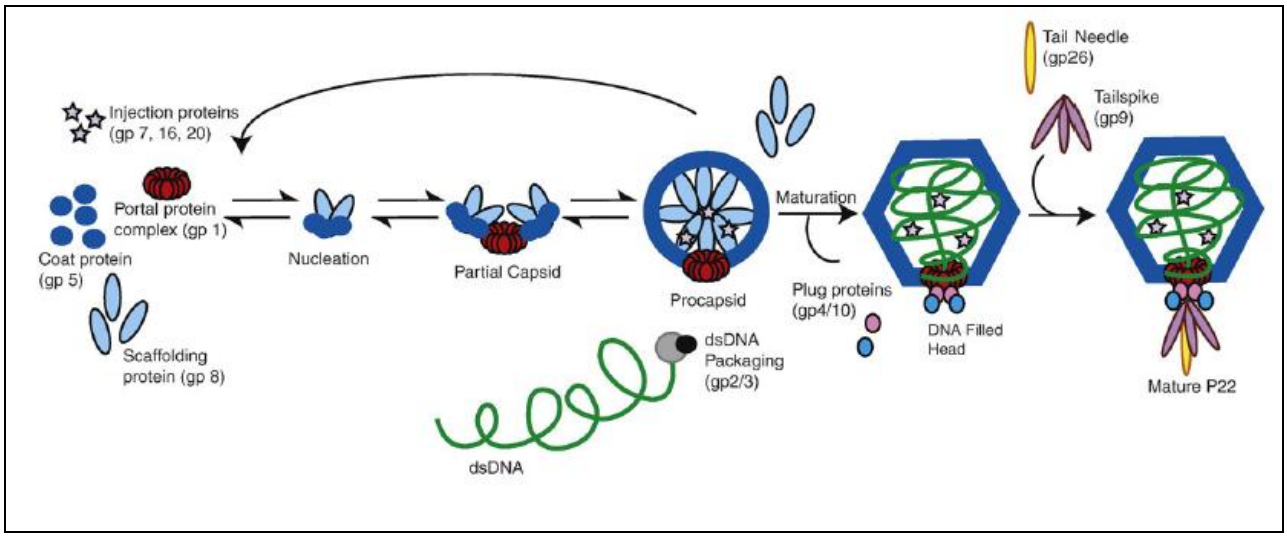
Η παρούσα μελέτη διεξήχθη στο εργαστήριο Φυσικών Βιοϋλικών, του Τμήματος Επιστήμης και Τεχνολογίας Υλικών στο Πανεπιστήμιο Κρήτης. Η έρευνα εστιάστηκε στη δημιουργία αυτο-οργανωμένων πρωτεϊνικών υλικών τα οποία μπορούν να χρησιμοποιηθούν ως ικριώματα στη νανοτεχνολογία. Το σύστημα που χρησιμοποιήθηκε για αυτό το σκοπό ήταν η β-ελικοειδής δομική περιοχή της πρωτεΐνης της ουράς του φάγου P22 (β-helical domain of the P22 phage tailspike - Bhx). Η αλληλουχία της Bhx είναι επιρρεπής στη συσσωμάτωση και θερμοαποκρινόμενη. Επιπλέον, προσδένεται σε ολιγοσακχαρίτες και έχει βρεθεί ότι σχηματίζει μη-αντιστρεπτά, ινώδη συσσωματώματα σε πειράματα αναδίπλωσης. Σε αυτή τη μελέτη, αναπτύχθηκαν σφαιρικές και ινώδεις μικρο-δομές της Bhx και παρατηρήθηκαν μέσω ηλεκτρονικής μικροσκοπίας. Οι μελέτες αυτές έδειξαν ότι υπάρχει ιεραρχία στο σχηματισμό των μικρο-δομών από νανο-ινίδια. Έχοντας έμπνευση από τον β-ελικοειδή παράγοντα μόλυνσης του βακτηρίου *Bordetella pertussis* ο οποίος μεσολαβεί για την προσκόλληση κυττάρων θηλαστικών, ένα μοτίβο RGD εισήχθη στην αλληλουχία της Bhx μέσω σημειακής μεταλλαξιγένεσης για την ανάπτυξη βιοσυμβατών ινώδων ικριωμάτων τα οποία θα μπορούσαν να προάγουν την κυτταρική προσκόλληση και ανάπτυξη. Επιπροσθέτως, η αλληλουχία πράσινης φθορίζουζας πρωτεΐνης (GFPem) και πεπτιδίου βιοτινυλίωσης εισήχθηκαν στην αλληλουχία της Bhx χρησιμοποιώντας τη μέθοδο κλωνοποίησης χωρίς ραφές (seamless cloning method). Η GFPem εισήχθη και στο αμινο- αλλά και στο καρβοξυ-τελικό άκρο της Bhx, παρέχοντας δεδομένα για την ανθεκτικότητά της σε αποδόμηση και τη λειτουργικότητα ενζύμων σε σωμάτια εγκλεισμού.

Chapter 2 – Introduction

2.1 The P22 Bacteriophage

The bacteriophage P22 belongs to the order Caudovirales, in the family of Podoviridae (Andres et al., 2010; Veesler and Cambillau, 2011). It is a ds-DNA phage with an icosahedral head assembly and a small, non-contractile tail.

Fourteen genes in the 42-kbp genome of P22 encode for the proteins that form its virion (Casjens and Thuman-Commike, 2011; Teschke and Parent, 2010) . The in vivo morphogenic pathway involves two basic steps for the formation of the virion (Fig. 1, top). First, a procapsid is assembled by the coat protein (gene product 5 - gp5), the scaffolding protein (gp8), the ejection proteins (gp7, 16, 20), and the portal protein complex (gp1). The procapsid formation is based solely on the non-covalent interactions of the proteins. DNA packaging by the terminase complex (gp2, 3) is ATP-dependent and leads to scaffolding protein release and expansion of the procapsid (Fig. 1, bottom) (Fuller and King, 1980). The plug proteins (gp4, 10) and the tail needle (gp26) bind to the portal protein complex, retaining the DNA inside the virion. The last step for the mature virion formation is the assembly of up to six trimers of the tailspike protein (gp9) to the plug.



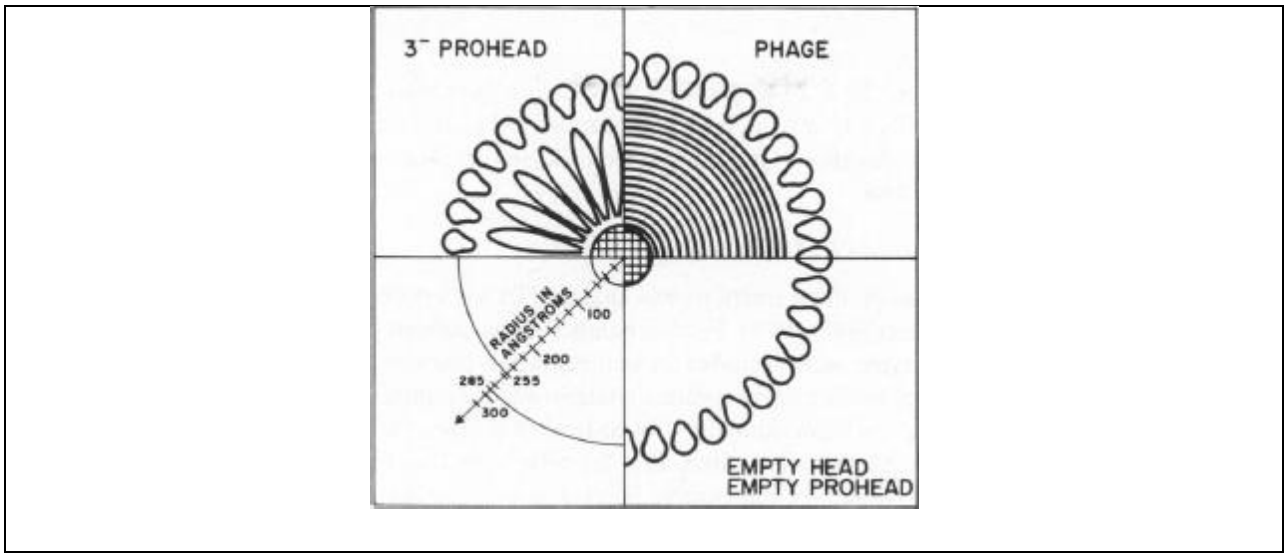


Figure 1: Top: Scheme of the assembly and maturation process of the P22 virion (Teschke and Parent, 2010). Bottom: Size comparison scheme of the procapsid and mature capsid head based on TEM and X-ray scattering observations (Fuller and King, 1980)

The phage P22 infects *Salmonella* strains. The infection is mediated by the attachment of the virus to the bacterial host lipopolysaccharide (LPS) receptor through the tailspike apparatus (Andres et al., 2010). The tailspike proteins bind and hydrolyze the sugar chain of the receptor, bringing the tail needle to contact with the cell envelope. Conformational changes of a C-terminal helical domain of the needle, named the lazo domain (Fig. 2), lead to the release of the needle from the virion, thus the DNA is released and transferred into the cell envelope with the support of ejection proteins (gp7, 16, 20) (Leavitt et al., 2013; Olia et al., 2007; Perez et al., 2009). The phage DNA is circularized in the cytoplasm, and the phage follows either a lytic or a lysogenic pathway (Harvey et al., 1981; Pipas and Reeves, 1979).

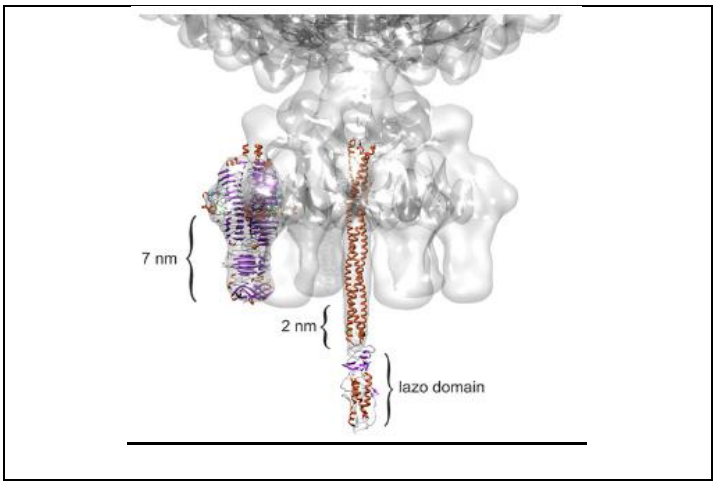


Figure 2: The P22 attachment apparatus (Andres et al., 2010)

2.2 The P22 Tailspike Protein

P22-like phage tailspikes share a similar β-solenoid fold, even though their amino acid sequence is not identical (Barbirz et al., 2009; Casjens and Thuman-Commike, 2011). In the case of P22, three to six 215 kDa homotrimers of the tailspike protein (gp9) are needed to form the attachment apparatus of an infectious phage.

The trimer is formed by identical chains of 666 aa and consists of three distinct domains; an N-terminal domain of around 110 aa that binds to the phage head, a right-handed β-helix domain (Bhx) in the center, and a C-terminal caudal fin and β-prism (Fig. 3, top).

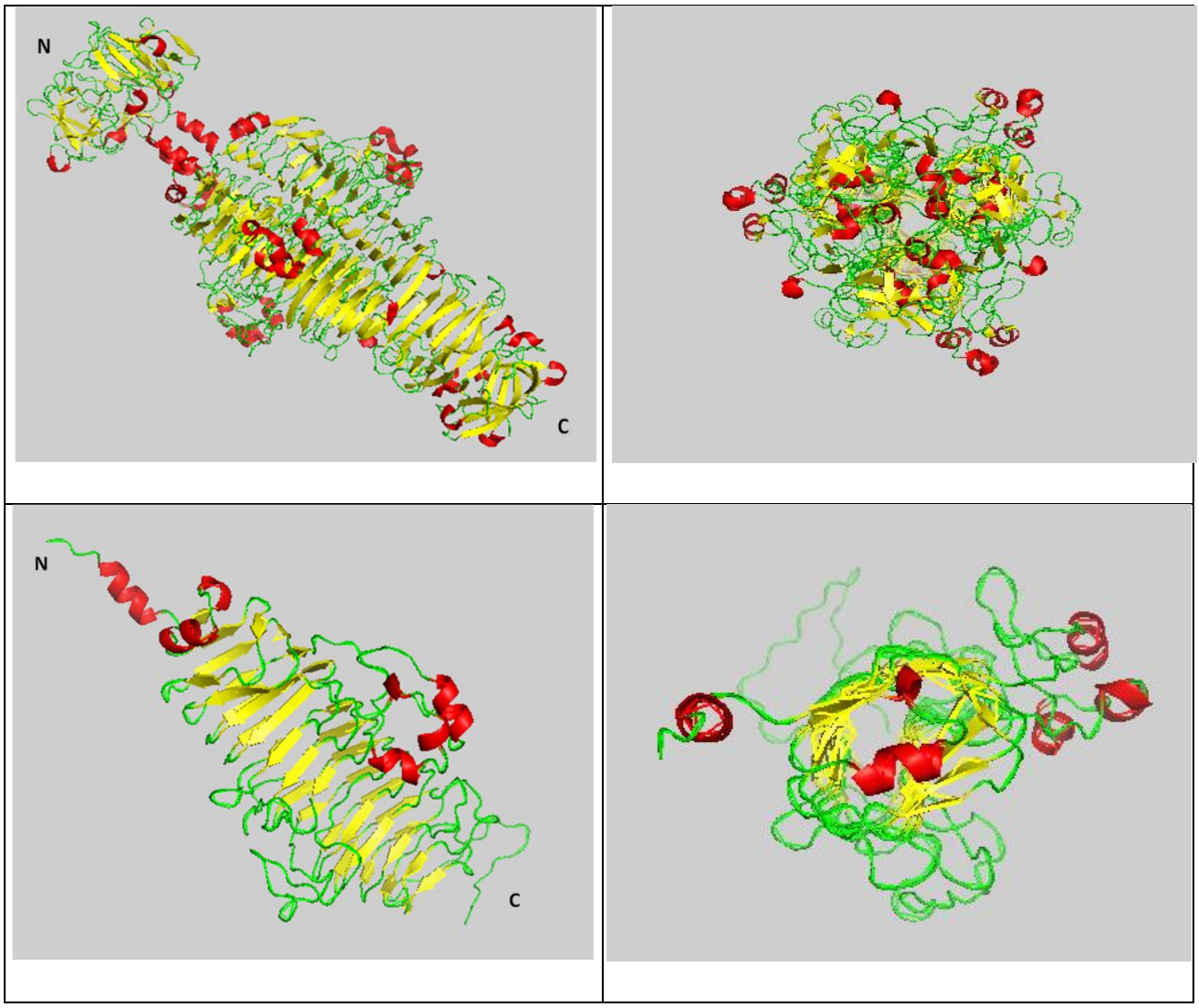


Figure 3: Top: The full-length P22 tailspike homotrimer side (left) and top view (right) (PDB entry: 2XC1). Bottom: The P22 tailspike β-helix domain (residues 108-544, PDB entry: 2XC1). The images were developed with PyMol.

The folding pathway of the tailspike is a temperature sensitive process, highly dependent on inter- and intramolecular interactions, and does not require a chaperone involvement. It involves intermediate states of partially folded monomers, dimers and protrimers (Fig. 4) (Mitraki, 2010). C-terminal transient disulfide bonds and hydrophobic interactions play a critical role in the assembly of the trimer (Gage and Robinson, 2003; Seckler, 1998; Seul et al., 2014; Takata et al., 2012). The mature full-length tailspike remains stable under denaturing conditions (i.e. heat, detergents). Furthermore, the part comprising the β-helix plus the C-terminal domain (residues 108-666) has been reported as stable under similar conditions (Barbirz et al., 2009; Mitraki, 2010). Nevertheless, the intermediate species have been found to be thermolabile and sensitive to detergents. Temperature-sensitive folding (*tsf*) mutations, which are located at the Bhx domain, are responsible for off-pathway aggregation in inclusion bodies. Protein aggregation is a result of intermolecular interactions of non-native folding intermediates, thus the aggregates consist of a majority of tailspike chains, with minor contaminants. Inclusion body formation occurs both during the natural infection process within Salmonella and during heterologous expression of recombinant proteins in E.coli and can be overcome with global suppressor mutations (Beissinger et al., 1995; King et al., 1996; Mitraki, 2010).

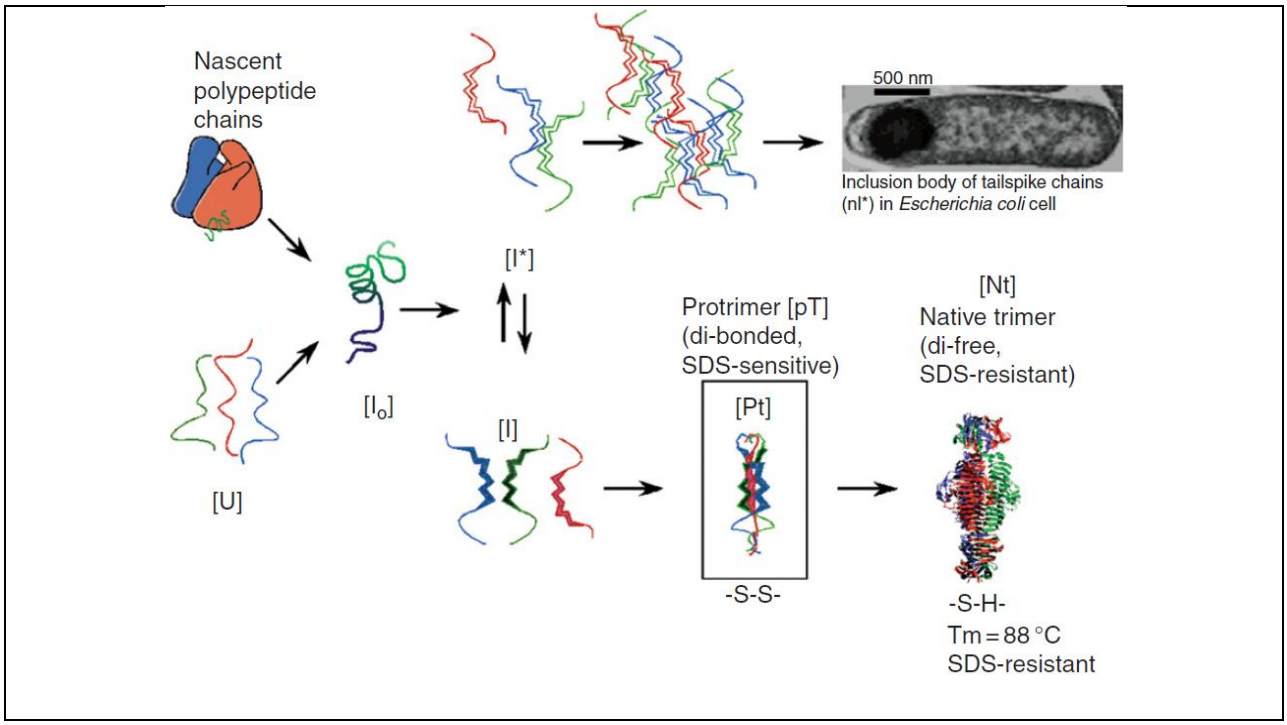


Figure 4: Scheme of the *in vivo* and *in vitro* folding pathway of the tailspike protein (Mitraki, 2010)

The tailspike binds to a variety of *O*-antigen LPS on the host surface that share a trisaccharide repeating unit (α-D-mannose-(1→4)-α-L-rhamnose-(1→3)-α-D-galactose-(1→2)) and cleaves the oligosaccharide chain by hydrolyzing at the Rha-Gal α(1→3)-glycosidic bond giving an octasaccharide product (Andres et al., 2013; Andres et al., 2010). Three catalytic residues (Asp392, Asp395 and Glu359) are responsible for the function of the protein (Fig. 4) (Baxa et al., 1996; Steinbacher et al., 1996; Steinbacher et al., 1997). The active site is located in the Bhx domain (Fig. 3, bottom).

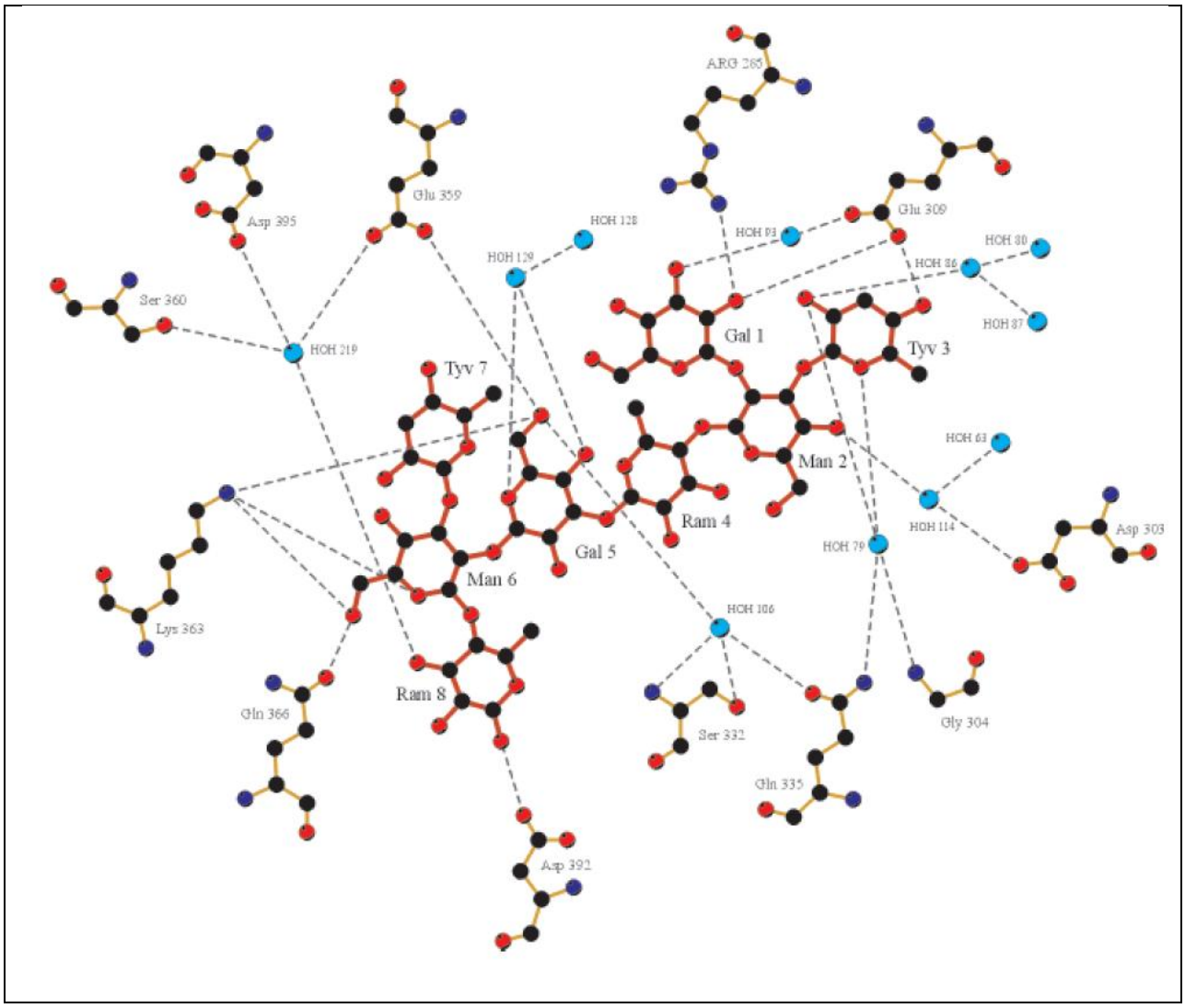


Figure 4: The tailspike binding site with an octasaccharide substrate (Baxa et al. 2001).

2.3 The Tailspike β-helix

The right-handed parallel β-helix, the main domain of the tailspike, is comprised of 13 three-strand rungs (or coils) that wrap around the long axis of the helix (Betts et al., 2004; Kreisberg et al., 2000). The strands are interrupted by turns or loops. The cross-section of the domain is kidney bean-shaped. Most hydrophobic residues are buried in the core of Bhx and support for correct folding (Simkovsky and King, 2006). The β-sheet A interacts with the β-sheet B of the neighboring subunit of the trimer, while the β-sheet C is exposed to the solvent (Fig. 5) (Seckler, 1998; Simkovsky and King, 2006).

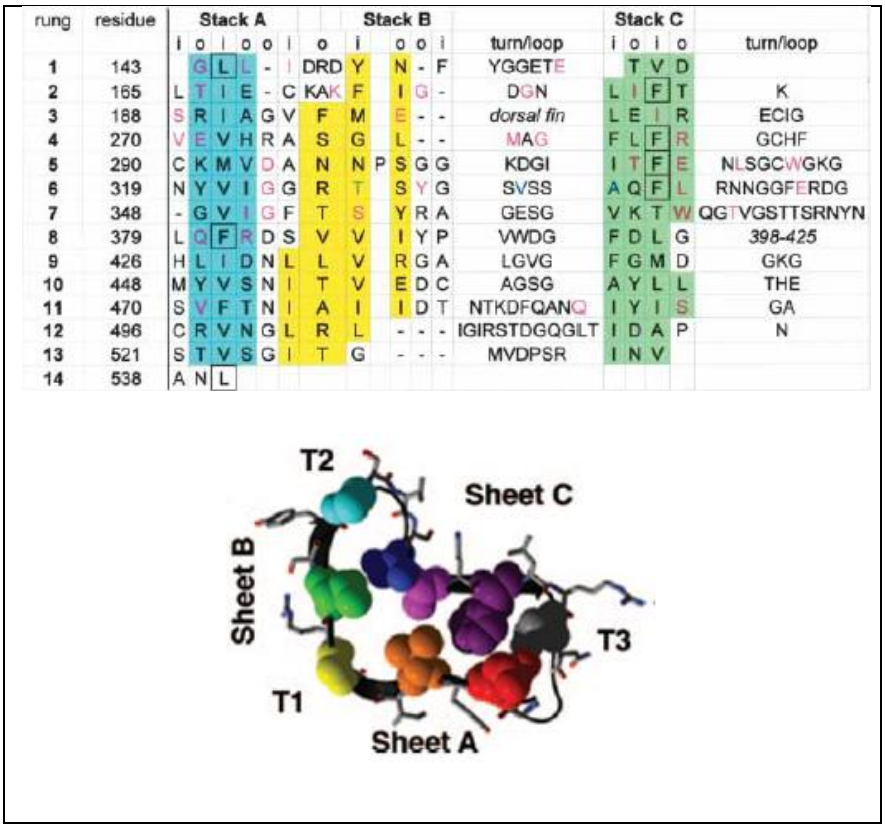


Figure 5: Top: Sequence alignment of Bhx based on structure data (Betts et al., 2004). Residues in the i-column point inwards, while residues in the o-column point outwards. Bottom: Cross-section of rung 6 (Simkovsky and King, 2006). T1, 2, 3 correspond to turns. Atoms of residues that point to the core are shown as spheres.

The β-helix domain of the tailspike protein (residues 108-544) was the model of study in this work. This isolated domain lacks the trimerization motif of the tailspike, but still adopts a native-like fold and can reversibly form trimers at high protein concentrations due to weak hydrophobic interactions (Miller et al., 1998; Schuler et al., 1999). Bhx forms inclusion bodies at elevated expression temperature (37 °C). The binding properties of the isolated domain are comparable to the native tailspike, whereas its cleavage capacity is significantly impaired.

Folding transition studies have shown that reversible unfolding of Bhx requires low ionic strength, low protein concentration, and low temperature. At higher salt or protein concentration, Bhx formed irreversible aggregates, some of which had a fibrillar structure (Fig. 6). Bhx in the

aggregates retains part of its secondary structure and Congo-red stained fibrils were birefringent under polarized light, indicating an amyloid-like assembly (Schuler et al., 1999).

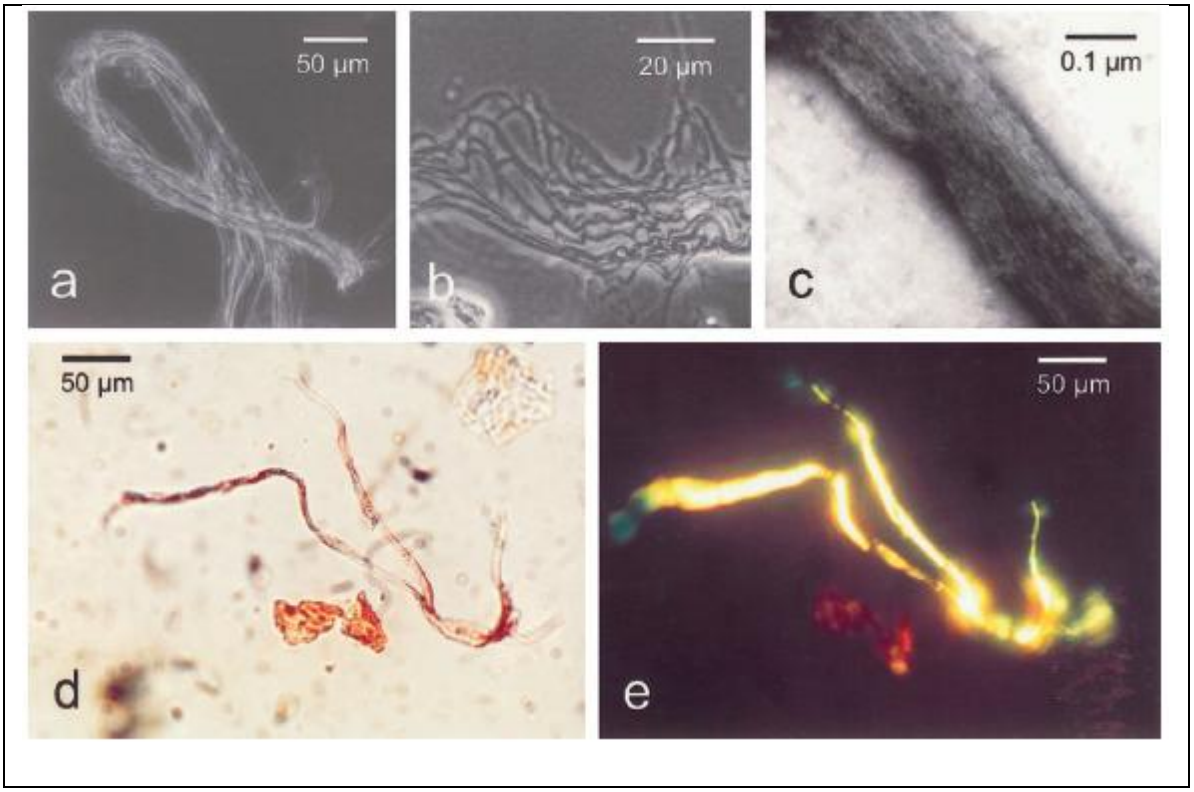


Figure 6: Bhx fibrillar aggregates under electron (a, b, c) and optical microscopy (d, e). (e) shows Congo-red stained fibrils under polarized light. (Schuler et al., 1999)

2.4 Green Fluorescent Protein - GFP

The Green Fluorescent Protein (GFP) was first purified from the photocytes of the bioluminescent jellyfish *Aequorea victoria* (Chalfie, 1995). The structure of this 27 kDa protein is a β-barrel (11 strands) with an α- helix inside, and short helices and loops at the ends of the barrel (Fig. 7) (Yang et al., 1996). GFP has the intrinsic property to fluoresce upon absorption of light in the ultraviolet to blue range. It has a major excitation peak at 395 nm and a minor at 475 nm, while its emission peak is at 508 nm. Fluorescence is a result of a Ser65-Tyr66-Gly67 sequence, located in the α-helix inside the β-barrel. This chromophore sequence undergoes a post-translational cyclization reaction between Ser65 and Tyr66 to form an imidazolinone ring

(Fig.7). Surrounding aminoacids also play an important role in the fluorescence of the protein (Barondeau et al., 2003; Ormo et al., 1996).

The fluorescence properties of this protein led to is utilization in cell biology as an imaging tool for protein trafficking, localization and protein-protein interactions (Snapp, 2005). Furthermore, GFP has inspired the design and development of more fluorescent proteins with enhanced properties. In this study a variant of GFP, emerald GFP (GFPem), was used. This variant has S65T and F64L mutations and four additional point mutations. These mutations enhance solubility and increase brightness (Day and Davidson, 2009).

The cDNA of fluorescent proteins is usually commercially available, thus fusion to the protein of interest is a relatively easy process. Nevertheless, fusion of fluorescent proteins can cause folding complications to the protein of interest (Hovmoller and Zhou, 2004). In that case, a strategy that can be followed is developing two different constructs, one with an N-terminal and one with a C-terminal fusion and also a peptide linker between the proteins (Snapp, 2005).

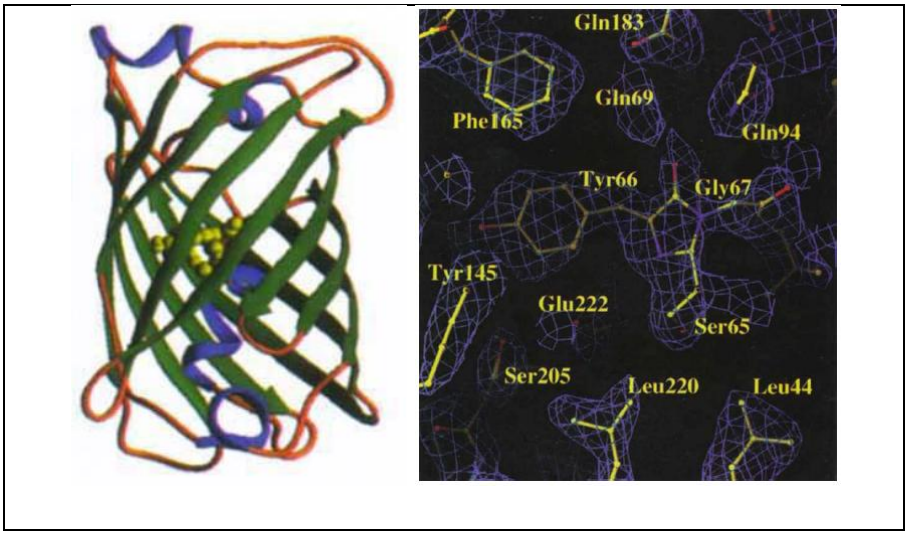


Figure 7: Left: The structure of GFP. The chromophore atoms are shown as spheres. Right: (Yang et al., 1996)

Aggregation-prone variants of GFP have been utilized for testing functional inclusion bodies in therapeutics. This work shows that inclusion bodies contained functional GFP and were biocompatible. They could penetrate cell membranes (even the nucleus) and were orally delivered to organisms without causing side-effects (Vazquez et al., 2012). Moreover, peptide tags can be introduced to the sequence of GFP for targeted delivery of GFP nanostructures with

controlled dimensions (Cespedes et al., 2014). These findings suggest that peptide tags and inclusion bodies can be exploited in order to generate novel protein nanostructures for therapeutic or diagnostic purposes.

2.5 The integrin recognition sequence - RGD

Integrin function is crucial for cell binding to proteins of the extracellular matrix (ECM). Fibronectin, an adhesion-promoting protein in the ECM, binds to integrin generating a cascade of reactions that promote cell adhesion (Alberts, 2002). The recognition site of fibronectin is a tripeptide of Arg-Gly-Asp (or RGD) located on a loop and followed by a Ser and Pro residue (Fig. 8).

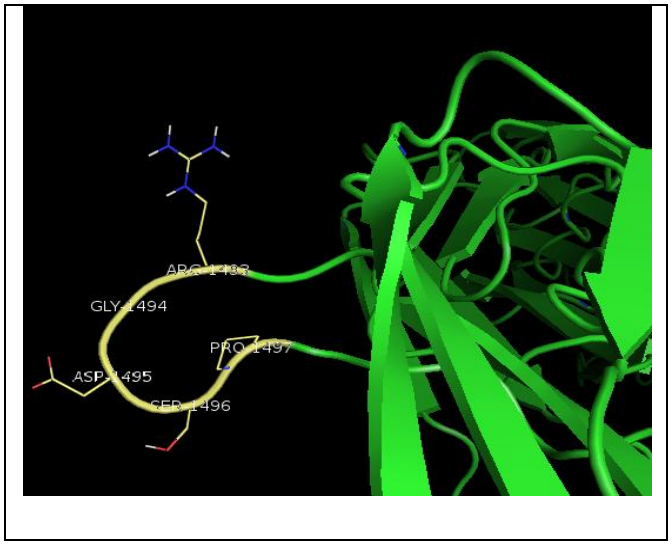


Figure 8: The RGD-loop of fibronectin (PDB entry: 1FNF). The image was developed with PyMol.

The RGD motif has had an important impact on tissue engineering research. It has been introduced (either chemically or genetically) to the sequence of many peptides and proteins for the formation of self-assembled nanostructures that promote cell adhesion and proliferation (Loo et al., 2015). In comparison with natural ECM proteins such as collagen, such systems can be derivatized more easily to possess multiple functions along with cell viability, and thus can also be used in theranostics and biosensor applications.

2.6 Thesis objectives

The aim of this thesis is to produce self-assembled nanostructures based on the P22 tailspike Bhx domain conjugated with the cell adhesion motif RGD or a model enzyme such as GFP. The long-term goal would be to use these as scaffolds for cell attachment, or as fibrous scaffolds for enzymatic catalysis.

Protein-based scaffolds gained interest because protein molecules have specific interactions and catalytic properties, while can be produced in large scale with a low cost. Protein nanostructures usually require ambient conditions for their formation, can be easily derivatized due to the variety of chemical moieties present in the amino acid residues, and their thermal properties can be modulated for desired applications (Bhattacharyya et al., 2015; Vazquez and Villaverde, 2010). Phage proteins have found multiple applications in energy-transfer nanodevices (Koshiyama et al., 2009), catalysis (Koshiyama et al., 2008), and composite materials (Kale et al., 2013). Tailspikes in specific can also serve as bacterial detection devices (Singh et al., 2010). Thus, inclusion bodies and nanostructures of Bhx can be used as tools for the detection of bacteria and as oligosaccharide binding materials.

Nature provides a good example of how β-solenoid proteins can be used in a biocompatible substrate that promotes cell adhesion. Filamentous hemagglutinin adhesin (FHA), the virulence factor of *Bordetella pertussis*, is a secreted β-helical protein that promotes mammalian cell adhesion via an RGD motif located on a loop (Kajava et al., 2001). The mechanical resistance of β-helical proteins like FHA (Alsteens et al., 2013) is another factor that renders them a good candidate for the development of scaffolds in tissue engineering. In order to enhance biocompatibility, an RGD tripeptide was inserted in the sequence of Bhx. RGD moieties need to be exposed or in loop configuration,this is why it was considered a proper design to insert it in the N-terminal region. The tripeptide motif was initially inserted before an Asp and Pro residue (RGDD-Bhx). The Asp was mutated to Ser (RGDS-Bhx) in order to mimic the native sequence of fibronectin, and RGDD-Bhx was kept for proof of principle studies.

GFPem was fused with Bhx in order to create fluorescent nanostructures. Furthermore, the folding behavior during recombinant expression was observed for the N-terminal and C-terminal GFPem fusion. Fluorescent proteins and enzymes fused with aggregation-prone

sequences have been shown to be catalytically active in inclusion bodies and amyloid filaments (Baxa et al., 2003; Garcia-Fruitos et al., 2005). Such aggregates have proven valuable in the field of protein scaffolds. For example, inclusion bodies of GFP fused with the VP1 capsid protein of the foot-and-mouth disease virus, these aggregates have been used for the functionalization of a silicon-based surface to enhance biocompatibility (Garcia-Fruitos et al., 2005; Garcia-Fruitos et al., 2009). Furthermore, protein fusions can provide additional data regarding the folding of the protein in the cellular environment as well as its resistance to degradation (Corchero et al., 1996). To that end, two constructs (N-GFPem-Bhx and Bhx-GFPem-C) were developed and studied for the folding behavior and structure formation.

Chapter 3 – Materials and methods

3.1 Materials

3.1.1. Plasmids and Primers

pASK40 was kindly provided by prof Robert Seckler. The plasmid codes for Bhx and also contains the ampicillin resistance gene and a lac-UV5 promoter (Skerra et al. 1991). Site directed mutageneses were performed on that vector and the Bhx insert gene was amplified and isolated for fusion with GFPem.

The nucleotide sequence coding for the biotinylation peptide was inserted in a pET28a vector by Dr. Tsitouroudi. This vector also contains the kanamycin resistance gene and a T7 promoter. pET28a was used for the fusion of GFPem to the C-terminus of Bhx (Bhx-GFPem-C).

The pT7-7 vector contains the ampicillin resistance gene and a T7 promoter. pT7-7 was used for the fusion of GFPem to the N-terminus of Bhx (N-GFPem-Bhx).

The pBirAcm vector contains a tac promoter, the gene coding for biotin ligase, and the chloramphenicol resistance gene. pBirAcm was used during the expression of N-GFPem-Bhx and Bhx-GFPem-C for biotin ligase-mediated biotinylation.

Vector maps are available in Appendix I.

3.1.2 Polymerase chain reaction - PCR

Reaction for Site-directed Mutageneses:

- 12.5 μL Q5 Hot Start High-Fidelity 2x Master Mix
- 1.25 μL 10 μM Forward Primer (5'-TGACGATCCAGATCAATATTCAATAGAAG-3' for RGDD-Bhx, 5'-GCGTGGTGACAGTCCAGATCAATATTC-3' for RGDS-Bhx)
- 1.25 μL 10 μM Reverse Primer (5'-CCACGCATTTTTTGCCCTCGTTAAC-3' for RGDD-Bhx, 5'-ATTTTTTGCCCTCGTTAAC-3' for RGDS-Bhx)
- 1 μL (1-25 ng/μL) Template DNA
- 9 μL water

Thermocycling Conditions (RGDD-Bhx):

- Initial Denaturation at 98 ^{o}C for 30 sec
- 25 cycles of Denaturation at 98 ^{o}C for 10 sec

Annealing at 60 ^{o}C for 30 sec

Extension at 72 ^{o}C for 3 min

- Final Extension at 72 ^{o}C for 2 min

Thermocycling Conditions (RGDS-Bhx):

- Initial Denaturation at 98 ^{o}C for 30 sec
- 25 cycles of Denaturation at 98 ^{o}C for 10 sec

  Annealing at 59 ^{o}C for 30 sec

  Extension at 72 ^{o}C for 3 min
- Final Extension at 72 ^{o}C for 2 min

Reaction for assembly and cloning into pUC19L:

- 5 μL 10x polymerase buffer
- 1 μL 100 mM $MgSO_4$
- 1 μL 10 mM dNTP mix
- 1 μL 20 μM Forward primer

(for the N-GFPem-Bhx construct 5'-AATTCGAGCTCGGTACATGTCCGGCCTGAACG-3' and 5'-GCCTGGTGCCGCGCGGCAGCATGGATCCAGATCAA-3' where used to amplify GFPem and Bhx respectively;
for the Bhx-GFPem-C construct 5'-AATTCGAGCTCGGTACATGGATCCAGAT-3' and 5'-AAGAAGGGAGCAGCGGCCTGGTGCCG-3' where used to amplify Bhx and GFPem respectively)

- 1 μL 20 μM Reverse primer

(for the N-GFPem-Bhx construct 5'-CCGCGCGGCACCAGGCCGCTCTTGTACAGCTC-3' and 5'-GCCAAGCTTGCATGCCCTAGTGGTGAGGTGAT-3' where used to amplify GFPem and Bhx respectively,
for the Bhx-GFPem-C construct 5'-ACCAGGCCGCTGCTCCCTTCTTCTG-3' and 5'-GCCAAGCTTGCATGCCTTAGTGATGGTGATGGTG-3' where used to amplify Bhx and GFPem respectively)

- 0.5 μL DNA
- 0.5 μL Deep Vent® DNA Polymerase (New England Biolabs)

- water was added to a final volume of 50 µL

Thermocycling Conditions (N-GFPem-Bhx):

- Initial Denaturation at 95 ^{o}C for 5 min
- 30 cycles of Denaturation at 95 ^{o}C for 30 sec
  Annealing at 66 ^{o}C for 30 sec
  Extension at 72 ^{o}C for 2 min
- Final Extension at 72 ^{o}C for 5 min

Thermocycling Conditions (Bhx-GFPem-C):

- Initial Denaturation at 95 ^{o}C for 5 min
- 30 cycles of Denaturation at 95 ^{o}C for 30 sec
  Annealing at 56 ^{o}C for 30 sec
  Extension at 72 ^{o}C for 2 min
- Final Extension at 72 ^{o}C for 5 min

<u>Reaction for Restriction Enzyme Cloning into pT7-7 and pET28a</u>

Amplification of N-GFPem-Bhx:

- 10 µL 5x Phusion HF buffer
- 1 µL 10 mM dNTP mix
- 2.5 µL 10 µM Forward primer (5'-CGGGCATATGTCCGGCCTGAACGA-3')
- 2.5 µL 10 µM Reverse primer (5'-CGCGCTGCAGCTAGTGGTGATGGTG-3')
- 0.5 µL DNA
- 0.5 µL Phusion® High-Fidelity DNA Polymerase (New England Biolabs)
- water was added to a final volume of 50 µL

Amplification of Bhx-GFPem-C:

- 5 µL 10x polymerase buffer
- 1 µL 100 mM $MgSO_4$
- 1 µL 10 mM dNTP mix

- 1 μL 20 μM Forward primer (5'-TATACCCGGGGAATGGATCCAGATCAATAT-3')
- 1 μL 20 μM Reverse primer (5'-TATACTGCAGCTAGTGGTGATGGTGATGGT-3')
- 0.5 μL DNA
- 0.5 μL Deep Vent® DNA Polymerase
- water was added to a final volume of 50 μL

Thermocycling Conditions (N-GFPem-Bhx):

- Initial Denaturation at 98 ^{o}C for 30 sec
- 35 cycles of Denaturation at 98 ^{o}C for 10 sec
  Annealing at 72 ^{o}C for 30 sec
  Extension at 72 ^{o}C for 30 sec
- Final Extension at 72 ^{o}C for 10 min

Thermocycling Conditions (Bhx-GFPem-C):

- Initial Denaturation at 95 ^{o}C for 5 min
- 30 cycles of Denaturation at 95 ^{o}C for 30 sec
  Annealing at 56 ^{o}C for 30 sec
  Extension at 72 ^{o}C for 2 min
- Final Extension at 72 ^{o}C for 5 min

3.1.3 Cloning into pUC19L reaction

- 50 ng pUC19L vector
- 100 ng of the amplified insert genes
- water was added if the gene-vector mixture was less than 5 μL
- 2x GeneArt enzyme mix was added to a 1x final concentration

The mixture was pipetted three times, tapped three to five times, and left for 15-30 min at room temperature. The mixture was subsequently left on ice for 2-5 min and half of it was used for the transformation of DH10B cells.

3.1.4 Double digest (cut) reaction

For a 50 μL reaction

10 μL DNA

33 μL water

The mixture was heated at 70 ^{o}C for 10 min, then cooled on ice for 1 min and spun down before the addition of the remaining reagents

5 μL of a 10x buffer for optimal performance (usually CutSmart® Buffer-New England Biolabs)

1 μL of each restriction endonuclease

3.1.5 Dephosphorylation reaction

18 μL vector

2.1 μL 10x Antarctic phosphatase buffer

1 μL Antarctic phosphatase (New England Biolabs)

Dephosphorylation was conducted at 37 ^{o}C for 15 min and the enzyme was heat-deactivated at 70 ^{o}C for 5 min. The mixture was then cooled on ice for 1 min and spun down.

3.1.6 Ligation reaction

50 ng dephosphorylated vector

37.5 ng insert gene

1 μL T4 ligase (New England Biolabs)

2 μL T4 ligase buffer

water was added to a final volume of 20 μL

The reaction was conducted overnight at 16 ^{o}C

3.1.7 TfbI buffer

30mM CH_3COOK (Sigma)

50mM $MnCl_2$ (Sigma)

100mM KCl (Sigma)

10mM $CaCl_2$ (Sigma)

15% glycerol (PENTA)

water was added to reach the final volume (usually a 100 mL stock solution was prepared)

3.1.8 TfbII buffer

MOPS (Sigma) ph7 10 mM

75mM $CaCl_2$

10mM KCl

15% glycerol

water was added to reach the final volume (usually a 100 mL stock solution was prepared)

3.1.9 Tris-Acetate-EDTA (TAE) Buffer (50x stock)

242 g/L Tris (Trizma Base - Sigma)

18.61 g/L Ethylenediaminetetraacetic acid (EDTA-Fluka)

pH is adjusted to 8 using Acetic Acid (Sigma)

3.1.10 Agarose Gel

1% w/v Agarose (Sigma) in TAE buffer

The mixture was heated in a microwave until agarose was dissolved. Gel red stain (GelRed™ Nucleic Acid Gel Stain, 10000x in water – BIOTIUM) was also diluted 10000-fold.

3.1.11 Antibiotics

Stock solutions of:

Ampicillin (Sigma-Aldrich) in water (100 mg/mL)

Kanamycin (Sigma-Aldrich) in water (50 mg/mL)

Chloramphenicol (Sigma-Aldrich) in ethanol (34 mg/mL)

were diluted 1000-fold in Luria Bertani (LB) medium or LB-agar for negative selection.

3.1.12 Luria Bertani (LB) Medium

For 1 L of LB broth:

10 g Tryptone

5 g yeast extract

10 g NaCl

are added in ca. 900 mL water, the pH is adjusted with NaOH at 7-7.5 and water is added to reach a final volume of 1 L. In this study, 20g of LB broth powder (Sigma) were dissolved in 1 L of water. LB broth was autoclaved before use.

3.1.13 LB-Agar plates

Agar is added to LB medium at a concentration of 15 g/L. In this study, 35 g of LB-agar powder (Sigma) was added to 1 L of water. The mixture was autoclaved and the powder was completely dissolved. The mixture was left to cool, the desired antibiotic(s) were added while still liquid. After the addition of the antibiotics, the mixture was poured into Petri-dishes and left to harden. The plates were stored at 4 ^{o}C.

For blue-white screening, LB agar plates were plated with 40 μL of a 40 mg/mL 5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside (X-gal – Thermo Scientific) solution in dimethylsulfoxide (DMSO-Sigma) and 40 μL 0.1 M isopropyl β-D-1-thiogalactopyranoside (IPTG - Sigma)

3.1.14 Lysis buffer:

For French press:

25 mM Tris-HCl pH 8

400 mM NaCl

1 mM EDTA

5 mM imidazole

For freeze-thaw lysis, 2% Triton X-100 (Sigma) was also added.

3.1.15 Nickel-Nitrilotriacetic Acid (Ni-NTA) column purification

Ni-Sepharose™ High Performance (Amersham Biosciences AB) affinity column was used.

The buffers prepared for the purification contained:

25 mM Tris-HCl

400 mM NaCl

and imidazole in increasing concentrations (35 and 50 mM for washes, 75, 100, and 500 mM for elution).

6 M of urea were also added for proteins solubilized from inclusion bodies

3.1.16 Streptavidin column purification

Streptavidin-Sepharose™ High Performance (Amersham Biosciences AB) affinity column was used.

The buffers prepared for the purification contained:

Binding Buffer: 20 mM sodium phosphate, 0.15 M NaCl, pH 7.5

6 M of urea were also added to the binding buffer for proteins solubilized from inclusion bodies

Elution buffer: 8 M guanidine-HCl (GdnCl), pH 1.5

The acid eluate is neutralized with 1M Tris-HCl pH 9

3.1.17 Lower Separating Buffer

187 g/L Tris-HCl pH 8.8

0.4% w/v Sodium Dodecyl Sulfate (SDS – Sigma)

The buffer was stored at 4 ^{o}C

3.1.18 Upper Stacking Buffer

60.5 g/L Tris-HCl pH 6.8

0.4% w/v Sodium Dodecyl Sulfate (SDS – Sigma)

The buffer was stored at 4 ^{o}C

3.1.19 Separating Gel for SDS-PAGE

For 15 mL of gel:

5 mL Acrylamide (Acrylamide:N,N'-Methylenebisacrylamide 5:1 solution – Fluka)

10 mL water

5 mL Lower Separating Buffer

100 μL 10% Ammonium Persulfate (APS - Sigma)

10 μL N,N,N',N'-Tetrametylethylene-diamine (TEMED - Sigma)

3.1.20 Stacking Gel for SDS-PAGE

For 10 mL of gel:

1 mL Acrylamide

6.5 mL water

2.5 mL Upper Stacking Buffer

50 μL 10% APS

10 μL TEMED

3.1.21 Running Buffer for SDS-PAGE (10x stock)

30.3 g/L Tris

144.1 g/L glycine (Sigma)

10 g/L SDS

The buffer was stored at 4 ^{o}C

3.1.22 Loading buffer for SDS-PAGE (3x stock)

For 100 mL stock solution

18.8 mL 1 M Tris-HCl pH 6.8

6 g SDS

15 mL β-mercaptoethanol (Sigma)

30 mL glycerol

a pinch of Bromophenol blue (Sigma)

3.1.23 Coomassie Blue Staining Solution

2.5 g/L Coomassie Brilliant Blue (Fluka)

10% Acetic acid

50% Methanol (Sigma)

3.1.24 Destaining Solution for Coomassie Blue Stain

10% Acetic acid

5% Methanol

3.1.25 Transfer Buffer (10x)

188 g/L glycine

30 g Tris

3.2 Methods

3.2.1 Competent cell preparation (for E. coli DH5α and BL21(DE3))

E. coli cells were prepared in order to yield transformed colonies. A variation of the protocol described by Sambrook (Sambrook and Russell, 2001) was followed. For the preparation of competent cells, 5 ml of LB medium were inoculated with the desired strain and the cells were incubated overnight at 37 ^{o}C. Then 50-100 ml LB medium were inoculated with 0.5-1 mL of the overnight cultures. The cells are incubated at 37 ^{o}C with shaking at 250 rpm until the optical density at λ=600 nm ($O.D._{600}$) was 0.4-0.6. The cultures were then centrifuged in falcon tubes at 2500 rpm for 1 min at 4 ^{o}C. Cells were gently resuspended in 15-30 ml TfbI buffer and incubated on ice for 30 min. The suspension was centrifuged again at 3500 rpm for 10 min at 4 ^{o}C and the cells were resuspended gently in 2-4 ml TfbII buffer. Aliquots of 100 μL were transferred in pre-cooled eppendorf tubes tubes, frozen in liquid N_2 or dry ice and stored at -80 ^{o}C.

3.2.2 Transformation of competent E. coli cells with plasmid DNA

The competent cells prepared were transfected with plasmid DNA. A variation of the protocol described by Sambrook (Sambrook and Russell, 2001) was followed. An aliquot of competent cells (~100 μL) was thawed on ice and mixed with 1-10 ng of plasmid DNA. For the biotinylation of the GFPem-fused proteins, an equal quantity of the pBirAcm plasmid, coding for biotin ligase, was also added. The mixture was left on ice for 30 min. A heat-shock was induced by incubation at 42 ^{o}C for 90 sec and then on ice for 2 min. After the heat-shock, 900 μL of LB medium were added and the cells were incubated for 1 h at 37 ^{o}C with shaking at 250 rpm. Subsequently, the suspension was centrifuged at 2000 rpm for 2 min and most of the supernatant was removed. The pelleted cells were resuspended in the remaining supernatant (ca. 100 μL) and plated onto LB/agar plates containing the selection antibiotic(s). The plates were incubated at 37 ^{o}C overnight. For the transformation of DHB10 cells, a slightly different protocol, provided by New England Biolabs, was followed.

3.2.3 Mini-preparation (mini-prep) of plasmid DNA

Mini-preparation of plasmid DNA is a method used for small-scale isolation of plasmid DNA by alkaline lysis with SDS (Sambrook and Russell, 2001). In the present study, the NucleoSpin® Plasmid Buffer Set kit (Macherey-Nagel) was used for the purification of plasmid DNA. DH5α or DH10B cells were inoculated in LB (usually 4 mL) cantaining the selection antibiotic and left to grow overnight at 37 ^{o}C with shaking at 250 rpm. The cells were pelleted and resuspended in an RNase containing buffer. Then the cells were lysed by adding an alkaline, SDS-containing buffer. The reaction was neutralized and the lysate was centrifuged at 11000 g for 10 min in order to pellet cell debris and genomic DNA. The supernatant was passed through an affinity column. Impurities were removed by washing with ethanolic buffers. The plasmid DNA was eluted using a low ionic strength, slightly alkaline buffer. DNA concentration was measured using a NanoDrop spectrophotometer.

3.2.4 PCR

PCR is a technique used for the amplification of a certain DNA fragment. The template DNA (usually plasmid DNA) was mixed with deoxynucleotides (dNTPs) and oligonucleotide primers in a buffer for the optimal function of a DNA polymerase. The fragment was amplified by cycles of denaturation of DNA at high temperature (usually 98 ^{o}C) – binding of the primers on the complementary strand (annealing, temperature depends on the primer sequence) – extension of the primers due to polymerase activity (usually 72 ^{o}C). PCR product concentration was measured using a NanoDrop spectrophotometer.

3.2.5 Agarose gel electrophoresis

Agarose gel electrophoresis is a method used for the separation of DNA fragments according to their size (Freifelder, 1987) was used in order to evaluate and isolate the desired products from PCR and cut reactions. A 1% w/v agarose gel was deposited in a horizontal electrophoresis apparatus and the apparatus was filled with 1x TAE buffer. A DNA sample was mixed with Gel Loading Dye (New England Biolabs) and loaded into the wells of the gel. A 2-log DNA ladder (New England Biolabs) was also mixed with the loading dye and loaded in the

gel. An electric field of 90 V was applied until the samples entered the gel and then the voltage was increased to 120 V. Agarose gels were viewed using a UV-lamp. UV light excites the gel red stain which is bound to DNA, rendering the samples visible.

3.2.6 Extraction from agarose gels

The desired DNA fragments were purified after agarose gel electrophoresis using the NucleoSpin® Gel and PCR clean-up kit (Macherey-Nagel). A cut-out gel band was mixed with binding buffer and was heated at 50 ^{o}C until agarose was dissolved and then passed through a silica membrane column. Impurities were removed using an ethanolic washing buffer. The DNA was eluted using a low ionic strength, slightly alkaline buffer.

3.2.7 Site-Directed mutagenesis

Site-directed mutagenesis is a method for the insertion, deletion, or substitution of bases in double-stranded plasmid DNA. The Q5® Site-Directed Mutagenesis Kit (New England biolabs) was used for the insertion of the N-terminal RGD motif in the sequence of Bhx and additionaly the substitution of the residue following the RGD motif from Asp to Ser (RGDD → RGDS). Using this method, up to 18 nucleotides can be inserted. The template DNA (pASK40 plasmid coding for Bhx) was mixed with with the primers designed for the insertion of the RGD motif and PCR-amplified using the Q5 Hot Start High-Fidelity DNA Polymerase. The PCR product was added to a Kinase-Ligase-DpnI (KLD) enzyme mix for circularization of the amplified plasmid and template removal. The KLD mix was used for the transformation of DH5α E. coli cells. Single colonies developed on the agar plate were inoculated in LB for mini-prep. The isolated plasmid containing the RDG motif was mixed with the primers designed to substitute the following Asp with Ser and the above protocol was performed again.

3.2.8 Seamless cloning method

The seamless cloning method is a method used for the insertion of DNA fragments into a plasmid vector. In the present study, GeneArt® Kit (Invitrogen) was employed for the fusion of

GFPem with Bhx.

For the fusion of GFPem to the N- and C-terminus of Bhx, the genes coding for Bhx and GFPem-thrombin cleavage site linker were PCR amplified using specific primers for the fusion of the genes and their insertion in a pUC19L linear vector. The insertion to pUC19L was achieved by using the enzyme mix of the kit. The enzyme mix contains a 5' exonuclease which generates terminal cohesive ends (overhangs), a polymerase which fills in the gaps of the annealed single-stranded regions, and a ligase which seals the nicks. DHB10 cells were transformed with the pUC19L and clones were selected by blue/white screening. The plasmids were isolated (mini-prep), and cut with restriction enzymes to determine the existence of the desired insert DNA. The genes of interest were PCR-amplified and inserted in the desired vector via restriction enzyme cloning.

3.2.9 Restriction enzyme cloning

Restriction enzyme cloning involves the utilization of restriction endonucleases to generate DNA fragments with specific ends in order to be inserted in the desired plasmid vector. The plasmid vector and PCR products from seamless cloning were cut using a specific pair of restriction enzymes. The ends of plasmid vector were dephosphorylated by a phosphatase and the gene was inserted using a ligase. The ligation product was used for transformation of DH5α cells and positive clones were determined by cutting the isolated plasmids with the same endonuclease pair. Using this method, N-GFPem-Bhx and Bhx-GFPem-C were cloned into a pT7-7 vector and pET28a vector respectively.

3.2.10 Recombinant Protein Expression

E. coli BL21(DE3) cells were used for recombinant protein overexpression. Precultures were prepared by resuspending a single colony from the LB/agar plates in liquid LB medium with antibiotic(s). Precultures were incubated overnight at 37 ^{o}C with shaking at 250 rpm. LB medium with antibiotic(s), ca. 25 times the preculture volume, was then inoculated with the preculture and incubated at 37^{o}C with shaking at 250 rpm until the optical density at λ=600 nm ($O.D._{600}$) was 0.5-0.8 . For the biotinylation of the GFPem-fused proteins, a solution of biotin in

Bicine buffer (see Materials section) is added in the culture to reach a final biotin concentration of 12 μg/mL. Protein expression was induced with IPTG at a final concentration of 1 mM and cells were incubated at 37 ^{o}C for 4 h or 30 ^{o}C overnight, aerated by shaking at 250 rpm. Cells were harvested by centrifugation of the culture at 7000 rpm for 15 min.

For a long term storage of the bacterial strains, bacterial stab cultures were prepared by mixing 500 μL of a grown preculture with an equal volume of 80% glycerol. The stab cultures were stored at -80 ^{o}C.

3.2.11 E. coli Cell Lysis

After harvesting, the cell paste was resuspended in four volumes of lysis buffer (see Materials section) and left on ice for 30 min.

Two methods proved efficient for lysis. The first method includes two freeze-thaw cycles at -80 ^{o}C and 37 ^{o}C respectively, the addition of 2 mg/mL lysozyme and 1x protease inhibitor coctail, another two freeze-thaw cycles and sonication on ice with six bursts of 30 s followed by cooling intervals of 30 s. The second method used was the French press, where the cell suspension is compressed by a piston at 1000 psi. French press could only be used for culture volumes above 1L. The suspension was passed 2-3 times through the French press.

After lysis, DNase I was added at a concentration of 100 μg/mL as well as $MgSO_4$ (Sigma) at a concentration of 100 mM, and the lysate was left for 30 min at room temperature with shaking at 80 rpm. The lysate was centrifuged at 7000 rpm for 30 min and the supernatant was separated from the pellet. The pellet was resuspended in the initial volume of lysis buffer. In the case of freeze-thaw lysis, the resuspended pellet was centrifuged again at 7000 rpm for 30 min, as a washing step to remove Triton X-100. Again, the pellet was resuspended in the initial volume of lysis buffer. For cultures grown at 37 ^{o}C, the lysis buffer for the resuspension of the pellet contained 6 M of urea (see Materials section) and was mixed overnight at 4 ^{o}C. After mixing, the remaining unsoluble products were separated by centrifugation at 7000 rpm for 30 min; the small amount of the pellet isolated was resuspended in 1-2 mL of lysis buffer.

3.2.12 Protein Purification

Cytosoluble proteins were purified from the supernatant obtained after the lysis, whereas insoluble proteins were purified from the supernatant obtained after resuspension of the pellet in 6M urea. All of the proteins studied, contained a C-terminal hexahistidyl tag, and thus were purified using a Ni-NTA affinity purification column. For every 2 mL of supernatant, 1 ml of Ni-NTA beads was used. The beads were deposited in a gravity-flow column and the supernatant was passed three times. Then, the beads were washed with 8 volumes of two washing buffers with increasing imidazole concentration (35 and 50 mM). The protein was eluted by washing with 5 volumes of three elution buffers with increasing imidazole concentration (75, 100 and 500 mM)

Biotinylated proteins were first purified using a Ni-NTA column and then using a streptavidin column. The buffer of the eluted protein from the Ni-NTA column was changed in order to match the binding buffer of the streptavidin column by dialysis. The column was equilibrated with 10 volumes of binding buffer and the protein solution was passed three times. Subsequently, the column was washed with 10 volumes of binding buffer and with 20 volumes of elution buffer. Every 2 mL of the eluate were gathered in 2 mL-eppendorf tubes and were neutralized with 200 μL 1M Tris-HCl pH 9.

3.2.13 Sodium dodecyl sulfate - Poly-acrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is a method used for the separation of proteins according to their molecular weight (Alberts, 2002). In this study, it was used to monitor the expression of the proteins as well as their condition after each processing step (e.g. possible impurities, degradation products). SDS, boiling, and β-mercaptoethanol are used in order to denature the protein. SDS binds to the linear protein molecule and confers a uniform negative charge. Then, the proteins are loaded on the SDS polyacrylamide gel and migrate across the gel in the presence of an electric field. The migration is size-dependent; larger proteins migrate less than smaller proteins, since they cannot fit easily into the pores of the gel. In this study, the separating and stacking polyacrylamide gels were used with an acrylamide content of 7.5% and 4.5% respectively. The separating gels were prepared by mixing acrylamide with water and separating buffer. Polymerization was induced

when APS and TEMED were added and the mixture was deposited between glass plates for electrophoresis. Isopropanol was added on top and the mixture was left to harden for approximately 1h. After hardening, isopropanol was removed and the stacking gel mixture was deposited on top of the separating gel. Gel combs were placed on top of the stacking gel mixture for the formation of wells. The stacking gel was also left to harden for 30 min and the gel was transferred in the electrophoresis apparatus, which was filled with running buffer (1x). The samples studied were mixed with loading buffer for SDS-PAGE, boiled for 5 min at 100 ^{o}C and loaded in the wells. Protein molecular weight marker (Precision Plus Protein™ Dual Color Standards – BIO-RAD) was also loaded in a well and electrophoresis was performed at 200 V for ca. 1h. For samples containing GdnCl, a precipitation protocol was followed before loading in the gel. An aliquot of 25 μL from each sample was transferred in 225 μL cold 100% ethanol, vortex-mixed, and left at -20 ^{o}C for 10 min. The mixture was centrifuged at 15000 g for 5 min and the supernatant was removed. The pellet was vortex-mixed in 90% cold ethanol and centrifuged at 15000 g for 5min. The supernatant was removed and the pellet was mixed in 25μL 1x loading buffer, boiled at 100 ^{o}C for 5 min, and loaded in the gel. The gels were stained with Coomassie Blue stain, which binds to protein molecules. After destaining the gels, only the protein bands remained colored.

3.2.14 Western Blot

Western blot is a technique that involves the immunological recognition of a protein using monoclonal antibodies. In the present study, monoclonal anti-His and anti-GFP antibodies were used. A PVDF membrane (Westran S - Whatman) was activated by submerging in methanol. An electrophoresed polyacrylamide gel and the membrane were sandwiched between paper and sponges soaked in Transfer buffer 1x (sponge/paper/gel/membrane/paper/sponge). All these were tightly packed and submerged in the transferring apparatus, which was filled with Transfer buffer 1x. An electrical field of 35 V was applied for ca. 2.5 h for the protein transfer onto the PVDF membrane. The membrane was blocked by mild shaking in 5% w/v skimmed milk in 1x phosphate-buffered saline (PBS - Sigma) for 1 h at room temperature. After blocking, the PBS-milk was removed and the membrane was soaked in a solution of PBS-milk containing the primary antibody (mouse anti-His (Qiagen) or anti-GFP (Sigma)) diluted 2000-fold. The

membrane remains in this solution overnight at 4 ^{o}C under mild shaking. The next day, the membrane was washed three times by shaking for 15 min in a solution of 1x PBS with 0.04% Tween 20 (Sigma). Then, the membrane was washed again with PBS-milk and soaked in a solution of PBS-milk containing the secondary antibody (anti-mouse IgG-alkaline phosphatase - Sigma) diluted 20000-fold. The membrane remains in the solution of the secondary antibody for 2 h at room temperature under mild shaking. After the addition of the secondary antibody, the membrane was washed three times by shaking for 15 min in PBS-Tween 20 and then quickly washed with 1x alkaline phosphatase buffer. The protein bands detected by the antibodies were colored when 1.76% NBT-BCIP in 1x alkaline phosphatase buffer was added. Alkaline phosphatase dephosphorylates BCIP which in turn reacts with NBT to form an indigo-blue precipitate.

3.2.15 Dialysis

A buffer exchange was achieved using dialysis. After the purification, the elution fractions were pooled and sealed in a dialysis tubing cellulose membrane (14 kDa cut-off – Sigma). The membrane was transferred in 2 L of the desired buffer (50 mM sodium phosphate pH 7 or 50 mM Tris-HCl pH 7.5) and dialyzed overnight at 4 ^{o}C. At high protein concentrations, aggregates where observed. If aggregates were not observed, the dialyzed samples were concentrated using an Amicon centrifugal filter (30 kDa cut-off) in order to achieve aggregation.

3.2.16 PROTEOSTAT® Aggresome detection dye

Proteostat is a dye that resembles Thioflavin T, exhibiting absorbance and emission maxima at ~500 and ~600 nm. Comparing to other dyes, Proteostat yields a brighter signal and provides a larger linear dynamic range for the detection of protein aggregates *in vitro*, and it also provided good results for the detection of inclusion bodies in eukaryotic and bacterial cells (Navarro and Ventura, 2014). In the present study, Proteostat was used for the detection of inclusion bodies of Bhx, RGDD-Bhx and RGDS-Bhx in *E. coli*. After overexpression, *E. coli* cells were resuspended in 1x PBS to an $O.D._{600}$ = 0.2 . Proteostat Aggresome Stain (Enzo Life

Sciences) was diluted 1:5000 in the cell suspension. After a 30 min incubation on ice, the cells were visualized via fluorescence microscopy.

3.2.17 Fluorescence Microscopy

Fluorescence is the intrinsic property of an organic or inorganic material to absorb and emit light. The electrons of an atom can absorb high energy irradiation and be excited to a higher energy level. The electron eventually falls back to its ground state, emitting radiation of lower energy. In a fluorescence microscope, a vertical illuminator directs the light from a high intensity source onto the sample through the microscope objective, acting as a condenser. Bands of wavelengths in the ultraviolet, blue or green regions of the visible spectrum are separated by an excitation filter and reflected from a dichroic mirror onto the sample (Alberts et al. 2008). The emission light of a fluorescent sample passes back from the dichroic mirror and is separated from other wavelengths by a barrier filter. The emitted light is captured by the illuminator through the same objective. The anatomy of a fluorescence microscope is shown in Figure.

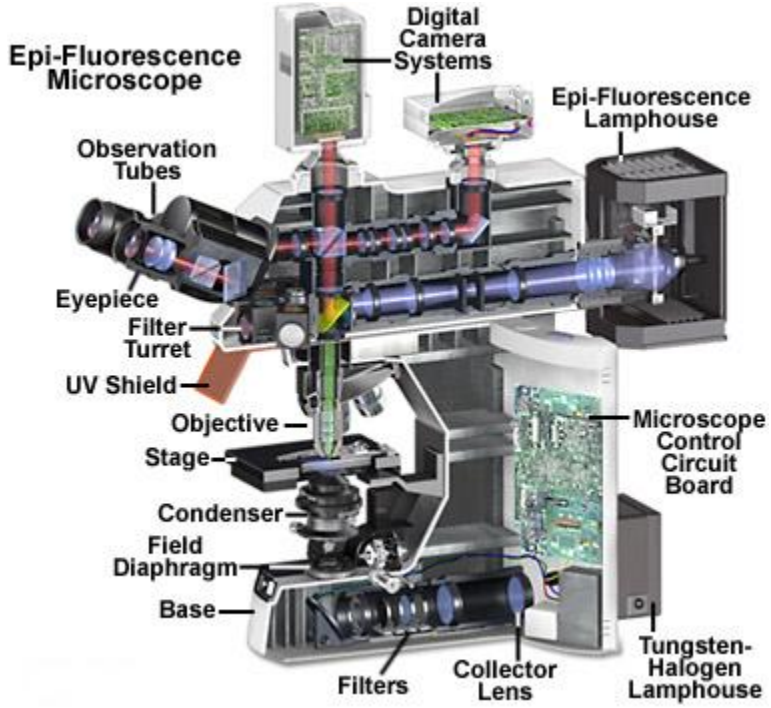


Figure: The anatomy of a fluoresce microscope
(https://www.microscopyu.com/articles/fluorescence/fluorescenceintro.html)

For fluorescence microscopy observation, a Nikon ECLIPSE E800 microscope was used. Bacteria in which inclusion bodies of the GFP-Bhx fusions were formed were resuspended in PBS to an $O.D._{600}$=0.2, deposited onto a glass slide, covered with a glass cover slip and observed directly. For the non-fluorescent Bhx, RGDD-Bhx and RGDS-Bhx inclusion bodies, bacteria were treated with Proteostat Aggresome Stain (Enzo Life Sciences).

3.2.18 Scanning Electron Microscopy – S.E.M.

S.E.M. was used in order to observe the 3D morphology of the protein nanostructures. In S.E.M., the specimen is scanned by an electron beam which is generated by a heated tungsten or lanthanum hexaboride (LaB_6) filament (Sawyer et al., 2008). The backscattered and secondary electrons, generated from the interaction of the primary electron beam with the surface of the sample, are collected by detectors and an image is produced. The S.E.M. resolution is around 5 nm. The samples are observed in vacuum and must be dry before observation.

In this study, protein nanostructures were deposited onto glass slides and left to dry in air. The dried samples were coated with 15 nm gold sputtering using a BALTEC SCD 050 SPUTTER COATER. Observation was carried out using a JEOL JSM-6390LV (Electron microscopy laboratory – University of Crete) scanning electron microscope operating at 15 kV.

3.2.19 Transmission Electron Microscopy – T.E.M.

In T.E.M., a thin specimen is scanned by an electron beam generated by a LaB_6 electron gun (Sawyer et al., 2008). The transmitted electrons are collected and used to produce a high resolution (even atomic) 2D image of the sample. As in S.E.M., the sample is observed in ultra-vacuum conditions and thus must be dry.

T.E.M. was used to observe protein molecules after purification as well as protein nanostructures. A sample solution of 8 μL was deposited onto a T.E.M. copper grid covered with Formvar. The sample solution was left onto the grid for 2 min and the excess solution was

removed using a filter paper. Subsequently, 8 μL of 1% Uranyl acetate (for sample pH≤7) or Phosphotungstic acid (for sample pH>7) negative stain was applied, left for 2 min and the excess solution was removed using a filter paper.

All samples were observed using a JEOL JEM-100 C transmission electron microscope (Electron microscopy laboratory – University of Crete) operating at 80 kV.

3.2.20 Cross-section preparation and Critical point drying

Cross-sections of protein nanostructures were prepared for TEM observation. For the preparation of the cross-sections, the nanostructures were centrifuged and 1.5 mL of a 2% glutaraldehyde and 2% para-formaldehyde solution in 0.1 M sodium cacodylate buffer (SCB) pH 7.4 was added to the pellet. The pellet was left in the fixation medium for 2 h at 4 ^{o}C. The structures were centrifuged again and 50 μL of 1 mM Congo red solution were added in order to render the sample visible when embedded in resin. The sample was left for 2 days at 4 ^{o}C. The sample was dehydrated through a series of graded ethanol. The dehydration steps included a change to 30% ethanol for 15 min, to 50% ethanol for 15 min, to 70% ethanol and Uranyl acetate for 1h, to 70% ethanol overnight, to 90% ethanol for 15 min, and to 100% ethanol for 15 min. After additional two changes of 30 min with dry ethanol and two changes of 30 min with propylene oxide, the embedding resin was added in three steps. The resin was added to a 1:3 dilution over propylene oxide for 1 h, to a 1:1 dilution for 1 h, and to a 3:1 dilution for 1h. The sample was then transferred to 100% resin for 1 h and to fresh 100% resin overnight. The specimen was placed in a vacuum chamber for 2 days and transferred to fresh 100% resin following 2 days in an oven at 60 ^{o}C for polymerization. The hard blocks were sectioned and sections were placed onto copper grids for T.E.M. observation.

Critical point drying was performed in order to dehydrate protein structures prior to SEM observation. The samples transferred in polypropylene barrels sealed with Millipore membrane. They were washed twice for 10 min with 0.1 M SCB pH 7.4 and left in 2% glutaraldehyde and 2% para-formaldehyde solution in 0.1 M sodium cacodylate buffer (SCB) pH 7.4 for 45 min. The samples were washed again twice for 10 min with 0.1 M SCB pH 7.4 and dehydrated with a series of graded ethanol (30-50-70-90-100% and dry ethanol, each change step lasted 10 min).

The barrels were transferred after the alcohol gradient to a BALTEC CPD 030 critical point dryer, where ethanol was gradually exchanged with liquid CO_2. The liquid was converted to gas by reaching the critical point of CO_2, i.e. 31 ^{o}C and 74 bar.

Chapter 4 – Results

4.1 Bhx

4.1.1 Expression-Lysis-Purification

Bhx overexpression was induced properly in the bacterial host at both 30 and 37 ^{o}C (Fig. 1a,b). An attempt of overexpressing Bhx at 22 ^{o}C overnight did not generate significant amounts of protein, and thus was considered unsuitable for the experimentation purposes (Fig. 1c).

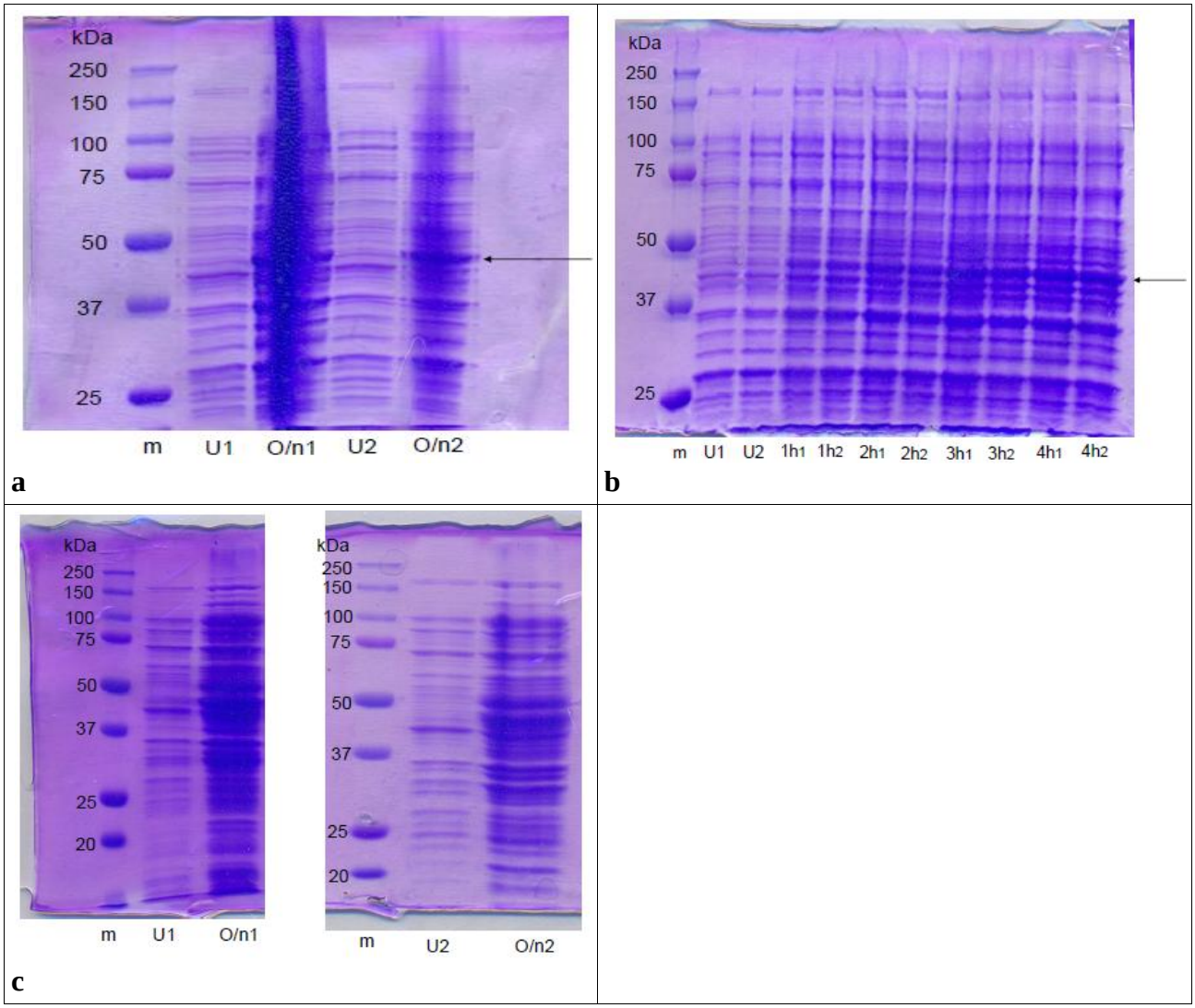


Figure 1: SDS-PAGE overexpression profile of Bhx at 30 (a) 37 (b) and 22 (c) ^{o}C in two separate flasks of 1 L culture. Arrows indicate the molecular weight of the protein, at 48 kDa.

Lane m: protein molecular weight marker

Lane U: uninducted bacteria

Lane O/n: protein overexpression overnight

Lane 1-4h: protein overexpression in the course of 4h

Fluorescence microscopy images of harvested bacteria indicate the formation of Bhx inclusion bodies after overexpression at 37 ^{o}C (Fig 2 b). On the other hand, protein inclusion bodies were significantly less, or not detected, when the protein was expressed at 30 ^{o}C (Fig 2 a).

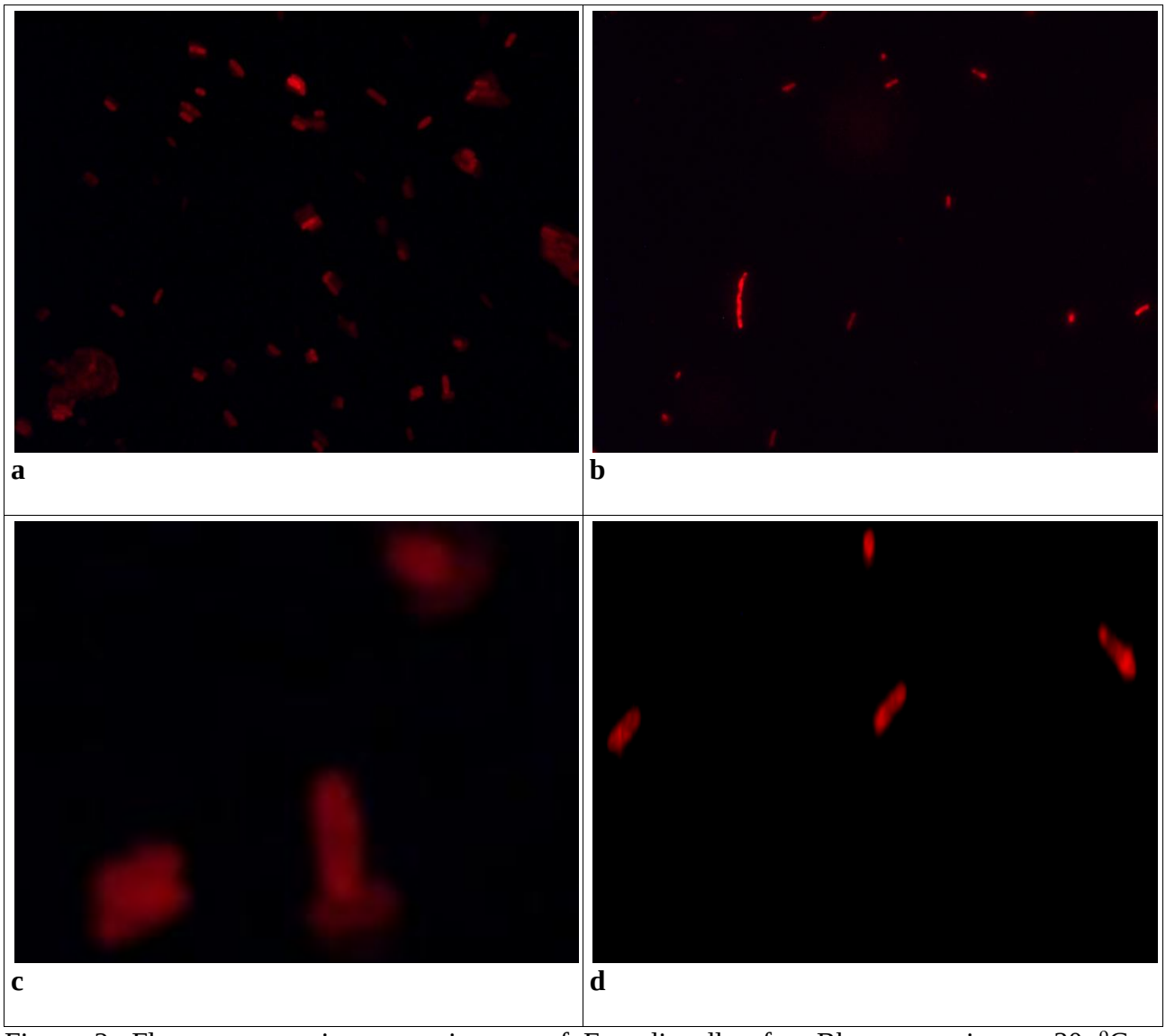


Figure 2: Fluorescence microscopy images of E. coli cells after Bhx expression at 30 ^{o}C overnight (a) and 37 ^{o}C for 4 h (b) stained with Proteostat Aggresome detection stain. The

inclusion bodies appear as bright red spots clustered inside the cells, often located at one or both poles of the cells. Images (a) and (b) were taken at 100x and 60x magnification respectively. (c) and (d) depict magnified areas from (a) and (b) respectively.

Lysis was effective using both French press (Fig. 3 a) and freeze-thaw cycles with sonication (Fig 3 b). The protein expressed at 30 ^{o}C can be purified from the supernatant obtained after lysis using Ni-NTA affinity chromatography (Fig 3 b,c). An amount of protein seems to be eluted in the washing fraction with 50 mM imidazole. The small amount of impurities observed in the elution fractions does not seem to be relevant to protein degradation, since they could not be detected in an anti-His Western blot (Fig. 3 d).

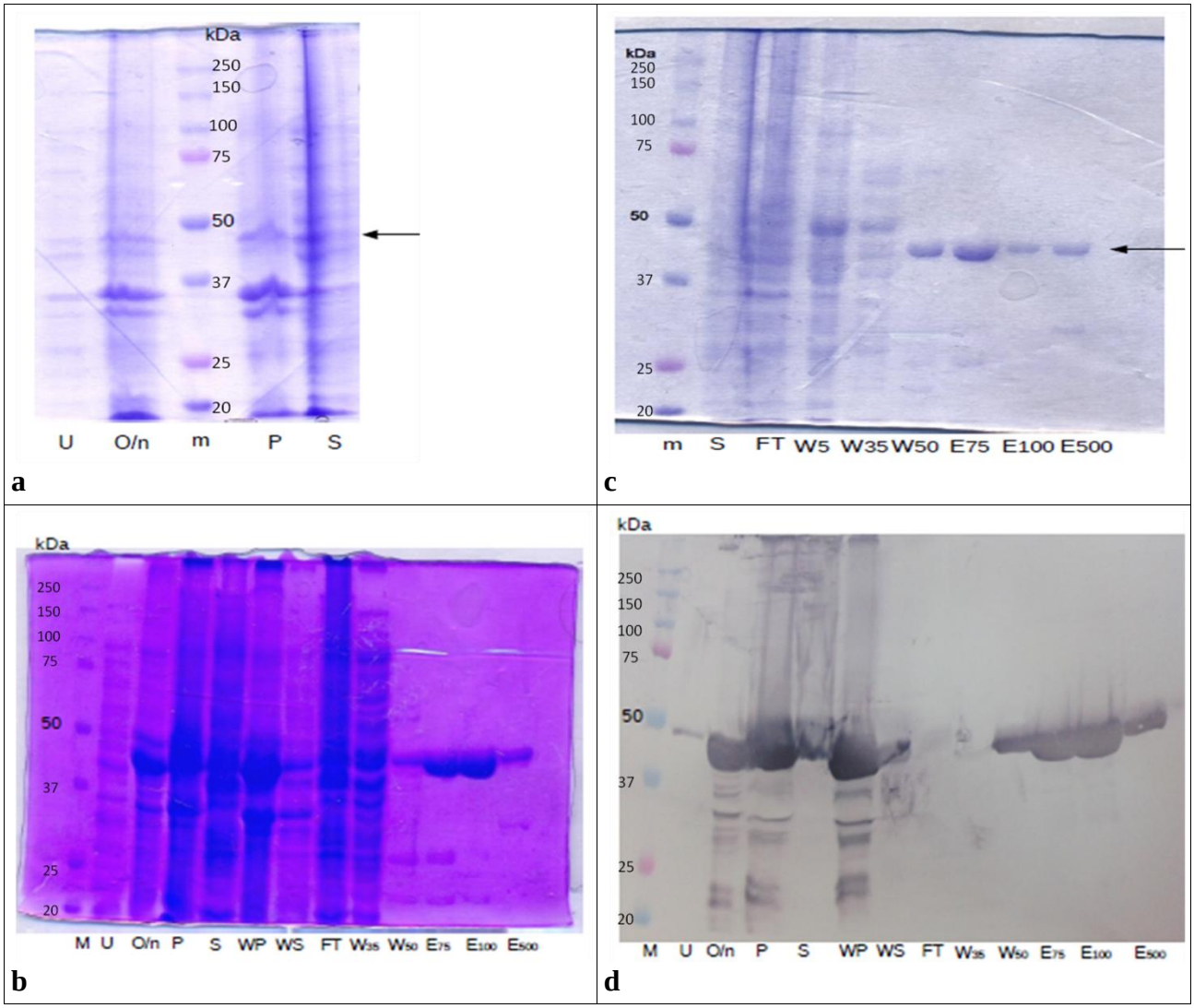


Figure 3: Expression at 30 ^{o}C, lysis, and purification of Bhx. (a) and (c) demonstrate the process after French press lysis, while (b) demonstrates the same process using freeze-thaw lysis with sonication. Western blot (d) was used in order to identify the impurities shown in (c).

Lane m: protein molecular weight marker

Lane U: uninducted bacteria

Lane O/n: protein overexpression overnight

Lane P: pellet obtained after lysis

Lane S: supernatant obtained after lysis

Lane WP: washing pellet

Lane WS: washing supernatant

Lane FT: supernatant after being flown through the chromatography column

Lane W35-50: washing fractions of 5, 35, and 50 mM of imidazole

Lane E75-100-500: elution fractions of 75, 100, and 500 mM of imidazole

Bhx expressed at 37 ^{o}C was solubilized in 6 M urea and purified in a Ni-NTA column (Fig 4 a). A larger amount of protein seems to be eluted in the washing fraction with 50 mM imidazole. The amount of impurities observed in the elution fractions seems to be a result of degradation beginning from the N-terminus of the protein; these degradation products were evident since the expression of the protein, as was shown by anti-His Western blot (Fig. 4 b).

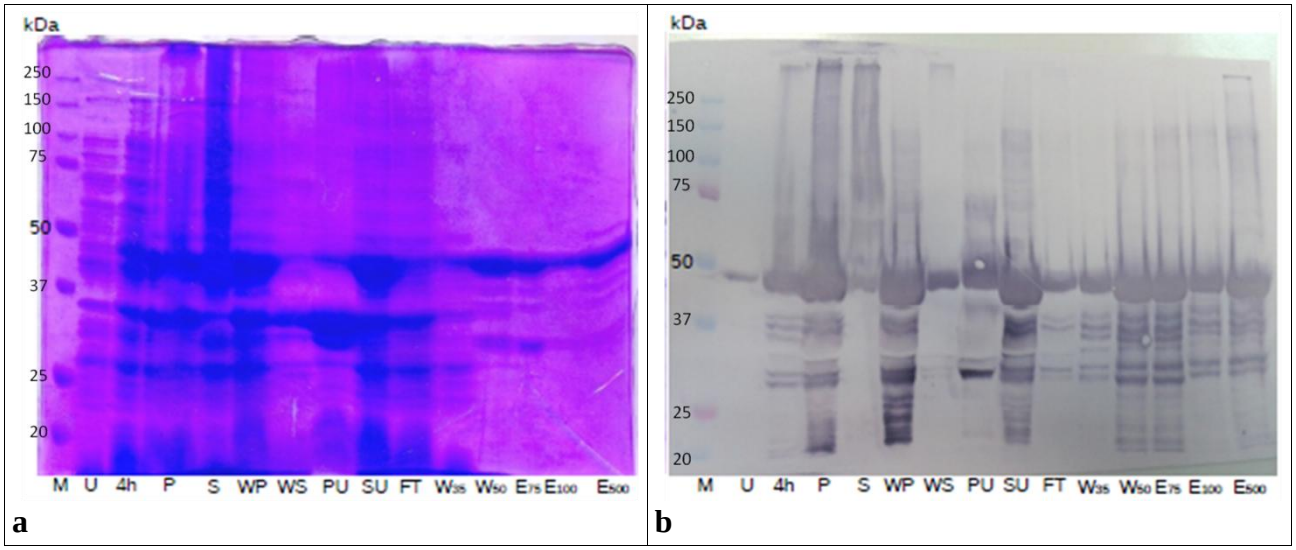


Figure 4: Expression at 37 ^{o}C, freeze-thaw lysis with sonication, and purification profile of Bhx in SDS-PAGE (a) and Western blot (b).

Lane m: protein molecular weight marker

Lane U: uninducted bacteria

Lane 4 h: protein overexpression after 4h at 37 ^{o}C

Lane P: pellet obtained after lysis

Lane S: supernatant obtained after lysis

Lane WP: washing pellet

Lane WS: washing supernatant

Lane PU: pellet obtained after resuspension in 6 M urea

Lane SU: supernatant obtained after resuspension in 6 M urea

Lane FT: SU after being flown through the chromatography column

Lane W35-50: washing fractions of 5, 35, and 50 mM of imidazole and 6 M of urea

Lane E75-100-500: elution fractions of 75, 100, and 500 mM of imidazole and 6 M of urea

4.1.2 Electron microscopy observation

The protein expressed at 30 ^{o}C could be observed via TEM directly after the purification step. Elution fractions were pooled and a sample solution was visualized in the microscope.Only a few,sparse nanofibrillar structures could be observed (Fig. 5).

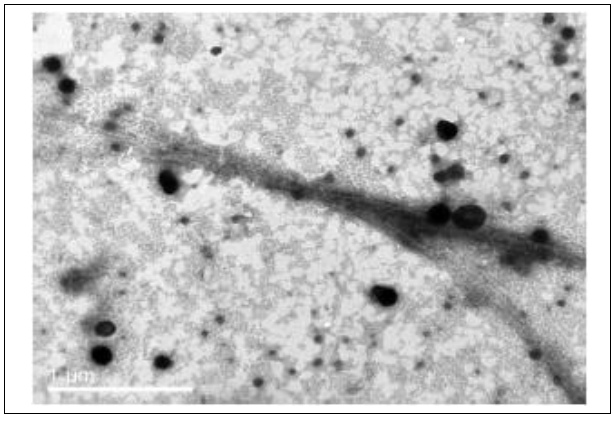


Figure 5: TEM image of the pooled elution fractions after purification of Bhx expressed at 30 ^{o}C. The black aggregates that appear were formed by the uranyl acetate negative stain.

After dialysis of the purified protein against either 50 mM sodium phosphate pH 7 or 50 mM Tris-HCl pH 7.5 in order to remove imidazole, protein aggregates in the solution could be observed with naked eye. These aggregates were collected by centrifugation and observed via electron microscopy. TEM results suggest that part of the protein aggregates into an amorphous state, yet there were large spherical microparticles observed with a diameter of 10-20 μm.

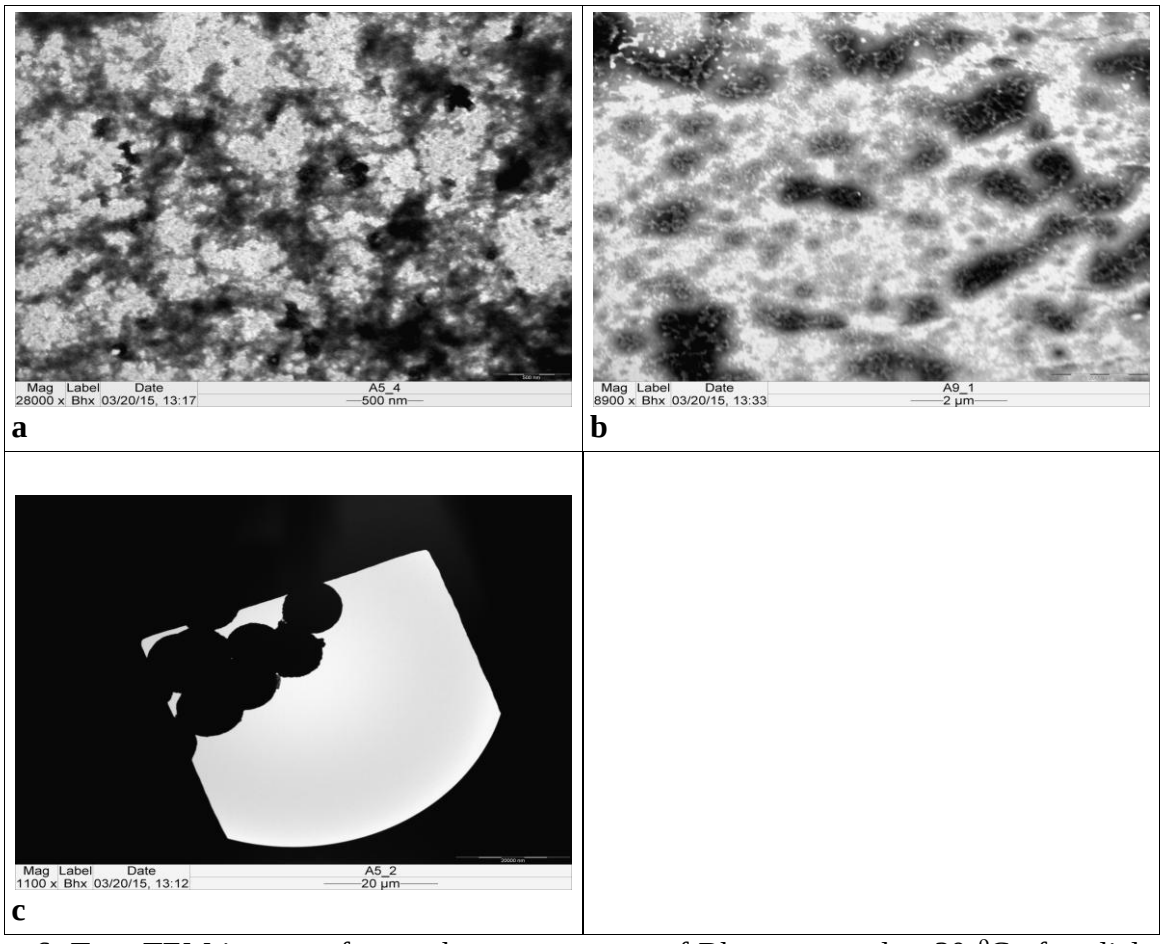


Figure 6: Top: TEM images of amorphous aggregates of Bhx expressed at 30 ^{o}C after dialysis against 50 mM sodium phosphate pH 7 (a) and 50 mM Tris-HCl pH 7.5 (b). Bottom: TEM image of spherical microparticles of Bhx expressed at 30 ^{o}C after dialysis against 50 mM sodium phosphate pH 7 (c).

Further characterization of these microparticles with SEM revealed that a better shape was retained after dialysis against phosphate and that the surface of particles was rough, indicating porosity (Fig. 7 a,b). Cross-sections of the fixated particles observed via TEM display a hierarchical formation of the microspheres by self-assembled cross-linked nano-fibrils (Fig. 7 c,d).

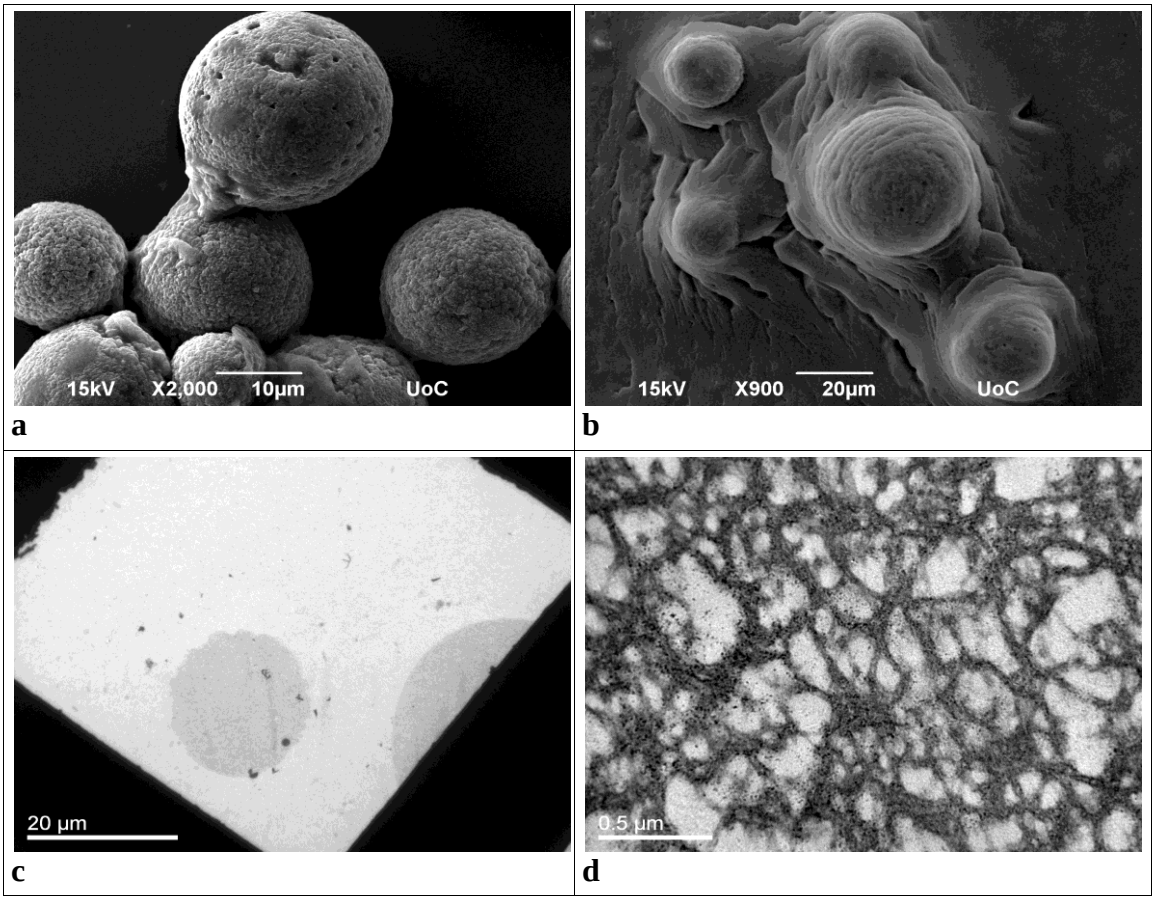


Figure 7: Top: SEM images of spherical microparticles of Bhx expressed at 30 ^{o}C after dialysis against 50 mM sodium phosphate pH 7 (a) and 50 mM Tris-HCl pH 7.5 (b). Bottom: TEM images of cross-sections of spherical microparticles of Bhx expressed at 30 ^{o}C after dialysis against 50 mM sodium phosphate pH 7 (c) and a magnified region upon the cross-section (d).

The protein expressed at 37 ^{o}C also formed visible aggregates after dialysis against either 50 mM sodium phosphate pH 7 or 50 mM Tris-HCl pH 7.5 in order to remove urea and imidazole. Part of these aggregates were millimeter-long fibers (Fig. 8). Aggregates were collected by centrifugation and further observed via electron microscopy.

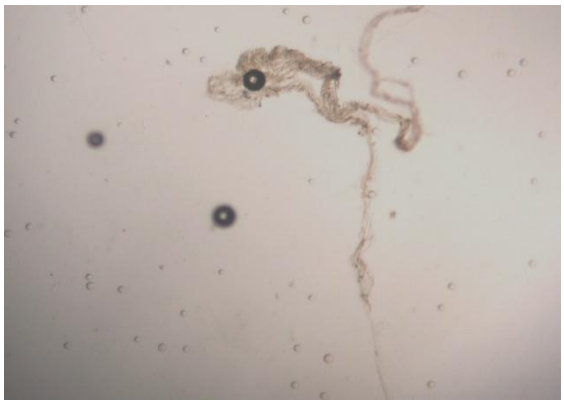
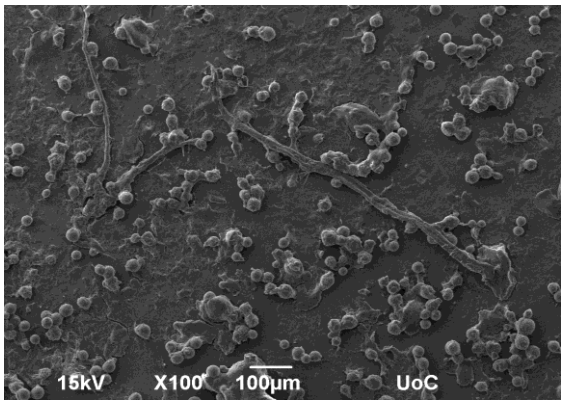
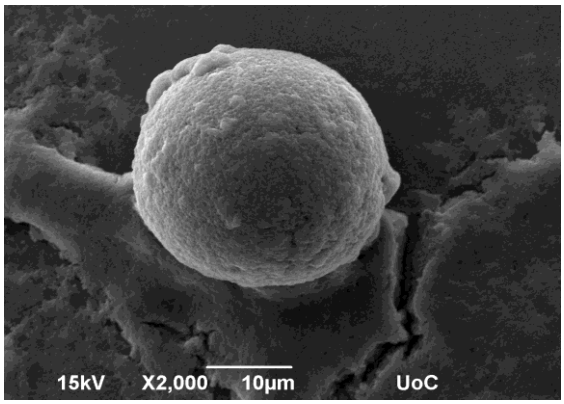


Figure 8: Stereoscope image of a fibrillar aggregate of Bhx expressed at 37 ^{o}C and dialyzed against 50 mM sodium phosphate pH 7. The image was taken in solution and the magnification is 5x.

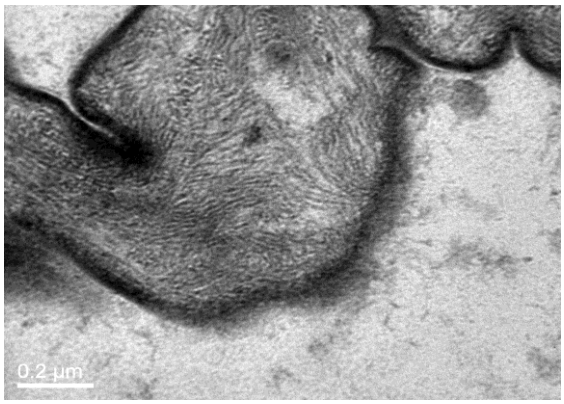
SEM results show the co-existence of the aforementioned fibrous aggregates with microspheres of ca. 20 μm in diameter (Fig. 9 a). The microspheres display the same characteristics as those formed by Bhx expressed at 30 ^{o}C (Fig. 9b). TEM observation of cross-sections of the fibrous aggregates show a hierarchical formation of the fibril by self-assembled nanofibrils (Fig. 9c).



a



b



c

Figure 9: Top: SEM images of aggregates formed by Bhx expressed at 37 ^{o}C after dialysis against 50 mM sodium phosphate pH 7 (a). A microsphere of the same sample visualised at higher magnification is displayed in (b) Bottom: TEM image of a cross-section of a fibrous aggregate (c).

Summary: Bhx was produced in bacteria at both 30 and 37^{o}C. The protein was soluble in the cytoplasmwhen expressed at 30 ^{o}C and was purified from the supernatant after lysisusing a Ni-NTA column. On the other hand, when expressed at 37 ^{o}C,Bhx aggregated in inclusion bodies and needed to be solubilized in urea in order to be purified in a Ni-NTA column. An interesting result is that a ladder of degradation products recognized by an anti-His antibody was quite pronouncedwhen the protein was expressed in at 37 ^{o}C and purified from urea, suggesting instability of the N-terminus that leads to gradual degradation of the protein. After the purification step, elution fractions were pooled and dialyzed in order to remove imidazole and/or urea. After dialysis of Bhx that was not solubilized in urea, spherical aggregates were observed.For Bhx which was denatured in urea, both spherical and fibrillar aggregates were observed. This indicates that aggregation is induced by a specific sequence which plays a key role in the refolding of Bhx.

4.2 RGDD-Bhx and RGDS-Bhx

4.2.1 Expression-Lysis-Purification

The molecular weight of pASK40 which contained the insert gene of Bhx was evaluated after the insertion and substitution mutagenesis reactions (Fig 10 a,b). The changes of the molecular weight of the plasmid were undetectable using agarose electrophoresis, thus the generated Bhx sequences were further validated via DNA-sequencing (see Appendix I).A schematic representation of the new constructs is available in Fig. 10.

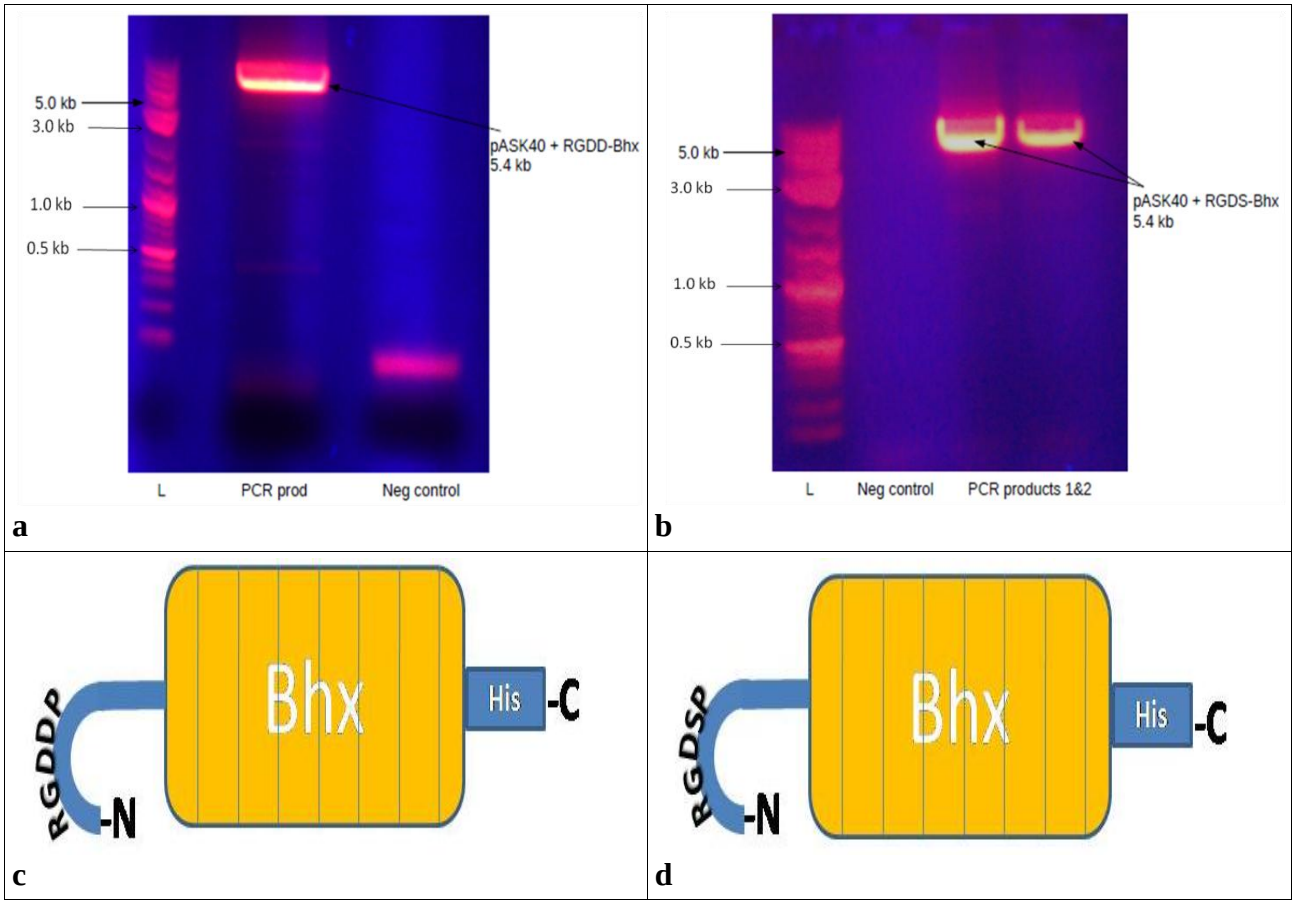


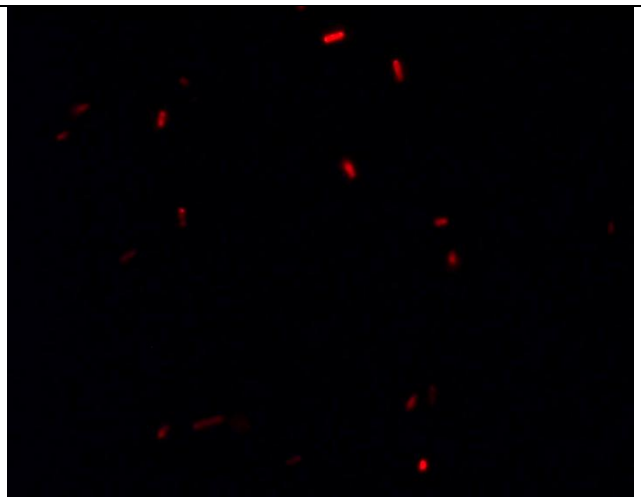
Figure 10: Top: Agarose gel electrophoresis of the PCR products after insertion (a) and substitution (b) mutageneses of pASK40. Bottom: Schematic representation of RGDD-Bhx (c) and RGDS-Bhx(d). Bhx stands for the tailspike β-helical domain and His for the C-terminal hexahistidyl tag.

Lane L: DNA Ladder

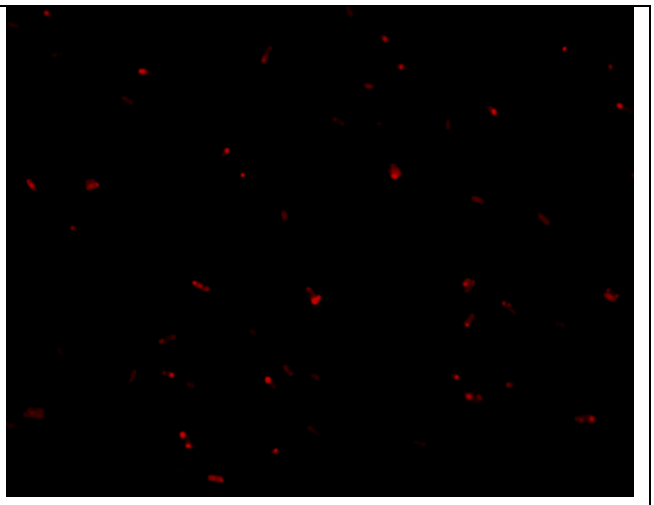
Lane Neg control: The negative control reaction, i.e. the same mutagenesis reaction without the use of template DNA

Lane PCR prod: The PCR product after the reaction

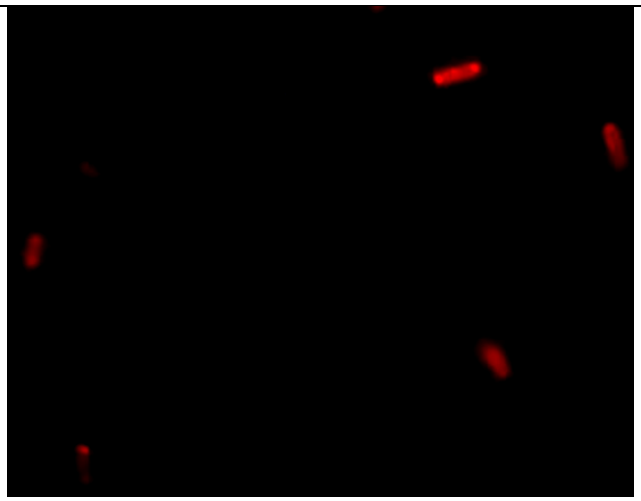
RGDD-Bhx and RGDS-Bhx overexpression was induced properly in the bacterial host at both 30 and 37 ^{o}C. Fluorescence microscopy results show the formation of inclusion bodies when the protein was expressed at 30 ^{o}C (Fig. 11 a,b), thus this temperature was considered as more appropriate for the experimentation purposes.



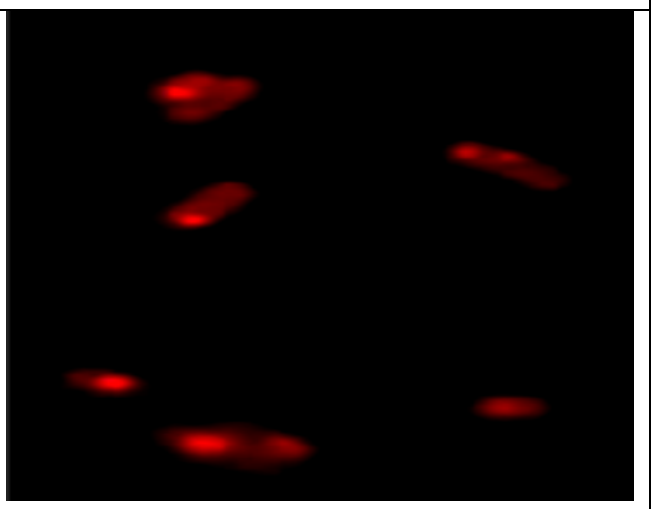
a



b



c



d

Figure 11: Top: Fluorescence microscopy images of E. coli cells after RGDD-Bhx (a) and RGDS-Bhx (b) expression at 30 ^{o}C overnight stained with Proteostat Aggresome detection stain. The inclusion bodies appear as bright red spots clustered inside the cells, often located at one or both poles of the cells. Images were taken at 100x magnification. (c) and (d) depict magnified areas from (a) and (b) respectively.

RGDD-Bhx expressed at 30 ^{o}C could be purified in a NI-NTA column after solubilization in 6 M urea (Fig 12). A large amount of protein seems to be eluted in the washing fraction with 50 mM imidazole.

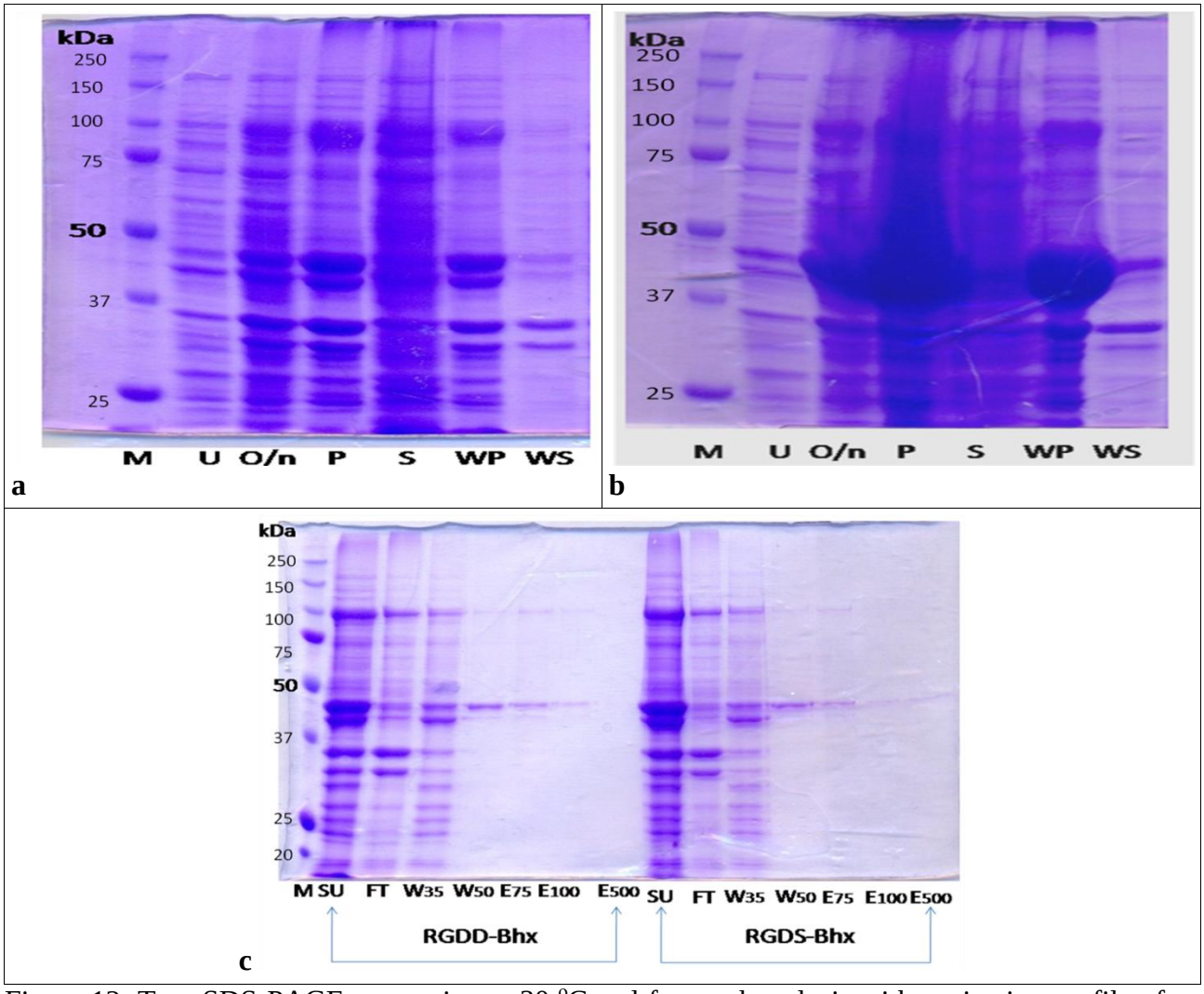


Figure 12: Top: SDS-PAGE expression at 30 ^{o}C and freeze-thaw lysis with sonication profile of RGDD-Bhx (a) and RGDS-Bhx. Bottom: SDS-PAGE purification profile of the same proteins (c).

Lane M: protein molecular weight marker

Lane U: uninducted bacteria

Lane O/n: protein overexpression overnight at 30 ^{o}C

Lane P: pellet obtained after lysis

Lane S: supernatant obtained after lysis

Lane WP: washing pellet

Lane WS: washing supernatant

Lane SU: supernatant obtained after resuspension in 6 M urea

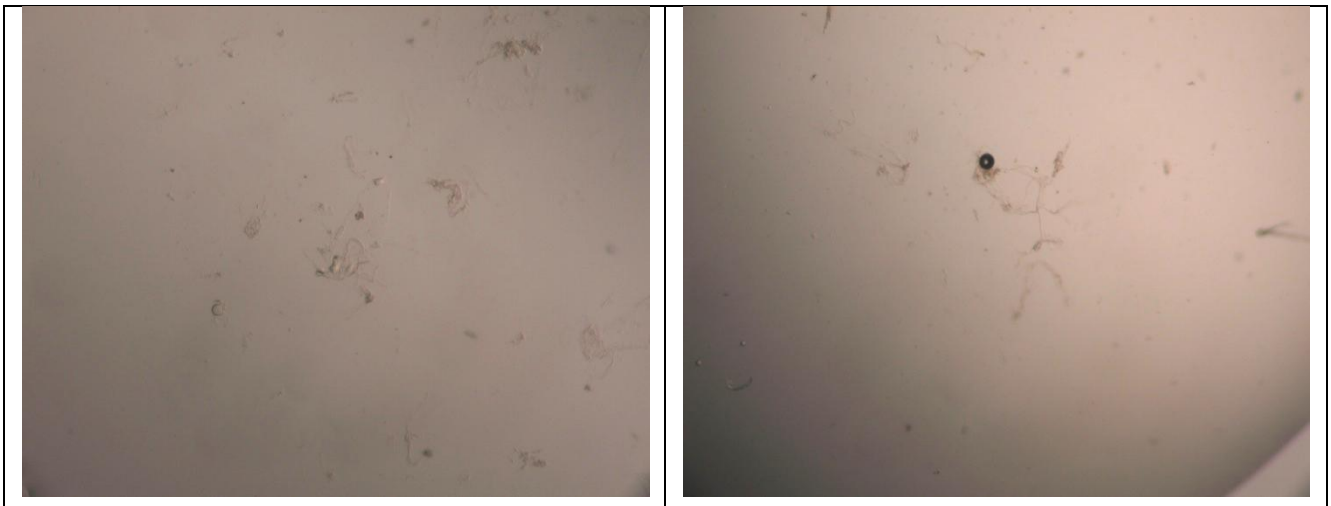
Lane FT: SU after being flown through the chromatography column

Lane W35-50: washing fractions of 35, and 50 mM of imidazole and 6 M of urea

Lane E75-100-500: elution fractions of 75, 100, and 500 mM of imidazole and 6 M of urea

4.2.2 Microscopy observation

Both proteins expressed at 30 ^{o}C formed visible aggregates at high concentration after dialysis against 50 mM sodium phosphate pH 7 to remove urea and imidazole. Visible millimeter-long fibers (Fig. 13) were observed amongst these aggregates. Aggregates were collected by centrifugation and further observed via optical microscopy.



a **b**

Figure 13: Stereoscope images of fibrillar aggregates of RGDD-Bhx (a) and RGDS-Bhx (b) expressed at 30 ^{o}C, solubilized in 6 M urea, purified in a Ni-NTA column, and dialyzed against 50 mM sodium phosphate pH 7. Both images were taken in solution and the magnification is 5x.

Summary: RGD insertion in the N-terminus of Bhxwas successful (RGDD-Bhx), as well as the substitution of the following Asp residue to Ser (RGDS-Bhx)in order to mimic the natural sequence of fibronectin which mediates cell adhesion. Both RGDD-Bhx and RGDS-Bhx analogues were insoluble when expressed at 30^{o}C, in contrast with Bhx which is mostly soluble in this temperature. After solubilization in urea, both RGD-containing analogues of Bhx were efficiently purified in a Ni-NTA column. After dialysis in order to remove imidazole and urea, fibrillar aggregates were observed for both RGDD-Bhx and RGDS-Bhx.

4.3 N-GFPem-Bhx

4.3.1 Construct Design and Insertion into Plasmid Vector

After the seamless cloning process, a cut reaction was performed to a number of isolated plasmids in order to check if the N-GFPem-Bhx gene was successfully constructed and inserted into the pUC19L vector. NdeI and ClaI restriction enzymes were chosen for the reaction; the enzyme reaction generated DNA fragments of 960 and 3600 bases, as expected.

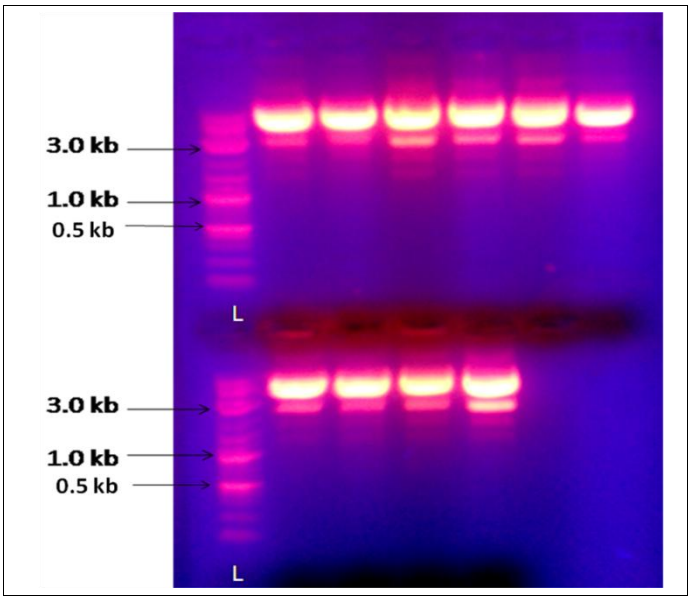


Figure 15: Agarose gel electrophoresis of pUC19L cut with NdeI and ClaI after seamless cloning. The clones in the second and sixth lane on the top row were not positive.

Lane L: DNA Ladder

The positive clones of the bottom row in Fig. 15 were selected and the insert gene was PCR-amplified in order to be cloned into a pT7-7 vector using NdeI and PstI restriction enzymes. After the ligation reaction was completed, a cut reaction was performed to the plasmids in order to check if the N-GFPem-Bhx gene was successfully inserted into the pT7-7 vector (Fig 16). After the selection of positive clones, the plasmids were further validated via DNA-sequencing (see Appendix I). A schematic representation of the new construct is available in Fig. 16.

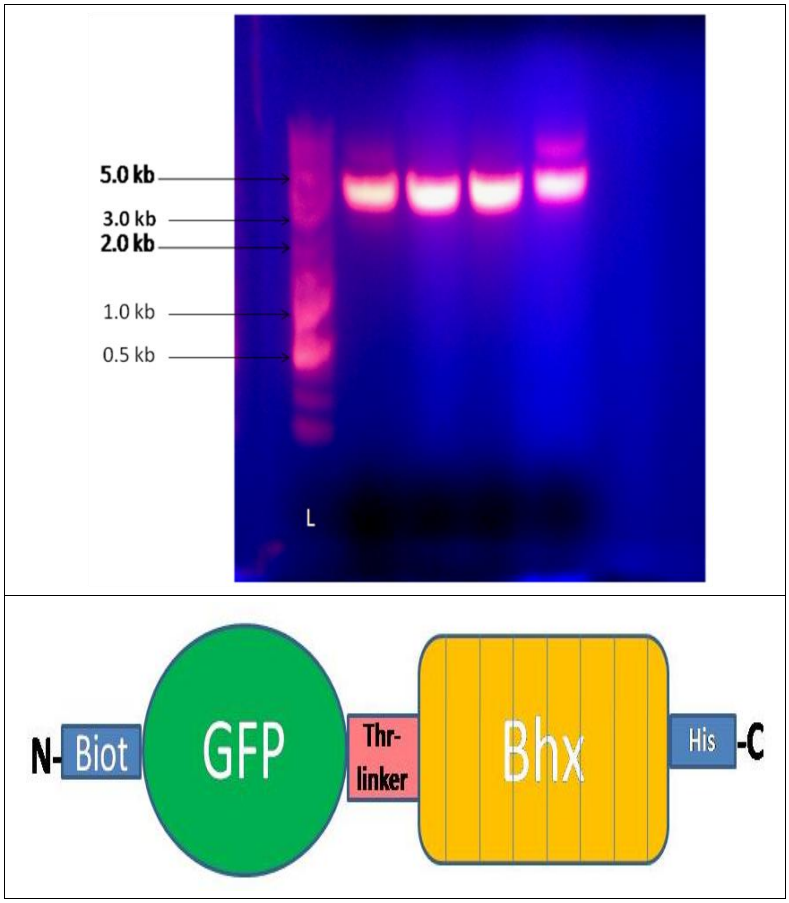


Figure 16: Top: Agarose gel electrophoresis of pT7-7 cut with NdeI and PstI after cloning the N-GFPem-Bhx gene into the vector. Positive clones are not easily visible due to bad analysis. The clones in the second and third lane were not positive. Bottom: Schematic representation of N-GFPem-Bhx. Biot stands for the biotinylation peptide, GFP stands for the emerald green fluorescent protein (GFPem) in the fusion, Thr-linker stands for the peptide linker that contains the thrombin recognition sequence (SGLVPRGS), Bhx stands for the tailspike β-helical domain and His for the C-terminal hexahistidyl tag.

Lane L: DNA Ladder

4.3.2 Expression-Lysis-Purification

N-GFPem-Bhx overexpression was induced properly in the bacterial host at both 30 and 37 ^{o}C, yielding a green bacterial pellet upon harvesting. Fluorescence microscopy results also show the formation of inclusion bodies when the protein fusion was expressed at 30 ^{o}C (Fig. 17 a,b), thus this temperature was considered as more appropriate for the experimentation purposes.

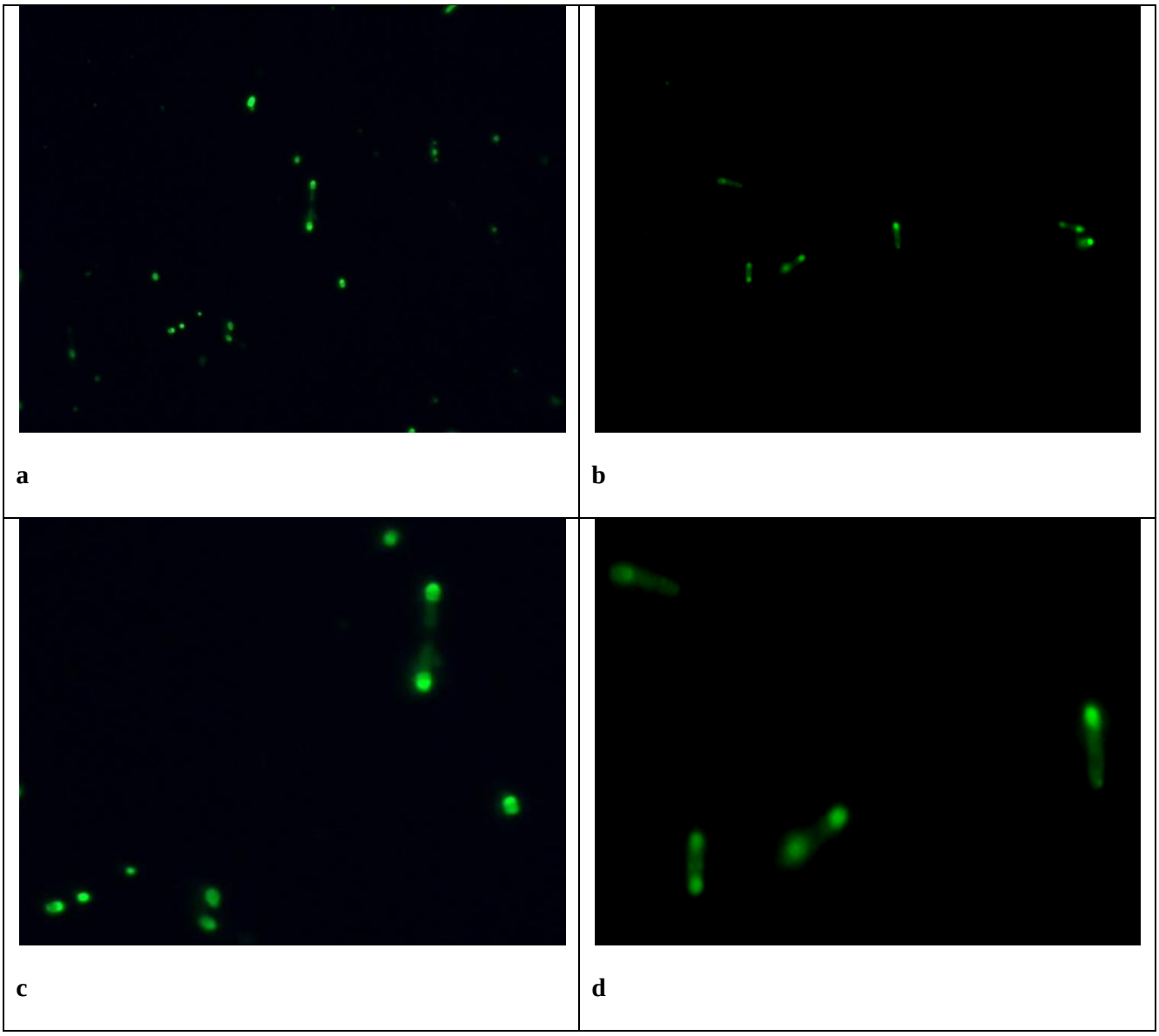


Figure 17: Top: Fluorescence microscopy images of E. coli cells after N-GFPem-Bhx expression at 30 ^{o}C overnight (a) and 37 ^{o}C for 4 h (b). The inclusion bodies appear as bright green spots due to GFPem fluorescence, often located at one or both poles of the cells. Images were taken at 100x magnification. (c) and (d) depict magnified areas from (a) and (b) respectively.

Apart from the fusion protein band (at ~78.5 kDa), there is a ladder of degradation products with a major band appearing since the early stage of overexpression at ~37 kDa (Fig. 18 a). This suggests that there is a proteolysis-sensitive site present approximately in the middle of the fusion. In Western blot, the degradation products were recognized by anti-His antibodies throughout the range of molecular weights (Fig. 18 b), whereas anti-GFP antibodies recognized only higher molecular weight bands (Fig. 18 c).This result indicated a complex degradation process, which involves the hydrolysis of the fusion protein in half, N-terminal degradation of

the whole protein fusion and/or the GFPem moiety, and further N-terminal degradation of the already hydrolyzed parts of the fusion.

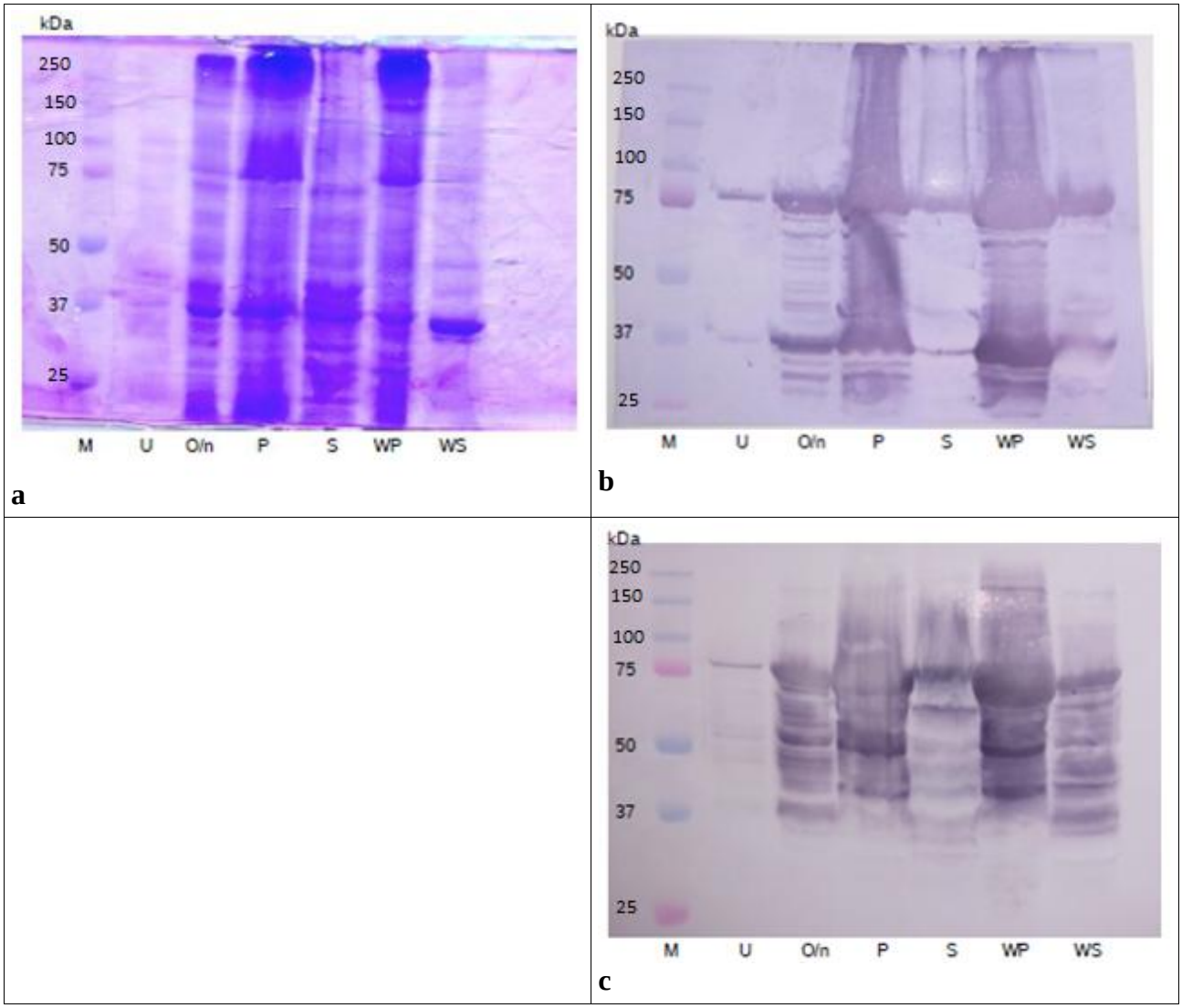


Figure 18: Top: SDS-PAGE expression at 30 ^{o}C and freeze-thaw lysis with sonication profile of N-GFPem-Bhx (a) and anti-His Western blot. Bottom: anti-GFP Western blot (c).

Lane M: protein molecular weight marker

Lane U: uninducted bacteria

Lane O/n: protein overexpression overnight at 30 ^{o}C

Lane P: pellet obtained after lysis

Lane S: supernatant obtained after lysis

Lane WP: washing pellet

Lane WS: washing supernatant

N-GFPem-Bhx expressed at 30 °C could not be purified in a Ni-NTA column after solubilization in 6 M urea due to the presence of the aforementioned degradation products (Fig. 19 a). In an anti-His Western blot, bands were recognized by the antibody throughout the range of molecular weights (Fig. 19 b). Again, anti-GFP antibodies recognized only higher molecular weight bands (Fig. 19 c).

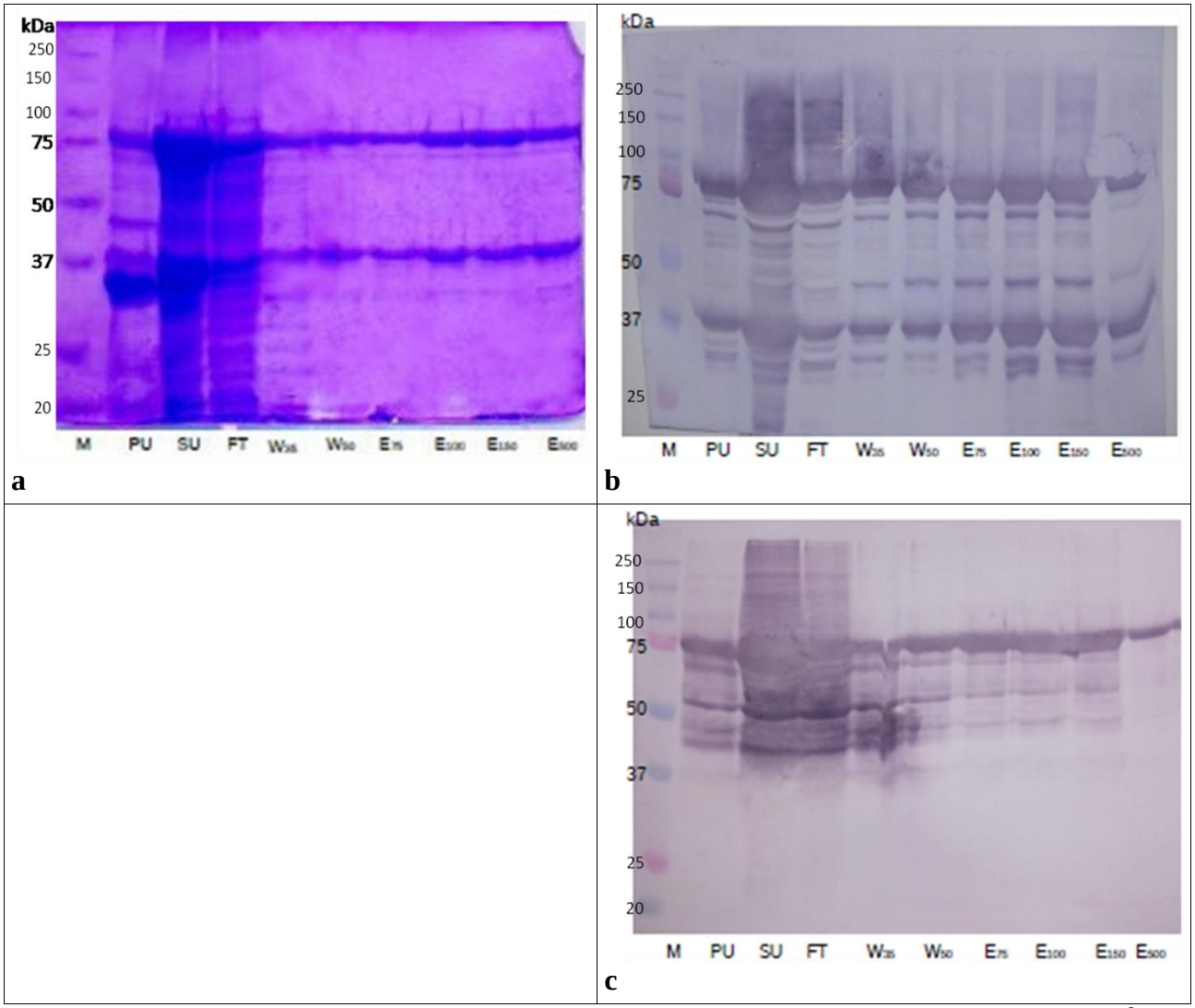


Figure 19: Top: SDS-PAGE Ni-NTA purification profile of N-GFPem-Bhx expressed at 30 °C (a) and anti-His Western blot. Bottom: anti-GFP Western blot (c)

Lane M: protein molecular weight marker

Lane PU: pellet obtained after resuspension in 6 M urea

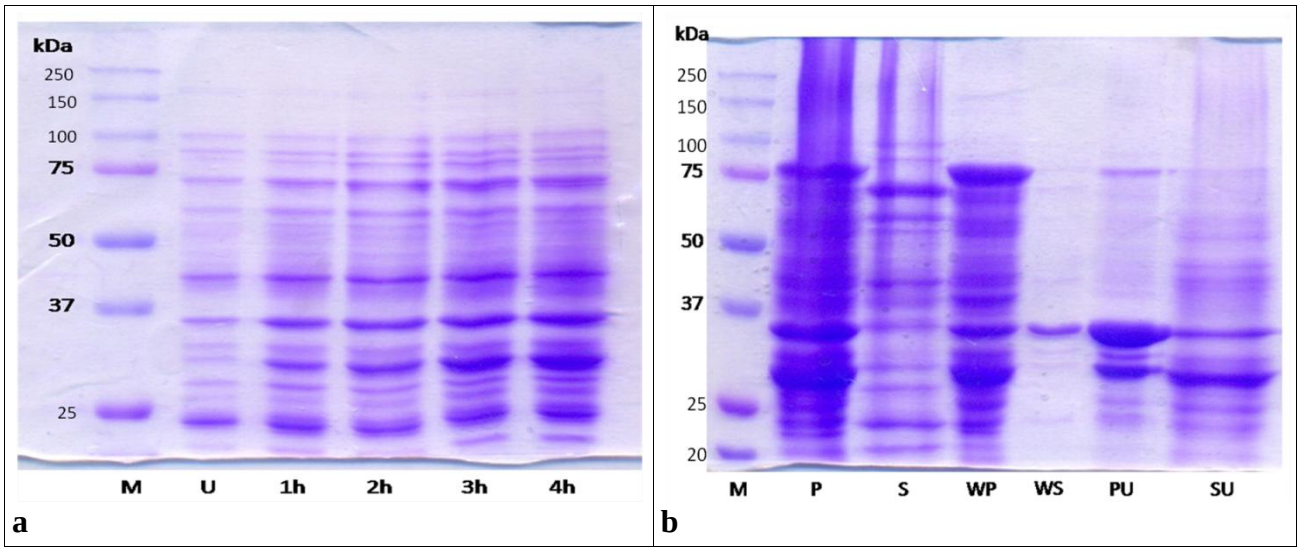
Lane SU: supernatant obtained after resuspension in 6 M urea

Lane FT: SU after being flown through the chromatography column

Lane W35-50: washing fractions of 5, 35, and 50 mM of imidazole and 6 M of urea

Lane E75-100-500: elution fractions of 75, 100, and 500 mM of imidazole and 6 M of urea

An attempt was made to purify the biotinylated protein fusion using both a Ni-NTA and a streptavidin column. N-GFPem-Bhx was co-expressed with biotin ligase for 4h at 30 ^{o}C in LB medium containing biotin (Fig. 20 a). The protein was purified using a Ni-NTA column from the supernatant obtained after resuspension of the lysis pellet in 6 M urea (Fig. 20 b,c). The elution fractions from the first purification step were pooled and dialyzed against 20 mM phosphate pH 7.5 and 150 mM NaCl in order to proceed to the second purification step. The results of the purification steps suggest that the protein was efficiently biotinylated and purified (Fig. 19 d). Nevertheless, the extreme denaturing conditions during elution from the strepavidin column render protein refolding difficult. GFPem retained its fluorescence properties in 6 M urea, whereas a loss of fluorescence was observed in 8 M Gdn-HCl. GFPem fluorescence could not be regained by dialyzing the sample against decreasing Gdn-Cl concentrations.



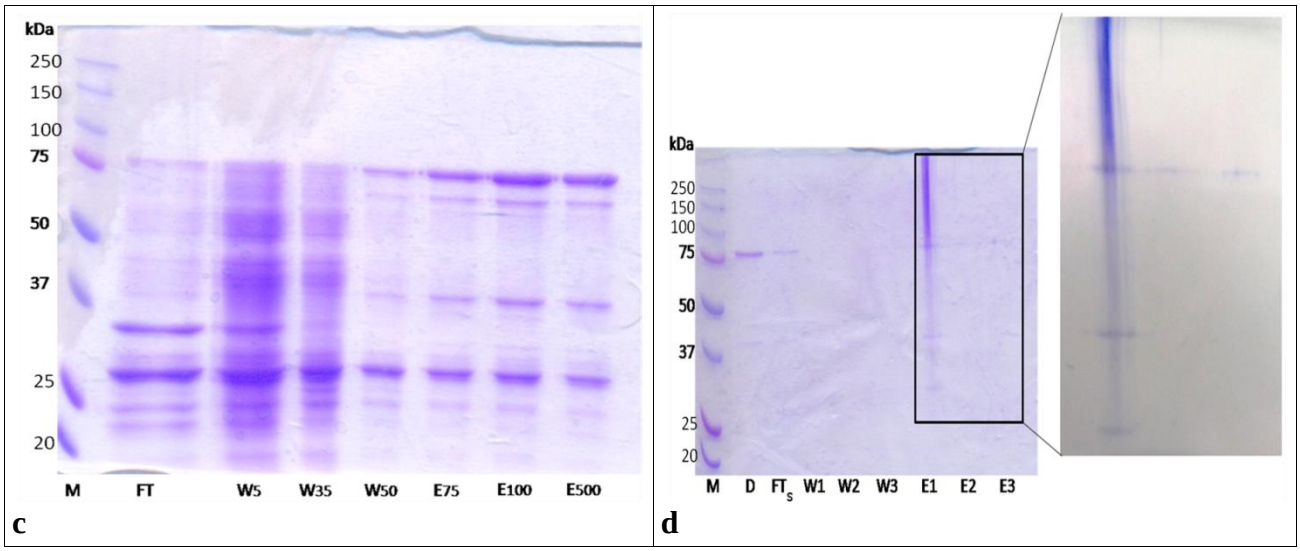


Figure 20: Top: SDS-PAGE expression at 30 ^{o}C (a), freeze-thaw lysis with sonication (b), Ni-NTA purification (C), and streptavidin purification (d) profile of N-GFPem-Bhx

Lane M: protein molecular weight marker

Lane U: uninducted bacteria

Lane 1-4h: protein overexpression in the course of 4h

Lane P: pellet obtained after lysis

Lane S: supernatant obtained after lysis

Lane WP: washing pellet

Lane WS: washing supernatant

Lane PU: pellet obtained after resuspension in 6 M urea

Lane SU: supernatant obtained after resuspension in 6 M urea

Lane FT: SU after being flown through the Ni-NTA chromatography column

Lane W5-35-50: washing fractions of 5, 35, and 50 mM of imidazole and 6 M of urea

Lane E75-100-500: elution fractions of 75, 100, and 500 mM of imidazole and 6 M of urea

Lane D: Ni-NTA column elution fractions after dialysis against 20 mM phosphate pH 7.5 and 150 mM NaCl

Lane FT_S: D after being flown through the chromatography column

Lane W1-2-3: subsequent washing fractions with 20 mM phosphate pH 7.5 and 150 mM NaCl

Lane E1-2-3: subsequent elution fractions with 8 M Gdn-HCl pH 1.5

Summary: The fusion protein N-GFPem-Bhx was efficiently constructed and cloned in the pT7-7 vector. The protein was not soluble when expressed at 30^{o}C. PAGE and Western blot results suggest that the fusion protein has a site which is sensitive to proteolysis located approximately in the middle of its sequence. Degradation is evident for the protein fusion and/or the GFPem moiety as well as for the N-terminus of the by-products which were cleaved in half. Thus, the protein fusion could not be efficiently purified in a Ni-NTA column. The biotinylated N-GFPem-Bhx could be efficiently purified in a streptavidin column, indicating that the biotinylation peptide sequence is exposed and fully functional, but the extreme elution conditions render refolding difficult.

4.4 Bhx-GFPem-C

4.4.1 Construct Design and Insertion into Plasmid Vector

After the seamless cloning process, a cut reaction was performed to a number of isolated plasmids in order to check if the Bhx-GFPem-C gene was successfully constructed and inserted into the pUC19L vector. NdeI and ClaI restriction enzymes were chosen for the reaction; the enzyme reaction generated DNA fragments of ~340 and ~3400 bases from positive clones (Fig. 21).

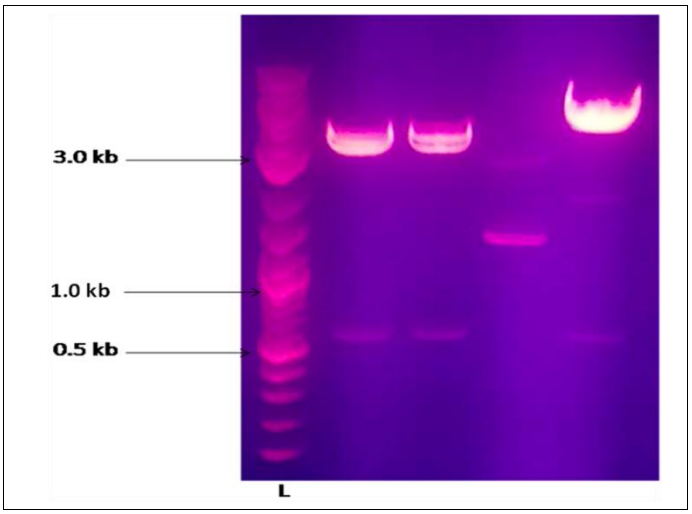


Figure 21: Agarose gel electrophoresis of pUC19L cut with NdeI and ClaI after seamless cloning. The clone in the third lane was not positive.

Lane L: DNA Ladder

The positive clones were selected and the insert gene was PCR-amplified in order to be cloned into a pET28a vector using EcoRI and XhoI restriction enzymes. After the ligation reaction was completed, a cut reaction was performed to the plasmids in order to check if the Bhx-GFPem-C gene was successfully inserted into the pET28a vector (Fig 22). All clones appeared positive, since the reaction generated a DNA fragment of 2076 bp. After an overexpression check, two of the clones were further validated via DNA-sequencing (see Appendix I). A schematic representation of the new construct is available in Fig. 22.

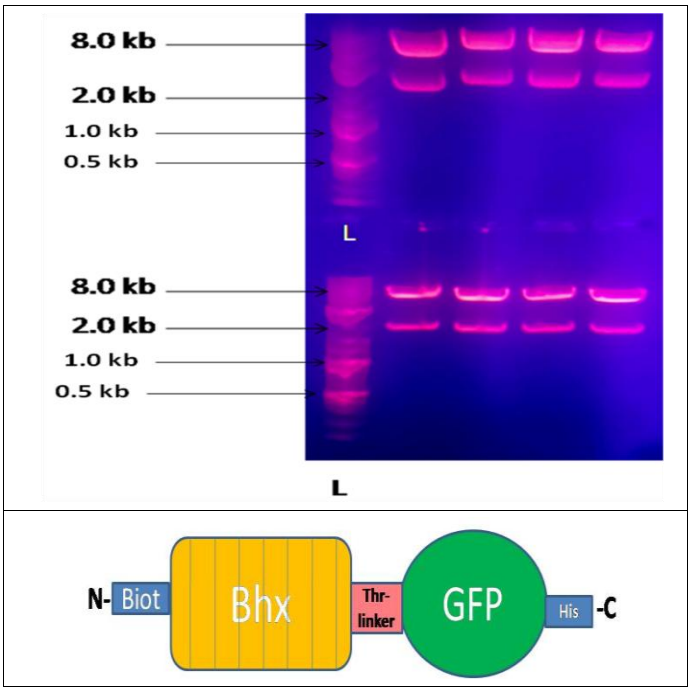


Figure 22: Top: Agarose gel electrophoresis of pET28a cut with EcoRI and XhoI after cloning the Bhx-GFPem-C gene into the vector. Bottom: Schematic representation of Bhx-GFPem-C. Biot stands for the biotinylation peptide, Bhx stands for the tailspike β-helical domain, Thr-linker stands for the peptide linker that contains the thrombin recognition sequence (SSGLVPRGS), GFP stands for the emerald green fluorescent protein (GFPem) in the fusion and His for the C-terminal hexahistidyl tag.

Lane L: DNA Ladder

4.4.2 Expression-Lysis-Purification

Bhx-GFPem-C overexpression was induced properly in the bacterial host at both 30 and 37 ^{o}C, yielding a green bacterial pellet upon harvesting (Fig. 23 a,b).

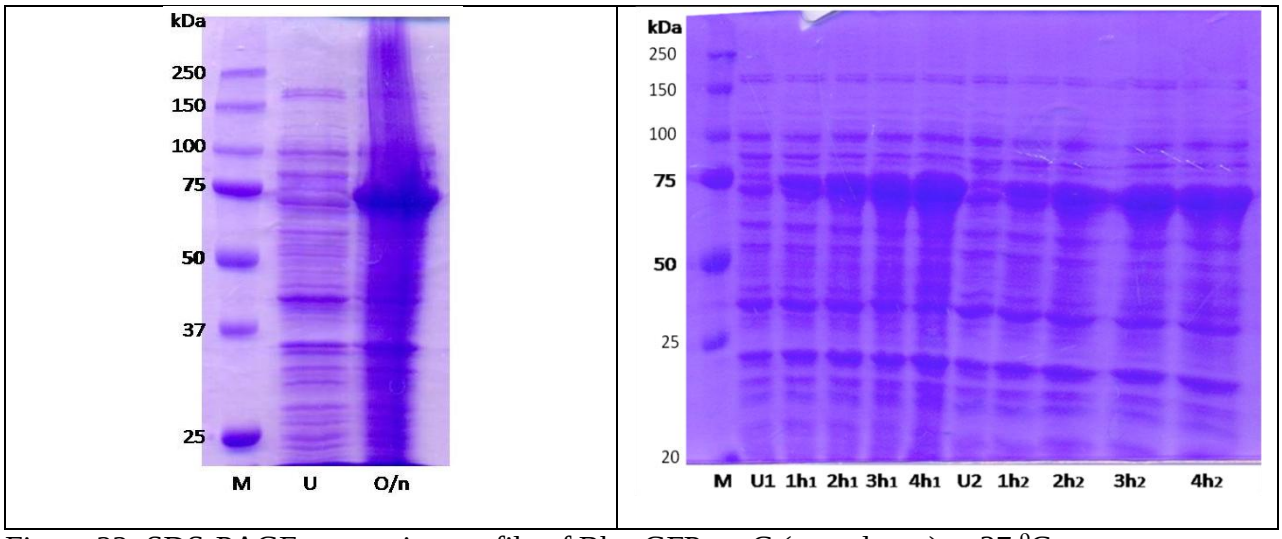


Figure 22: SDS-PAGE expression profile of Bhx-GFPem-C (two clones) at 37 ^{o}C

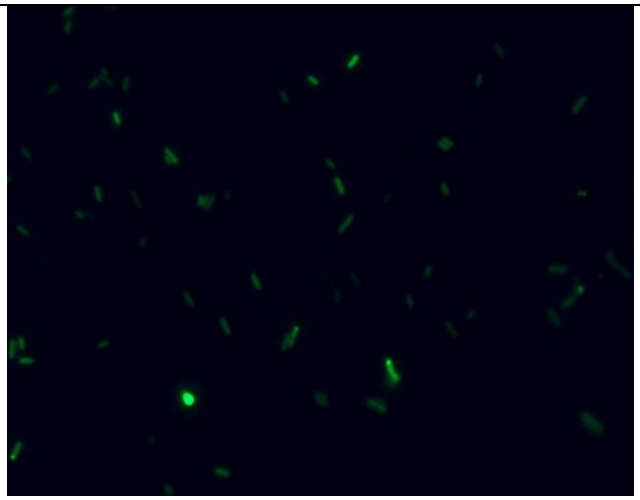
Lane M protein molecular weight marker

Lane U: uninducted bacteria

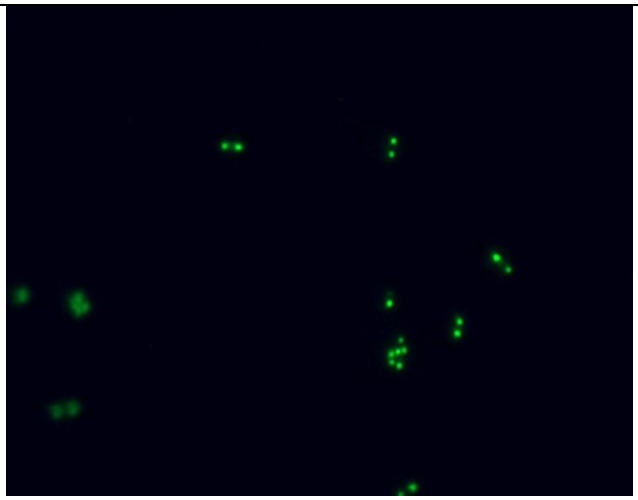
Lane O/n: protein overexpression overnight

Lane 1-4h: protein overexpression in the course of 4h

Fluorescence microscopy results show the formation of inclusion bodies when the protein fusion was expressed at 37 ^{o}C, while at 30 ^{o}C a partition of the fusion protein in the cytoplasm and inclusion bodies was observed (Fig. 24 a,b). PAGE results suggest that at both temperatures the protein fusion is found as aggregates in the pellet of the lysate (Fig. 25).



a



b

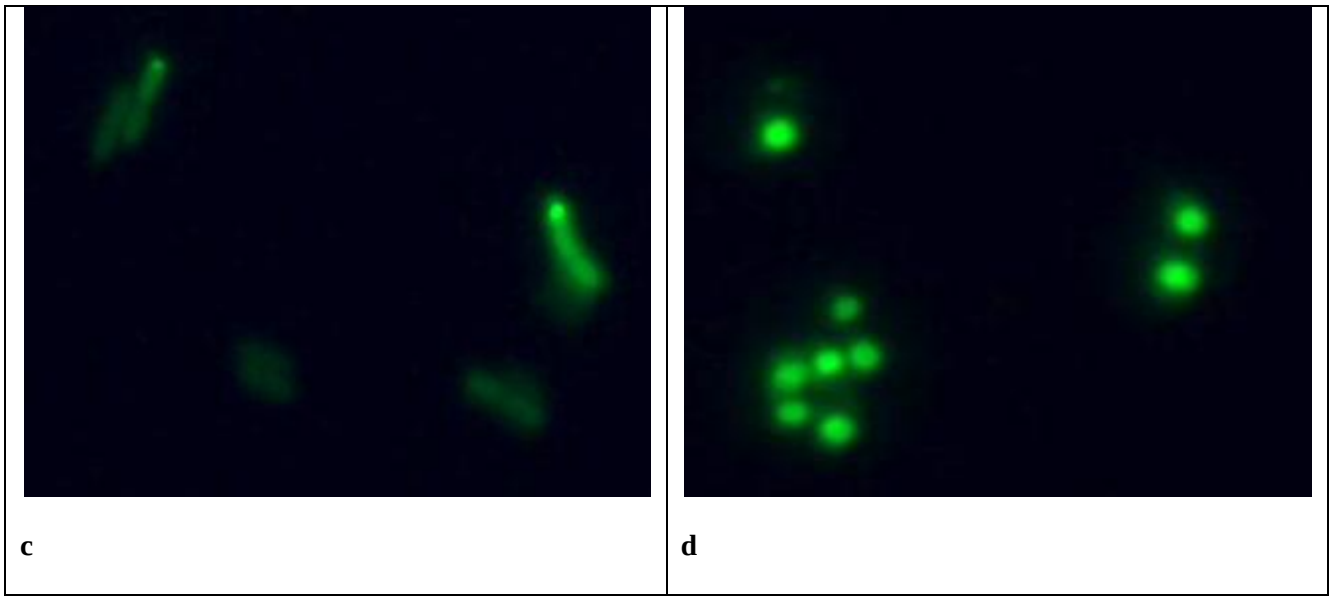
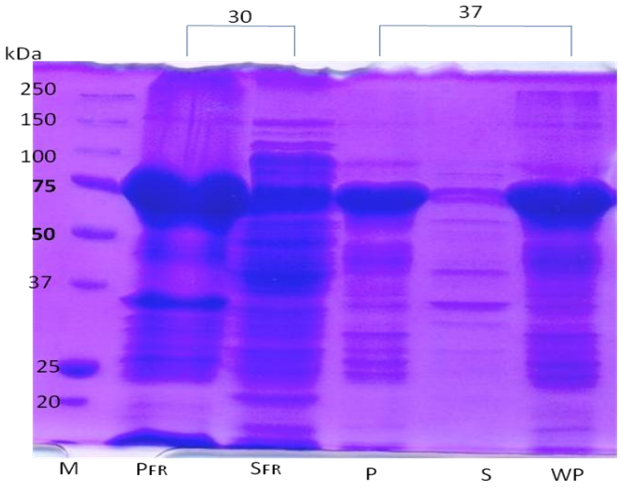


Figure 24: Fluorescence microscopy images of E. coli cells after N-GFPem-Bhx expression at 30 ^{o}C overnight (a) and 37 ^{o}C for 4 h (b). The inclusion bodies appear as bright green spots due to GFPem fluorescence, often located at one or both poles of the cells. Images were taken at 100x magnification. (c) and (d) depict magnified areas from (a) and (b) respectively.

Bhx-GFPem-C expressed at 30 and 37 ^{o}C could be purified in a Ni-NTA column after solubilization in 6 M urea (Fig 25). A large amount of the protein fusion seems to be eluted in the washing fraction with 50 mM imidazole. In comparison with N-GFPem-Bhx, the present of degradation products is less evident and the impurities can probably become less with further optimization of the purification process.



a

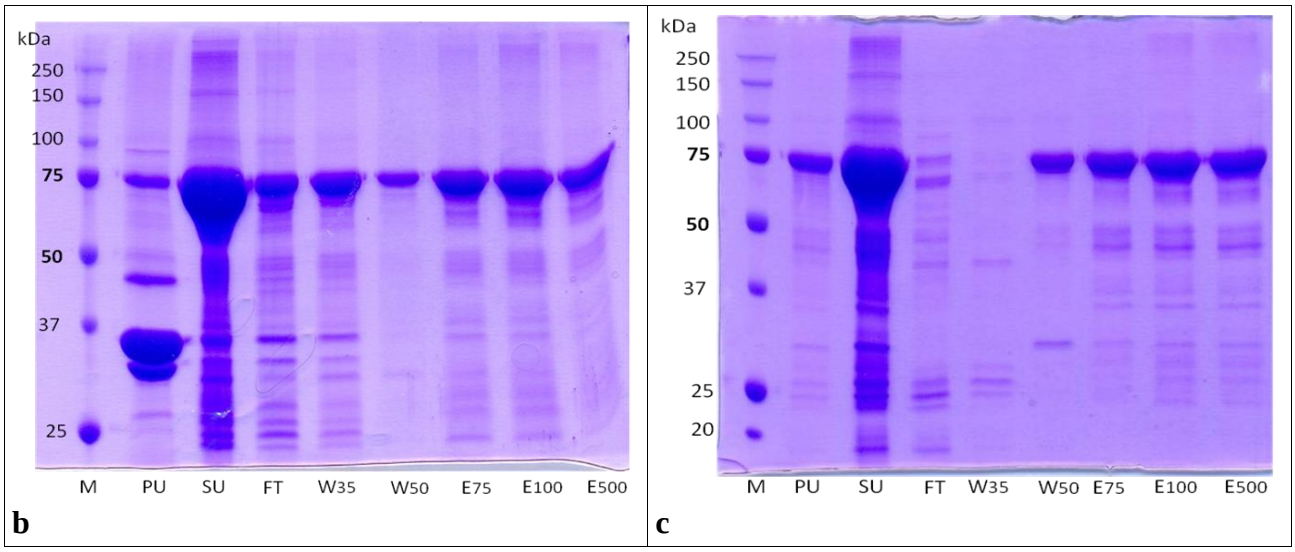


Figure 25: Top: SDS-PAGE profile of French press and freeze-thaw lysis with sonication of Bhx-GFPem-C expressed at 30 and 37 ^{o}C (a). Bottom: Ni-NTA purification profile of Bhx-GFPem-C expressed at 30 (b) and 37 ^{o}C (c)

Lane M: protein molecular weight marker

Lane PFR: pellet obtained after French press lysis

Lane SFR: supernatant obtained after French press lysis

Lane P: pellet obtained after freeze-thaw lysis

Lane S: supernatant obtained after freeze-thaw lysis

Lane WP: washing pellet

Lane PU: pellet obtained after resuspension in 6 M urea

Lane SU: supernatant obtained after resuspension in 6 M urea

Lane FT: SU after being flown through the Ni-NTA chromatography column

Lane W35-50: washing fractions of 35, and 50 mM of imidazole and 6 M of urea

Lane E75-100-500: elution fractions of 75, 100, and 500 mM of imidazole and 6 M of urea

Summary: The fusion protein Bhx-GFPem-C was efficiently constructed and cloned in the pET28a vector. When expressed at 30 ^{o}C, the protein partitioned between a soluble state and inclusion bodies. PAGE and Western blot results suggest that the fusion protein is far less sensitive to degradation in comparison to the N-terminal GFPem fusion. Thus the protein was purified more effectively in a Ni-NTA column. In Bhx-GFPem-C, the C-terminus of the Bhx

domain and the N-terminus of GFPem are protected. As a result, of Bhx-GFPem folds in a form that is resistant to proteolysis.

Chapter 5 – Discussion

Overexpression of Bhx at 30 °C and 37 °C provided high amounts of the protein in a cytosoluble and aggregated form, respectively. A ladder of degradation products recognized by the anti-His antibody were evident when the protein was expressed at 37 °C, solubilized in urea and purified. This finding suggests that the N-terminus is not stable, leading to gradual degradation of the protein. Dialysis of the protein expressed at 30 °C resulted in the formation of spherical aggregates, whereas the protein expressed at 37 °C formed both fibrous and spherical microstructures. Cross-sections of the microstructures were observed via TEM and revealed the hierarchical formation of these microstructures from cross-linked nano-fibrils. The lack of the natural C-terminal trimerization motif plays probably a determining role in the formation of these amyloid-type nano-fibrils along with the dialysis conditions (low ionic strength, high protein concentration). Trimerization motifs of proteins with repetitive sequences such as the adenovirus fiber shaft protein have C-terminal trimerization motifs that are highly important for efficient folding, while their absence leads to amyloid fibril formation (Papanikolopoulou et al., 2004). Furthermore, sequences in the Bhx domain itself that probably correspond to the sites of *tsf* mutations remain to be determined in future studies and must play a key role in the thermoresponsiveness and the effective refolding of the domain.

The insertion of RGD in the sequence of Bhx resulted in the formation of inclusion bodies during recombinant expression at 30 °C. This result indicates the significance of N- and C-terminal functional peptides in the thermal response and aggregation of soluble proteins (Zhou et al., 2012). Dialysis in this case also resulted in the formation of fibrillar aggregates. This result indicates that renaturation from the unfolded state in urea is the major factor affecting the formation of macroscopic fibrils and it is a good start for their further evaluation as scaffolds for tissue engineering.

Fusion of GFPem with Bhx revealed two basic findings. The first is the importance of the location of a protein in the fusion construct regarding its resistance to proteolysis. N-terminal fusion of GFPem resulted in extensive degradation of the protein fusion and/or the GFPem moiety, whereas for the C-terminal fusion, degradation was significantly less. Fluorescence

microscopy results revealed that when expressed at 37 °C both N-GFPem-Bhx and Bhx-GFPem-C form inclusion bodies during overexpression, while at 30 °C Bhx-GFPem-C seemed to partition between the cytoplasm and inclusion bodies. This result indicates that the C-terminus of Bhx and the N-terminus of GFPem need to be protected in the fusion order to fold into a protease resistant form. The second finding is that GFPem can be efficiently biotinylated and retain its native fluorescence in inclusion bodies in the fusion protein. This result is consistent with previous findings which suggest that proteins in inclusion bodies retain their activity (Vazquez et al., 2012). Fusion proteins of GFP with aggregation-prone sequences such as the VP1 capsid protein of the foot-and-mouth virus or the Aβ peptide from Alzheimer's disease are have been reported to form inclusion bodies, with the GFP moiety remaining active (Garcia-Fruitos et al., 2005; Garcia-Fruitos et al., 2009). The presence of an active fluorescent protein on an aggregation-prone sequence can provide data regarding folding both *in vivo* and *in vitro*, and determine the shape of these aggregates (Ignatova and Gierasch, 2006; Ignatova et al., 2007). It should be noted that in these protein fusions the fluorescent protein moiety was located in the C-terminus. The existence of inclusion bodies of GFPem fused with Bhx generates interesting questions regarding the structure of these aggregates. Therefore, there should be further investigation in order to determine if the aggregation-prone moiety drives the assembly in a way that a central spine is formed with GFPem protruding. If Bhx can form a fibrous matrix that can be used as support for immobilized active enzymes, such scaffolds would be good candidates for catalytical purposes. GFPem was used as a model enzyme, but other enzymes could be used in its place for the desired properties. For example, considering that Bhx has a both a binding and endorhamnosidase function, it is possible to fuse Bhx with another functional enzyme for the development of protein aggregates that can facilitate complex carbohydrate reactions.

In modern research, protein aggregates are no longer discarded as useless. They have been re-evaluated and find applications in medicine and industry as novel materials. Bhx has proven useful for the construction of self-assembling protein scaffolds. It can be produced in a large scale in bacteria, aggregate effectively in ambient conditions, and be further modulated by inserting peptides or proteins that retain their functionality. Bhx aggregates need to be further explored for their biocompatibility and catalytic activity, but the results so far are promising for future applications.

References

Alberts, B., 2002. Molecular biology of the cell, 4th ed. Garland Science, New York.

Alsteens, D., Martinez, N., Jamin, M., Jacob-Dubuisson, F., 2013. Sequential unfolding of beta helical protein by single-molecule atomic force microscopy. PloS one 8, e73572.

Andres, D., Gohlke, U., Broeker, N.K., Schulze, S., Rabsch, W., Heinemann, U., Barbirz, S., Seckler, R., 2013. An essential serotype recognition pocket on phage P22 tailspike protein forces Salmonella enterica serovar Paratyphi A O-antigen fragments to bind as nonsolution conformers. Glycobiology 23, 486-494.

Andres, D., Hanke, C., Baxa, U., Seul, A., Barbirz, S., Seckler, R., 2010. Tailspike interactions with lipopolysaccharide effect DNA ejection from phage P22 particles in vitro. The Journal of biological chemistry 285, 36768-36775.

Barbirz, S., Becker, M., Freiberg, A., Seckler, R., 2009. Phage tailspike proteins with beta-solenoid fold as thermostable carbohydrate binding materials. Macromolecular bioscience 9, 169-173.

Barondeau, D.P., Putnam, C.D., Kassmann, C.J., Tainer, J.A., Getzoff, E.D., 2003. Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. Proceedings of the National Academy of Sciences of the United States of America 100, 12111-12116.

Baxa, U., Steinbacher, S., Miller, S., Weintraub, A., Huber, R., Seckler, R., 1996. Interactions of phage P22 tails with their cellular receptor, Salmonella O-antigen polysaccharide. Biophysical journal 71, 2040-2048.

Baxa, U., Taylor, K.L., Wall, J.S., Simon, M.N., Cheng, N.Q., Wickner, R.B., Steven, A.C., 2003. Architecture of Ure2p prion filaments - The N-terminal domains form a central core fiber. Journal of Biological Chemistry 278, 43717-43727.

Beissinger, M., Lee, S.C., Steinbacher, S., Reinemer, P., Huber, R., Yu, M.H., Seckler, R., 1995. Mutations that stabilize folding intermediates of phage P22 tailspike protein: folding in vivo and in vitro, stability, and structural context. Journal of molecular biology 249, 185-194.

Betts, S., Haase-Pettingell, C., Cook, K., King, J., 2004. Buried hydrophobic side-chains essential for the folding of the parallel beta-helix domains of the P22 tailspike. Protein science : a publication of the Protein Society 13, 2291-2303.

Bhattacharyya, J., Bellucci, J.J., Weitzhandler, I., McDaniel, J.R., Spasojevic, I., Li, X., Lin, C.C., Chi, J.T., Chilkoti, A., 2015. A paclitaxel-loaded recombinant polypeptide nanoparticle outperforms Abraxane in multiple murine cancer models. Nature communications 6, 7939.

Casjens, S.R., Thuman-Commike, P.A., 2011. Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. Virology 411, 393-415.

Cespedes, M.V., Unzueta, U., Tatkiewicz, W., Sanchez-Chardi, A., Conchillo-Sole, O., Alamo, P., Xu, Z., Casanova, I., Corchero, J.L., Pesarrodona, M., Cedano, J., Daura, X., Ratera, I., Veciana, J., Ferrer-Miralles, N., Vazquez, E., Villaverde, A., Mangues, R., 2014. In vivo architectonic stability of fully de novo designed protein-only nanoparticles. ACS nano 8, 4166-4176.

Chalfie, M., 1995. Green fluorescent protein. Photochem Photobiol 62, 651-656.

Corchero, J.L., Viaplana, E., Benito, A., Villaverde, A., 1996. The position of the heterologous domain can influence the solubility and proteolysis of beta-galactosidase fusion proteins in E-coli. J Biotechnol 48, 191-200.

Day, R.N., Davidson, M.W., 2009. The fluorescent protein palette: tools for cellular imaging.

Chemical Society reviews 38, 2887-2921.
Freifelder, D., 1987. Microbial genetics. Jones and Bartlett, Boston, MA.
Fuller, M.T., King, J., 1980. Regulation of coat protein polymerization by the scaffolding protein of bacteriophage P22. Biophysical journal 32, 381-401.
Gage, M.J., Robinson, A.S., 2003. C-terminal hydrophobic interactions play a critical role in oligomeric assembly of the P22 tailspike trimer. Protein science : a publication of the Protein Society 12, 2732-2747.
Garcia-Fruitos, E., Gonzalez-Montalban, N., Morell, M., Vera, A., Ferraz, R.M., Aris, A., Ventura, S., Villaverde, A., 2005. Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. Microbial cell factories 4.
Garcia-Fruitos, E., Rodriguez-Carmona, E., Diez-Gil, C., Ferraz, R.M., Vazquez, E., Corchero, J.L., Cano-Sarabia, M., Ratera, I., Ventosa, N., Veciana, J., Villaverde, A., 2009. Surface Cell Growth Engineering Assisted by a Novel Bacterial Nanomaterial. Advanced Materials 21, 4249-+.
Harvey, A.M., Hava, P., Oppenheim, A.B., Prell, H.H., Soska, J., 1981. Repression of ant synthesis early in the lytic cycle of phage P22. MGG Molecular & General Genetics 181, 74-81.
Hovmoller, S., Zhou, T., 2004. Why are both ends of the polypeptide chain on the outside of proteins? Proteins 55, 219-222.
Ignatova, Z., Gierasch, L.M., 2006. Extended polyglutamine tracts cause aggregation and structural perturbation of an adjacent beta barrel protein. The Journal of biological chemistry 281, 12959-12967.
Ignatova, Z., Thakur, A.K., Wetzel, R., Gierasch, L.M., 2007. In-cell aggregation of a polyglutamine-containing chimera is a multistep process initiated by the flanking sequence. The Journal of biological chemistry 282, 36736-36743.
Kajava, A.V., Cheng, N., Cleaver, R., Kessel, M., Simon, M.N., Willery, E., Jacob-Dubuisson, F., Locht, C., Steven, A.C., 2001. Beta-helix model for the filamentous haemagglutinin adhesin of Bordetella pertussis and related bacterial secretory proteins. Molecular microbiology 42, 279-292.
Kale, A., Bao, Y., Zhou, Z., Prevelige, P.E., Gupta, A., 2013. Directed self-assembly of CdS quantum dots on bacteriophage P22 coat protein templates. Nanotechnology 24, 045603.
King, J., Haase-Pettingell, C., Robinson, A.S., Speed, M., Mitraki, A., 1996. Thermolabile folding intermediates: inclusion body precursors and chaperonin substrates. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 10, 57-66.
Koshiyama, T., Ueno, T., Kanamaru, S., Arisaka, F., Watanabe, Y., 2009. Construction of an energy transfer system in the bio-nanocup space by heteromeric assembly of gp27 and gp5 proteins isolated from bacteriophage T4. Organic & biomolecular chemistry 7, 2649-2654.
Koshiyama, T., Yokoi, N., Ueno, T., Kanamaru, S., Nagano, S., Shiro, Y., Arisaka, F., Watanabe, Y., 2008. Molecular design of heteroprotein assemblies providing a bionanocup as a chemical reactor. Small 4, 50-54.
Kreisberg, J.F., Betts, S.D., King, J., 2000. Beta-helix core packing within the triple-stranded oligomerization domain of the P22 tailspike. Protein science : a publication of the Protein Society 9, 2338-2343.
Leavitt, J.C., Gogokhia, L., Gilcrease, E.B., Bhardwaj, A., Cingolani, G., Casjens, S.R., 2013. The tip of the tail needle affects the rate of DNA delivery by bacteriophage P22. PloS one 8, e70936.
Loo, Y., Goktas, M., Tekinay, A.B., Guler, M.O., Hauser, C.A., Mitraki, A., 2015. Self-Assembled Proteins and Peptides as Scaffolds for Tissue Regeneration. Advanced healthcare

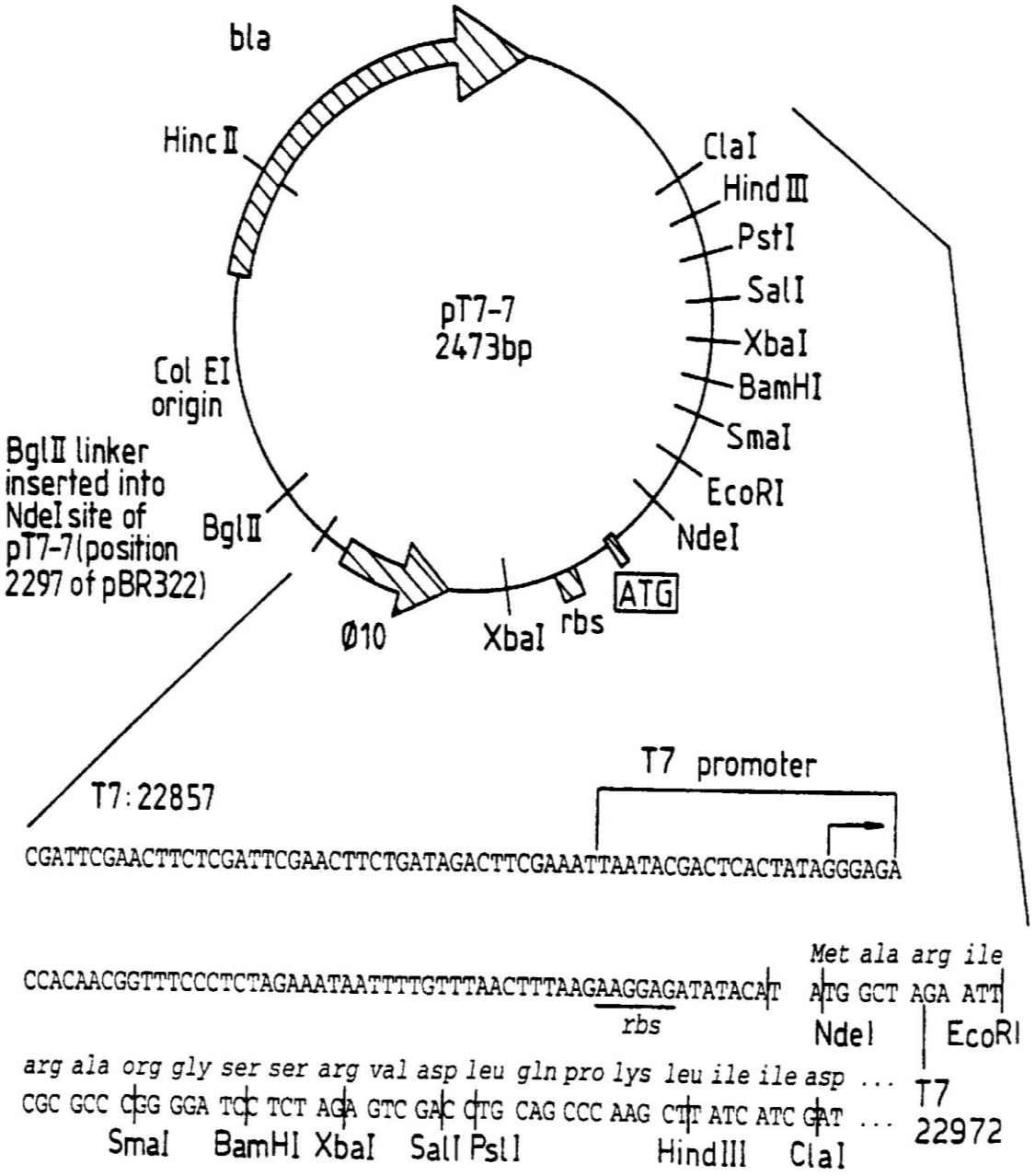
materials 4, 2557-2586.
Miller, S., Schuler, B., Seckler, R., 1998. A reversibly unfolding fragment of P22 tailspike protein with native structure: the isolated beta-helix domain. Biochemistry 37, 9160-9168.
Mitraki, A., 2010. Protein aggregation from inclusion bodies to amyloid and biomaterials. Advances in protein chemistry and structural biology 79, 89-125.
Navarro, S., Ventura, S., 2014. Fluorescent dye ProteoStat to detect and discriminate intracellular amyloid-like aggregates in Escherichia coli. Biotechnology journal 9, 1259-1266.
Olia, A.S., Casjens, S., Cingolani, G., 2007. Structure of phage P22 cell envelope-penetrating needle. Nature structural & molecular biology 14, 1221-1226.
Ormo, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., Remington, S.J., 1996. Crystal structure of the Aequorea victoria green fluorescent protein. Science 273, 1392-1395.
Papanikolopoulou, K., Forge, V., Goeltz, P., Mitraki, A., 2004. Formation of highly stable chimeric trimers by fusion of an adenovirus fiber shaft fragment with the foldon domain of bacteriophage t4 fibritin. The Journal of biological chemistry 279, 8991-8998.
Perez, G.L., Huynh, B., Slater, M., Maloy, S., 2009. Transport of phage P22 DNA across the cytoplasmic membrane. Journal of bacteriology 191, 135-140.
Pipas, J.M., Reeves, R.H., 1979. Host transcription in bacteriophage P22-infected Salmonella typhimurium. Journal of virology 32, 822-831.
Sambrook, J., Russell, D.W., 2001. Molecular cloning : a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
Sawyer, L., Grubb, D., Meyers, G.F., 2008. Polymer Microscopy. Springer New York.
Schuler, B., Rachel, R., Seckler, R., 1999. Formation of Fibrous Aggregates from a Non-native Intermediate: The Isolated P22 Tailspike -Helix Domain. Journal of Biological Chemistry 274, 18589-18596.
Seckler, R., 1998. Folding and function of repetitive structure in the homotrimeric phage P22 tailspike protein. Journal of structural biology 122, 216-222.
Seul, A., Muller, J.J., Andres, D., Stettner, E., Heinemann, U., Seckler, R., 2014. Bacteriophage P22 tailspike: structure of the complete protein and function of the interdomain linker. Acta crystallographica. Section D, Biological crystallography 70, 1336-1345.
Simkovsky, R., King, J., 2006. An elongated spine of buried core residues necessary for in vivo folding of the parallel beta-helix of P22 tailspike adhesin. Proceedings of the National Academy of Sciences of the United States of America 103, 3575-3580.
Singh, A., Arya, S.K., Glass, N., Hanifi-Moghaddam, P., Naidoo, R., Szymanski, C.M., Tanha, J., Evoy, S., 2010. Bacteriophage tailspike proteins as molecular probes for sensitive and selective bacterial detection. Biosensors & bioelectronics 26, 131-138.
Skerra, A., Pfitzinger, I., Plückthun, A., 1991. The Functional Expression of Antibody Fv Fragments in Ischhuchia coli: Improved Vectors and a Generally Applicable Purification Technique. Bio/Technology 9, 273-278.
Snapp, E., 2005. Design and use of fluorescent fusion proteins in cell biology. Current protocols in cell biology / editorial board, Juan S. Bonifacino ... [et al.] Chapter 21, Unit 21 24.
Steinbacher, S., Baxa, U., Miller, S., Weintraub, A., Seckler, R., Huber, R., 1996. Crystal structure of phage P22 tailspike protein complexed with Salmonella sp. O-antigen receptors. Proceedings of the National Academy of Sciences of the United States of America 93, 10584-10588.
Steinbacher, S., Miller, S., Baxa, U., Budisa, N., Weintraub, A., Seckler, R., Huber, R., 1997. Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 A, fully refined structure of the endorhamnosidase at 1.56 A resolution, and the molecular basis of O-antigen

recognition and cleavage. Journal of molecular biology 267, 865-880.
Takata, T., Haase-Pettingell, C., King, J., 2012. The C-terminal cysteine annulus participates in auto-chaperone function for Salmonella phage P22 tailspike folding and assembly. Bacteriophage 2, 36-49.
Teschke, C.M., Parent, K.N., 2010. 'Let the phage do the work': using the phage P22 coat protein structures as a framework to understand its folding and assembly mutants. Virology 401, 119-130.
Vazquez, E., Corchero, J.L., Burgueno, J.F., Seras-Franzoso, J., Kosoy, A., Bosser, R., Mendoza, R., Martinez-Lainez, J.M., Rinas, U., Fernandez, E., Ruiz-Avila, L., Garcia-Fruitos, E., Villaverde, A., 2012. Functional inclusion bodies produced in bacteria as naturally occurring nanopills for advanced cell therapies. Advanced materials (Deerfield Beach, Fla.) 24, 1742-1747.
Vazquez, E., Villaverde, A., 2010. Engineering building blocks for self-assembling protein nanoparticles. Microbial cell factories 9.
Veesler, D., Cambillau, C., 2011. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. Microbiology and molecular biology reviews : MMBR 75, 423-433, first page of table of contents.
Yang, F., Moss, L.G., Phillips, G.N., Jr., 1996. The molecular structure of green fluorescent protein. Nature biotechnology 14, 1246-1251.
Zhou, B., Xing, L., Wu, W., Zhang, X.E., Lin, Z., 2012. Small surfactant-like peptides can drive soluble proteins into active aggregates. Microbial cell factories 11, 10.

Appendix I

Vector maps:

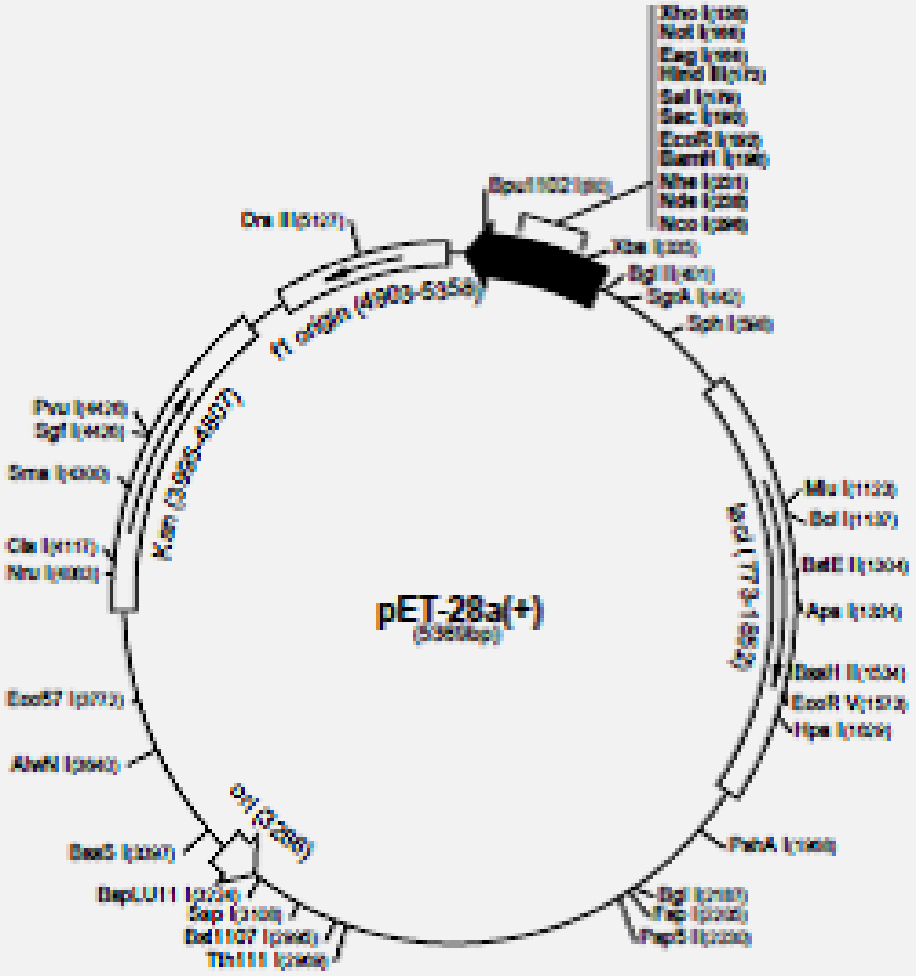
pT7-7 map (Stan Tabor, Dept of Biol. Chem., Harvard Med. School, Boston Mass. 02115):



pET28a map (Novagen):

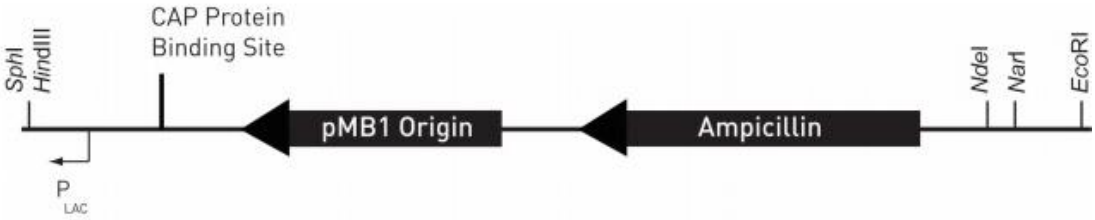
pET-28a(+) sequence landmarks	
T7 promoter	310-326
T7 transcription start	329
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites (*Nco*I - *Xho*I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	773-3852
pBR322 origin	3298-
Kan coding sequence	3995-4807
f1 origin	4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid; subtract 1bp from each site beyond *Nco*I at 158. pET-28c(+) is a 5362bp plasmid; subtract 2bp from each site beyond *Nco*I at 158.



pET-28a-c(+) cloning/expression region

pUC19L vector map (Invitrogen):

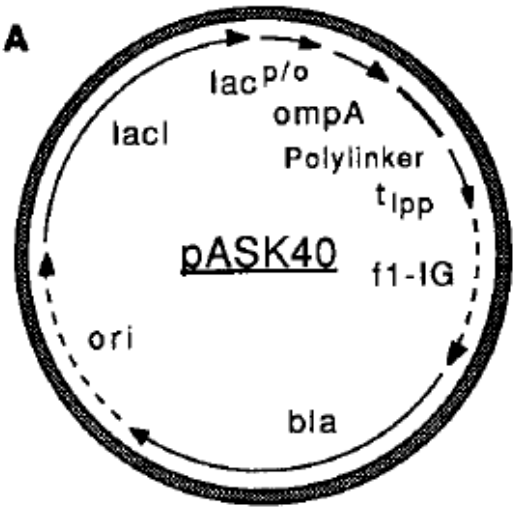


Comments for pUC19L:
2,659 nucleotides

P_{LAC} promoter: bases 75–104 (c)
CAP protein binding site: bases 124–136
pMB1 origin of replication: bases 428–1,016 (c)
Ampicillin (bla) resistance gene: bases 1,187–2,047 (c)

(c) = complementary strand

pASK40 vector map (Skerra et al., 1991):



B

TTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCA
-35 -10 → mRNA

XbaI
CACAGGAAACAGCTATGACCATGATTACGAATTTCTAGATAACGAGGGCAAAAAATGAAAAAGACAGCTATCGCG
LacZ: *MetThrMetIleThrAsnPheEnd* *OmpA:* *MetLysLysThrAlaIleAla*
-21

EcoRI KpnI BamHI XbaI
ATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAGGCCGCGAATTCGAGCTCGGTACCCGGGGATCCTCTA
IleAlaValAlaLeuAlaGlyPheAlaThrValAlaGlnAla...
-1

PstI HindIII
GAGTCGACCTGCAGGCATGCAAGCTTGACCTGTGAAGTGAAAAATGGCGCACATTGTGCGACATTTTTTTGTCT
* mRNA →

BamHI
GCCGTTTACCGCTACTGCGTCACGGATCC

pASK40-Bhx sequence:

ACCCGACACCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATT
CAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGAC
CGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGA
TGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTG
GCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCG
ATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGC
GGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGC
CATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCA
ACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACC
AGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATA
AATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTT
TTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGA
TGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTCGGTAGTG
GGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCC
TGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAG
CTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCG
CGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAA
CGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAA
TGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGGATTCACTG
GAACTCTAGTTAACGAGGGCAAAAAATGGATCCAGATCAATATTCAATAGAAGCTGATAAAAAATTTAA
GTATTCAGTAAAATTATCAGATTATCCAACATTGCAGGATGCAGCATCTGCTGCGGTTGATGGCCTTCTT
ATCGATCGAGATTATAATTTTTATGGTGGAGAGACAGTTGATTTTGGCGGAAAGGTTCTGACTATAGAA
TGTAAAGCTAAATTTATAGGAGATGGAAATCTTATTTTTACGAAATTAGGCAAAGGTTCCCGCATTGCCG
GGGTTTTTATGGAAAGCACTACAACACCATGGGTTATCAAGCCTTGGACGGATGACAATCAGTGGCTAA
CGGATGCCGCAGCGGTCGTTGCCACTTTAAAACAATCTAAAACTGATGGGTATCAGCCAACCGTAAGCG
ATTACGTTAAATTCCCAGGAATAGAAACGTTACTCCCACCTAATGCAAAAGGGCAAAACATAACGTCTAC
GTTAGAAATTAGAGAATGTATAGGGGTCGAAGTTCATCGGGCTAGCGGTCTAATGGCTGGTTTTTTGTT
TAGAGGGTGTCACTTCTGCAAGATGGTAGACGCCAATAATCCAAGCGGAGGTAAAGATGGCATTATAA
CCTTCGAAAACCTTAGCGGCGATTGGGGGAAGGGTAACTATGTCATTGGCGGACGAACCAGCTATGGG
TCAGTAAGTAGCGCCCAGTTTTTACGTAATAATGGTGGCTTTGAACGTGATGGTGGAGTTATTGGGTTTA
CTTCATATCGCGCTGGGGAGAGTGGCGTTAAAACTTGGCAAGGTACTGTGGGCTCGACAACCTCTCGCA
ACTATAATCTGCAATTCCGCGACTCGGTCGTTATTTACCCCGTATGGGACGGATTCGATTTAGGTGCTGA
CACTGACATGAATCCGGAGTTGGACAGGCCAGGGGACTACCCTATAACCCAATACCCACTGCATCAGTT
ACCCCTAAATCACCTGATTGATAATCTTCTGGTTCGCGGGGCGTTAGGTGTAGGTTTTGGTATGGATGGT
AAGGGCATGTATGTGTCTAATATTACCGTAGAAGATTGCGCTGGGTCTGGCGCGTACCTACTCACCCAC

GAATCAGTATTTACCAATATAGCCATAATTGACACCAATACTAAGGATTTCCAGGCGAATCAGATTTATA
TATCTGGGGCTTGCCGTGTGAACGGTTTACGTTTAATTGGGATCCGCTCAACCGATGGGCAGAGTCTAA
CCATAGACGCCCCTAACTCTACCGTAAGCGGTATAACCGGGATGGTAGACCCCTCTAGAATTAATGTTGC
TAATTTGGCAGAAGAAGGGCACCATCACCATCACCACTAGTAATTCGGCTATGATAGCGCAGCGATTAA
ACTGCGGATTCATAAGTTATCAAAGACATTAGATAGCGGAGCATTGTACTCCCACATTAACGGGGGGGC
CGGTTCTGGCTCAGCGTATACTCAACTTACTGCTATTTCAGGTAGCACACCTGACGCTGTATCATTAAAA
GTTAACCACAAAGATTGCAGGGGGGCAGAGATACCATTTGTTCCTGACATCGCGTCAGATGATTTTATA
AAGGATTCCTCATGTTTTTTGCCATATTGGGAAAATAATTCTACTTCTTTAAAGGCTTTAGTGAAAAAACC
CAATGGAGAATTAGTTAGATTAACCTTGGCAACACTTTAGTAAGCTTGACCTGTGAAGTGAAAAATGGC
GCACATTGTGCGACATTTTTTTTGTCTGCCGTTTACCGCTACTGCGTCACGGATCCCCACGCGCCCTGTAG
CGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAG
CGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATC
GGGGCATCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGA
TGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTT
AATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGG
GATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAAC
AAAATATTAACGTTTACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTAT
TTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGA
AAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCC
TGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGG
GTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAAT
GATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACT
CGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACG
GATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTA
CTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACT
CGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCC
TGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAA
TTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGG
TTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGAT
GGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGA
CAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATAC
TTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATG
ACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTT
CTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGT
TTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCA
AATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCT
CGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCA
AGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTT
GGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG

AAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGC
TTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCT
GGCCTTTTGCTGGCCTTTTGCTCACAT

Yellow: LacUV5 promoter
Pink: RBS (ribosome binding site)
Red: Beginning of reading frame, first Met residue
Green: His-tag
Grey: Stop codon

RGDD-Bhx gene:

ATGCGTGGTGACGATCCAGATCAATATTCAATAGAAGCTGATAAAAAATTTAAGTATTCAGTAAAATTAT
CAGATTATCCAACATTGCAGGATGCAGCATCTGCTGCGGTTGATGGCCTTCTTATCGATCGAGATTATAA
TTTTTATGGTGGAGAGACAGTTGATTTTGGCGGAAAGGTTCTGACTATAGAATGTAAAGCTAAATTTATA
GGAGATGGAAATCTTATTTTTACGAAATTAGGCAAAGGTTCCCGCATTGCCGGGGTTTTTATGGAAAGC
ACTACAACACCATGGGTTATCAAGCCTTGGACGGATGACAATCAGTGGCTAACGGATGCCGCAGCGGTC
GTTGCCACTTTAAAACAATCTAAAACTGATGGGTATCAGCCAACCGTAAGCGATTACGTTAAATTCCCAG
GAATAGAAACGTTACTCCCACCTAATGCAAAAGGGCAAAACATAACGTCTACGTTAGAAATTAGAGAAT
GTATAGGGGTCGAAGTTCATCGGGCTAGCGGTCTAATGGCTGGTTTTTTGTTTAGAGGGTGTCACTTCT
GCAAGATGGTAGACGCCAATAATCCAAGCGGAGGTAAAGATGGCATTATAACCTTCGAAAACCTTAGCG
GCGATTGGGGGAAGGGTAACTATGTCATTGGCGGACGAACCAGCTATGGGTCAGTAAGTAGCGCCCAG
TTTTTACGTAATAATGGTGGCTTTGAACGTGATGGTGGAGTTATTGGGTTTACTTCATATCGCGCTGGGG
AGAGTGGCGTTAAAACTTGGCAAGGTACTGTGGGCTCGACAACCTCTCGCAACTATAATCTGCAATTCC
GCGACTCGGTCGTTATTTACCCCGTATGGGACGGATTCGATTTAGGTGCTGACACTGACATGAATCCGG
AGTTGGACAGGCCAGGGGACTACCCTATAACCCAATACCCACTGCATCAGTTACCCCTAAATCACCTGAT
TGATAATCTTCTGGTTCGCGGGGCGTTAGGTGTAGGTTTTGGTATGGATGGTAAGGGCATGTATGTGTC
TAATATTACCGTAGAAGATTGCGCTGGGTCTGGCGCGTACCTACTCACCCACGAATCAGTATTTACCAAT
ATAGCCATAATTGACACCAATACTAAGGATTTCCAGGCGAATCAGATTTATATATCTGGGGCTTGCCGTG
TGAACGGTTTACGTTTAATTGGGATCCGCTCAACCGATGGGCAGAGTCTAACCATAGACGCCCCTAACTC
TACCGTAAGCGGTATAACCGGGATGGTAGACCCCTCTAGAATTAATGTTGCTAATTTGGCAGAAGAAGG
GCACCATCACCATCACCACTAG

Red: Beginning of reading frame, first Met residue
Blue: Sequence coding for RGD
Green: His-tag
Grey: Stop codon

RGDS-Bhx gene:

ATGCGTGGTGACAGTCCAGATCAATATTCAATAGAAGCTGATAAAAAATTTAAGTATTCAGTAAAATTAT
CAGATTATCCAACATTGCAGGATGCAGCATCTGCTGCGGTTGATGGCCTTCTTATCGATCGAGATTATAA
TTTTTATGGTGGAGAGACAGTTGATTTTGGCGGAAAGGTTCTGACTATAGAATGTAAAGCTAAATTTATA
GGAGATGGAAATCTTATTTTTACGAAATTAGGCAAAGGTTCCCGCATTGCCGGGGTTTTTATGGAAAGC
ACTACAACACCATGGGTTATCAAGCCTTGGACGGATGACAATCAGTGGCTAACGGATGCCGCAGCGGTC
GTTGCCACTTTAAAACAATCTAAAACTGATGGGTATCAGCCAACCGTAAGCGATTACGTTAAATTCCCAG
GAATAGAAACGTTACTCCCACCTAATGCAAAAGGGCAAAACATAACGTCTACGTTAGAAATTAGAGAAT
GTATAGGGGTCGAAGTTCATCGGGCTAGCGGTCTAATGGCTGGTTTTTTGTTTAGAGGGTGTCACTTCT
GCAAGATGGTAGACGCCAATAATCCAAGCGGAGGTAAAGATGGCATTATAACCTTCGAAAACCTTAGCG
GCGATTGGGGGAAGGGTAACTATGTCATTGGCGGACGAACCAGCTATGGGTCAGTAAGTAGCGCCCAG
TTTTTACGTAATAATGGTGGCTTTGAACGTGATGGTGGAGTTATTGGGTTTACTTCATATCGCGCTGGGG
AGAGTGGCGTTAAAACTTGGCAAGGTACTGTGGGCTCGACAACCTCTCGCAACTATAATCTGCAATTCC
GCGACTCGGTCGTTATTTACCCCGTATGGGACGGATTCGATTTAGGTGCTGACACTGACATGAATCCGG
AGTTGGACAGGCCAGGGGACTACCCTATAACCCAATACCCACTGCATCAGTTACCCCTAAATCACCTGAT
TGATAATCTTCTGGTTCGCGGGGCGTTAGGTGTAGGTTTTGGTATGGATGGTAAGGGCATGTATGTGTC
TAATATTACCGTAGAAGATTGCGCTGGGTCTGGCGCGTACCTACTCACCCACGAATCAGTATTTACCAAT
ATAGCCATAATTGACACCAATACTAAGGATTTCCAGGCGAATCAGATTTATATATCTGGGGCTTGCCGTG
TGAACGGTTTACGTTTAATTGGGATCCGCTCAACCGATGGGCAGAGTCTAACCATAGACGCCCCTAACTC
TACCGTAAGCGGTATAACCGGGATGGTAGACCCCTCTAGAATTAATGTTGCTAATTTGGCAGAAGAAGG
GCACCATCACCATCACCACTAG

Red: Beginning of reading frame, first Met residue
Blue: Sequence coding for RGD
Yellow: Serine codon
Green: His-tag
Grey: Stop codon

N-GFPem-Bhx gene:

ATGTCCGGCCTGAACGACATCTTCGAGGCTCAGAAAATCGAATGGCACGAAGGCGCGCCGGAGCTCGA
GGATCCGATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGAC
GGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGC
TGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCTTGAC
CTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATG
CCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG
GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACG
GCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAAGGTCTATATCACCGCCGACAAGC

AGAAGAACGGCATCAAGGTGAACTTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCC
GACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGC
ACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACC
GCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGAGCGGCCTGGTGCCGCGCGGCAGCATGGA
TCCAGATCAATATTCAATAGAAGCTGATAAAAAATTTAAGTATTCAGTAAAATTATCAGATTATCCAACAT
TGCAGGATGCAGCATCTGCTGCGGTTGATGGCCTTCTTATCGATCGAGATTATAATTTTTATGGTGGAGA
GACAGTTGATTTTGGCGGAAAGGTTCTGACTATAGAATGTAAAGCTAAATTTATAGGAGATGGAAATCT
TATTTTTACGAAATTAGGCAAAGGTTCCCGCATTGCCGGGGTTTTTATGGAAAGCACTACAACACCATGG
GTTATCAAGCCTTGGACGGATGACAATCAGTGGCTAACGGATGCCGCAGCGGTCGTTGCCACTTTAAAA
CAATCTAAAACTGATGGGTATCAGCCAACCGTAAGCGATTACGTTAAATTCCCAGGAATAGAAACGTTA
CTCCCACCTAATGCAAAAGGGCAAAACATAACGTCTACGTTAGAAATTAGAGAATGTATAGGGGTCGAA
GTTCATCGGGCTAGCGGTCTAATGGCTGGTTTTTTGTTTAGAGGGTGTCACTTCTGCAAGATGGTAGAC
GCCAATAATCCAAGCGGAGGTAAAGATGGCATTATAACCTTCGAAAACCTTAGCGGCGATTGGGGGAA
GGGTAACTATGTCATTGGCGGACGAACCAGCTATGGGTCAGTAAGTAGCGCCCAGTTTTTACGTAATAA
TGGTGGCTTTGAACGTGATGGTGGAGTTATTGGGTTTACTTCATATCGCGCTGGGGAGAGTGGCGTTAA
AACTTGGCAAGGTACTGTGGGCTCGACAACCTCTCGCAACTATAATCTGCAATTCCGCGACTCGGTCGTT
ATTTACCCCGTATGGGACGGATTCGATTTAGGTGCTGACACTGACATGAATCCGGAGTTGGACAGGCCA
GGGGACTACCCTATAACCCAATACCCACTGCATCAGTTACCCCTAAATCACCTGATTGATAATCTTCTGGT
TCGCGGGGCGTTAGGTGTAGGTTTTGGTATGGATGGTAAGGGCATGTATGTGTCTAATATTACCGTAGA
AGATTGCGCTGGGTCTGGCGCGTACCTACTCACCCACGAATCAGTATTTACCAATATAGCCATAATTGAC
ACCAATACTAAGGATTTCCAGGCGAATCAGATTTATATATCTGGGGCTTGCCGTGTGAACGGTTTACGTT
TAATTGGGATCCGCTCAACCGATGGGCAGAGTCTAACCATAGACGCCCCTAACTCTACCGTAAGCGGTA
TAACCGGGATGGTAGACCCCTCTAGAATTAATGTTGCTAATTTGGCAGAAGAAGGGCACCATCACCATC
ACCACTAG

Blue: Biotinylation sequence
Yellow: GFPem sequence
Red: Linker (thrombin site is underlined)
Cyan: Bhx sequence
Green: His-Tag
Grey: Stop codon

Bhx-GFPem-C gene:

ATGGATCCAGATCAATATTCAATAGAAGCTGATAAAAAATTTAAGTATTCAGTAAAATTATCAGATTATCCAACATT
GCAGGATGCAGCATCTGCTGCGGTTGATGGCCTTCTTATCGATCGAGATTATAATTTTTATGGTGGAGAGACAGTT
GATTTTGGCGGAAAGGTTCTGACTATAGAATGTAAAGCTAAATTTATAGGAGATGGAAATCTTATTTTTACGAAAT
TAGGCAAAGGTTCCCGCATTGCCGGGGTTTTTATGGAAAGCACTACAACACCATGGGTTATCAAGCCTTGGACGG
ATGACAATCAGTGGCTAACGGATGCCGCAGCGGTCGTTGCCACTTTAAAACAATCTAAAACTGATGGGTATCAGCC
AACCGTAAGCGATTACGTTAAATTCCCAGGAATAGAAACGTTACTCCCACCTAATGCAAAAGGGCAAAACATAAC

GTCTACGTTAGAAATTAGAGAATGTATAGGGGTCGAAGTTCATCGGGCTAGCGGTCTAATGGCTGGTTTTTTGTTT
AGAGGGTGTCACTTCTGCAAGATGGTAGACGCCAATAATCCAAGCGGAGGTAAAGATGGCATTATAACCTTCGAA
AACCTTAGCGGCGATTGGGGGAAGGGTAACTATGTCATTGGCGGACGAACCAGCTATGGGTCAGTAAGTAGCGC
CCAGTTTTTACGTAATAATGGTGGCTTTGAACGTGATGGTGGAGTTATTGGGTTTACTTCATATCGCGCTGGGGAG
AGTGGCGTTAAAACTTGGCAAGGTACTGTGGGCTCGACAACCTCTCGCAACTATAATCTGCAATTCCGCGACTCGG
TCGTTATTTACCCCGTATGGGACGGATTCGATTTAGGTGCTGACACTGACATGAATCCGGAGTTGGACAGGCCAG
GGGACTACCCTATAACCCAATACCCACTGCATCAGTTACCCCTAAATCACCTGATTGATAATCTTCTGGTTCGCGGG
GCGTTAGGTGTAGGTTTTGGTATGGATGGTAAGGGCATGTATGTGTCTAATATTACCGTAGAAGATTGCGCTGGG
TCTGGCGCGTACCTACTCACCCACGAATCAGTATTTACCAATATAGCCATAATTGACACCAATACTAAGGATTTCCA
GGCGAATCAGATTTATATATCTGGGGCTTGCCGTGTGAACGGTTTACGTTTAATTGGGATCCGCTCAACCGATGGG
CAGAGTCTAACCATAGACGCCCCTAACTCTACCGTAAGCGGTATAACCGGGATGGTAGACCCCTCTAGAATTAATG
TTGCTAATTTGGCAGAAGAAGGGAGCAGCGGCCTGGTGCCGCGCGGCAGCATGGTGAGCAAGGGCGAGGAGCT
GTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCG
AGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCT
GGCCCACCCTCGTGACCACCTTGACCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAG
ACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGA
GGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAAGGTCTATATCACCGCCGACAAGCA
GAAGAACGGCATCAAGGTGAACTTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACT
ACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCT
GAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGG
CATGGACGAGCTGTACAAGCATCACCATCACCATCACTAA

Cyan: Bhx sequence
Red: Linker (thrombin site is underlined)
Yellow: GFPem sequence
Green: His-Tag
Grey: Stop codon