ABSTRACT

Title of Thesis: OPTIMIZATION OF RECOMBINANT PROTEIN EXPRESSION FOR CELL-PENETRATING PEPTIDE FUSIONS TO PROTEIN CARGO

Sayanee Adhikari, Master of Science, 2017

Thesis directed by: Prof. Amy J. Karlsson
Department of Chemical and Biomolecular Engineering

Recombinant production of cell-penetrating peptides (CPPs) as fusions to protein "cargo" leads to low yields for some CPP-cargo fusions; thus, ways to enhance the recombinant expression of peptide-cargo fusions need to be identified. We optimized expression conditions for fusions of five CPPs (NPFSD, pVEC, SynB, histatin-5 and MPG) to the cargo proteins biotin carboxyl carrier protein (BCCP), maltose binding protein (MBP) and green fluorescent protein GFP. Glutathione-S-transferase was incorporated as a fusion partner to improve expression. In general, expression at 37 ^{o}C for 6 h and 10 h led

to the highest levels of expression for the different CPP-cargo constructs. The fusion of histatin-5 to GFP was purified, and its translocation into the fungal pathogen *Candida albicans* was studied. The purified protein translocated into the nearly 3% of *C. albicans* cells. These results provide the foundation for future studies to improve translocation of varied CPP-cargo fusions into *C. albicans* cells.

OPTIMIZATION OF RECOMBINANT PROTEIN EXPRESSION FOR CELL-PENETRATING PEPTIDE FUSIONS TO PROTEIN CARGO

by

Sayanee Adhikari

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2017

Advisory Committee:
Professor Amy J. Karlsson, Chair
Professor Jeffery B. Klauda
Professor Taylor J. Woehl

Dedication

This thesis is dedicated to my parents for their constant love, inspiration and guidance and my husband for his constant support. I thank God for the countless blessings enabling me to make it this far.

Acknowledgements

The past 2 years has been a rapid roller coaster ride. The change from being a petrochemical engineer to a biological field sure had its chain of challenges. My limited knowledge in the biological field needed me to put in a lot of hard work and sincerity to learn what each process and term meant. I would like to acknowledge Dr. Amy J. Karlsson for her relentless support and patience throughout the project and helping me deliver. Her continuous encouragement, mentoring and faith helped build my confidence to work independently and harder.

I would also like to my team members especially Zifan Gong for training me in the initial months and his continual insights into my project for which I am very thankful. Svetlana Ikonomova and Parisa Moghaddam-Taaheri are my go-to people in the lab, for every small doubt I have, they are always the first people I run too. Till date there hasn't been a moment when I have tired them off and here's hoping to more of those times!

I would also like to thank Turki Alahamadi for his help with experiments in the lab. I would like to thank the National Science Foundation (CBET Award #1511718) for making this work possible. I would like to thank Dr. Xiangbin Zeng for his help with the flow cytometry training.

The best thing though is having family and friends to depend on. My parents, my elder sister, my uncle and my best friend, who happens to be my husband, have forever been my greatest strength and they continue to be so. Their criticisms, their

support, their unending faith in me and their encouragement have all been crucial in everything I have ever done. Above all I would really like to thank God, for giving me these wonderful people and also this opportunity to make it thus far.

Table of Contents

List of Tables

List of Figures

Appendix A

List of Abbreviations

AMP	Antimicrobial peptides
BCCP	Biotin carboxyl carrier protein
CDC	Centers for Disease Control and Prevention
CFU	Colony-forming unit
CPP	Cell-penetrating peptide
FAM	Carboxyfluorescein
FPLC	Fast protein purification liquid chromatography
GFP	Green fluorescent protein
G_4S	Glycine-serine linker
GST	Glutathione S-transferase
Hst-5	Histatin 5
HRP	Horse-radish peroxidase
IEX	Ion exchange chromatography
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth
MBP	Maltose binding protein
NaCl	Sodium Chloride
NaPB	Sodium Phosphate Buffer

OD	Optical density
PBS	Phosphate-buffered saline
PI	Propidium iodide
SEC	Size exclusion chromatography
YPD	Yeast extract peptone dextrose medium

Chapter 1. Introduction

Fungal pathogens are a reason for concern as infections caused by them contribute to ill health and death in healthy and immunocompromised individuals suffering from AIDS, HIV, cancer and other illnesses [8,9]. One of the most prevalent opportunistic fungal pathogens is *Candida albicans*[8,9]. Infection due to *C. albicans* can be fatal, and growing drug resistance has contributed to ineffective treatment[8,9]. Thus, new therapeutic approaches are needed. In considering new treatments, one important feature is the ability of new antifungal molecules to be effectively and specifically delivered across cell membranes to intracellular targets.

Over the years, numerous approaches have been recognized to deliver therapeutic agents through cellular membranes. These include microinjection, electroporation, and liposome and viral-based vectors[2,5], but these methods have limitations like low efficiency, high toxicity, poor bioavailability and poor specificity[5]. Thus, an alternate approach is needed that could help solve these issues and help deliver non-cell permeable drugs to treat infections caused by *C. albicans*. A promising and novel approach for intracellular cargo delivery is the use of cell-penetrating peptides (CPPs). CPPs are a class of peptides that can cross the phospholipid bilayer of the cell membrane[3,5], and therefore, could be a potentially powerful option. CPPs like TAT and penetratin have been shown to effectively translocate through the cell membranes and enter cells[10-12]. CPPs are able to deliver molecular cargoes that may consist of DNA, RNA, proteins, and nanoparticles [3,5]into

cells. CPPs have shown translocation into *C. albicans* with molecular cargo, for example, NPFSD and Hst-5 can transport an N-terminal fluorescein isothiocyanate (FITC)[13,14] and NPFSD has also translocated GFP[13].

In this study, we focused on enabling the use CPPs to deliver bioactive cargoes into *Candida* cells by better understanding how to produce CPPs fused to cargo at high levels.

1.1 C. albicans

C. albicans is a commensal organism (lives in the body and causes no harm) and an opportunistic pathogen of humans that causes infections in the form of oral rashes and can lead to serious bloodstream infections as well as life-threatening disseminated and organ infections[9]. Pathogens such as *C. albicans* are of concern because of the growing occurrence of immunosuppression caused by AIDS, diabetes, cancer therapies, organ transplantation, and other conditions[9] .

Antifungal agents like amphotericin B and fluconazole help treat candidiasis[15] but they can have toxic side-effects or infections can be resistant to them. For example, amphotericin B has severe toxicity to the kidneys[16]. In terms of resistance, the Centers for Disease Control and Prevention (CDC) lists fluconazole-resistant *Candida* species as a "serious threat" to public health and estimated an annual healthcare cost of $6000-$9000 for every infection [17]. Antifungal resistance is a difficult issue mostly because the initial diagnosis of systemic fungal infection can be late and there are very few treatments available[16]. Various strains of fluconazole- and

amphotericin B-resistant *C. albicans* have been found in HIV-infected patients with prolonged treatment with azoles[18]. Because of the limitations of current antifungal agents and increasing resistance to these agents, new therapeutic approaches are needed to treat infections caused by *C. albicans.* This leads to our motivation to study cell-penetrating peptides (CPPs) that can act as potential drug delivery vehicles and help develop new treatments for candidiasis.

1.2 Cell penetrating peptides

CPPs are known for their ability to cross membranes and have been shown as potential drug delivery vehicle for intracellular delivery of a diverse selection of bioactive cargoes that includes plasmid DNA, small interfering RNAs (siRNAs), proteins and peptides[1,3] (Figure 1.1) inside cells for high levels of gene expression, gene silencing, or tumor targeting[1,5]. CPPs are short peptides, frequently positively charged or polar with several lysine or arginine residues in the sequence[3,5] (Table 1). The polar/charged residues are often alternately arranged with non-polar/hydrophobic residues leading to an amphipathic secondary structure[19]. These flexible peptides are simple to synthesize, functionalize, and characterize[5]. The design of clinically

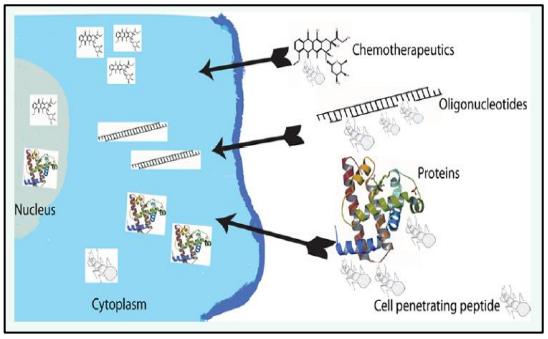

Figure 1.1 Schematic representation of CPPs entering the cell membrane (Taken from Copolovici *et al.*)[5]

operational delivery systems involves design of both the delivered cargo and the CPP used to deliver it[5].

There are many proposed models for cellular internalization of CPPs (Figure 1.2). Uptake of the peptide–protein complex on the plasma membrane can be through endocytic or non-endocytic routes. Endocytic routes include receptor-based endocytosis like clathrin-mediated endocytosis, or non-receptor-based endocytosis like macropinocytosis. Using clathrin-mediated endocytosis, cells absorb proteins by the inward folding and budding of the plasma membrane vesicles enclosing proteins with receptor sites which are specific to the molecules being absorbed[20,21].

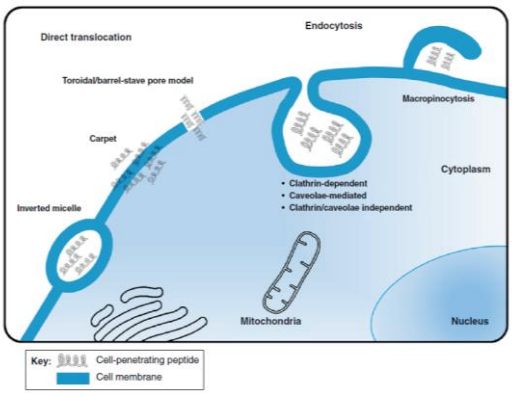

Figure 1.2 Schematic representation of proposed models and mechanisms for cellular internalization of CPPs (Adapted from Koren *et al.*)[3]

Electron microscopy and biochemical studies show that endocytosis occurs via the nucleation of a clathrin-coated pit, including cargo capture and multimerization of clathrin followed by propagation of a coated pit involving membrane invagination and then is followed by budding or scission of the clathrin-coated vesicle [20,21]. This is followed by completion of the clathrin cage and the action of dynamin, and inside the cell, the coat disassembles and the uncoated vesicle is transferred to the target[20,21]. Macropinocytosis is a clathrin-independent endocytic mechanism that is induced for a very short time due to cell-surface receptor activation by specific cargoes [22,23]. This process leads to formation of ruffles or protrusions and

can collapse back onto the cell, entrapping fluid into vesicles called macropinosomes[20-22]. Non-endocytic routes are described by the inverted micelle model, the carpet model and the pore formation model. The inverted micelle model requires the CPPs to have hydrophobic amino acids that are essential for the translocation process. Electrostatic interactions between the CPP and negatively charged membrane phospholipids enables insertion into the membrane, and interaction between hydrophobic CPP residues and the membrane core led to formation of vesicles that captures the CPP[24]. The carpet model includes binding of the peptide with negatively charged phospholipids, followed by rotation of the peptide that allows interaction between hydrophobic residues of the peptide and the membrane[5]. This is followed by a disruption in the packing of the membrane thus allowing internalization of the peptide. Another model is the pore formation or barrel-stave model that is due to formation of bundles by amphipathic α-helical peptides[5]. Pores form due to interaction between hydrophobic residues that face outward and hydrophilic surfaces that face inward, when the concentration of the peptides is higher than a specific threshold concentration that differs for each peptide.

The translocation of CPPs across a membrane depends on a number of factors. These factors include the specific sequence of the CPP, the concentration of CPP and the cell-type[24]. Additionally, the cargo that is fused to the CPP and the design of the CPP-cargo fusion may influence the method of translocation [24,25] For example, TAT has been suggested to use a caveolae-mediated endocytosis with protein cargo, but uses a clathrin-mediated delivery when attached to small molecules[21,25]. Caveolae are invaginated, flask-shaped plasma membrane domains, enriched in cholesterol and

sphingolipids and can internalize large molecular complexes[21]. At high CPP concentrations of >10 μM direct penetration seems to be the predominant mechanism[5]. At low concentrations of CPPs, though, endocytosis is the main mechanism of uptake[5] (Table 2).

CPPs have been used to deliver protein cargoes with various sizes ranging from 25 kDa to 150 kDa[26-28]. β-Galactosidase (120 kDa) has been delivered into mouse tissues, even in the brain, while maintaining its biological activity[29]. CPPs like TAT or penetratin have yielded substantial tissue localization (in mouse and rats) *in vivo* with antibody fragments as the cargoes[30]. Given their utility in delivering a range of cargo, further study of methods to improve their potential as therapeutics against *C. albicans* are warranted.

Table 1. CPPs, their molecular weight, lengths and charge

Peptide	Sequence	Length (a.a.)	MW (Da)	Charge*	Ref.
SynB	RGGRLSYSRRRFSTSTGR	18	2100.34	+6	[31]
MPG	GALFLGFLGAAGSTMGAWSQPKKKRKV	27	2807.36	+5	[31]
NPFSD	VLTNENPFSDP	11	1232.31	-2	[31]
$(KFF)_3K$	KFFKFFKFFK	10	1413.77	+4	[31]
Penetratin	RQIKIWFQNRRMKWKK	16	2246.75	+7	[31]
Pep-1	KETWWETWWTEWSQPKKKRKV	21	2848.26	+2	[31]
hCT	LGTYTQDFNKTFPQTAIGVGAP	22	2326.59	0	[31]
PAF26	RKKWFW	6	950.15	+3	[31]

TP-10	AGYLLGKINLKALAA LAKKIL	21	2182.77	+4	[31]
CecropinB	KWKVFKKIEKMGRNI RNGIVKAGPAIAVLGE AKAL	35	3835.71	+12	
Histatin 5	DSHAKRHHGYKRKFH EKHHSHRGY	24	3036.33	+12	[31]
pVEC	LLIILRRRIRKQAHAHS K	18	2209.7	+8	[2,13 ,32]

***Includes only charges due to amino acid side chains (pH 7)**

Table 2. Examples of CPPs, their structures and proposed mechanisms

CPP	Structure	Proposed mechanism	References
TAT	Random coil/ helix	Direct penetration, pore formation	[3]
Penetratin (pAntp)	Amphipathic, α-helical/β-sheet (higher concentration)	Direct penetration, endocytosis	[3]
Polyarginines	Random coil, α-helical	Direct penetration, endocytosis	[3]
pVEC	Amphipathic, β-sheet	Direct penetration, transporter-mediated	[3]
Pep-1	Amphipathic, α-helical	Direct penetration, pore formation	[3]
Transportan	Amphipathic, α-helical	Endocytosis, direct penetration	[3]
MAP	Amphipathic, α-helical	Multiple mechanisms	[3]
SynB	β-sheet, α-helical	Endocytosis	[33]
MPG	Amphipathic	Endocytosis	[33]
NPFSD	No data	Endocytosis	[13,34,35]

1.3 Challenges in production of CPP fusions

To enable delivery of protein cargo by CPPs, a method to link the protein and cargo is needed. Two main approaches exist: (1) chemical coupling of purified peptides to protein cargo and (2) recombinant production of a genetic fusion of a CPP to the cargo protein. Chemical coupling of CPPs involves using a chemical reaction

to attach protein cargo to a peptide that has been synthesized and purified. There are various advantages of using the chemical coupling method. Peptide toxicity to cells is not an issue, and any cargo (commercial or recombinant) can be attached. Another advantage of this method is the availability of different methods of conjugation to cargo. Each protein and peptide comprises of various amino acids with different sequences and thus availability of the reactive groups is abundant. Most important amino acids that can be used for modification and conjugation purposes, as they contain ionizable side chains are aspartic acid, glutamic acid, lysine, arginine, cysteine, histidine, and tyrosine. Derivatization of the side chain sulfhydryl of cysteine is a particularly important reaction for modification and conjugation of proteins[36].

Although chemical coupling has some advantages, it also has important limitations. The peptide synthesis itself is one limitation, due to the high cost, the time lag from ordering the peptide commercially to receive it (3 weeks or more). The commercially synthesized peptides may lack a carboxyl group thus making it difficult to chemically conjugate it to a protein cargo. Additionally, although the number of available reactive groups in proteins and peptides is an advantage in terms of versatility, it also makes site-specific conjugation of peptides to protein cargo challenging.

Rather than using chemistry approaches to couple CPPs and cargo, recombinant protein production techniques can be used. Advantages of such a technique are many. They are less costly, allow easy control of cargo fusion to N- or C-terminus of the peptide, facilitate easy and quick design changes of the fusion, and

are already well established for production in *E. coli.* Recombinant protein expression of CPPs as fusions to protein cargo is a promising approach, but it still can present various challenges. One challenge is that of no or low expression, which could be due to toxicity of the peptides and proteins before or after induction could be solved by addition of glucose, controlling the level of induction and codon bias can be solved by optimization of codon frequency[37]. Optimization of temperature and induction time could also help in overcoming challenges in fusion production[37 38]. Another potential challenge is that of inclusion body formation which could be due to incorrect disulfide bond formation, incorrect folding or low solubility of the CPP-cargo fusion protein. Using *E. coli* strains with oxidative cytoplasmic environment, co-expressing the protein along with molecular chaperones or using a fusion to a solubility enhancing protein could help alleviate the problems[37]. Protein inactivity due to incomplete folding or mutations in the cDNA could be other possible issues during recombinant protein production[37]. The fusion may not be sufficiently soluble in a biological buffer, or the fusion may not successfully enter cells[39].

1.4 Overview of thesis

This thesis describes the work to produce CPP-cargo fusions and optimize their expression. Various strategies to enhance production need to be applied, and we believe the combination of right peptide and the right cargo size, linkers and the perfect soluble partner should enable us to achieve targeted drug delivery. In Chapter 2, I describe the detailed design for all my fusions of CPPs to the cargoes used in this

study. I describe in detail the characteristic features of all the CPPs (NPFSD, pVEC, SynB, histatin-5 and MPG) and cargoes (BCCP, GFP and BCCP) used. In Chapter 3, I describe the detailed experimental methods and the materials used for expression and translocation studies. In Chapter 4, I present my results of the experiments evaluating various expression conditions and fusion designs. In Chapter 5, I describe the purification of a CPP-cargo fusion for preliminary translocation experiments into *C. albicans*. Finally, my conclusions and recommendations for future work are summarized in Chapter 6.

Chapter 2. Design of constructs

To study the expression of CPP fusions to cargo, we designed a modular system that includes a CPP genetically fused to a protein cargo (Figure 2.1). A modular design ensures that the CPPs and the cargoes can be easily exchanged. A glycine-serine linker (G_4S) was included to improve expression, since it was previously found to be effective in enhancing soluble protein expression[13,40]. This flexible linker allows interaction between domains or can increase the separation between domains thus allowing flexibility[41]. Glutathione-S-transferase (GST) was included as a fusion partner at the N-terminus or C-terminus in the constructs to evaluate its effect on expression. GST has been used as a fusion partner since it enhances expression[37]. GST can also act as an affinity tag for affinity purification of proteins, increasing its utility in recombinant expression[42,43].

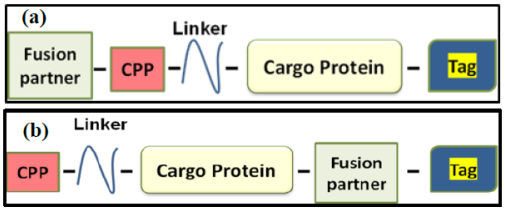

Figure 2.1. Fusion constructs designed for this work with the fusion partner at the **(a)** N-terminus and **(b)** C-terminus.

2.1 Peptides

We selected several CPPs to study in our system: NPFSD, pVEC, SynB, histatin-5 and MPG. These peptides were selected because they had been previously shown to be translocated into *C. albicans*[13,31] and thus they look promising for future

studies of translocation of protein cargo. They have varied mechanisms and properties (Table 1 & 2 in Chapter 1).

2.1.1 SynB

SynB (RGGRLSYSRRRFSTSTGR) is a CPP obtained from the antimicrobial peptide (AMP) protegrin 1 (PG-1), an 18 amino acid peptide initially isolated from porcine leukocytes[33]. Numerous linear analogues of PG-1 without cysteine residues were designed to remove cyclization and prevent pore formation. The linear SynB peptide interacts with the cell surface and crosses the plasma membrane but does not disrupt the membrane. It was suggested that SynB adopts an energy-dependent endocytic mechanisms to translocate into cells[33,44].

2.1.2 Histatin 5

Hst-5 which is found in human saliva, is an important component of the human immune system due to its candidacidal and to some extent bactericidal effects[45-48]. Hst-5 is effective in killing both the yeast and hyphal forms of *C. albicans*[46]. The killing process involves binding to the cell membrane of *C. albicans*, translocating into the cytoplasm and then targeting the mitochondria to lead to both membrane damage and cell death[47,49,50]. Unlike many other AMPs, membrane disruption alone is not responsible for the fungicidal activity of Hst-5 and it must cross the cell membrane to exert its activity[14,45-50]. Thus, although it is best known for being an AMP, it is also a CPP.

2.1.3 MPG

Figure 2.2 Sequence alignment and secondary structure of MPG; residues involved in cargo binding and cellular uptake are in red and yellow respectively. (Adapted from Morris *et al.*)[7].
MPG sequence is acetylated at the N-terminus and a cysteamide group at the C-terminus; these are necessary for stability of the peptide and its transduction mechanism.

MPG is a short amphipathic peptide[7] (Figure 2.2). It consists of three domains: an N-terminal hydrophobic motif; a hydrophilic lysine-rich domain derived from the nuclear localization sequence (NLS) of the SV40 (simian virus 40) large T-antigen (KKKRKV) that is necessary for interaction with nucleic acids and cellular uptake, intracellular trafficking of cargo and solubility of the peptide vector; a linker domain (WSQP) between the two other domains, which contains a proline residue to improve flexibility and the integrity of both the hydrophobic and hydrophilic domains[7]. MPG can form stable nanoparticles with protein cargoes without any cross-linking or chemical modifications[7], though we will use it covalently linked to cargo for consistency with our other peptides. MPG (either bound to or free of cargo), has been shown to strongly interact with membrane lipids and spontaneously enters the lipid-phase and inserts into natural membranes[7].

2.1.4 NPFSD

The peptide NPFSD was originally found as an endocytosis signal in *Saccharomyces cerevisiae*[34,35], and this signal sequence can transport both large and

small molecules into *S. cerevisiae* and *C. albicans*[13,34,35]. NPFSD uptake is a clathrin-dependent endocytosis process[35].

2.1.5 pVEC

pVEC is a peptide that originates from a murine vascular endothelium (VE)-cadherin. pVEC contains a sequence of 18 amino acids, out of which 13 amino acids are from the cytosolic domain of VE-cadherin closest to the membrane and 5 amino acids are from the C-terminus of the transmembrane region of VE-cadherin[2](Figure 2.3).

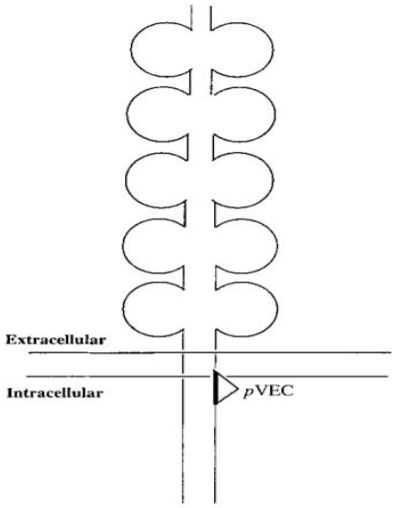

Figure 2.3 Schematic representation of the origin of pVEC from amino acids that traverse the transmembrane part to the cytosolic tail of the murine VE-cadherin (Taken from Elmquist *et al*)[2].

The presence of four arginine residues and two lysine residues gives pVEC its positive charge. The N-terminus of the pVEC sequence is hydrophobic, with a charged middle part and a hydrophilic C-terminus. Clathrin-dependent endocytosis and macropinocytosis have been proposed as mechanisms of translocation for pVEC[2,31,32].

2.2 *Protein cargoes*

Three cargoes were included in our constructs to provide a variety of sizes: biotin carboxyl carrier protein (BCCP, 10 kDa), maltose-binding protein (MBP, 42 kDa) and green fluorescent protein (GFP, 27 kDa) (Table 3). This will ultimately offer the opportunity to explore whether there is a size limitation in translocation. All these cargoes have been shown to be expressed recombinantly in *E. coli*[51-54].

Table 3. Protein cargoes

Cargo	Sequence	Isoelectric point	Length (a.a.)	Mass (kDa)	Ref.
BCCP	MEAPAAAEISGHIVRSPMVGTF YRTPSPDAKAFIEVGQKVNVGD TLCIVEAMKMMNQIEADKSGT VKAILVESGQPVEFDEPLVVIE	4.5	156	10.0	[52,55,56]
GFP	SKGEELFTGVVPILVELDGDVN GHKFSVSGEGEGDATYGKLTL KFICTTGKLPVPWPTLVTTFAY GLQCFARYPDHMKQHDFFKSA MPEGYVQERTIFFKDDGNYKT RAEVKFEGDTLVNRIELKGIDF KEDGNILGHKLEYNYNSHNVYI MADKQKNGIKVNFKIRHNIEDG SVQLADHYQQNTPIGDGPVLLP DNHYLSTQSALSKDPNEKRDH MVLLEFVTAAGITHGMDELYK	5.80	238	26.9	[51,57,58]
MBP	MKIKTGARILALSALTTMMFSA SALAKIEEGKLVIWINGDKGYN GLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFW AHDRFGGYAQSGLLAEITPDKA FQDKLYPFTWDAVRYNGKLIA YPIAVEALSLIYNKDLLPNPPKT WEEIPALDKELKAKGKSALMF NLQEPYFTWPLIAADGGYAFK YENGKYDIKDVGVDNAGAKA GLTFLVDLIKNKHMNADTDYSI AEAAFNKGETAMTINGPWAWS NIDTSKVNYGVTVLPTFKGQPS	4.9	387	42.0	[53,59-62]

KPFVGVLSAGINAASPNKELAK
EFLENYLLTDEGLEAVNKDKPL
GAVALKSYEEELVKDPRIAAT
MENAQKGEIMPNIPQMSAFWY
AVRTAVINAASGRQTVDEALK
DAQT

2.2.1 Biotin carboxyl carrier protein

E. coli acetyl-CoA carboxylase (ACC) catalyzes the first step of fatty acid synthesis, to form malonyl-CoA[52]. For these reactions, ACC needs two protein subassemblies, biotin carboxylase (BC) and carboxyl-transferase (CT). A third protein, BCCP (Figure 2.4), is known to carry the biotin cofactor covalently bound to the lysine residue. All three ACC proteins in *E. coli* form a complex. Previous studies demonstrate that biotinylation stabilizes *E. coli* BCCP. BCCP is inherently produced by *E. coli* and is also a small-sized protein, thus making it a suitable choice for our expression studies.

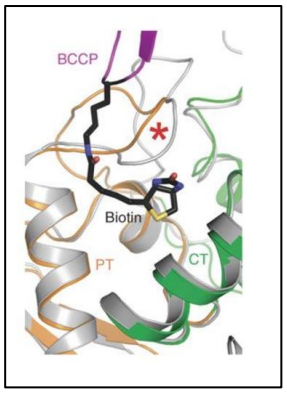

FIGURE 2.4 Biotin carboxyl carrier protein (pyruvate carboxylase (PC) tetramerization (PT) and carboxyltransferase (CT)) (Adapted from Xiang *et al.*)[6]

2.2.2 Green fluorescent protein

FIGURE 2.5 Green fluorescent protein from Aequorea victoria *(*Adapted from Ormo et al., 1996)[4] (PDB ID: 1EMA)

GFP is a protein with 238 amino acid residues (26.9 kDa; Figure 2.5). It absorbs blue light at 395 nm with a minor peak at 470 nm and exhibits bright green fluorescence at 509 nm[4,57]. The green fluorescence is stable and almost no photo bleaching is seen[57]. This protein has been selected due to its fluorescence property, which will help us to study the translocation into cells when it is fused to an appropriate CPP. At the same time, the choice is fitting, since it is bigger in size compared to BCCP but smaller than MBP.

2.2.3 Maltose-binding protein

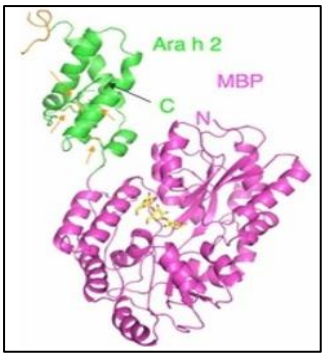

FIGURE 2.6 Maltose binding protein (Mueller *et al.*)[1]

E. coli MBP is a 370 amino acid protein (42 kDa) with an ellipsoidal structure that is divided into distinct globular structures separated by a deep groove (Figure 2.6). Each domain is built from amino- and carboxyl-terminal halves, and both domains show similar supersecondary structure, consisting of a central α-pleated sheet flanked on both sides with parallel α-helices[59]. E. coli MBP is soluble in the E. coli cytoplasm, and its binding to amylose resin can be used to aid in its purification [53,54]. MBP does not have any cysteine residues so it cannot affect the formation of disulfide bonds with the CPPs in our fusion constructs[54]. MBP has been chosen as a model cargo since it is a soluble partner and is a large sized protein, thus allowing us to study the size-limitation in translocation studies.

Chapter 3. Methods and materials

3.1 Plasmid construction

We used recombinant cloning techniques to produce CPP-cargo fusion complexes. Plasmids containing MBP, GFP and BCCP and their genetic fusions to NPFSD, pVEC, SynB, MPG and Hst-5 were constructed based on the pET21a and pET-42a vectors (Novagen). Plasmids used in this work are described in Table 4.

3.1.1 pNGST and pCGST vectors

The pCGST and pNGST vectors differ mainly by the position of their soluble fusion partner, GST. The pCGST design has a C-terminal GST whereas pNGST has an N-terminal GST. This change in design is suggested because it is believed that with the soluble partner at the N-terminus, the construct will fold properly during expression.

The pCGST plasmids were designed based on the pET-21(a) vector with the N-terminal T7 tag (Figure 3.1). The pET-42a vector has glutathione-S-transferase (GST) attached as a fusion partner. GST was PCR amplified using NotI-Ala-KpnI-GST-F and GST-XhoI-R, and inserted into pET-21 digested by NotI and XhoI, resulting in pGST. FLAG tag and TEV cut-sites were created by oligonucleotides dimer insertion with NotI-FLAG-HindIII-F/NotI-FLAG-HindIII-R, and HindIII-TEV-KpnI-F/HindIII-TEV-KpnI-R. Two primer dimers were inserted into pGST digested by NotI and Kpn I to form pCGST plasmid (Figure 3.1).

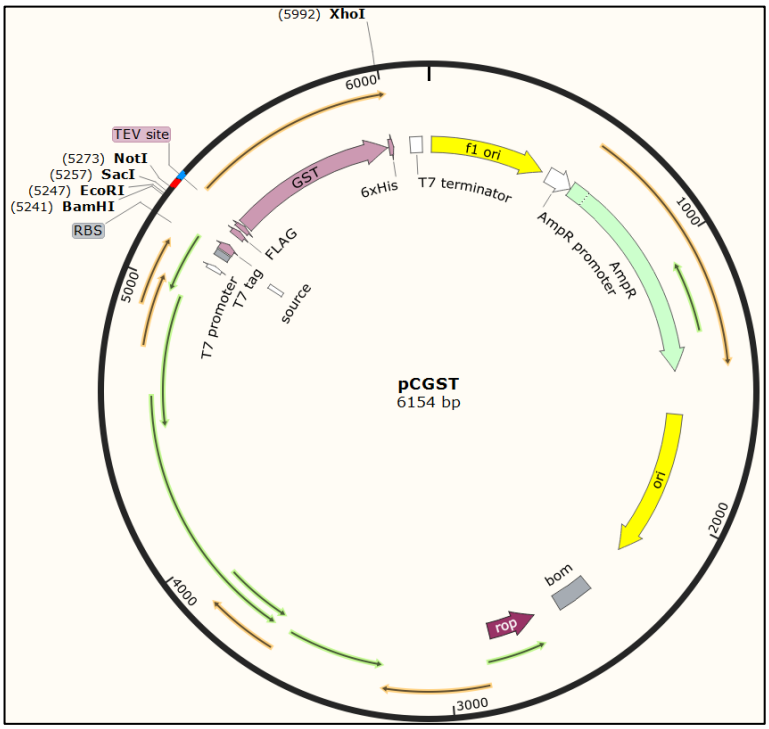

Figure 3.1 Plasmid map of pCGST. This plasmid has been designed with appropriate cut sites. pCGST21 was designed based on pET-21(a) vector with the N-terminal T7 tag. The plasmid map shows the cut sits for the insertion of the peptides (EcorI and SacI) and the cargoes (SacI and NotI); FLAG, TEV and 6XHis sites are also shown along with a C-terminal GST.

The pNGST plasmid was constructed with pET-21(a) as the backbone and GST from the pET-42a vectors on the N-terminus (Figure 3.2). The GST fusion partner was first amplified using the primers BamHI-GST-F and GST-EcoRI-R and inserted into pET-21(a) with BamHI and EcoRI digestion. To allow removal of the GST, a Factor Xa cleavage site (ATTGAGGGACGC) was also included.

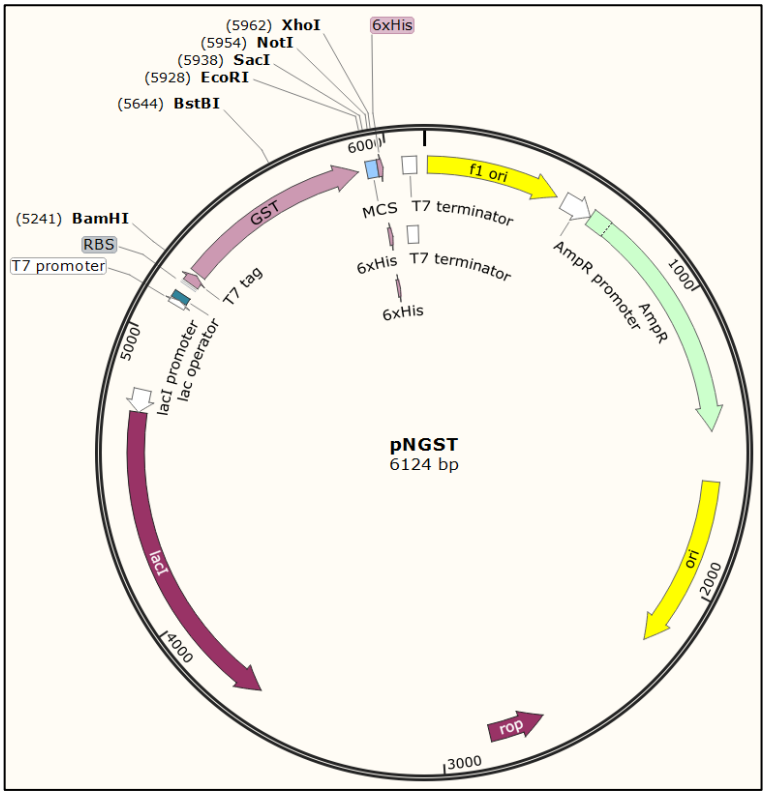

Figure 3.2 Plasmid map of pNGST. The pNGST plasmid was constructed with pET-21(a) as the backbone and GST from the pET-42a vectors on the N-terminus (Figure 3.2). 6XHis site is also shown with an N-terminal GST followed by the cut sites for the insertion of the peptides (EcorI and SacI) and the cargoes (SacI and NotI).

3.1.2 Construction of fusion constructs

Plasmids containing various combinations of peptides (or no peptide) and cargo proteins (Table 4) were constructed in pCGST (NPFSD and pVEC) and pNGST (SynB, Hst-5 and MPG).

Plasmids containing BCCP and genetic fusions of NPFSD to BCCP were created using the pCGST plasmid for expression with a C-terminal GST tag. All primers are listed in Table 3. DNA encoding BCCP with an N-terminal NPFSD and pVEC peptide were amplified using a PCR reaction with a forward primer that

contained the peptide sequence followed by a C-terminal G4S peptide linker sequence or followed by no linker sequence (Table 3) and with the reverse primer BCCP-NotI-R. These PCR products were then introduced between the EcoRI and NotI sites of the pCGST plasmid forming the plasmids pCGST-CPP-G4S-BCCP and pCGST-CPP-BCCP, where the CPP was either NPFSD or pVEC).

Peptides were inserted into the pNGST plasmid for expression with an N-terminal GST tag using pairs of annealed oligonucleotides. The pairs of oligonucleotides also contained the coding sequence for proteolytic cleavage by Factor Xa. The oligonucleotide pairs were synthesized commercially to result in EcoRI and SacI sticky ends upon annealing to form dimers (SynB: Xa-SynB-top-1 and Xa-SynB-bottom-1; Xa-SynB-top-2 and Xa-SynB-bottom-2; Xa-SynB-G_4S-top and Xa-SynB-G_4S-bottom) (Table 3). For example, the pairs EcoRI-FactorXa-MPG-Top1 and EcoRI-FactorXa-MPG-Bottom1, and the pairs SacI-G_4S-MPG-Top2 and SacI-G_4S-MPG-Bottom2 were annealed to form the MPG dimers. These dimers were inserted between the EcoRI and SacI sites of pNGST.

The DNA encoding the cargoes was PCR-amplified from template plasmids in our lab stocks using the primers SacI-[Cargo]-F and [Cargo]-NotI-R (Table 3), where [Cargo] represents MBP, GFP and BCCP. The resulting products were inserted between the SacI and NotI sites of pCGST and pNGST.

Table 4. Oligonucleotide sequences

Oligonucleotide sequence name	Sequence (5' to 3')	References
SacI-ATG-GFP-F	GCGATGGAGCTCAGTAAAGGAGAAGAACTTTTC	This study
EcoRI-NPFSD-SacI-GFP-F	TTCGGAATTCAATGGTGCTGACCAACGAAAACCCGTTTTCTGATCCGGAGCTCAGTAAAGGAGAAGAACTTTTC	[13]
EcoRI-NPFSD-G4S-SacI-GFP-F	TTCGGAATTCAATGGTGCTGACCAACGAAAACCCGTTTTCTGATCCGGGAGGCGGTGGAAGCGAGCTCGAGCTCAGTAAAGGAGAAGAACTTTTC	[13]
GFP-NotI-R	AATAAAGCGGCCGCTTTGTATAGTTCATCCATGC	[13]
sacI-MBP-fwd	ATTGGGAGCTCAAAATCGAAGAAGGTAAACTGG	This study
NotI-MBP-Rev	ATTGGGCGGCCGCCAGTCTGCGCGTCTTT	This study
SacI-BCCP-F	GCGATGGAGCTCATGGAAGCGCCAGCAGCAG	This study
BCCP-NotI-R	AATAAAGCGGCCGCCTCGATGACGACCAGCGGCTC	This study
EcorI-FactorXa-MPG-Top1	AATTCATTGAGGGACGCGGCGCACTTTTCTTAGGGTTCCTTGGAGCCGCCGGGAG	This study
EcorI-FactorXa-MPG-Bottom1	CGGCGGCTCCAAGGAACCCTAAGAAAAGTGCGCCGCGTCCCTCAATG	This study
SacI-G4S-MPG-Top2	GATGGGTGCCTGGTCCCAGCCAAAGAAGAAACGTAAAGTAGGAGGCGGTGGAAGCGAGCT	This study
SacI-G4S-MPG-Bottom2	CGCTTCCACCGCCTCCTACTTTACGTTTCTTCTTTGGCTGGGACCAGG	This study
SacI-G4S-Hst-5-Bottom2	CGCTTCCACCGCCTCCGTACCCGCGATGACTGTGGTGCTTTTCGTG	This study
SacI-G4S-Hst-5-Top2	CAAGTTCCACGAAAAGCACCACAGTCATCGCGGGTACGGAGGCGGTGGAAGCGAGCT	This study
EcorI-FactorXa-Hst-5-Bottom1	GAACTTGCGCTTATACCCGTGATGGCGTTTTGCATGAGAATCGCGTCCCTCAATG	This study
EcorI-FactorXa-	AATTCATTGAGGGACGCGATTCTCATGCA	This study

Hst-5-Top1	AAACGCCATCACGGGTATAAGCG	
Xa-SynB-top-1	AATTCATTGAGGGACGCCGCGGGGGGCGGCTTAGCTATTCAAGACGGC	This study
Xa-SynB-bottom-1	GTCTTGAATAGCTAAGCCGCCCCCCGCGGCGTCCCTCAATG	This study
Xa-SynB-top-2	GGTTCTCCACTTCCACGGGCCGTGAGCT	This study
Xa-SynB-bottom-2	CACGGCCCGTGGAAGTGGAGAACCGCC	This study
Xa-SynB-G4S-top	GGTTCTCCACTTCCACGGGCCGTGGAGGCGGTGGAAGCGAGCT	This study
Xa-SynB-G4S-bottom	CGCTTCCACCGCCTCCACGGCCCGTGGAAGTGGAGAACCGCC	This study
EcoRI-SynB-F	AATAAGAATTCATGCGCGGGGGGCGGCTTAGCTATTCAAG	This study
pVEC-1-top	AATTCAATGCTGTTGATCATCCTGCGCCGC	[13]
pVEC-1-bottom	TGCGGCGGCGCAGGATGATCAACAGCATTG	[13]
pVEC-2-top	CGCATCCGTAAGCAGGCCCACGCGCATAGTAAAGAGCT	[13]
pVEC-2-bottom	CTTTACTATGCGCGTGGGCCTGCTTACGGA	[13]
pVEC-G_4S-2-top	CGCATCCGTAAGCAGGCCCACGCGCATAGTAAAGGAGGCGGTGGAAGCGAGCT	[13]
pVEC-G4S-2-bottom	CGCTTCCACCGCCTCCTTTACTATGCGCGTGGGCCTGCTTACGGA	[13]
BamHI-GST-F	GCGATGGGATCCATGTCCCCTATACTAGGTTATTG	This study
GST-EcoRI-R	AATAAAGAATTCACCAGAACCACTAGTTGAAC	This study
NotI-FLAG-HindIII-F	GGCCGCAGATTACAAGGATGACGACGATAAGA	This study
NotI-FLAG-HindIII-R	AGCTTCTTATCGTCGTCATCCTTGTAATCTGC	This study
HindIII-TEV-KpnI-F	AGCTTGAGAACCTGTACTTCCAGGGCGGTAC	This study
HindIII-TEV-KpnI-R	CGCCCTGGAAGTACAGGTTCTCA	This study
NotI-Ala-KpnI-GST-F	ATTAGCGGCCGCAGCTGGTACCATGTCCCCTATACTAGGTTATTGG	This study
GST-XhoI-R	ATTTCTCGAGTTTTGGAGGATGGTCGCCAC	This study
NPFSD-top-1	AGCTTGTGCTGACCAACGAAAAC	This study

NPFSD-Bottom-1	ACGGGTTTTCGTTGGTCAGCACA	This study
NPFSD-top-2	CCGTTTTCTGATCCGTGATGAC	This study
NPFSD-Bottom-2	TCGAGTCATCACGGATCAGAAA	This study
NPFSD-linker-top-1	AGCTTGGAGGCGGTGGAAGCGTGCTGAC CAACGAAAAC	This study
NPFSD-linker-bottom-1	ACGGGTTTTCGTTGGTCAGCACGCTTCCA CCGCCTCCA	This study

Table 5. Plasmids used in this study

Plasmid Name	Backbone	CPP	Linker	Cargo
pCGST-NPFSD-BCCP	pCGST	NPFSD	No	BCCP
pCGST-NPFSD-G_4S-BCCP	pCGST	NPFSD	Yes	BCCP
pCGST-pVEC-BCCP	pCGST	pVEC	No	BCCP
pCGST-pVEC-G_4S-BCCP	pCGST	pVEC	Yes	BCCP
pNGST-SynB-MBP	pNGST	SynB	No	MBP
pNGST-SynB-G_4S-MBP	pNGST	SynB	Yes	MBP
pNGST-SynB-G_4S-GFP	pNGST	SynB	Yes	GFP
pNGST-SynB-G_4S-BCCP	pNGST	SynB	Yes	BCCP
pNGST-Hst-5-G_4S-MBP	pNGST	Hst-5	Yes	MBP
pNGST-Hst-5-G_4S-GFP	pNGST	Hst-5	Yes	GFP
pNGST-Hst-5-G_4S-BCCP	pNGST	Hst-5	Yes	BCCP
pNGST-MPG -G_4S-MBP	pNGST	MPG	Yes	MBP
pNGST-MPG -G_4S-GFP	pNGST	MPG	Yes	GFP
pNGST-MPG-G_4S-BCCP	pNGST	MPG	Yes	BCCP
pNGST-MBP	pNGST	-	-	MBP
pNGST-GFP	pNGST	-	-	GFP
pET21-GFP	pET21	-	-	GFP[13]

3.2 Bacterial strains and culture conditions

All plasmid construction was done using *E. coli* DH5α (Novagen). Following successful plasmid construction, plasmids were transformed into *E. coli* BL21 (DE3) (Novagen) for all expression studies. BL21 (DE3) is the most extensively used *E. coli* strain for high-yield expression of recombinant proteins and uses the promoter selectivity and transcriptional action of the T7 RNA polymerase[63].

To compare expression between fusion constructs with different cargoes and CPPs, 25 mL of overnight culture for each construct was sub-cultured into 400 mL of fresh Luria Bertani (LB) broth (10 mg/mL tryptone**,** 5 mg/mL yeast extract and 5 mg/mL NaCl) at an optical density of OD_{600} = 0.05 and grown at 37 °C for 2 h. After incubation, the culture was aliquoted into three 100 mL cultures. Expression of the fusion proteins was then induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG, Fisher BioReagents) to each culture aliquot at a final concentration of 0.01 mM. The three separate cultures were then incubated at 37 °C, 30 °C or 20 °C, and samples were aliquoted from each flask at 6 h, 10 h and 24 h. For the BCCP constructs, biotin was added at 5 mM along with IPTG to allow detection of BCCP via the biotin modification.

3.3 Protein expression level examined by Western blot

The expression level of the peptide-cargo fusions at different conditions was compared using Western blotting. Following induction of the protein fusions, the cells in each culture were pelleted by centrifugation at 4300×g for 15 min at 4 °C and lysed with BugBuster Master Mix (EMD Millipore) according to the manufacturer's protocol. Proteins in the soluble cell lysates were recovered by centrifugation and

were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) on 7.5% Mini-PROTEAN® TGX™ gels for MBP fusion constructs or Any kD™ Mini-PROTEAN® TGX™ gels for the remaining constructs (Bio-Rad). Standard Western blotting protocols were used to transfer and detect the proteins. Briefly, separated proteins were transferred to a polyvinyl difluoride (PVDF) membrane and stained with the horseradish peroxidase (HRP)-conjugated anti-6X His tag antibody (Abcam) and/or HRP-conjugated anti-GST antibody (Abcam) to compare expression levels. The BCCP constructs were stained with an HRP-conjugated anti-biotin antibody (Abcam) for detecting the biotinylated protein cargoes. After staining the membranes, the images were developed using Clarity Western ECL Substrate (Bio-Rad), and the chemiluminescence was imaged on a ChemiDoc MP documentation system (Bio-Rad). Western blots were performed for three independent biological replicates for each construct.

3.3.1 Densitometry analysis

Densitometry analysis was done on the Western blots to assess the relative expression level based on the signal of the detected bands. Densitometry analysis of all gels was done using ImageLab software (Version 5.2; Bio-Rad). The software reads the raw data in three dimensions by using the length and width of the band. The chemiluminescent signal emitted from the blot is recorded in the third dimension as a peak. The density of any given band was quantified as the total volume under the three-dimensional peak. Within the expression data for each construct, the data were normalized by dividing the signal for each sample by the signal for the average

intensities of all samples. A one-way (for Figure 4.9) and two-way (all other data) ANOVA analysis (GraphPad Prism 7.03) was used for statistical analysis to calculate the statistical significance of each data point across each induction time and each temperature.

3.4 Protein purification

The Hst-5-G_4S-GFP fusion construct was purified for translocation studies. A culture of 500 mL of BL21 (DE3) cells containing pNGST-Hst-5-G_4S-GFP was incubated at 37 °C for 6 h. Following harvesting of the cells by centrifugation, the cells were resuspended in the buffer for purification (125 mM Tris (pH 7.5), 150 mM NaCl, 1 mM EDTA). The cells were then lysed using a homogenizer cell disruption system (Avestin). After lysis of each sample, the whole-cell lysates were pelleted by centrifugation at 11,400×g for 50 min at 4 °C, and the supernatants containing the soluble fraction of the cell lysates were collected.

The soluble fraction of each lysate was passed through a 0.2 μm filter and applied to an IMAC Profinity Ni-NTA resin column (Bio-Rad). After thoroughly washing the column, Factor Xa (10 μL) was added to the column, and the column was shaken overnight to cleave the GST expression partner from the constructs bound to the column. After overnight incubation, the flow-through and washes were collected. GFP and Hst-5-G_4S-GFP were eluted in a buffer containing 20 mM sodium phosphate (pH 7.4), 0.5 M potassium chloride (KCl), and 500 mM imidazole.

For downstream ion-exchange purification to remove the Factor Xa, the elution was dialyzed against 20 mM imidazole and applied to an Enrich-Q anion-exchange column (Bio-Rad) attached to an NGC liquid chromatography system (Bio-Rad). The bound Hst-5–G_4S–GFP fusion protein was washed with 5 column volumes of 20 mM imidazole and eluted by applying a gradient of NaCl, from 10% to 40% over 25 min. Following ion exchange, the purified proteins were dialyzed against 10 mM sodium phosphate buffer (NaPB) to further prepare for translocation studies.

The product at each step of the purification was stored at 4 °C, and the total protein concentration was estimated using a NanoDrop instrument (Thermo Scientific) to measure the absorbance at 280 nm. SDS-PAGE was used to estimate protein purity. Samples from each step of the purification (lysate, washes and elutes) were normalized by culture volume, and separated by SDS-PAGE on Any kD Mini-PROTEAN TGX gels. The protein gels were stained with Bio-Safe Coomassie stain (Bio-Rad) and imaged on the ChemiDoc MP documentation system. For quantifying the yield of purified protein, three biological replicates (three separate cultures) were prepared on three separate days.

GFP was produced using the pET21-GFP plasmid (construct lacked GST)[13]and purified using both IMAC and IEX to further use the purified product for translocation studies.

3.5 Candida strain and culture conditions

The *C. albicans* clinical isolate strain SC5314 (American Type Culture Collection) was used for experiments involving *C. albicans.* Cells were inoculated

into 5 mL of yeast–peptone–dextrose (YPD) liquid medium (1 % yeast extract, 2 % peptone and 2 % glucose) from an YPD agar plate (1 % yeast extract, 2 % peptone, 2 % glucose and 2 % agar). The liquid culture was then grown overnight at 30 °C while shaking at 230 rpm. The cells in the overnight culture were subcultured into 5 mL of fresh YPD medium at an OD_{600} = 0.1 (approximately 2×10^6 cells/mL). The culture was then grown at 30 °C to OD_{600} = 0.5 (approximately 1×10^7 cells/mL) while shaking. Cells were harvested by centrifugation at 4000×g at 4 °C for 10 min and washed twice with 10 mM sodium phosphate buffer (NaPB) before use in the translocation study.

3.6 Cellular uptake studies using flow cytometry

A solution of purified GFP or Hst5-G4S-GFP (100 µL of each, 1 µM) was made in 10 mM NaPB, mixed with 100 µL of cell suspension that contained 5×105 cells in 10 mM NaPB and incubated at 30 °C for 1 h. Cells were pelleted by centrifugation at 5,000 ×g for 10 min at 4 °C and washed with 10 mM NaPB. The cell pellet was then incubated with 200 µL of trypsin (Invitrogen) at 37 °C in a water bath for 5 min to remove surface-bound protein[64]. After trypsin treatment, cells were washed with 200 µL of 10 mM NaPB and re-suspended in 250 µL of 10 mM NaPB. The intracellular fluorescence level of GFP was quantified by flow cytometry (BD FACSCantoII). Single cells were selected and the percentage of GFP fluorescence-positive cells was used to evaluate the translocation efficacy. To examine the potential membrane damage caused by CPPs, propidium iodide (PI) was added into

the suspension to a final concentration of 0.1 mg/ml before flow cytometry measurement. Three replicates were performed and consisted of three different batches of purified protein samples being incubated with three separate cell cultures (on different days).

Chapter 4. Expression optimization studies

4.1 Objective of expression studies

Our goal was to evaluate the recombinant expression of genetic fusions of CPPs to protein cargoes to identify optimal expression conditions for the fusions and determine whether a standardized set of expression conditions can be used. We evaluated five CPPs (NPFSD, pVEC, SynB, Hst-5 and MPG) and three cargoes (BCCP, GFP and MBP). These CPPs and cargoes allowed us to explore whether expression conditions can be generalized for a variety of constructs. We also explored the utility of including a flexible G_4S linker in constructs containing the soluble fusion partner GST (at the N-terminus or C-terminus). Gong and Karlsson previously showed that the linker helped in expression of NPFSD constructs genetically fused to GFP, without reducing translocation into the fungal pathogen *C. albicans* or affecting toxicity toward *C. albicans*[13].

4.2 Effect of linker on expression in constructs containing GST

NPFSD and pVEC have shown translocation into *C. albicans* with molecular cargo. For example, NPFSD can transport an N-terminal fluorescein isothiocyanate (FITC)[14], C-terminal green fluorescent protein (GFP)[13], and C-terminal ricin A chain toxin into the cytosol of *C. albicans*[35], and pVEC can carry an N-terminal

carboxyfluorescein and C-terminal GFP into *C. albicans*[13,65]. These CPPs thus were a perfect starting point for improving recombinant expression.

GST was fused to the C-terminus to allow enhancement in expression, because earlier attempts to express these constructs resulted in low yields[13]. A 6XHis tag was included in the constructs to allow detection by Western blotting. While evaluating the effect of the linker, we also expressed the constructs at various temperatures and induction times to concurrently evaluate expression conditions.

To study the significance of the linker on expression of CPP–cargo protein fusions, the NPFSD and pVEC were genetically fused to BCCP with and without a G_4S linker between the CPP and BCCP (Figure 4.1). Temperatures of 37 °C and 20 °C and induction times of 4 h, 6 h, 8 h, 10 h and 24 h were chosen to evaluate a variety of conditions in which to express our fusion constructs in *E. coli* BL21(DE3) cells. The soluble cell lysate was analyzed by Western blotting.

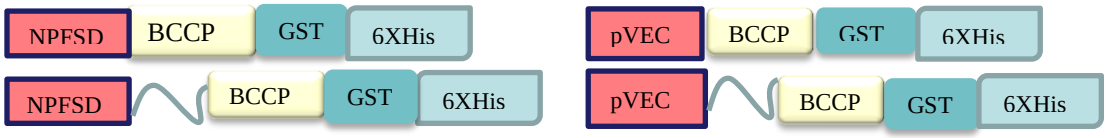

Figure 4.1 Design of CPP-BCCP constructs with a C-terminal GST and with or without a flexible linker. The flexible linker was a glycine-serine linker (G_4S).

Expression results varied depending on the construct. We first evaluated expression at 20 °C and 37 °C for 8 h and observed better expression at 37 °C compared to 20 °C for the NPFSD constructs (Figure 4.2). However, pVEC-G_4S-BCCP-GST failed to express well at either temperature (Figure 4.2).

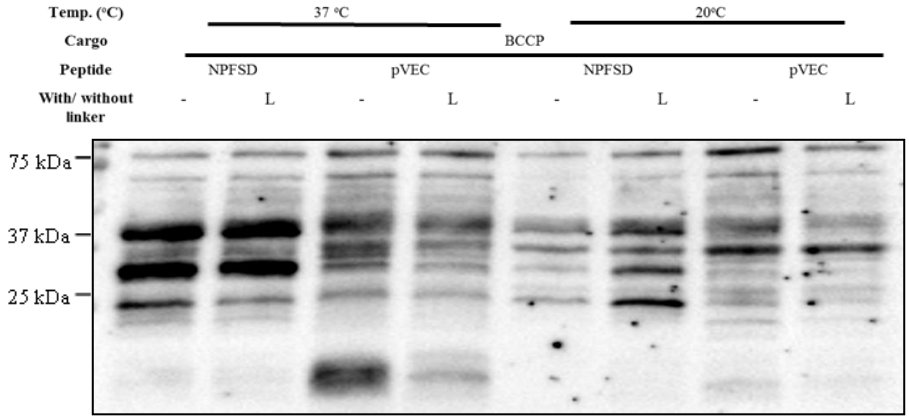

Figure 4.2 Expression of CPP-G_4S-BCCP-GST and CPP-BCCP-GST at different temperatures. CPP fusions to BCCP were expressed in BL21(DE3) cells at 37 °C and 20 °C at 8 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. Proteins were detected using an anti-His antibody. Molecular weights of fusion proteins are as follows: NPFSD-BCCP-GST = 36.7 kDa; pVEC-BCCP-GST = 37.6 kDa; NPFSD-G_4S-BCCP-GST= 37.3 kDa; pVEC-G_4S-BCCP-GST = 38.2 kDa.

We then evaluated expression of the constructs at induction times ranging from 4 h – 24 h at 37 °C and still did not observe substantial expression of the pVEC constructs (Figure 4.3). This poor expression could be due to the fact that pVEC shows antimicrobial activity[13,32], and the intracellular accumulation of pVEC fusions results in insoluble expression or even cell death. For NPFSD constructs, the data for induction times of 4 h – 24 h (Figure 4.3) showed no significant expression at 4 h and obvious degradation at 24 h. For this reason, 4 h was not included in subsequent experiments to evaluate the effect of the linker.

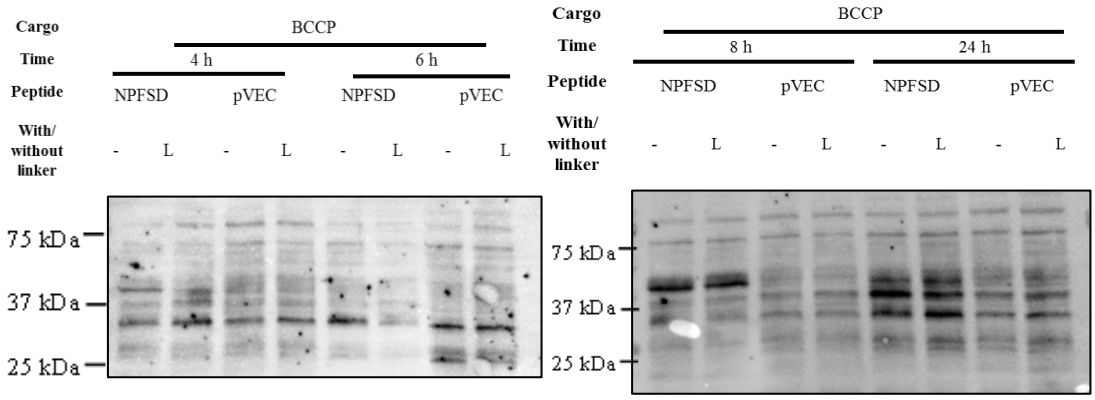

Figure 4.3 Expression of CPP-G_4S-BCCP-GST and CPP-BCCP-GST at different induction times. CPP fusion to BCCP was expressed in BL21 (DE3) cells at 37 °C for 4 h, 6 h, 8 h and 24 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. Proteins were detected using an anti-His antibody. Molecular weights of fusion proteins are as follows: NPFSD-BCCP-GST = 36.7 kDa; pVEC-BCCP-GST = 37.6 kDa; NPFSD-G4S-BCCP-GST= 37.3 kDa; pVEC-G4S-BCCP-GST = 38.2 kDa.

The addition of a linker did not enhance expression for these constructs as it did previously[13]. Densitometry analysis of expression data at 4, 6, and 8 h for the NPFSD constructs showed that the constructs without the linker were similar to those with the linker (Figure 4.4), and no statistically significant effect of the linker (or the induction time) was observed. Several differences between the constructs in this study and those in the previous study may contribute to this discrepancy. The two major differences are the cargo protein (BCCP in this study vs. GFP in the previous study) and the presence of the solubility enhancer GST in the current study. This suggests that fusion proteins—either solubility enhancer or cargo proteins—may play a larger role in expression than the linker. Thus, to see how a different cargo affects the expression, we compared between different combinations of cargoes and peptides. Although the addition of a linker had no clear positive or negative effect on

expression for the fusion constructs, subsequent studies have included the linker between the CPP and cargo due to its positive effect in the previous study.

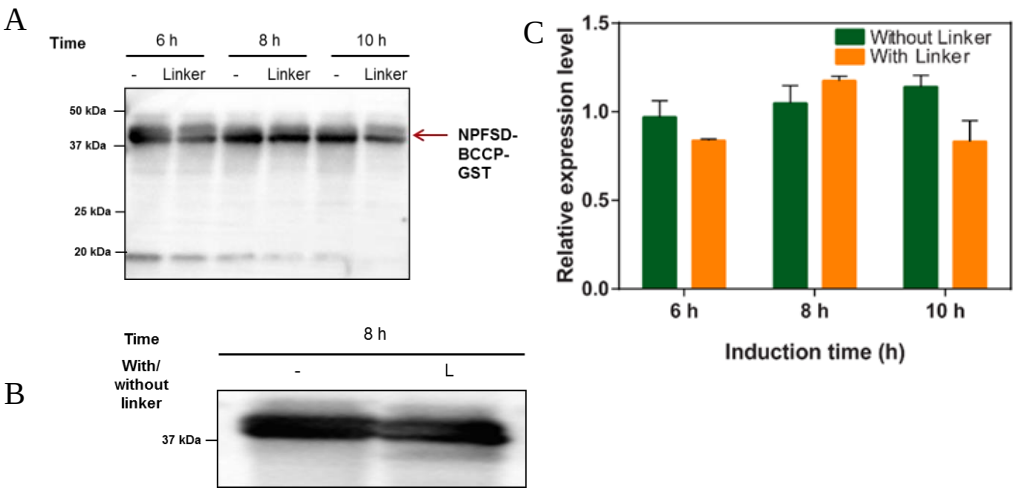

Figure 4.4 Quantification of expression of NPFSD-G$_4$S-BCCP-GST and NPFSD-BCCP-GST at different induction times. CPP fusions to BCCP were expressed in BL21 (DE3) cells at 37 °C for 6 h, 8 h and 10 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an (A) anti-His (mouse) primary antibody and (B) an anti-biotin HRP-conjugated anti-mouse antibody (to confirm biotinylation). (C) The Western blot data were quantified by densitometry, and the error bars represent the standard error of the average of three biological replicates (N=3). In this case neither the temperature nor the addition of a linker was statistically significant. Refer to Figure A7 in Appendix A for raw data. Molecular weights of fusion proteins are as follows: NPFSD-BCCP-GST = 36.7 kDa; NPFSD-G$_4$S-BCCP-GST= 37.3 kDa.

4.3 Expression of CPP constructs with MBP as the cargo

One cargo protein that we evaluated was MBP, the well-known solubility enhancer. MBP is the largest cargo protein we evaluated. Each of the MBP constructs, as well as the constructs for the other cargo proteins, included GST at the N-terminus, a G$_4$S linker

Figure 4.5 Design of constructs with CPP fused to cargo.

between the CPP and cargo, and a C-terminal 6XHis tag. Our previous constructs (with GST at the C-terminus) showed expression, but other work in our lab showed that moving the soluble partner to the N-terminus allows better expression, consistent with the standard location in the literature[66].The peptides chosen here were SynB, Hst-5 and MPG, each of which has previously shown translocation of cargo into cells[33 5,7,19,24,25,33,42,45,67]. The expression of the CPP-cargo fusions was evaluated at induction temperatures of 20 °C, 30 °C and 37 °C and induction times of 6 h, 10 h and 24 h to cover a wide range. Following Western blotting to detect expression of the fusion proteins, the expression level was quantified by densitometry.

4.3.1 Expression of GST-SynB-G_4S-MBP

The expression of the fusion construct containing SynB and MBP was evaluated at different expression conditions (Figure 4.6). This construct expressed well at all temperatures and induction times, as the strong band at about the size of the GST-SynB-MBP-G_4S fusion (70.2 kDa). The relative expression level confirmed that expression at 37 ^oC and 6 h and 10 h was better than at 24 h ($p \leq 0.01$). At 6 h, both the constructs at 37 ^oC and 30 ^oC expressed better than those at 20 ^oC ($p \leq 0.05$ and $p \leq 0.001$ respectively). 37 ^oC expressed better than 20 ^oC at 10 h ($p \leq 0.05$) and 37 ^oC expressed better than both 30 ^oC and 20 ^oC at 24 h ($p \leq 0.05$). For this construct, both temperature and induction time was significant for expression ($p \leq 0.01$), with 37 ^oC and shorter times were the best expression conditions.

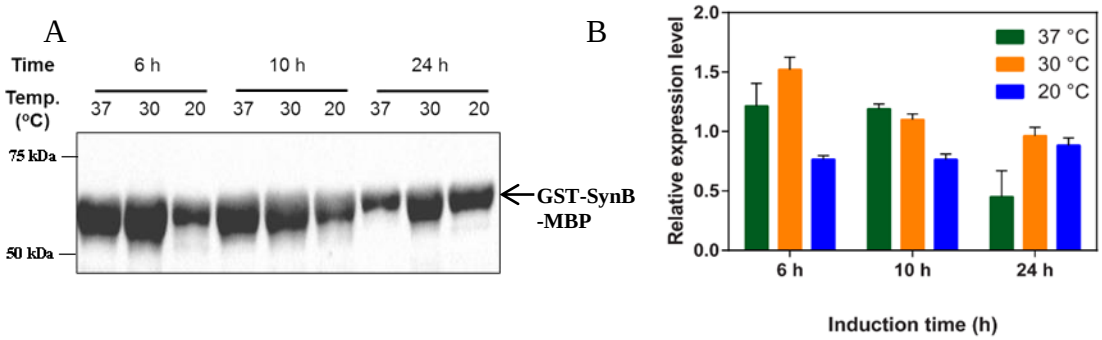

Figure 4.6 Expression of GST-SynB-G_4S-MBP. CPP fusion to MBP was expressed in BL21 (DE3) cells at 37 °C, 30 °C and 20 °C for 6 h, 10 h and 24 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an (A) HRP-conjugated anti-GST antibody. (B) The Western blot data were quantified by densitometry, and the error bars represent the standard error of the average of three replicates (N=3). Both temperature and induction time had a statistically significant effect ($p<0.01$). Refer to Figure A1 in Appendix A for raw data. Molecular weight of fusion protein is: GST-SynB-G_4S-MBP = 70.2 kDa.

4.3.2 Expression of GST-Hst-5-G_4S-MBP

The construct containing MBP with the CPP Hst-5 was also evaluated. The expected size for this constructs is 71.1 kDa (GST-G_4S-Hst-5-MBP), and it expressed well at all expression conditions (Figure 4.7). The highest level of expression occurred at 37 ^{o}C and 6 h (Figure 4.7). While expression of the constructs at 37 ^{o}C was higher at 6 h than 10 h or 24 h ($p \leq 0.05$ & $p \leq 0.01$ respectively), induction time did not have a significant effect on expression at 30 ^{o}C or 20 ^{o}C, whereas for 6 h induction time expression of the constructs was higher at 37 ^{o}C than 30 ^{o}C and 20 ^{o}C ($p \leq 0.01$ and $p \leq 0.0001$ respectively). For this construct, the expression temperature had a clear effect for the 6 h induction time, with expression increasing as the induction temperature was increased.

Expression studies in literature report that induction times longer than 6 h are associated with low productivity whereas induction times between 4 h and 6 h

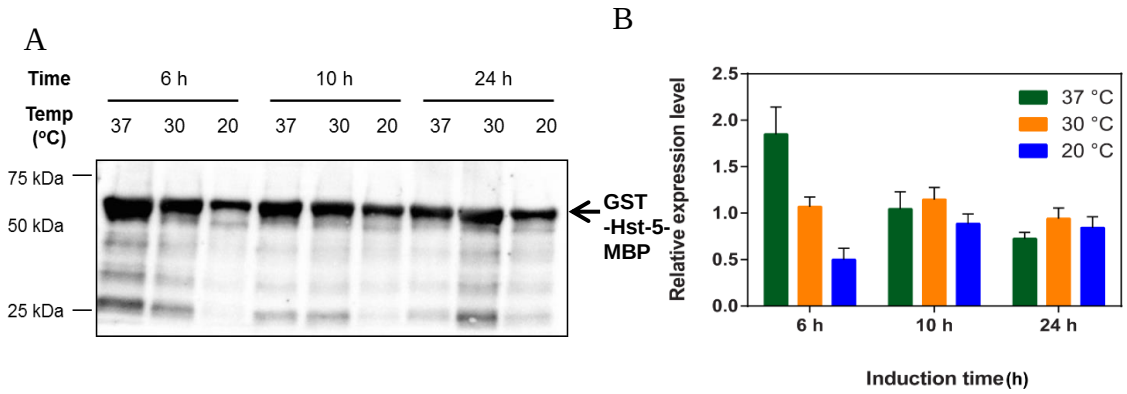

Figure 4.7 Expression of GST-Hst-5-G_4S-MBP. CPP fused to MBP were expressed in BL21 (DE3) cells at 37 °C, 30 °C and 20 °C for 6 h, 10 h and 24 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an (A) HRP-conjugated anti-GST antibody. (B) The Western blot data were quantified by densitometry, and the error bars represent the standard error of the average of three replicates (N=3). In our case temperature had significant implication ($p < 0.001$). Refer to Figure A2 in Appendix A for raw data. Molecular weight of fusion protein is: GST-Hst-5-G_4S-MBP = 71.1 kDa.

presented comparable levels of productivity[40,60,61]. This is consistent with our results showing better protein production at shorter times for our many constructs. It was also observed that growth rate reduces with expression temperature reduction from 37 °C to 25 °C[40,60]. This may explain our improved expression at higher temperatures, since an increased growth rate would result in more cells and increase the overall level of protein production for a given culture volume.

4.3.3 Expression of GST-MPG-G_4S-MBP

Similar to the constructs containing SynB and Hst-5 as the CPPs, the constructs containing MPG with MBP (GST-G_4S-MPG-MBP, 71.1 kDa) expressed at all conditions tested (Figure 4.8). Expression of constructs at 37 ^{o}C was higher than at 20 ^{o}C at both 6 h and 24 h ($p \leq 0.05$).

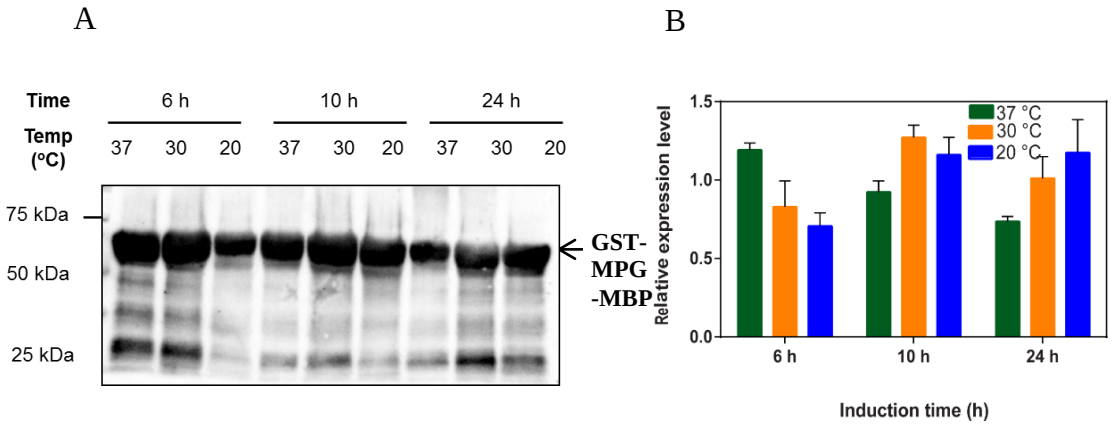

Figure 4.8 Expression of GST-MPG-G_4S-MBP. CPP fused to MBP were expressed in BL21 (DE3) cells at 37 °C, 30 °C and 20 °C for 6 h, 10 h and 24 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an (A) HRP-conjugated anti-GST antibody. (B) The Western blot data were quantified by densitometry, and the error bars represent the standard error of the average of three replicates (N=3). In our case neither temperature nor induction time had any significant implication. Refer to Figure A3 in Appendix A for raw data. Molecular weight of fusion protein is: GST-MPG-G_4S-MBP = 71.1 kDa.

4.3.4 Comparison of all MBP constructs at same expression conditions

To directly compare expression of the different CPPs with MBP as the cargo and MBP with no CPP, samples were induced using the same conditions and analyzed on the same Western blot (Figure 4.9). Since 37 ^{o}C was generally the best expression condition for fusions of MBP, expression was induced at 37 ^{o}C for 10 h. GST-MBP was not significantly different from GST-SynB-G_4S-MBP and GST-MPG-G_4S-MBP and these constructs were similar in expression (p=0.6045 and p=0.9970 respectively). In contrast, GST-Hst-5-G_4S-MBP expresses better than both SynB constructs (with and without linker) and the MPG construct (p=0.0125,

p=0.0168, p=0.0328). It even expresses at a higher level than MBP alone (p=0.0437). Thus, Hst-5 helps in better expression of MBP.

These data also show that the CPPs do not reduce expression as has been seen before[13]. Further supporting our previous result that the linker does not affect the expression of constructs containing GST (Section 4.2), we also found that removing the linker from the GST-SynB-G_4S-MBP to yield GST-SynB-MBP has no effect on expression, as this construct also expresses similarly to GST-MBP (p=0.9940).

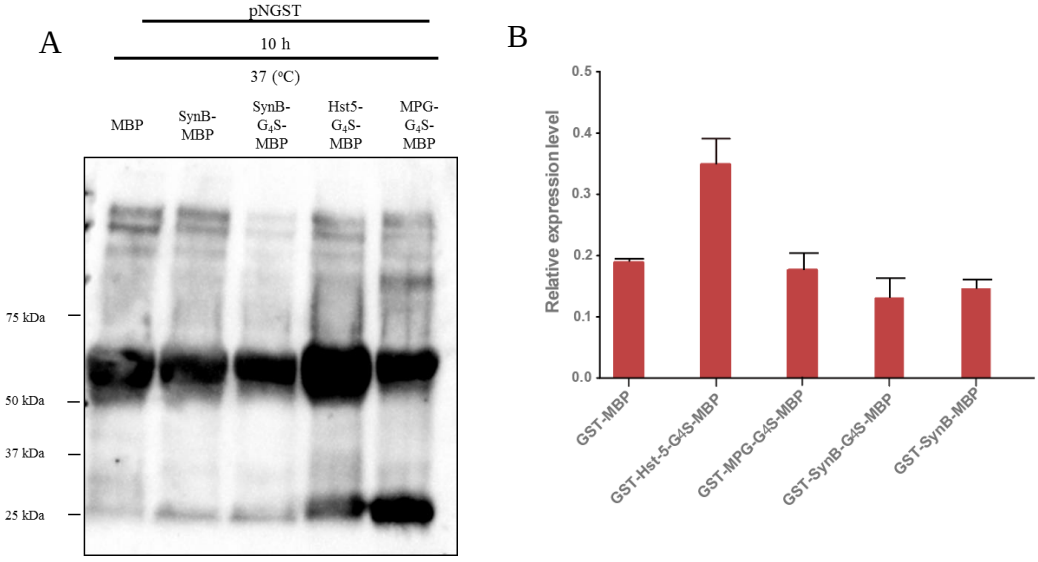

Figure 4.9 Expression of GST-MBP and all GST-CPP-G_4S-MBP constructs. CPP fused to MBP were expressed in BL21 (DE3) cells at 37 °C for 10 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an (A) HRP-conjugated anti-GST antibody. (B) The Western blot data were quantified by densitometry, and the error bars represent the standard error of the average of three replicates (N=2). In our case there was significant difference ($p < 0.05$). Refer to Figure A8 in Appendix A for raw data.

Based on the expression data for all of the different MBP constructs, the fusion constructs with MBP express very well. This could be due to the fact that MBP

is a very soluble fusion partner and thus enhances expression. MBP has been used widely to circumvent inclusion body formation, particularly in *E. coli* where the reduced solubility of recombinant proteins is a serious problem[77]. It has been suggested that solubility enhancers may have an innate, inactive chaperone-like quality that displays itself as iterative cycles of temporary intramolecular binding to cargo such that it averts their self-association and aggregation[61].

Since MBP is known to be a better fusion partner than GST[36, 60,61] and our results show that each of our CPPs express well with MBP, MBP could serve as an alternative solubility partner to replace GST at the N-terminus (rather than serving as a cargo at the C-terminus). This could lead to enhanced expression for constructs containing cargo that are more difficult to express. Proteins like TEV protease and GFP fold spontaneously when their tendency to form insoluble aggregates is blocked by fusion to MBP [53,60].

4.4 Expression of CPP constructs with GFP as the cargo

CPP-cargo fusion constructs with GFP as the cargo were studied to determine promising induction conditions for their expression. GFP is smaller in size than MBP and its characteristic fluorescence helps in detection of translocation. Like the constructs with MBP, the constructs with GFP had an N-terminal GST, a glycine-serine linker between the peptide and cargo, and a C-terminal 6XHis tag. Expression was also evaluated with the same peptides (SynB, Hst-5 and MPG).

4.4.1 Expression of GST-SynB-G_4S-GFP

Expression of GST-SynB-G_4S-GFP was inconsistent over several replicate experiments, with most experiments resulting in no clear expression of the desired product. The expected size of the GST-SynB-G_4S-GFP construct is 55.2 kDa and we observed a band at this size for each of the induction conditions evaluated in one of our experiments (Figure 4.10). However, the band(s) around 25 kDa (degradation product) are stronger than the expected band for desired product.

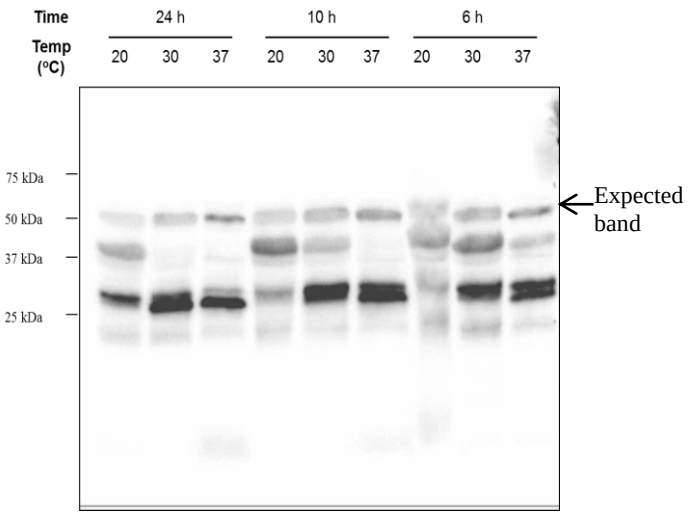

Figure 4.10 Expression of GST-SynB-G_4S-GFP. CPP fused to GFP were expressed in BL21 (DE3) cells at 37 °C, 30 °C and 20 °C for 6 h, 10 h and 24 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an anti-His primary antibody. Refer to Figure A4 in Appendix A for raw data. Molecular weight of fusion protein is: GST-SynB-G_4S-GFP = 55.2 kDa

This degradation product could be either the cargo or the fusion partner but it is difficult to exactly point out since these are relatively of the same sizes. This blot

could also mean that this construct just does not express well, and thus it is difficult to know the best conditions of expression time and temperatures.

4.4.2 Expression of GST-Hst-5-G_4S-GFP

The GST-Hst-5-G_4S-GFP constructs (56.1 kDa) showed significant expression under all induction conditions, without the degradation observed with GST-SynB-G_4S-GFP (Figure 4.11). Expression at 30 °C is good but not significantly better than any of the other conditions. Temperature though was more significant than induction time ($p \leq 0.01$).

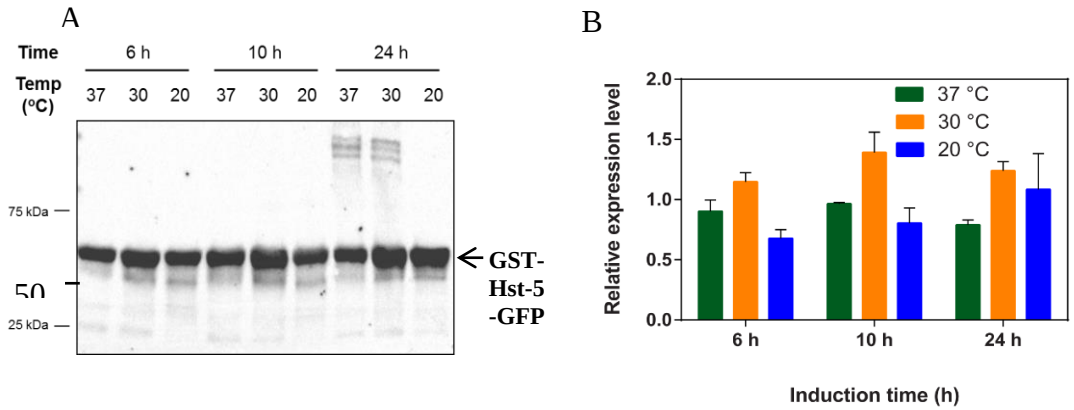

Figure 4.11 Expression of GST-Hst-5-G_4S-GFP. CPP fused to GFP were expressed in BL21 (DE3) cells at 37 °C, 30 °C and 20 °C for 6 h, 10 h and 24 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an (A) HRP-conjugated anti-GST antibody. (B) The Western blot data were quantified by densitometry, and the error bars represent the standard error of the average of three replicates (N=3). In our case temperature had significant implication ($p < 0.01$). Refer to Figure A6 in Appendix A for raw data. Molecular weight of fusion protein is: GST-Hst-5-G_4S-GFP = 56.1 kDa

4.4.3 Expression of GST-MPG-G_4S-GFP

GST-MPG-G_4S-GFP constructs ran lower than the size expected when expressed at different temperatures of 37 °C, 30 °C and 20 °C and induction times of 6 h, 10 h and 24 h (Figure 4.12). The blot was stained with 1:8000 anti-GST. The expected size of this construct is 56.1 kDa, though the construct didn't express at the size expected.

30 ^{o}C was higher at 6 h than 10 h or 24 h ($p \leq 0.05$ & $p \leq 0.01$ respectively) and 20 ^{o}C was higher at 6 h and 10 h than 24 h ($p \leq 0.05$) whereas none of the other conditions were significantly effective. For this construct, the expression temperature had a clear effect for the 6 h induction time but comparative higher temperature conditions didn't express at all. In general temperature was more significant ($p \leq 0.0001$) compared to induction time.

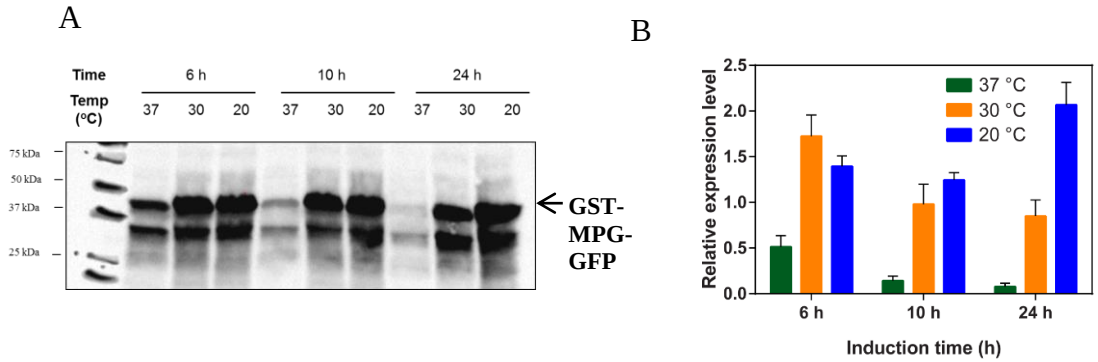

Figure 4.12 Expression of GST-MPG-G_4S-GFP. CPP fused to GFP were expressed in BL21 (DE3) cells at 37 °C, 30 °C and 20 °C for 6 h, 10 h and 24 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an (A) HRP-conjugated anti-GST antibody. (B) The Western blot data were quantified by densitometry, and the error bars represent the standard error of the average of three replicates (N=3). In our case temperature had significant implication ($p < 0.0001$). Refer to Figure A6 in Appendix A for raw data. Molecular weight of fusion protein is: GST-MPG-G_4S-GFP = 56.1 kDa

There might be various reasons for this construct to run at lower size. The fusion protein could be truncated and thus not getting fully expressed or just unable to be fully expressed. Additionally, some proteins simply run at different sizes than expected based on the molecular weight standards.

To determine whether the entire MPG fusion was being expressed, a Western blot to detect GST-MPG-G_4S-GFP was performed again, but the protein was detected on the blot using an anti-6XHis antibody instead of the anti-GST antibody used in Figure 4.12 (Figure 4.13). The anti-6XHis antibody recognizes the C-terminal tag on the protein, so detection by both the anti-GST and anti-6XHis antibody indicates the protein contains the expected domains for a full-length fusion.

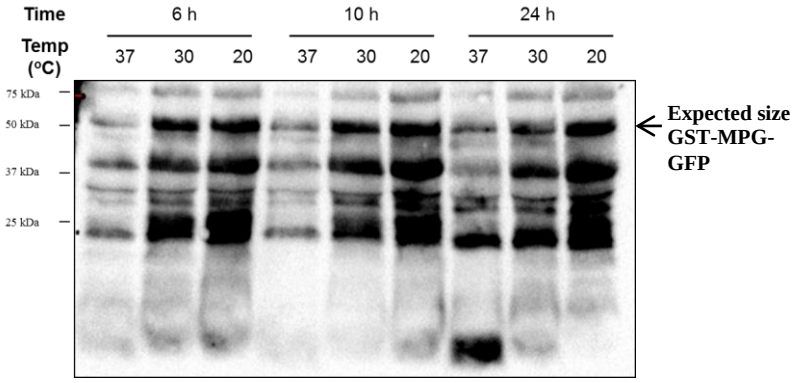

Figure 4.13 Expression of MPG-G_4S-GFP with anti-His Ab. CPP fused to GFP were expressed in BL21 (DE3) cells at 37 °C for 6 h, 8 h and 10 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an anti-His primary antibody.

4.4.4 Comparison of all GFP constructs at same expression conditions

Comparison was done between expression of different CPPs and between construct with GFP only (Figure 4.13). These were all run on the same blot to be able

to compare between them. 20 °C and 24 h were chosen as the condition to express all the fusions of GFP.

SynB-G_4S-GFP fails to express compared to the GFP only, whereas Hst-5-G_4S-GFP expresses really well. It is unclear whether MPG-G_4S-GFP construct is expressing compared than the GFP only and Hst-5- G_4S-GFP constructs.

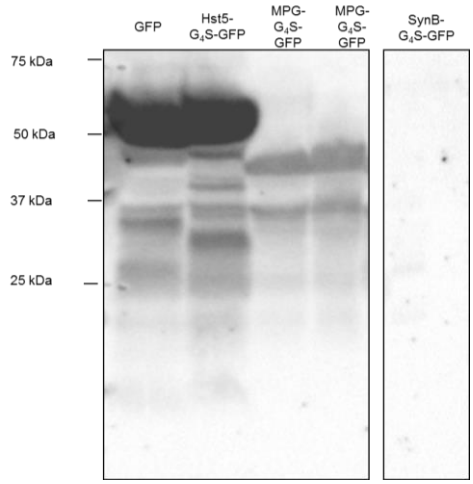

Figure 4.14 Expression of all GFP constructs. CPP fused to GFP were expressed in BL21 (DE3) cells at 20 °C for 24 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an anti-GST primary antibody. The SynB-G_4S-GFP image has been cropped from the same blot. This image has been exposed for the same time but failed to produce reliable data.

4.5 Expression of CPP constructs with BCCP as the cargo

BCCP was chosen as one of the cargoes to study because it is significantly smaller (10 kDa) than the other cargoes, which will eventually allow comparison of size effects on translocation. Additionally, our lab has previously shown that this protein can be produced in *E. coli.* Previous work in our lab with CPP-BCCP constructs containing a C-terminal GST showed that the constructs expressed best at

37 °C, so we focused on this temperature to express the BCCP fusions with the CPPs. The GST-CPP-G_4S-BCCP constructs were expressed at 37 °C with induction times of 6 h, 8 h and 24 h (Figure 4.14). Since BCCP is naturally biotinylated by *E. coli*, biotin was supplied during induction to allow detection of biotin as a complementary detection method (data not shown).

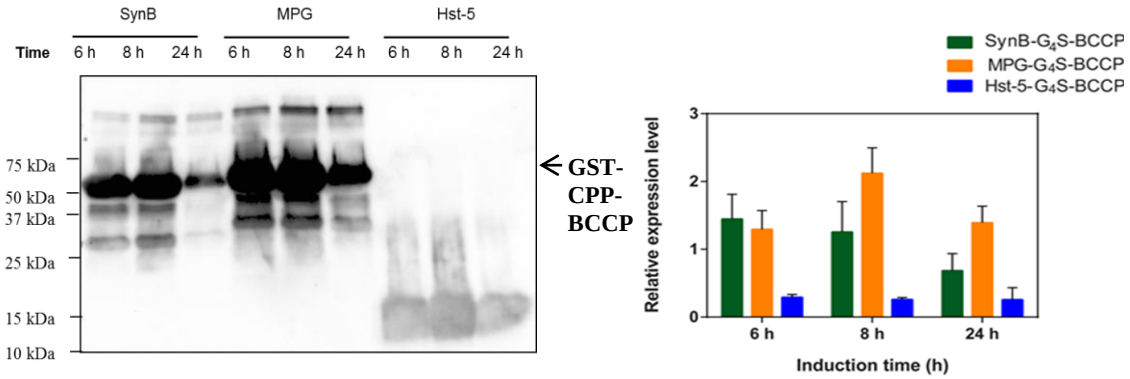

Figure 4.15 Expression of GST-CPP-G_4S-BCCP constructs. CPP fused to BCCP were expressed in BL21 (DE3) cells at 37 °C for 6 h, 8 h and 24 h with 0.01 mM IPTG. The soluble cell lysate was analyzed by Western blotting, with samples normalized by culture volume. The Western blot was stained using an (A) HRP-conjugated anti-His antibody. (B) The Western blot data were quantified by densitometry, and the error bars represent the standard error of the average of three replicates (N=2). In our case time had no significant implication. Molecular weight of fusion protein is: GST-SynB-G_4S-BCCP = 38.2 kDa, GST-MPG-G_4S-BCCP = 39.1 kDa, GST-Hst-5-G_4S-BCCP = 39.1 kDa

GST-SynB-G_4S-BCCP expressed at the expected size of 38.2 kDa (Figure 4.14). It expressed better at 6 h and 8 h than GST-Hst-5-G_4S-BCCP ($p \leq 0.05$). The expression of this construct itself had no significant difference in expression level at different times. GST-MPG-G_4S-BCCP constructs expressed at the expected size of 39.1 kDa and expressed significantly better than GST-Hst-5-G_4S-BCCP at 6 h, 8 h and 24 h ($p \leq 0.05$, $p \leq 0.001$ & $p \leq 0.05$ respectively). The expression of this construct showed no significant difference at different induction times. Unlike the

other BCCP fusions, GST-Hst-5-G_4S-BCCP failed to express at the expected size (39.1 kDa). Bands are present at around 10 – 15 kDa. It is unclear what proteins these bands represent, but the size is similar to where BCCP alone runs on a gel.

4.6 Summary of CPP-cargo expression data

Expression of CPP-cargo fusions depends on both the cargo and the CPP it is fused to. MBP as a cargo consistently expressed well no matter which CPP was fused to it. Hst-5 when fused to MBP and GFP expressed well, though the same was not true for expression of its fusion to BCCP. Interestingly, Hst-5 when fused to MBP is better than even the protein alone thus showing Hst-5 helps in improving the expression of MBP.

Fusions with MBP and GFP as cargoes show better expression at 30 °C compared to the other temperatures, though no specific temperature or time pattern was apparent for the peptides. In general, expression at 37 °C for 6 h and 10 h gave good levels of protein expression for many of the fusion proteins, and thus, these conditions represent a good starting point when expressing new combinations of peptide and cargo.

In general, the cargo present in a fusion construct was the best indication of the optimal temperature and induction time. When the same peptide was fused to different cargoes, the optimal expression conditions changed, while the same cargo fused to different peptides had similar optimal expression conditions.

Chapter 5. Purification and translocation of histatin-5-G_4S-GFP

This chapter focuses on studying the purification and translocation of Hst-5-G_4S-GFP. As described in Chapter 4, fusion constructs containing Hst-5 as the CPP expressed well with both MBP and GFP as cargo, and, thus, it is likely that it can be produced in large enough quantities to help study their translocation[14]. Hst-5 is present in the saliva and thus is safe and also specifically exhibits anti-candidal properties[47], making it a promising CPP for translocating cargo into *C. albicans*. Due to its fluorescence, translocation of GFP into cells can easily be followed using Flow cytometry (FC). Because of these characteristics, the Hst-5-G_4S-GFP fusion construct was selected for purification and evaluation of translocation into *C. albicans.*

5.1 Purification of GFP and Hst-5-G_4S-GFP

Hst-5-G_4S-GFP and unconjugated GFP were purified for studying translocation into cells using IMAC and IEX. The constructs were expressed in BL21 (DE3) cells at 37 °C for 6 h, since these conditions yielded high levels of Hst-5-G_4S-GFP in the expression study (Figure 4.11).

Unconjugated GFP was purified using IMAC and IEX to obtain a pure product (Figure 5.1). Fractions E1, E2 and E3 were combined as the product. The yield was 1.36 mg/L of culture which was higher than that seen in a previous study[13].

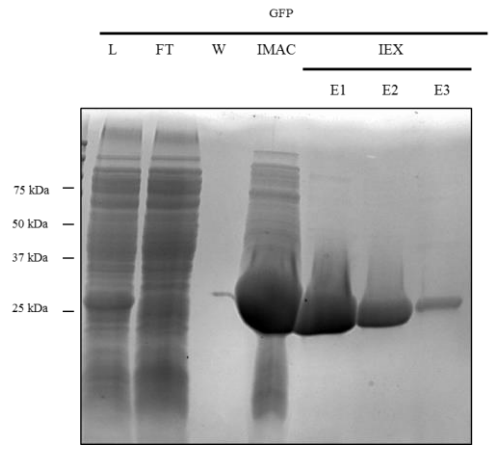

Figure 5.1 Purification of GFP. CPP fused to GFP were expressed in BL21 (DE3) cells at 37 °C for 6 h with 0.01 mM IPTG. Coomassie stain was used to estimate purity of the fusion proteins. The crude soluble lysates, washes and elution from immobilized metal affinity chromatography are shown. The expected sizes of the proteins are 27 kDa for GFP, 56.055 kDa for Hst-5-G_4S-GFP. The crude soluble lysate (L) contains the whole construct; flow through (FT) doesn't bind to the column and so flows through, Wash (W), Elutes (E1, E2 and E3) are eluted using a gradient of 10 - 40% of buffer A.

Following expression and recovery of the soluble lysate, IMAC was used to bind GST-Hst-5-G_4S-GFP to the column and eliminate most impurities (Figure 5.2).

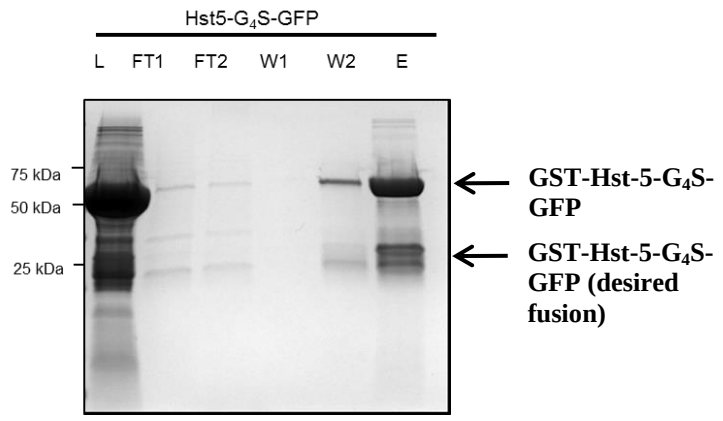

Figure 5.2 Purification of Hst-5-G_4S-GFP with cleaved Factor Xa. CPP fused to GFP were expressed in BL21 (DE3) cells at 37 °C for 6 h with 0.01 mM IPTG. Coomassie stain was used to estimate purity of the fusion proteins. The crude soluble lysates, washes and elution from immobilized metal affinity chromatography are shown. On-column cleavage with Factor Xa to remove GST was done. The expected sizes of the proteins are 30.1 kDa for our protein. The crude soluble lysate (L) contains the whole construct; flow through (FT) doesn't bind to the column and so flows through, Wash (W), Elutes (E6 contains our un-cleaved product at around 56 kDa and our cleaved protein at around 25 kDa) are eluted using a gradient of

GST-Hst-5-G_4S-GFP contained a Factor Xa enzyme cleavage site between the GST and the Hst-5, so Factor Xa was applied to the IMAC column to remove GST while Hst-5-G_4S-GFP remained bound (Figure 5.2). The protein lysate was applied to the column and Factor Xa was added to the column after washing. The column was shaken overnight for over 16 h with the Factor Xa for the cleavage to occur. Following Factor Xa removal, the sample contained substantial fusion protein with GST still attached (56.1 kDa). This indicates that we will need to improve the cleavage conditions in the future to achieve better removal of the GST. However, our desired product without the GST was present (30.1 kDa), so we continued the purification process with IEX to remove the fusion with uncleaved GST (Figure 5.3). The GST-Hst-5-G_4S-GFP can be separated from the Hst-5-G_4S-GFP by ion

exchange, though the purity isn't perfect. The desired protein is present at a reasonable purity in fractions E2 and E3, so these fractions were combined as the purified Hst-5-G_4S-GFP (63 % purity).

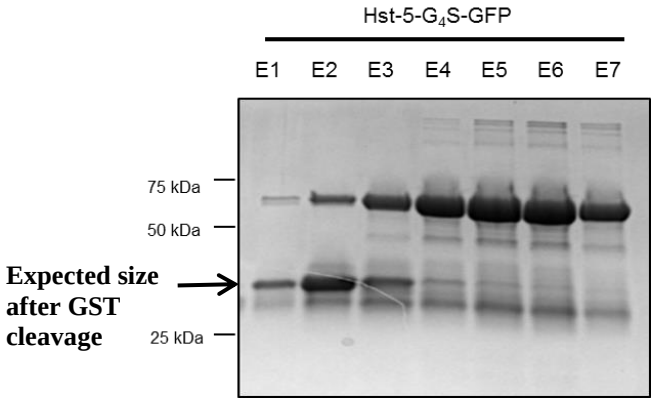

Figure 5.3 Purification of Hst-5-G_4S-GFP constructs. CPP fused to GFP were expressed in BL21 (DE3) cells at 37 °C for 6 h with 0.01 mM IPTG. Coomassie stain was used to estimate purity of the fusion proteins. The crude soluble lysates, washes and elution from anion-exchange chromatography are shown. The expected sizes of the proteins are 27 kDa for GFP, 56.1 kDa for Hst-5-G_4S-GFP – without GST our expected band should be at 31 kDa. These data illustrate the quality of the purification only.

Even without optimization, the purification process produced sufficient protein for translocation studies. The yield of Hst-5-G_4S-GFP was 36 mg per liter of culture. In a previous study, Gong *et al*. purified NPFSD-G_4S-GFP and pVEC-G_4S-GFP produced in *E. coli* with no GST as an expression partner and only achieved yields of 0.264 ± 0.013 mg protein/Liter of culture and 0.159 ± 0.053 mg protein/Liter of culture, respectively. The yield of Hst-5-G_4S-GFP in this work was two orders higher in magnitude than these earlier yields. This substantial difference in expression is likely due to a combination of Hst-5 expressing better than the peptides NPFSD and pVEC and to the addition of the soluble fusion partner GST in our

constructs. These higher yields of CPP-cargo fusion will facilitate studying translocation into cells. By evaluating and optimizing our purification protocols, we will be able to further improve both the purity and the yield of our fusion. One key aspect will be optimizing the amount of Factor Xa to fully cleave the GST from our fusion partner, which has the potential to substantially increase the yield and purity of our desired construct without any additional modifications to the protocol.

5.2 Translocation of CPPs fused to cargoes into C. albicans

We showed we could purify Hst-5-G_4S-GFP fusion protein with a good yield using straightforward purification steps. As the main reason of creating the CPP-GFP fusion is to utilize CPPs to help deliver cargoes into fungal cells, we next explored the intracellular delivery of this purified fusion construct into the important fungal pathogen *C. albicans*.

To screen the CPPs for translocation into *Candida* species, *C. albicans* cells were incubated with both Hst-5-G_4S-GFP and GFP lacking a CPP. We examined and quantified translocation of the proteins using flow cytometry. All single *C. albicans* cells were selected, and the percentage of fluorescence-positive cells was used to evaluate the GFP delivery efficacy. About 3% of the fungal cells exhibited a green fluorescence signal (Figure 5.4c) when treated with 0.5 µM purified Hst-5-G_4S-GFP fusion protein (Figure 5.4). (This concentration reflects the total protein concentration of the sample and does not take the purity into consideration). When the concentration was reduced to 0.05 µM Hst-5-G_4S-GFP, no significant uptake was observed compared with GFP lacking Hst-5 (Figure 5.4d). The concentration-

dependence is expected based on previous data for CPP delivery of fluorescein [31]. The study here has been done with lower end of concentration (0.5 μM vs. 1 μM in literature). We would expect higher translocation with higher concentration based on studies by Gong *et al*., 2017. An interesting fact to note here is that fact that we are using a large protein and not the small-molecule fluorescein when comparing our results.

To further explore the effect of the Hst-5-G_4S-GFP fusion protein on *C. albicans*, we evaluated the permeability of the cells after treatment with the fusion protein using PI. Cells are normally impermeable to PI, but PI fluorescence can be detected inside cells with destabilized membranes. Destabilization can occur with cell death or with pore formation on the membrane[78]. We observed substantially low, <1% PI-positive cells following incubation with Hst-5-G_4S-GFP, indicating the fusion protein does not significantly affect the membrane integrity. This was similar to the level of toxicity of GFP alone. This likely indicates that Hst-5 does not have antifungal activity under these conditions. This is consistent with a previous study about Hst-5 (with no cargo) that showed no significant antifungal activity towards *C. albicans* for Hst-5 concentrations below 1.63 μM[47].

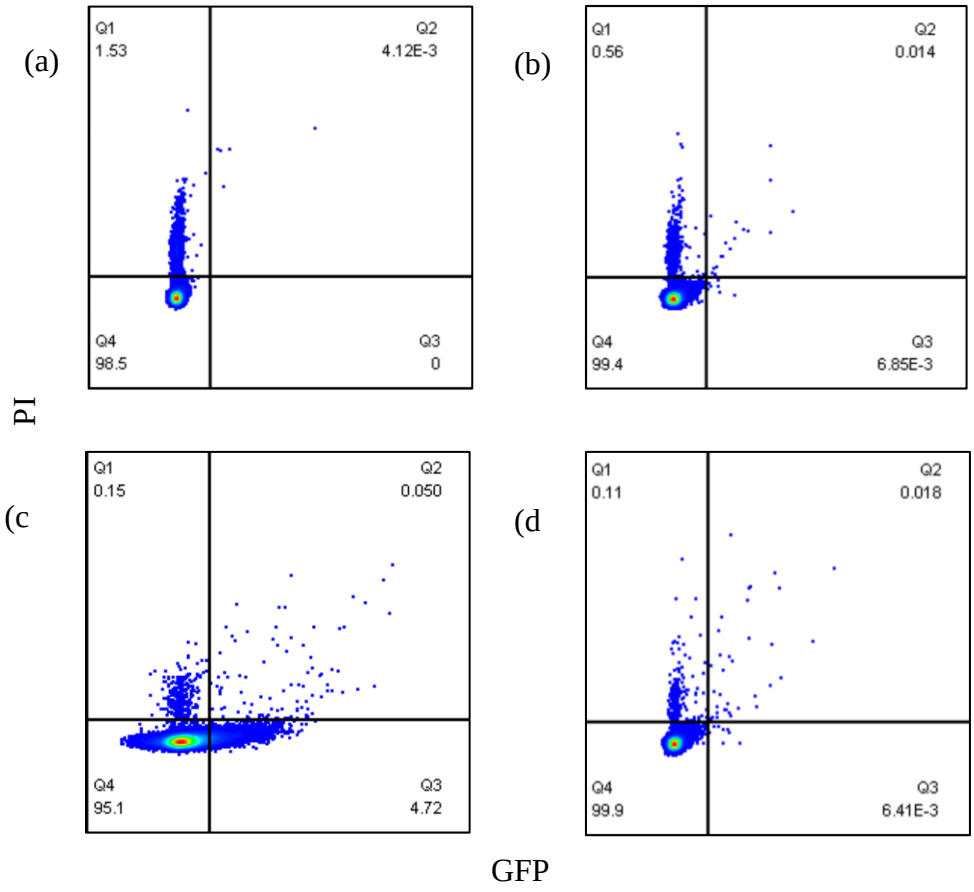

Figure 5.4 Cellular Uptake studies. Flow cytometry was used to quantify translocation and membrane permeabilization in *C. albicans* for (a) cells incubated with buffer only and no stain, (b) cells incubated with GFP and PI, (c) cells with 0.5 µM Hst-5-G_4S-GFP and PI, and (d) cells incubated with 0.05 µM Hst-5-G_4S-GFP. The translocation studies were done for 3 replicates but a, b, c and d is a representative set of data. Average data for each group are found in Table 5.1. Only single cells were selected for analysis. In each plot, quadrant Q1 represents cell that are **PI +/GFP -**, Q2 represents those that are **PI +/GFP +**, Q3 represents those that are **PI -/GFP +**, and Q4 represents those that are **PI -/GFP -**. The numbers under each quadrant label provide the percentage of cells in the experiment that fall into each quadrant. The quadrant boundaries were chosen such that there were no cells with translocation of GFP only. Each cell in the sample is represented by a dot, and the color on the dot plots indicate the density of cells (red represents high density and blue represents low density).

Table 6. Cellular Uptake studies using FLOWJO

Concentration of the Hst-5-GFP (μM)	% GFP-positive*	% of cells with GFP that are dead
0.5	3.12±2.46	0.365±0.233
0.05	0.03±0.001	0.11±0.02

*Values are given as averages of 2 biological replicates with standard error

These results show that Hst-5 (an antimicrobial peptide) can be used to deliver protein cargo. Hst-5 shows specificity for fungal cells, and thus is promising to be used as a specific delivery tool. Increasing concentration of fusion is likely to improve results based on the results above and previous work. To understand how the peptide and cargo are translocated and any intracellular effect of the cargo and the CPP, further studies will need to be done. Additionally, similar studies with other CPPs or cargoes will help better understand the capabilities and limitations of CPP-cargo fusions in delivering cargo to fungal cells.

Chapter 6. Conclusions and future work

6.1 Conclusion

This study was done to increase the expression and production and form a generalized platform for expression of CPP-cargo fusions. We showed that recombinant expression technology can be used to conjugate CPPs and cargo proteins *in vivo*.

We tried to improve recombinant production of CPPs as fusions to protein cargo by considering numerous variables: induction temperatures, induction times, CPPs, cargoes, linkers and a soluble partner. We were able to express fusion proteins with different combinations of CPPs and cargoes. In general, induction at 37 °C for 6 h or 10 h resulted in comparatively high levels of expression; these conditions would be a good starting point for future studies of new CPPs or cargo proteins. Hst-5 as a CPP expressed well with both MBP and GFP but failed to express at the expected size with BCCP as a cargo.

We were able to produce sufficient Hst-5-G_4S-GFP for performing a translocation experiment in *C. albicans*. At a concentration of 10 µM, Hst-5-G_4S-GFP translocated into *C. albicans* cells with minimal killing activity.

Now that we have a better understanding of conditions that improve expression of CPP-cargo fusions, we can focus on improving purification protocols to further improve the yields of purified CPP-cargo fusions. Higher yields of purified

fusions will be needed to for more complete studies of their biological activity and mechanisms of translocation and their functions once inside the target cell.

6.2 Path forward

6.2.1 Toxicity of CPPs toward Candida cells

To study the toxicity of CPPs toward *C. albicans*, we incubated the cells with serial dilutions of the GFP and Hst5-G_4S-GFP. Little to no antifungal activity could be detected for CPP fusions that exhibited translocation (Chapter 5). The PI data has already shown no killing activity but a standard antifungal assay can be done similar to what is used for *Candida* to further show the killing activity.

Though toxicity of CPPs is required for delivering antifungal molecules, CPPs can also act as delivery vehicles for non-toxic biomolecules or the CPPs may be toxic to the target cells – thus lower toxicity is needed. Thus the knowledge of toxicity of CPPs and their attached cargo are of prime importance.

This leads us to our second future goal of understanding function of CPPs inside cells.

6.2.2 CPPs and their interaction with fungal cells

To understand how the peptide acts inside the cell and how long it is actively delivering the cargo, further studies will need to be done that allows us to know the cellular localization of the fusion construct tested in Chapter 5. Further studies also need to be done to know the stability of the membrane after the specific cargo delivery.

At different concentrations of the fusions, the uptake has shown to be varied, and thus its needs to be seen how the cells behave at even higher concentrations of the peptide if the cells retain their vacuole. In yeast, vacuoles are the most acidic part of the cell that has a role in protein degradation, ion and metabolite storage, and detoxification[68].

It has been shown for the results for pVEC[31] where flow cytometry was used to evaluate the intracellular trafficking in *C. albicans* and *C. glabrata* and there was substantial uptake. Intracellular trafficking through vacuoles has also been detected for all peptides studied that suggest that vacuoles are important in the translocation mechanism of CPPs for fungal cells. Mechanisms for translocation of the fusion protein into fungal cells also need to be explored.

6.2.3 Replacement of GST by MBP

As is well known MBP is a better soluble fusion partner than GST[43,60], so a next logical step would be to replace GST at the N-terminus by MBP thus enhancing expression of fusion cargoes. One limitation, though, is that MBP binds to an amylose resin which is an expensive technique.

6.2.4 Design changes in CPPs to make cargo delivery better

Specific amino acids can be changed/mutated in the sequences of the CPPs tested in Chapter 4, which would change their charge/hydrophobicity (as has been

done by Gong *et al.)*. These modified sequences along with the combination of right cargoes and better soluble partners might help increase expression and delivery of multiple constructs can be studied (as opposed to only a single construct studied in this thesis.

6.2.5 Detection of fusions inside cells

In the present work, we have used a fluorescent cargo to study translocation. Very limited number of fluorescent cargoes exists and we would like to extend our studies to look at other cargoes for e.g., MBP and BCCP. To be able to do this, we need a method(s) to detect non-fluorescent proteins. The expression of the fusion construct can be tested in *C. albicans* by using Western Blot and/or ELISA and the efficacy of the methods could be tested using GFP constructs, since we already have a technique to detect them.

Appendix A – Densitometry and WB raw results (Chapter 4)

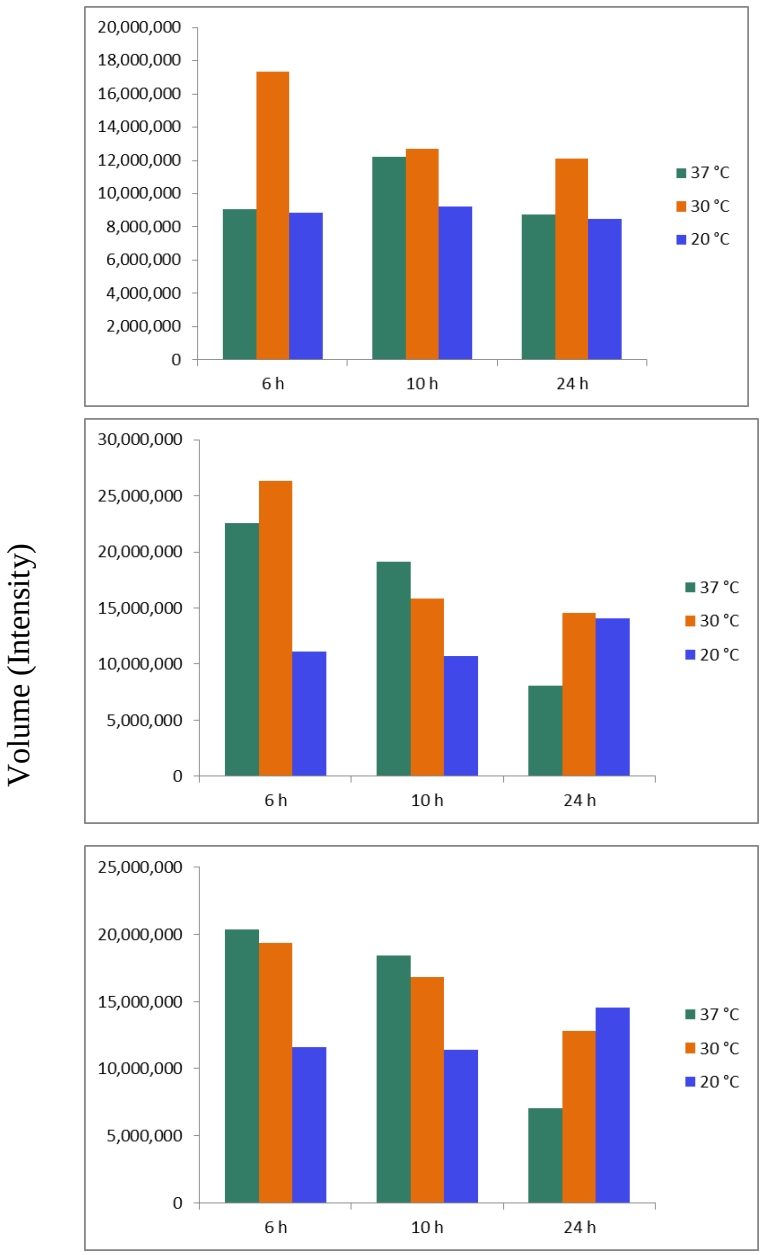

Figure A1 SynB-G$_4$S-MBP raw data. The y-axis represents the volume (intensity) and the X-axis represents the induction time in hours

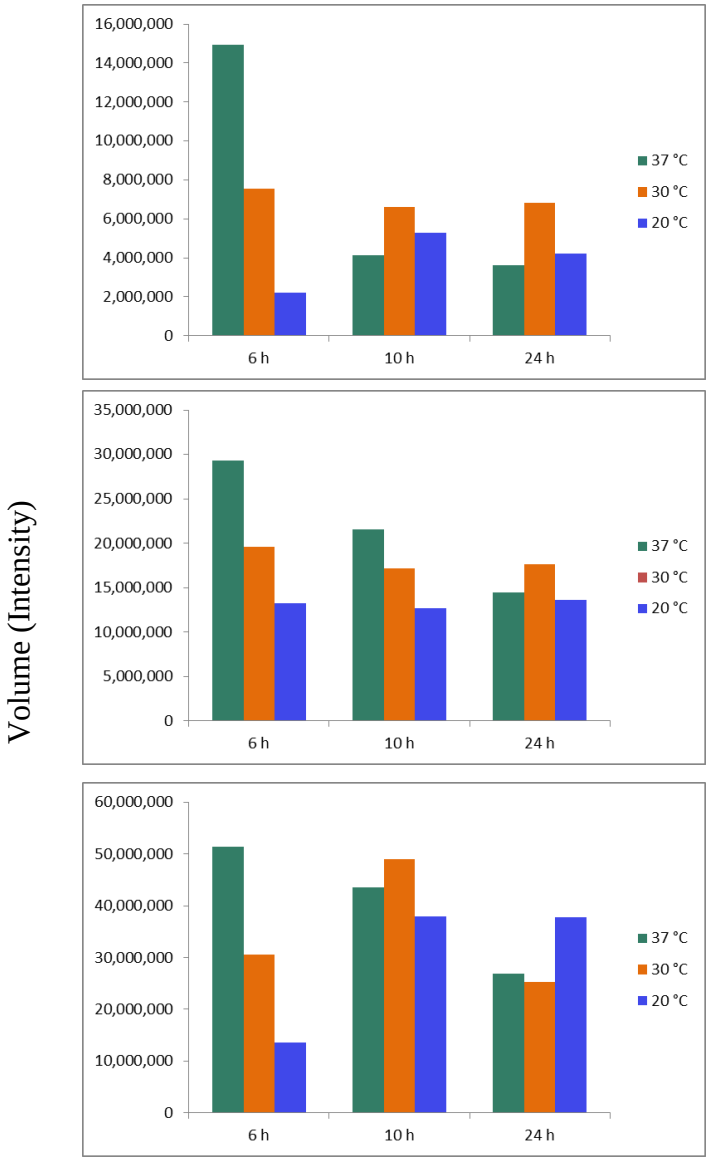

Figure A2 Hst-5-G_4S-MBP raw data. The Y-axis represents the volume (intensity) and the X-axis represents the induction time in hours

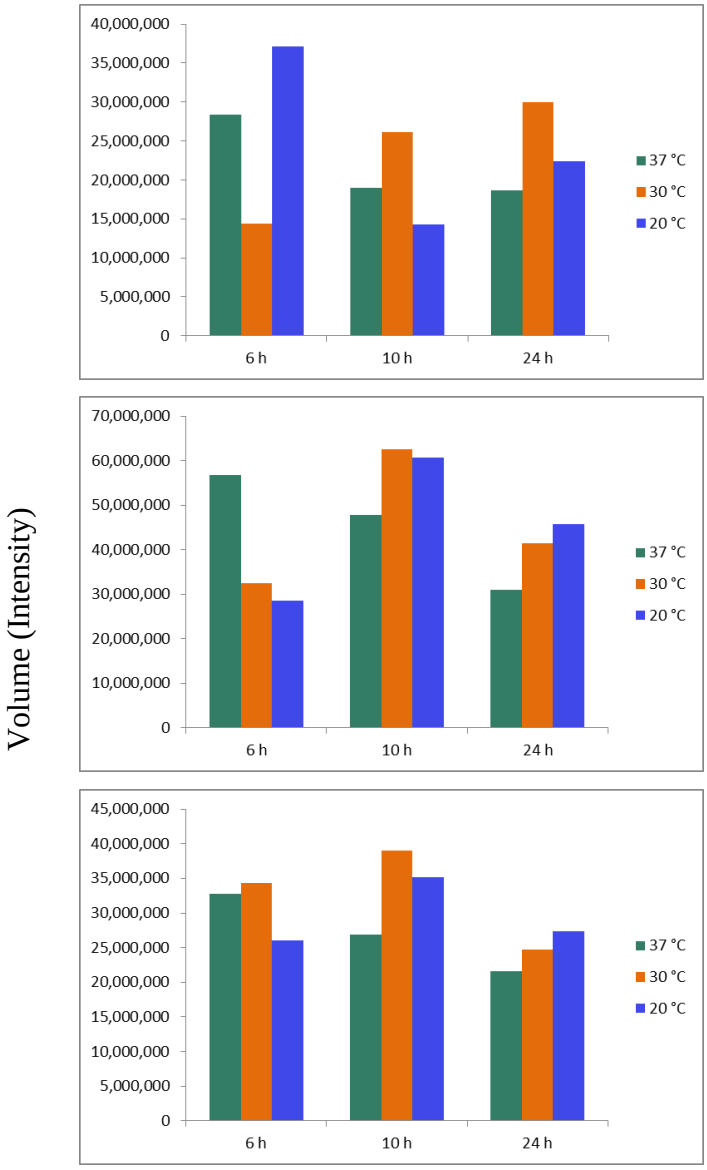

Figure A3 MPG-5-G_4S-MBP raw data. The Y-axis represents the volume (intensity) and the X-axis represents the induction time in hours

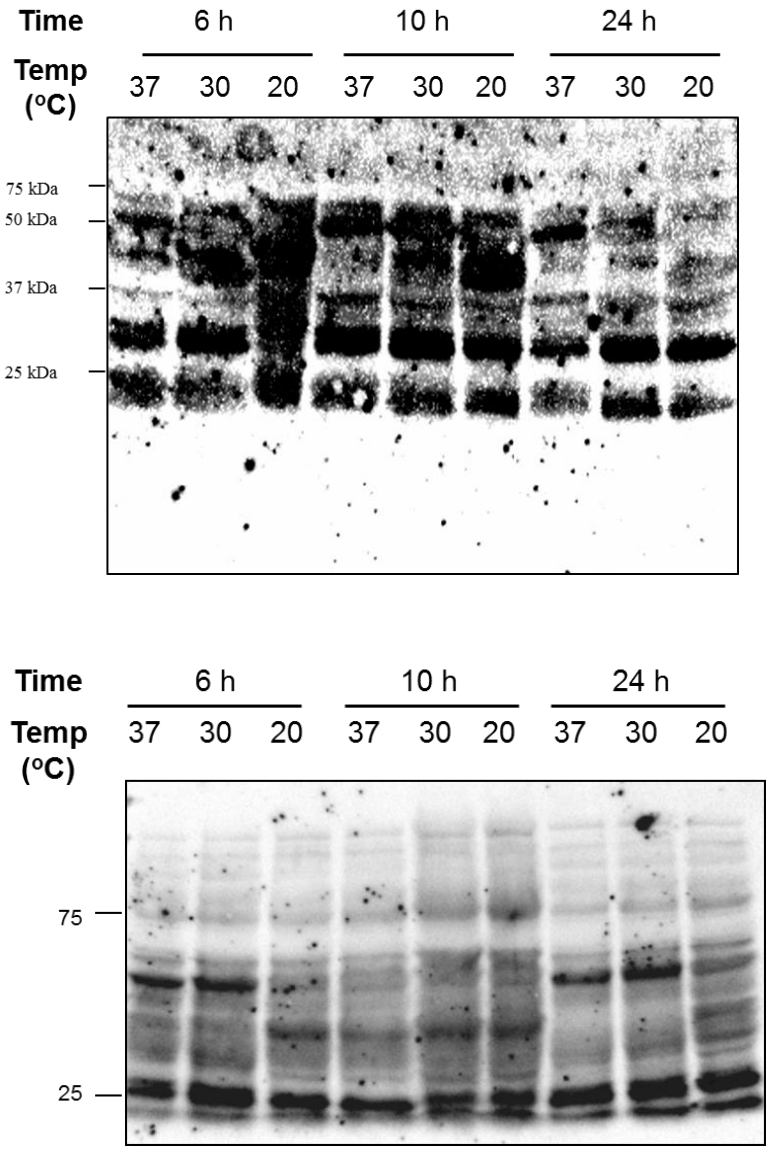

Figure A4 SynB-G_4S-GFP raw data.

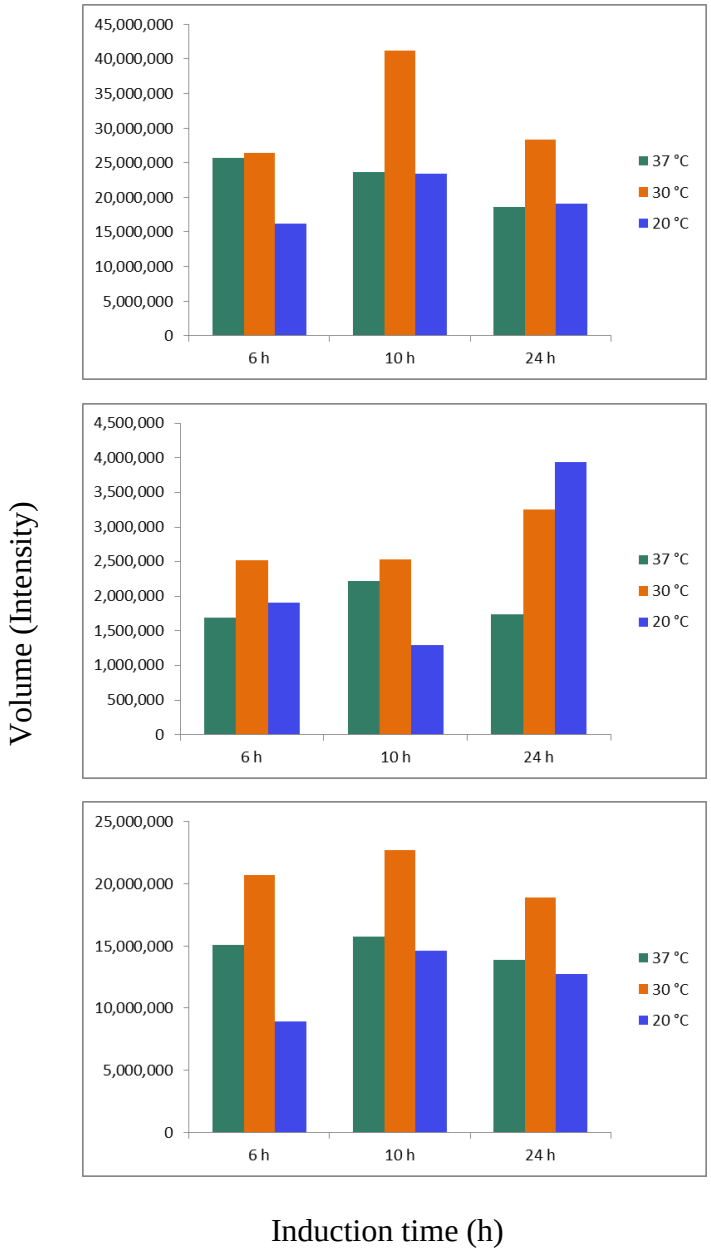

Figure A5 Hst-5-G_4S-GFP raw data. The Y-axis represents the volume (intensity) and the X-axis represents the induction time in hours

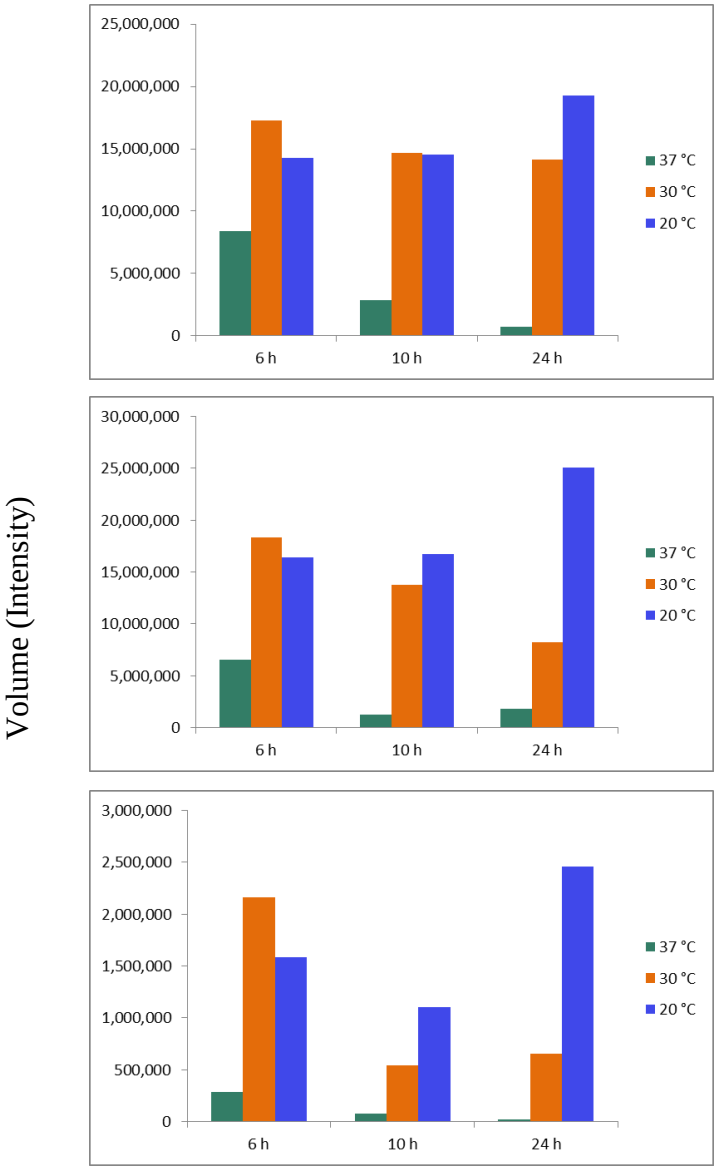

Figure A6 MPG-G_4S-GFP raw data. The Y-axis represents the volume (intensity) and the X-axis represents the induction time in hours

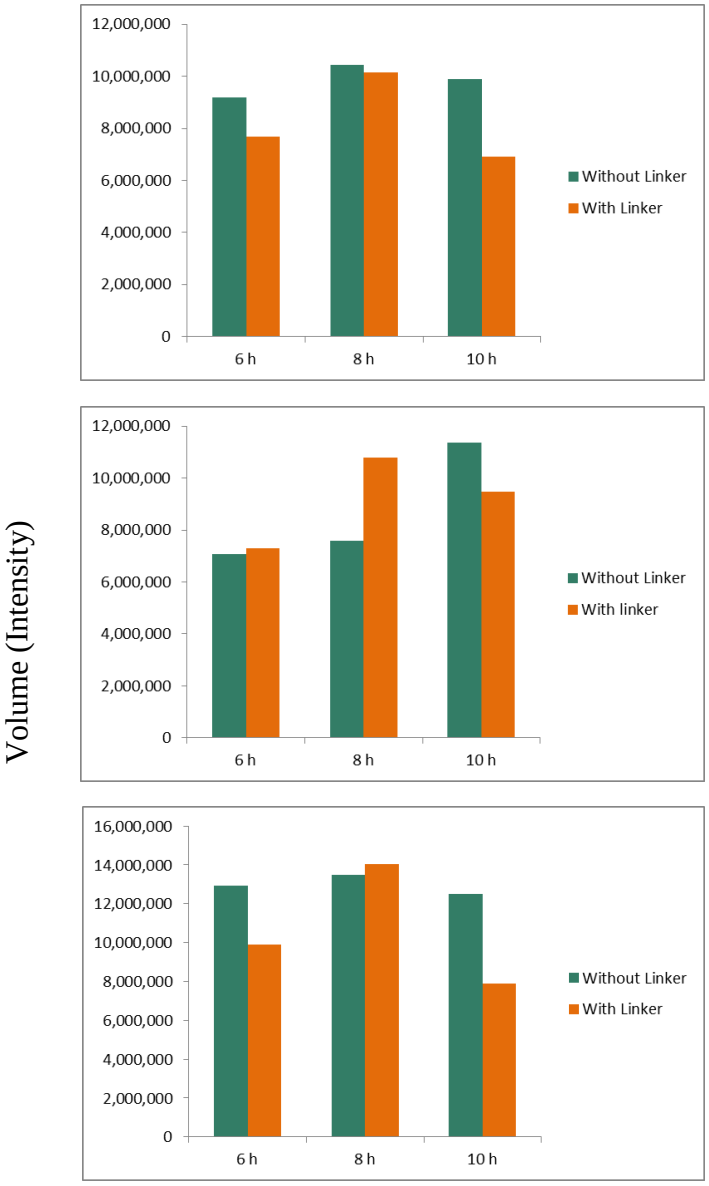

Figure A7 NPFSD-G_4S-BCCP raw data. The Y-axis represents the volume (intensity) and the X-axis represents the induction time in hours

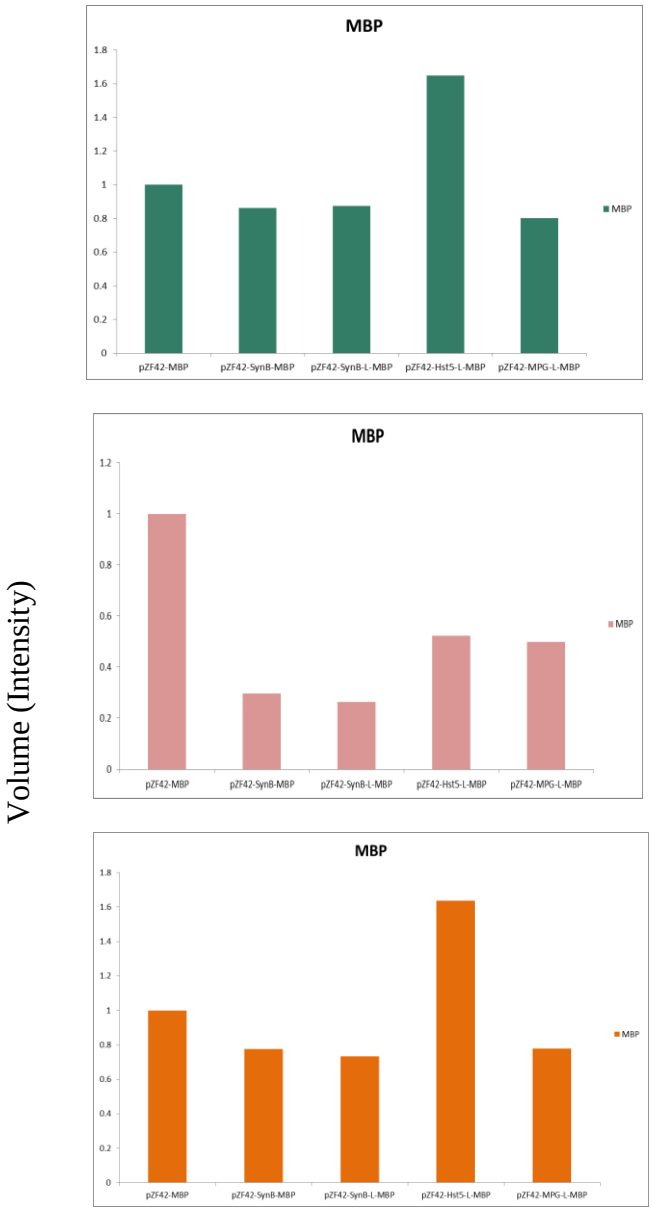

Figure A8 All MBP constructs raw data. The Y-axis represents the volume (intensity) and the X-axis represents the induction time in hours.

References

1 Mueller, N. H., Ammar, D. A. & Petrash, J. M. Cell penetration peptides for enhanced entry of alphaB-crystallin into lens cells. *Invest Ophthalmol Vis Sci* **54**, 2-8, doi:10.1167/iovs.12-10947 (2013).

2 Elmquist, A., Lindgren, M., Bartfai, T. & Langel, U. VE-cadherin-derived cell-penetrating peptide, pVEC, with carrier functions. *Exp Cell Res* **269**, 237-244, doi:10.1006/excr.2001.5316 (2001).

3 Koren, E. & Torchilin, V. P. Cell-penetrating peptides: breaking through to the other side. *Trends Mol Med* **18**, 385-393, doi:10.1016/j.molmed.2012.04.012 (2012).

4 Ormo, M. *et al.* Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **273**, 1392-1395 (1996).

5 Copolovici, D. M., Langel, K., Eriste, E. & Langel, U. Cell-penetrating peptides: design, synthesis, and applications. *ACS Nano* **8**, 1972-1994, doi:10.1021/nn4057269 (2014).

6 Xiang, S. & Tong, L. Crystal structures of human and Staphylococcus aureus pyruvate carboxylase and molecular insights into the carboxyltransfer reaction. *Nat Struct Mol Biol* **15**, 295-302, doi:10.1038/nsmb.1393 (2008).

7 Morris, M. C., Deshayes, S., Heitz, F. & Divita, G. Cell-penetrating peptides: from molecular mechanisms to therapeutics. *Biol Cell* **100**, 201-217, doi:10.1042/BC20070116 (2008).

8 Richardson, J. P. & Moyes, D. L. Adaptive immune responses to Candida albicans infection. *Virulence* **6**, 327-337, doi:10.1080/21505594.2015.1004977 (2015).

9 Braun, B. R. & Johnson, A. D. Control of filament formation in Candida albicans by the transcriptional repressor TUP1. *Science* **277**, 105-109 (1997).

10 Frankel, A. D. & Pabo, C. O. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **55**, 1189-1193 (1988).

11 Green, M. & Loewenstein, P. M. Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell* **55**, 1179-1188 (1988).

12 Derossi, D., Joliot, A. H., Chassaing, G. & Prochiantz, A. The 3rd Helix of the Antennapedia Homeodomain Translocates through Biological-Membranes. *Journal of Biological Chemistry* **269**, 10444-10450 (1994).

13 Gong, Z., Walls, M. T., Karley, A. N. & Karlsson, A. J. Effect of a Flexible Linker on Recombinant Expression of Cell-Penetrating Peptide Fusion Proteins and Their Translocation into Fungal Cells. *Mol Biotechnol* **58**, 838-849, doi:10.1007/s12033-016-9983-5 (2016).

14 Helmerhorst, E. J. *et al.* The cellular target of histatin 5 on Candida albicans is the energized mitochondrion. *J Biol Chem* **274**, 7286-7291 (1999).

15 Peter G. Pappas *et al.* Clinical Practice Guideline for the Management of Candidiasis: 2016 Update by the Infectious Diseases Society of America. *Clinical Infectious Diseases* **62** (2016).
16 Cannon, R. D. *et al.* Candida albicans drug resistance another way to cope with stress. *Microbiology* **153**, 3211-3217, doi:10.1099/mic.0.2007/010405-0 (2007).
17 Prevention., C. f. D. C. a. & States, A. R. T. i. t. U. *http://www.cdc.gov/drugresistance/threat-report-2013/Accessed July 2014*, 2013).
18 White, T. C., Marr, K. A. & Bowden, R. A. Clinical, cellular, and molecular factors that contribute to antifungal drug resistance. *Clinical Microbiology Reviews* **11**, 382-+ (1998).
19 Alhakamy, N. A., Nigatu, A. S., Berkland, C. J. & Ramsey, J. D. Noncovalently associated cell-penetrating peptides for gene delivery applications. *Ther Deliv* **4**, 741-757, doi:10.4155/tde.13.44 (2013).
20 Royle, S. J. The cellular functions of clathrin. *Cell Mol Life Sci* **63**, 1823-1832, doi:10.1007/s00018-005-5587-0 (2006).
21 Rejman, J., Oberle, V., Zuhorn, I. S. & Hoekstra, D. Size-dependent internalization of particles via the pathways of clathrin- and caveolae-mediated endocytosis. *Biochem J* **377**, 159-169, doi:10.1042/BJ20031253 (2004).
22 Alberts, B. *et al. Molecular Biology of the Cell*. Sixth edn, 732-734 (2014).
23 Raagel, H., Saalik, P. & Pooga, M. Peptide-mediated protein delivery-which pathways are penetrable? *Biochim Biophys Acta* **1798**, 2240-2248, doi:10.1016/j.bbamem.2010.02.013 (2010).
24 Kristensen, M., Birch, D. & Morck Nielsen, H. Applications and Challenges for Use of Cell-Penetrating Peptides as Delivery Vectors for Peptide and Protein Cargos. *Int J Mol Sci* **17**, doi:10.3390/ijms17020185 (2016).
25 Shin, M. C. *et al.* Cell-penetrating peptides: achievements and challenges in application for cancer treatment. *J Biomed Mater Res A* **102**, 575-587, doi:10.1002/jbm.a.34859 (2014).
26 Fawell, S. *et al.* Tat-Mediated Delivery of Heterologous Proteins into Cells. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 664-668, doi:DOI 10.1073/pnas.91.2.664 (1994).
27 Lim, K. J. *et al.* A cancer specific cell-penetrating peptide, BR2, for the efficient delivery of an scFv into cancer cells. *PLoS One* **8**, e66084, doi:10.1371/journal.pone.0066084 (2013).
28 Masuda, R., Yamamoto, K. & Koide, T. Cellular Uptake of IgG Using Collagen-Like Cell-Penetrating Peptides. *Biol Pharm Bull* **39**, 130-134, doi:10.1248/bpb.b15-00548 (2016).
29 Schwarze, S. R., Ho, A., Vocero-Akbani, A. & Dowdy, S. F. In vivo protein transduction: delivery of a biologically active protein into the mouse. *Science* **285**, 1569-1572 (1999).

30 Kameyama, S. *et al.* Effects of cell-permeating peptide binding on the distribution of 125I-labeled Fab fragment in rats. *Bioconjug Chem* **17**, 597-602, doi:10.1021/bc050258k (2006).
31 Gong Z. & Karlsson AJ. Translocation of cell-penetrating peptides into Ca*ndida* fungal pathogens. *Protein science* (2017).
32 Elmquist, A., Hansen, M. & Langel, U. Structure-activity relationship study of the cell-penetrating peptide pVEC. *Biochim Biophys Acta* **1758**, 721-729, doi:10.1016/j.bbamem.2006.05.013 (2006).
33 Jamal Temsamani & Laruelle, C. SynB peptide vectors: A new approach to drug delivery. *Chimica Oggi / Chemistry Today* **28** (2010).
34 Tan, P. K., Howard, J. P. & Payne, G. S. The sequence NPFXD defines a new class of endocytosis signal in Saccharomyces cerevisiae. *J Cell Biol* **135**, 1789-1800 (1996).
35 Rajarao, G. K., Nekhotiaeva, N. & Good, L. The signal peptide NPFSD fused to ricin A chain enhances cell uptake and cytotoxicity in Candida albicans. *Biochem Biophys Res Commun* **301**, 529-534 (2003).
36 Hermanson, G. T. in *Bioconjugate Techniques* Ch. 1, 1233 (Academic Press, 2008).
37 Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol* **5**, 172, doi:10.3389/fmicb.2014.00172 (2014).
38 Marini, G. *et al.* Experimental design approach in recombinant protein expression: determining medium composition and induction conditions for expression of pneumolysin from Streptococcus pneumoniae in Escherichia coli and preliminary purification process. *BMC Biotechnol* **14**, 1, doi:10.1186/1472-6750-14-1 (2014).
39 Michiue, H. *et al.* The NH2 terminus of influenza virus hemagglutinin-2 subunit peptides enhances the antitumor potency of polyarginine-mediated p53 protein transduction. *J Biol Chem* **280**, 8285-8289, doi:10.1074/jbc.M412430200 (2005).
40 Zhang, J. H., Yun, J., Shang, Z. G., Zhang, X. H. & Pan, B. R. Design and optimization of a linker for fusion protein construction. *Progress in Natural Science* **19**, 1197-1200, doi:10.1016/j.pnsc.2008.12.007 (2009).
41 Chen, X., Zaro, J. L. & Shen, W. C. Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev* **65**, 1357-1369, doi:10.1016/j.addr.2012.09.039 (2013).
42 Yu, K., Liu, C., Kim, B. G. & Lee, D. Y. Synthetic fusion protein design and applications. *Biotechnol Adv* **33**, 155-164, doi:10.1016/j.biotechadv.2014.11.005 (2015).
43 Esposito, D. & Chatterjee, D. K. Enhancement of soluble protein expression through the use of fusion tags. *Curr Opin Biotechnol* **17**, 353-358, doi:10.1016/j.copbio.2006.06.003 (2006).
44 Drin, G., Cottin, S., Blanc, E., Rees, A. R. & Temsamani, J. Studies on the internalization mechanism of cationic cell-penetrating peptides. *J Biol Chem* **278**, 31192-31201, doi:10.1074/jbc.M303938200 (2003).

45 Du, H. *et al.* Human Salivary Protein Histatin 5 Has Potent Bactericidal Activity against ESKAPE Pathogens. *Front Cell Infect Microbiol* **7**, 41, doi:10.3389/fcimb.2017.00041 (2017).
46 Puri, S. & Edgerton, M. How does it kill?: understanding the candidacidal mechanism of salivary histatin 5. *Eukaryot Cell* **13**, 958-964, doi:10.1128/EC.00095-14 (2014).
47 Khan, S. A. *et al.* Impaired Histatin-5 Levels and Salivary Antimicrobial Activity against C. albicans in HIV Infected Individuals. *J AIDS Clin Res* **4**, doi:10.4172/2155-6113.1000193 (2013).
48 Mochon, A. B. & Liu, H. The antimicrobial peptide histatin-5 causes a spatially restricted disruption on the Candida albicans surface, allowing rapid entry of the peptide into the cytoplasm. *PLoS Pathog* **4**, e1000190, doi:10.1371/journal.ppat.1000190 (2008).
49 Jang, W. S., Li, X. S., Sun, J. N. & Edgerton, M. The P-113 fragment of histatin 5 requires a specific peptide sequence for intracellular translocation in Candida albicans, which is independent of cell wall binding. *Antimicrob Agents Chemother* **52**, 497-504, doi:10.1128/AAC.01199-07 (2008).
50 Dong, J., Vylkova, S., Li, X. S. & Edgerton, M. Calcium blocks fungicidal activity of human salivary histatin 5 through disruption of binding with Candida albicans. *J Dent Res* **82**, 748-752, doi:10.1177/154405910308200917 (2003).
51 Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802-805 (1994).
52 Choi-Rhee, E. & Cronan, J. E. The biotin carboxylase-biotin carboxyl carrier protein complex of Escherichia coli acetyl-CoA carboxylase. *J Biol Chem* **278**, 30806-30812, doi:10.1074/jbc.M302507200 (2003).
53 Kapust, R. B. & Waugh, D. S. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8**, 1668-1674, doi:10.1110/ps.8.8.1668 (1999).
54 di Guan, C., Li, P., Riggs, P. D. & Inouye, H. Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. *Gene* **67**, 21-30 (1988).
55 Fal, R. R. & Vagelos, P. R. Acetyl Coenzyme A Carboxylase - Molecular Forms And Subunit Composition Of Biotin Carboxyl Carrier Protein*. *The Journal Of Biological Chemistry* **247**, 8005-8015 (1972).
56 UniProt. *http://www.uniprot.org/uniprot/P0ABD8.*
57 Tsien, R. Y. The green fluorescent protein. *Annu Rev Biochem* **67**, 509-544, doi:10.1146/annurev.biochem.67.1.509 (1998).
58 Uniprot. *http://www.uniprot.org/uniprot/P42212*.
59 Spurlino, J. C., Lu, G. Y. & Quiocho, F. A. The 2.3-a Resolution Structure of the Maltose-Binding or Maltodextrin-Binding Protein, a Primary Receptor of Bacterial Active-Transport and Chemotaxis. *Journal of Biological Chemistry* **266**, 5202-5219 (1991).

60 Raran-Kurussi, S. & Waugh, D. S. The ability to enhance the solubility of its fusion partners is an intrinsic property of maltose-binding protein but their folding is either spontaneous or chaperone-mediated. *PLoS One* **7**, e49589, doi:10.1371/journal.pone.0049589 (2012).

61 Nallamsetty, S., Austin, B. P., Penrose, K. J. & Waugh, D. S. Gateway vectors for the production of combinatorially-tagged His6-MBP fusion proteins in the cytoplasm and periplasm of Escherichia coli. *Protein Sci* **14**, 2964-2971, doi:10.1110/ps.051718605 (2005).

62 MyBioSource. *https://www.mybiosource.com/prods/Recombinant-Protein/Maltose-Binding-Protein/MBP/datasheet.php?products_id=203159*.

63 Robichon, C., Luo, J., Causey, T. B., Benner, J. S. & Samuelson, J. C. Engineering Escherichia coli BL21(DE3) derivative strains to minimize E. coli protein contamination after purification by immobilized metal affinity chromatography. *Appl Environ Microbiol* **77**, 4634-4646, doi:10.1128/AEM.00119-11 (2011).

64 Richard, J. P. *et al.* Cell-penetrating peptides: a re-evaluation of the mechanism of cellular uptake. *The Journal Of Biological Chemistry* **278**, 585–590 (2003).

65 Holm, T., Netzereab, S., Hansen, M., Langel, U. & Hallbrink, M. Uptake of cell-penetrating peptides in yeasts. *FEBS Lett* **579**, 5217-5222, doi:10.1016/j.febslet.2005.07.099 (2005).

66 Lee, C. H., Chou, C. C., Hsu, M. F. & Wang, A. H. Determining the N-terminal orientations of recombinant transmembrane proteins in the Escherichia coli plasma membrane. *Sci Rep* **5**, 15086, doi:10.1038/srep15086 (2015).

67 Ilsen, C. Y. T. Z. & Onsan, B. K. Optimization of Starting Time and Period of Induction and Inducer Concentration in the Production of the Restriction Enzyme EcoRI from Recombinant *E. coli*. *Turkish Journal of Chemistry* **22**, 221 – 226 (1998).

68 Li, S. C. & Kane, P. M. The yeast lysosome-like vacuole: endpoint and crossroads. *Biochim Biophys Acta* **1793**, 650-663, doi:10.1016/j.bbamcr.2008.08.003 (2009).