Title	Production of a recombinant, 40 amino acid helix-loop-helix peptide (EF2-GGC) for incorporation into a novel affinity purification system
Authors(s)	Vesey, Orla
Publication date	2016
Publication information	Vesey, Orla. "Production of a Recombinant, 40 Amino Acid Helix-Loop-Helix Peptide (EF2-GGC) for Incorporation into a Novel Affinity Purification System." University College Dublin. School of Biomolecular and Biomedical Science, 2016.
Publisher	University College Dublin. School of Biomolecular and Biomedical Science
Item record/more information	http://hdl.handle.net/10197/8570

Production of a Recombinant, 40 Amino Acid Helix-Loop-Helix Peptide (EF2-GGC) for Incorporation into a Novel Affinity Purification System

Orla Vesey BSc (Hons)

Student Number: 14207657

This thesis is submitted to University College Dublin in fulfilment of the requirements for the degree of Masters by Research in the College of Science.

January 2016

Based on research carried out in the Conway Institute, School of Biomolecular and Biomedical Science, University College Dublin

Dr. Keith Murphy, Head of School

Dr. David O'Connell, Principal Supervisor

Table of Contents

Abstract

Fragment complementation between EF1 and EF2, calcium binding protein subdomains of the protein calbindin D9k, forms the basis behind a novel affinity protein purification system enabling rapid, highly specific capture and elution of pure EF1-tagged protein in physiological conditions. Construction of an affinity resin for this purification system relies on efficient production of the EF2 peptide as the affinity ligand to be displayed on a nanoscale scaffold.

In this thesis a protocol was developed for production of the EF2 peptide via a recombinant fusion protein construct incorporating the EF2 peptide with an added terminal GGC. Isolation of pure EF2-GGC peptide from this fusion protein followed. This protocol involved expression of this recombinant protein through the bacterial cell host, *Escherichia coli (E.coli),* and purification via native cell lysis and immobilized metal affinity chromatography (IMAC), both optimized for yield and purity. Subsequent processing steps of this recombinant fusion protein consisted of TEV protease cleavage and additional affinity purification steps to obtain the peptide in a pure formulation for further steps in affinity resin manufacture.

Initial protein expression investigations demonstrated that the EF2-GGC recombinant fusion protein was soluble and expressed well in both autoinduction and isopropyl β-D-1-thiogalactopyranoside (IPTG) induction systems. Ni-NTA IMAC purification of this recombinant fusion protein was successful with a yield of approximately ~50 mg/L. Proteolytic cleavage of the EF2-GGC recombinant fusion protein successfully releasing the EF2-GGC peptide of interest was optimised. A second round of Ni-NTA IMAC, ion exchange chromatography (IEX) and a final polishing step of size exclusion chromatography (SEC) retrieved the peptide in a highly pure form, suitable for subsequent resin coupling steps.

While this protocol was optimised to provide 6 - 12 mg/L of recombinant EF2-GGC peptide, SEC analysis confirmed that the peptide was prone to significant aggregation. Further investigations into the peptide buffer formulation failed to completely reverse aggregation. Therefore, while the expression and purification protocols developed during this project have been effective in isolating a pure EF2-GGC peptide, further work to identify an effective formulation for the purified peptide is required.

Declaration

I declare that this thesis is my own original work, it contains nothing which is the outcome of work done by others or in collaboration with others. All work was completed while registered as a student for the degree of Masters by Research in the School of Biomolecular and Biomedical Science, University College Dublin, from September 2014 to January 2016.

I also declare that this work was not submitted, in whole or as part to any other university or college for any other degree or qualification.

Signed: _____   Date: _____

       Orla Vesey

Acknowledgements

Firstly, I would like to express my sincerest gratitude to my supervisor Dr. David O'Connell. For lending me his guidance, his knowledge and his energetic motivational speeches which has steered and encouraged me throughout every step of this project, good and bad times, which without, this thesis would not have come together. I immensely appreciate everything, thank you.

I want to thank all members of the O'Connell lab group, past and present. "Dr. to be" Niamh Murphy for welcoming me so warmly into the lab and for not only being a wonderfully eccentric friend but also for being the lab sister I never knew I could have. "Dr. to be" Gavin (McGosling) McGauran for answering and laughing at the many, small, yet important questions I struggled with during this project, wise words I genuinely would not have survived without. Dr. Darragh O'Donovan for being the person to aspire to not only scientifically but also in your unique quick smart witty nature. And Dr. Eugene Mahon for during your brief time gave so much skill and raw intelligence, it was not overlooked.

I would also like to thank my "Nutgrove Avenue Huns", Grainne, Amanda and Aoife for making my time in Dublin a crazy and brilliant experience and for being the best people to ever have lived with this past year.

Finally I would like to thank my family. Their never failing support, love and encouragement even when they hadn't the slightest notion what I was talking about is something I will never take for granted.

Abbreviations

2X Tryptone Yeast extract	2X TY
6xHis tagged Human Carbonic Anhydrase	6xHis CA
Bicinchoninic Acid	BCA
Bovine Serum Albumin	BSA
Calmodulin Binding Peptide	CBP
Deoxyribonucleic Acid	DNA
Diethylaminoethanol	DEAE
Dithiothreitol	DTT
Escherichia coli	*E.coli*
Ethylene Glycol Tetraacetic Acid	EGTA
Ethylenediaminetetraacetic Acid	EDTA
Flag-Acidic-Target Tag	FATT
Flow-Through	FT
Glutathione S-transferase	GST
Green Fluorescent Protein	GFP
Guanidine hydrochloric acid	GuHCl
Hydrophobic Interaction Chromatography	HIC
Immobilized Metal Affinity Chromatography	IMAC
Ion Exchange Chromatography	IEX
Isoelectric Point	pI
Isopropyl β-D-1-thiogalactopyranoside	IPTG
Kilo Daltons	kDa
Lameilli Sample Buffer	LSB
Maltose Binding Protein	MBP
Molecular Weight	mw

Molecular Weight Cut Off	MWCO
Monoclonal Antibody	Ab
Nanoparticles	NPs
Nitrilotriacetic Acid	NTA
Optical Density	OD
Polyacrylamide Gel Electrophoresis	PAGE
Polyethersulfone	PES
Polyethyleneimine	PEI
Size Exclusion Chromatography	SEC
Sodium Dodecyl Sulphate	SDS
Solid-Phase Peptide Synthesis	SPPS
Starch-Binding Domain	SBD
Surface Plasmon Resonance	SPR
Tandem Affinity Purification	TAP
Tobacco Etch Virus	TEV
Transfer Ribonucleic Acids	tRNAs
Tris(2-carboxyethyl)phosphine	TCEP
Tyrptone Yeast Extract	TYE

List of Figures

List of Tables

Chapter 1

Introduction to Research

1.1 Methods in Protein Purification

The discovery and development of proteins as therapeutics, or as targets for therapeutics to treat disease, continues to drive research into their isolation and preparation in the native form. To facilitate this need for protein isolation and purification, there is an extensive range of methods existing today including size exclusion chromatography (SEC), affinity chromatography, hydrophobic interaction chromatography (HIC) and ion exchange chromatography (IEX). Some of these techniques exploit the protein of interest's individual, native characteristics e.g. molecular weight, electrical charge, in order to achieve separation and others may require the incorporation of a tag in order to obtain proteins in pure fractions. These techniques are routinely used in both academic and industrial research from initial protein investigation studies to large scale production of therapeutic proteins for treatment of disease. Pure preparations of a protein are valuable to facilitate subsequent investigations such as functional assays, protein-protein interaction studies, or mass spectrometry analysis.

In this thesis, a novel affinity chromatography system for the purification of proteins, which has not been previously published, will be discussed. The mechanism behind this affinity system lies in the specific and high affinity binding interaction between two helix-loop-helix protein domains, EF1 and EF2. The production of the EF2 protein domain as an affinity ligand for building a nanoscale affinity purification resin was investigated in this thesis and tested for its ability to purify EF1 tagged recombinant proteins from highly complex mixtures.

1.1.1 Affinity Chromatography and Tags

Protein purification by means of affinity chromatography is a well-established area. There are numerous types of affinity chromatography in existence with popular methods including ligand binding, immuno-affinity binding and metal binding. These methods require the presence of an affinity tag, or multiple tags, namely combinatorial tagging, located at the N or C-terminus of the protein to facilitate capture and purification of the protein via a corresponding affinity resin (Wood, 2014). Essentially the principle of this approach lies behind the specific affinity of the tag on a protein to a corresponding resin. For example, the Flag tag has a high affinity to anti-flag monoclonal antibodies immobilized on a solid resin support (Table 1.1). The attraction of this tag to this antibody affinity ligand results in binding of the tagged protein to the

solid support. In this way other, non-tagged proteins will not bind and flow past the resin, essentially separating this tagged protein from a complex mixture of proteins. The bound protein can be released from the affinity resin via introduction of a competitive binding agent of a higher affinity, also known as an affinity ligand/eluting agent (Lichty *et al.,* 2005). In the example of the Flag tag, the eluting agent can be either Flag peptide or EDTA (Table 1.1). The Flag tag system is an example of immune-affinity chromatography. Other types of affinity chromatography are based on the same principles as described here.

Affinity tags are effective tools not only for purification purposes but can also play roles in enhancing protein solubility, facilitating detection e.g. antibody binding in Western blotting and characterization of recombinant proteins, e.g. immunoprecipitation techniques (Young *et al.,* 2012).

Maltose binding protein (MBP) is one of the larger sized affinity tags, 396 amino acids in length which can bind amylose resins but is more often employed to improve solubility and stability of a protein instead of for affinity purification purposes (Kapust & Waugh, 1999; Lichty *et al.,* 2005) (Table 1.1). This solubilisation function aids in solving the challenge posed by inclusion body formation which occurs when the protein aggregates forming insoluble, inactive protein accumulations intracellularly (Rudolph & Lilie, 1996). Other affinity tags, glutathione S-transferase (GST) and thioredoxin are also known to impart solubilisation properties, however MBP has previously been shown by Kapust & Waugh (1999) to be the most effective.

Strep II is an 8 amino acid affinity tag that has been used not only for purification but also for protein detection and immobilization (Terpe, 2003), as shown in Table 1.1. A striking difference between MBP and *Strep* II is in the number of amino acids that they are comprised of in that MBP has 388 more amino acids than *Strep* II. Indeed, this highlights one of many advantages that affinity tags have to offer in that there is quite a broad range of different sized tags, enabling selection of a specific affinity tag most suited to the specific requirements of the protein to be tagged. Affinity chromatography using *Strep* II operates via streptavidin immobilized resins also known commercially as S*trep*-Tactin®, to purify tagged proteins (Skerra & Schmidt, 1999). There are antibodies against the *Strep* II tag available which may have an enzymatic or fluorescent marker facilitating protein detection in Western blotting. Non-conjugated antibodies against *Strep* II can be used to immobilize *Strep* II tagged proteins on solid surfaces for example gold microchips for SPR analysis (Schmidt & Skerra, 2007). In this way, *Strep* II tagged proteins can be characterized for protein-to-protein binding affinities or *in vivo* protein complex formation/interactions.

Some novel tags to have emerged more recently in the literature include the starch-binding domain (SBD) tag, a heptamer tag derived from phage display, a silkworm β-1, 3-glucan recognition protein tag and Flag-Acidic-Target Tag (FATT) (Wood, 2014) (Islam *et al.,* 2014) (Table 1.1). The development of these new affinity protein tags are part of a continuous effort to improve protein purification for simplified, less expensive and time consuming processes. Particularly the SBD tag, heptamer tag and the silkworm β-1, 3-glucan recognition protein tag bind to inexpensive materials namely raw corn or other vegetable starches, ceramic fluorapatite and curdlan, a high molecular weight (mw) polymer of glucose, respectively. These cheap raw materials are attractive for larger scale protein purification procedures often carried out in industrial settings where the need for materials in bulk can be quite expensive. These new tag options could significantly reduce protein purification costs overall and/or increase production for the same price, either way equating to a greater reduction in cost spent and thus a greater profit.

The EF1 tag (Table 1.1) is another novel affinity tag that also aims to improve current affinity protein purification methods. The highly specific binding of the EF1 protein domain to the EF2 protein domain occurs only in the presence of calcium and with a very high affinity of binding (picomolar) (Dell'Orco *et al.*, 2005). Chelating calcium ions from the system will therefore cause dissociation of these two domains. This calcium switch is a key step in exploiting this EF1-EF2 system for protein purification procedures. Tagging a protein with EF1 and introducing an EF2 affinity resin in calcium containing buffers will allow for highly specific binding of only EF1 tagged proteins to this affinity resin. Washing away unbound contaminating proteins will purify the bound protein from unwanted proteins. Following washing steps, removal of calcium will disrupt EF1-EF2 binding and elute the EF1 tagged protein in a pure preparation. Operation of this system in physiological buffers means that the eluted protein is ready for downstream analytical studies immediately. This eliminates the need for extra steps such as dialysis to prepare the protein in a working form, considerably reducing experimental timeframes comparing to current protein purification methods.

Table 1.1 Comparative summary of affinity tags in common practice. Table adapted from: (Young *et al.,* (2012), Wood, (2014), Waugh, (2005), Lichty *et al.,* (2005), Terpe, (2003)

Tag	Size (aa)	Resin	Eluting Agent	Advantages	Limitations
EF1	40	EF2 peptide ligand resins, agarose or NP	~1 mM EDTA	Elution in physiological buffer	May not be suitable for some metal binding proteins
His	6 - 10	Ni-NTA agarose/nanoparticles (other metals also e.g. Cu^{2+}, Fe^{2+}, Zn^{2+}	~250 mM imidazole	Can be performed in native or denaturing conditions,	Requires subsequent processing e.g. dialysis/buffer exchange
MBP	396	Amylose	Maltose	Multi-functional tag; enhanced solubility and purification	Large tag size may need to be cleaved following purification
Strep II	8	Strepavidin	Biotin derivatives e.g. desthiobiotin	Binds with very high affinity	May not be accessible due to small size
CBP	26	Calmodulin	EGTA	Elution in physiological buffer	May not be suitable for metal binding proteins
GST	~218	GSH-Sepharose	Free Glutathione	Often used to Identify protein interactions	Large tag may affect native protein state
c-myc	10	Monoclonal Ab	Low pH	Allows for Ab detection	Low pH elution
FLAG	8	Anti-FLAG monoclonal Ab	FLAG peptide/ pH 3.0/2-5 mM EDTA	Tag removal by enterokinase	Ab matrix may not be as stable as others
S-tag	15	S-protein	Low pH (2)	Colorimetric detection via WB or assay	Low pH elution conditions
Chitin Binding Domain	51	Chitin	30-50 mM DTT, Cysteine or β-mercap	inexpensive resin material	High DTT elutions may be unsuitable for native protein prep and proteins with disulphide bonds

1.1.2 Polyhistidine Tag

Perhaps one of the best known and most commonly utilised affinity purification systems is the polyhistidine tag or His-tag in immobilized metal affinity chromatography (IMAC). The principle of this method lies in the high affinity of the histidine residues (6 - 10) of the tag to metal ions, e.g. nickel (Ni^{2+}) immobilized on a solid support e.g. agarose, or magnetic beads (Figure 1.1). Metal ions can be immobilized by metal chelators, for example Ni^{2+} can be immobilized by nitrilotriacetic acid (NTA) or iminodiacetic acid (Kuo & Chase, 2011; Lindner *et al.,*1997). Other metals such as cobalt (Co^{2+}), copper (Cu^{2+}) and zinc (Zn^{2+}) can also be used to purify his tagged proteins, all compatible under native or denaturing conditions (Block *et al.,* 2009). The his-tagged proteins will bind exclusively to the metal ion of the affinity resin while other non-his-tagged proteins will not and are therefore easily removed (Figure 1.1).
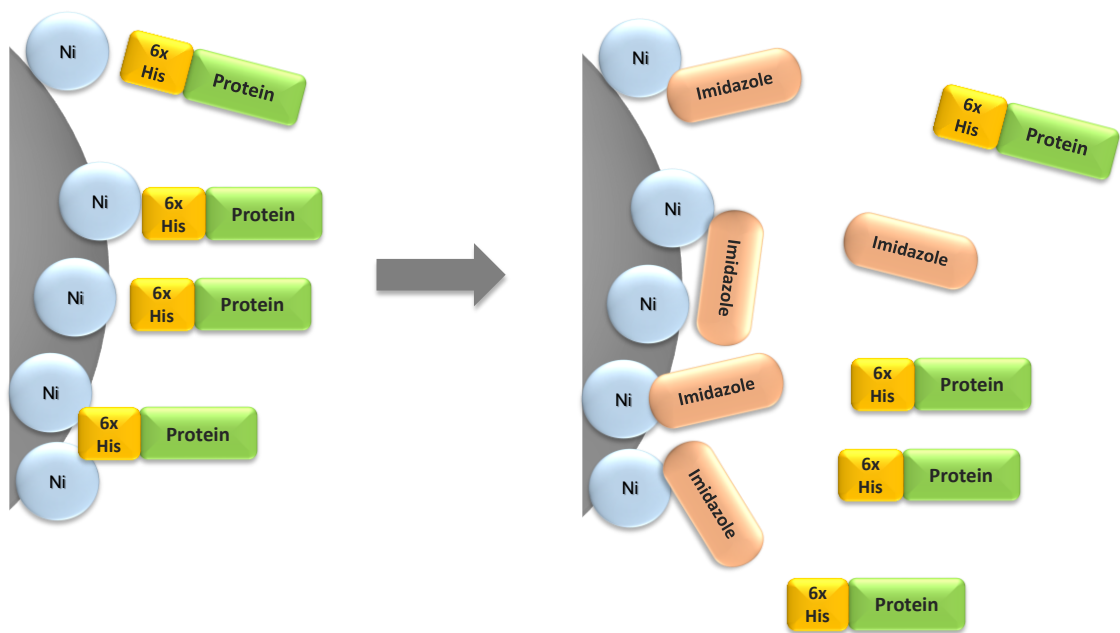
Figure 1.1 Schematic overview of Ni-NTA IMAC binding and elution of a polyhistidine tagged protein

Immediate elution of the bound His tagged proteins can be achieved by introducing high concentrations of imidazole, ~250 mM (Figure 1.1). Imidazole is a planar 5-membred ring, naturally present as a side chain of the histidine residue. It is this ring that is responsible for the binding of the histidine residues to the metal ions of the resin. Therefore, upon addition of free imidazole salts at high concentrations, they will

compete for binding with the tagged protein and displacement of the protein occurs easily. Elution can also be achieved by lowering the pH. Binding of tags proteins occurs optimally at pH 8.0. Lowering the pH during washes will wash away weakly bound proteins and for elution usually pH ~4.5 is used (Bornhorst & Falke, 2000). This works due to the lower pH protonating the imidazole ring of the histidine residues, disrupting bond coordination and therefore binding. This is generally only performed under denaturing protein purifications as lower pH may have adverse effects on the purified protein, i.e. disruption of protein conformation. The durability of the metal ions give IMAC an advantage over more sensitive affinity matrices, immune-affinity resins for example, that may not have reliable storage stability or reusability (Kuo & Chase, 2011).

Despite these attractive attributes IMAC is a method of purification that should be chosen carefully depending on the type of protein to be purified and what the end use is destined to be. Limitations in IMAC purifications can include unsatisfactory purity levels and necessary subsequent processing steps such as dialysis or buffer exchange following purification in order to obtain a suitable working formulation for downstream processing.

1.1.3 Calmodulin Binding Peptide

Other affinity purification systems in use include the calmodulin binding peptide (CBP) system (Table 1.1). CBP is derived from the protein kinase, muscle myosin light chain kinase, a 26 amino acid peptide (4 kilo Daltons (kDa)) from its C-terminus that demonstrates tight binding to calmodulin under low concentrations of calcium (Zheng *et al.,* 1997). Using CBP as an affinity tag proteins can bind to calmodulin resins and are effectively eluted with ethylene glycol tetraacetic acid (EGTA), a chelating agent that will chelate calcium ions important in holding/binding the tagged protein to the calmodulin resin. In this way the eluted protein is then housed in a mild, native, EGTA containing physiological buffer. This is helpful for subsequent assay steps of the protein to be performed without delay as demonstrated by Rigaut *et al.* (1999) where purified protein by CBP was assayed immediately in a gel-shift analysis to test activity of a purified yeast cap binding complex binding to radiolabelled capped RNA. In this manner this purification technique circumvents the limitations of conventional IMAC purifications whereby dialysis of the purified protein may be necessary to achieve the protein in a suitable formulation for immediate assay steps.

1.1.4 Tandem Affinity Purification

Tandem affinity purification (TAP) describes the process by which a recombinant protein may be purified in two or more rounds of affinity chromatography purification procedures. This requires combinatorial tagging of the recombinant protein whereby two or more tags are present. TAP may be sought after for a number of reasons; to ensure a highly pure protein fraction is obtained from extremely complex mixtures, to use one or both tags to confirm the identity of the protein that has been purified for example detection via Western blotting, or to use each tag to confer different properties, i.e. enhance the recombinant protein's solubility using a MBP tag and purify the protein using a His-tag (Waugh, 2005).

1.2 EF1-EF2 Interaction for Affinity Protein Purification

This study describes how the EF1-EF2 fragment complementation interaction may be exploited as a powerful tool for affinity protein purification systems. EF1 and EF2 are EF hand protein domains that originate from the protein calbindin D9k. Calbindin D9k belongs to an evolutionarily conserved family of calcium binding proteins (Linse *et al.,* 1987; Belkacemi *et al.,* 2002). These proteins are important in helping to regulate calcium ion concentrations maintaining correct cellular homeostasis (Belkacemi *et al.,* 2002). Each EF-hand motif is approximately 4.3 kDa in size and with two EF hands making up calbindin D9k this accounts for calbindin D9k's total molecular weight of ~8,500 kDa (Figure 1.2). The EF hand motif is a well-studied structure that comprises of two alpha helix motifs separated by a twelve amino acid loop where one calcium ion may bind to corresponding carboxyl and carbonyl oxygens located in this loop (Linse *et al.,* 1987; Julenius *et al.,* 2002), thus enabling calbindin D9k to bind two calcium ions. Through a mutation of the calbindin D9k protein sequence the two EF hand motifs were separated by chemical cleavage (Dell'Orco *et al.* 2005). The N-terminal of calbindin D9k, amino acids 1 - 43, are known as EF1 and the second half of calbindin D9k, or the second EF hand, amino acids 44 - 75, are known as EF2 (Lindman *et al.*, 2009).
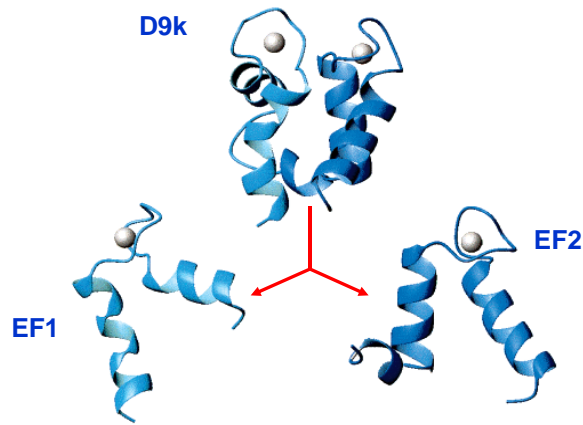
Figure 1.2. Schematic image of Calbindin D9k protein structure, with its two EF hands, EF1 and EF2, separated beneath

The separated EF1 and EF2 hand motifs of calbindin D9k were found to reconstitute with very high affinity (picomolar range) in the presence of the metal ion calcium (Dell'Orco *et al.*, 2005; Julenius *et al.,* 2002). As this specific binding requires the presence of calcium ions, this lends itself to act as a sort of "molecular switch" between fragment binding in its presence and fragment dissociation in its absence (Lindman *et al.*, 2009). This highly specific fragment complementation interaction can be exploited for a number of applications in proteomic research including protein immobilization for surface plasmon resonance, enzyme linked immunoassays, protein purification as an affinity chromatography technique and possibly therapeutic drug delivery through cell penetrating peptide-drug conjugation/loading.

This EF hand fragment complementation technology has been patented for use as a novel affinity protein purification system, O'Connell *et al.,* (2012), and is under development and optimization in the O'Connell lab in the Conway Institute, University College Dublin. Protein purification is achieved through incorporating the EF1 tag into a protein of interest through recombinant deoxyribonucleic acid (DNA) technology. This tag will then facilitate binding of the recombinant protein to an EF2 affinity resin in the presence of calcium (Figure 1.3, purple spheres indicate calcium ions (Ca^{2+})). Following EF1 recombinant protein capture, the affinity resin is washed to remove unbound contaminating proteins. The tagged recombinant protein can be eluted from the affinity resin through removal of calcium from the system. Calcium is removed by ethylenediaminetetraacetic acid (EDTA) containing buffers which chelate the calcium ions, disrupting the binding of EF1 and EF2. This will result in dissociation of the EF1 and EF2 hands and therefore release of the EF1 tagged protein in a pure preparation (Figure 1.3).
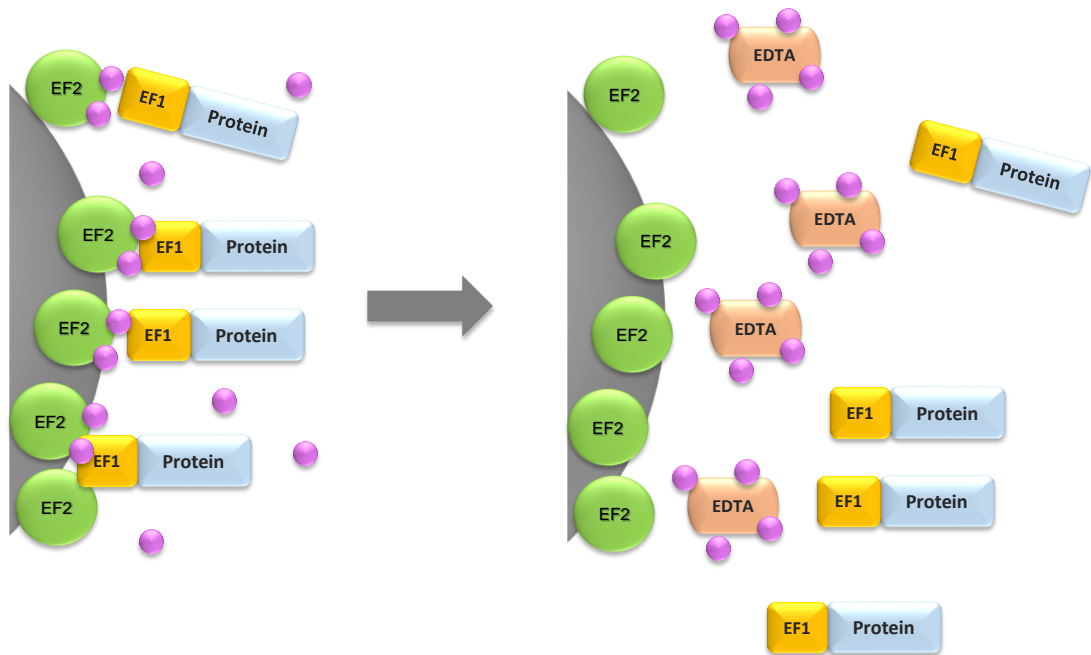
Figure 1.3. Overview of EF hand technology in the purification of EF1 tagged recombinant proteins. Green = EF2 affinity ligand, Yellow = EF1 recombinant protein tag, Blue = recombinant protein, Purple = Ca^{2+} ions, Orange = EDTA molecule, Grey = inert solid resin support.

1.3 The EF2-GGC Peptide

To construct a functional affinity resin for the EF1-EF2 protein purification system, a functional EF2 peptide ligand needs to be produced. For a sustainable manufacturing process of this peptide this thesis discusses in detail development of a standard operating procedure for EF2 peptide production.

A peptide is a short chain of amino acids, generally understood and seen as less than fifty amino acids in length (Furman *et al.*, 2015). Two amino acids can be linked via a peptide bond between the carboxyl group of one amino acid to the amino group of another (Sheehan & Hess, 1995; Jad *et al.,* 2015). In this condensation reaction the carboxyl group loses a hydrogen and oxygen, and the amino group loses a hydrogen, thus producing a water molecule and a peptide bond. As few as two amino acids (dipeptides) can be joined, to ten amino acids (decapeptides) to hundreds (polypeptides). Peptides and polypeptides are unbranched chains in contrast to proteins which form from arrangements of these peptides and polypeptides linking with

10

one another and branching out to form specific tertiary and quaternary structures critical for the proteins biological function.

Peptides are valuable entities for natural biological homeostasis, for example in the form of peptide hormones, but also have applications in clinical research. Peptides as therapeutic agents for treatment of disease is a current area of major investigation (Bommarius *et al.*, 2010). As aforementioned, some commonly known affinity protein tags are peptides. Incorporation of peptide affinity tags to a protein can facilitate binding of the protein to a corresponding affinity resin for protein purification, protein immobilization and/or protein detection. This study provides an example of such peptide affinity tags in action. In this system, not only is the affinity tag a peptide but the corresponding affinity ligand is also a peptide, namely the EF2-GGC peptide. This peptide will be investigated for its use as a novel affinity ligand in a nanoscale protein purification system.

The EF2-GGC peptide is 40 amino acids in length generating an overall molecular weight of 4.382 kDa (Figure 1.4);

Gly – Leu – Lys – Gly – Pro – Ser – Thr – Leu – Asp – Glu – Leu – Phe – Glu – Glu – Leu – Asp – Lys – Asp – Gly – Asp – Gly – Gln – Val – Ser – Phe – Glu – Glu – Trp – Gln – Val – Leu – Val – Lys – Lys – Ile – Ser – Gln – Gly – Gly – Cys

Figure 1.4. Complete amino acid sequence of the EF2-GGC peptide. Highlighted residues; in green a tryptophan, aromatic residue for absorbance, in blue a glycine-glycine hinge like region and in yellow a terminal cysteine residue, all contributing important structural elements.

The sequence shown in figure 1.4 is a slightly modified version of the naturally occurring EF2 hand sequence and has been done so in order to incorporate some additional features. These features translate into important structural elements of the peptide which play a role in downstream processing steps and in affinity resin manufacture. The C-terminus of the EF2 peptide was designed to have a terminal cysteine residue (highlighted in yellow, Figure 1.4). As cysteine residues have sulfhydryl side chains, the presence of this side chain can facilitate directional thiol coupling of the EF2 peptide ligand to nanoparticle scaffolds in downstream resin manufacturing steps. In this manner the EF2 ligand can be optimally displayed from the nanoparticle surface so it is effectively available for binding to EF1. Other elements of the EF2 peptide sequence that provide important structural functions include the

presence of two glycine (GG) residues (highlighted in blue, Figure 1.4) which act as a small spacer between the peptide ligand and the nanoparticle surface. This "hinge" like region gives the EF2 ligand some leeway for movement to optimally bind EF1 tagged proteins. A tryptophan residue (highlighted in green, Figure 1.4) was also incorporated into the peptide amino acid sequence because it was not naturally occurring. This was done so in order to provide a means of detection. Tryptophan is an aromatic amino acid known to strongly absorb light at 280 nm with an emission peak from 300 to 350 nm, so it is useful to allow spectroscopy methods to detect and quantify the presence and amount of EF2-GGC peptide in a sample.

Finally the EF2 ligand itself, as described earlier, has not been modified as it already functions to bind one calcium ion as well as the other half of its original structure, the EF1 hand in high affinity. This peptide has an extinction coefficient of 5,500 and a pI of 4, determined by ExPASy bioinformatics resource tool (Table 1.3, section 1.6).

1.3.1 Chemical Peptide Synthesis

Other than the natural biologically occurring reactions, peptides can be made by chemical synthesis through organic compounds (Bray, 2003). Liquid-phase and solid-phase peptide synthesis exist however liquid-phase is slightly more complicated as it requires more peptide isolation steps versus solid-phase peptide synthesis (SPPS) pioneered by Merrifield, (1963), which has been developed considerably over the years and has become the method of choice for peptide synthesis (Thundimadathil 2013). Small porous beads are used in SPPS as an inert support from which peptide chains can be built in the direction of C to N terminus with non-native amino acids (Figure 1.5).

The first amino acid is bound to an inert resin, via its carboxyl group through a linker intermediate using standard coupling reagents (Pipkorn *et al.,* 2014). The linker intermediate has an important function in that it enables cleavage of the peptide from the resin once the peptide is fully synthesized. Each amino acid used in SPPS have cleavable protecting groups. This means that some amino and carboxyl groups as well as other reactive side chain groups of the amino acid must be "blocked" to ensure ligation of the amino acid occur in the correct direction and that no other possible ligations occur between non-specific reactive side chains of one amino acid to another. The next residue of the desired sequence must therefore be de-protected to reveal a reactive group so it is available for ligation to the previous residue. These steps of de-protection and ligation are repeated until all residues of the desired peptide sequence are ligated and it is at this point that the peptide can be cleaved from the support.

The choice of inert resin support used and amino terminal de-protecting agent used is decisive to the conditions, and thus reagents required for peptide synthesis. For example fluorenemethyloxycarbonyl (Fmoc), is an N-protecting group of an amino acid and can be easily removed by piperidine (basic solution). Fmoc is good as it is stable in uronium and phosphonium based coupling reagents often used for aa coupling steps e.g. PyBopw 62 (Prasad *et al.*, 2011; Montalbetti *et al.,* 2005; Jad *et al.,* 2015). Once the peptide is fully synthesised, acidic peptide cleavage using trifluoroacetic acid (TFA) can cleave the peptide from hydroxybenzyl-based resins such as Wang resin. Premature peptide cleavage or inefficient coupling may arise from improper reaction conditions, which may be caused by incompatible reagents.
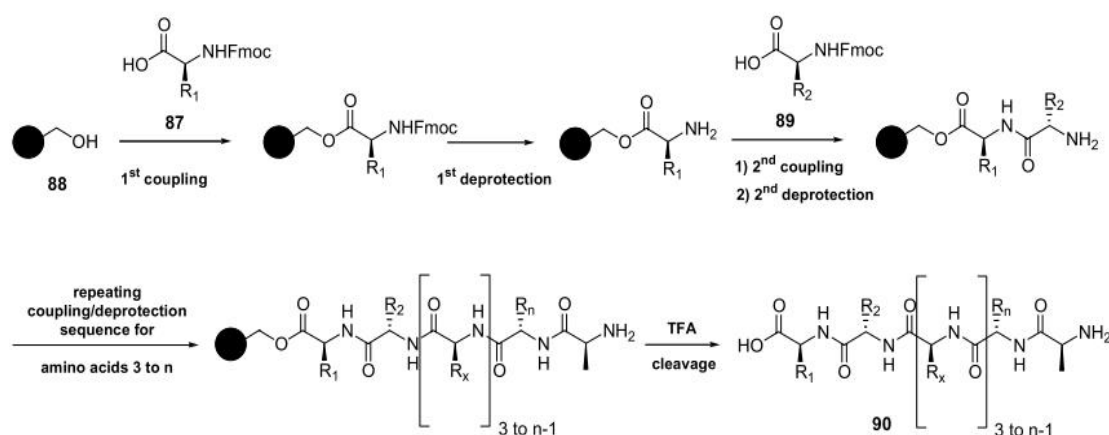
Scheme 53. Solid-supported peptide synthesis.

Figure 1.5. SPPS taken from Montalbetti et al., 2005, illustrating an inert resin supporting coupling and de-protecting reactions to build a peptide chain until final TFA cleavage releases the peptide for purification.

SPPS as a method to produce the EF2-GGC peptide is unsustainable for long term affinity resin manufacture as this system is very expensive (Andersson *et al.,* 2000). Another concern would be due to the longer peptide length of EF2-GGC as it would bring with it more cleavage and coupling steps and therefore longer manufacturing time and increased probability for undesirable coupling of the incorrect amino acid or premature peptide cleavage. In this way more reagents and amino acid derivatives would be required to carry out this process and would incur an increased cost to the lab. And so, for a practical, maintainable production line of the EF2-GGC peptide for affinity resin manufacture, this study investigates production of the EF2-GGC peptide in the form of a recombinant fusion protein.

1.4 Recombinant Fusion Proteins for Peptide Production

Recombinant fusion proteins are a product of recombinant DNA technology. Genetic modifications of DNA through molecular cloning techniques such as polymerase chain reactions, restriction enzyme digestion and ligation, allow for gene sequences that do not naturally occur, to be incorporated and confer their features in the fully translated recombinant fusion protein. In this way affinity tags can be incorporated into the translated protein or two proteins can be translated in the form of one fusion protein. This area of recombinant DNA technology has been advanced by the use of plasmid DNA vectors which already encode the sequence for a particular affinity tag and/or fusion protein. This approach simplifies efforts as a specific gene sequence may be inserted directly into the vector alongside the desired tag/fusion protein sequence so that it will be expressed as one fusion protein. For example vectors with the EF1 affinity tag have been used to produce EF1 tagged recombinant proteins for purification. Optimization of these plasmid vectors in terms of their primer regions, codon optimization, and incorporation of strong promoters builds the foundations for high expression of the tagged recombinant protein. Today there are numerous tools available online, free of charge, providing methods to engineer plasmid vectors for these applications, for example ExPASy bioinformatics resource tool (http://www.expasy.org/) and Gene Designer (https://www.dna20.com/resources/genedesigner) provided by the gene synthesis company DNA 2.0.

In this study, the EF2-GGC peptide was expressed in the form of a recombinant fusion protein. This was because peptides can be particularly difficult to express by themselves with low expression levels and poor stability reported (Rodríguez *et al.*, 2014). Peptides are also prone to proteolytic degradation *in vivo* in comparison to other larger entities and it was because of this that expression of the EF2-GGC peptide in the form of a recombinant fusion protein is helpful (Kyle *et al.*, 2012). Previous attempts to produce peptides through recombinant fusion proteins have been successful in the production of antimicrobial peptides (Li, 2011). Antimicrobial peptides are a part of the innate immune system armoury present in all lifeforms. The name might suggest otherwise, however there are antimicrobial peptides produced in bacteria, fungi and yeast, as well as mammalian hosts and can attack plasma membranes of foreign entities causing membrane permeabilization and also target intracellular components such as DNA and cytoplasmic proteins. LL-37 is an antimicrobial peptide that was investigated for expression via recombinant fusion protein. A review by Li, 2011 discussed and compared LL-37 expression and purification via a number of different

fusion partners including thioredoxin, GST and SUMO. The highest yield, by a significant margin, of LL-37 peptide retrieved was achieved by expression with thioredoxin as the fusion protein partner and an enterokinase cleavage site. This method reported a 1000 mg/L return of fusion protein and following cleavage a 40 mg/L return of LL-37 peptide (Table 1.2). The authors suggest reasons as to the high yields were owed to the high density cultivation of the *E.coli* culture and efficient cleavage of fusion protein. Other studies did not have such success with 1-2 mg/L of peptide produced (Table 1.2). Common problems encountered were identified in low expression levels, poor cleavage of the recombinant fusion protein and peptide aggregation.

Table 1.2. Example table taken from Li, 2011 summarizing different approaches tested for production of a recombinant peptide, LL-37, in *E.coli*

Fusion Carrier	Cleavage Reagent	Chromatographic Steps	Yield (mg/L)
GST	Thrombin	Affinity AU-PAGE	NA[a]
Thioredoxin[c]	Thrombin	Affinity, cation exchange	24.3/1.1[b]
Thioredoxin	Formic acid	Affinity, gel filtration, affinity, HPLC	28.4/1.7
GST	Factor Xa	Affinity, HPLC	7.6/0.3
Thioredoxin	Formic acid[d]	Affinity, gel filtration, HPLC	27.1/2.6
Family III CBM	Formic acid	Affinity, HPLC	NA/1
Thioredoxin	Enterokinase	Affinity, affinity, HPLC	1000/40
SUMO	SUMO Protease	Affinity, HPLC	NA

[a] Not Available

[b] Amount of purified fusion and peptide, respectively.

[c] In addition to thioredoxin, the fusion also contains the cathelin-like prosequence.

[d] The fusion was first cleaved with thrombin to release LL-37-containing fragment.

1.4.1 EF2-GGC Recombinant Fusion Protein

In an effort to obtain a functional EF2-GGC peptide, this study outlines the steps involved for the peptide to be produced through a recombinant fusion protein. The EF2-GGC recombinant fusion protein was expressed in *E.coli* consisting of a 6xHis tagged human carbonic anhydrase protein (6xHis CA), an internal TEV cleavage sequence/site and the EF2-GGC peptide located at the C-terminal (Figure 1.6), generating a total molecular weight of 36.44 kDa.
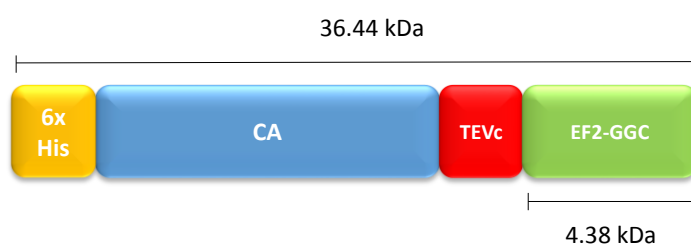
Figure 1.6. EF2-GGC Recombinant Fusion Protein

Human Carbonic Anhydrase I (CA) is a 32 kDa eukaryotic, metalloenzyme that has in previous work carried out by the O'Connell lab group, University College Dublin demonstrated high expression in bacterial host systems with good solubility (Krishnamurthy *et al.* 2008). The presence of a 6x histidine (6xHis) tag on this recombinant fusion protein also provided a means of protein purification. These factors of high expression, good solubility and a means of native protein purification were sought after for maximum EF2-GGC peptide production and so this protein was chosen as the partner fusion protein to which the EF2-GGC peptide would be expressed alongside in the form of one recombinant fusion protein (Figure 1.6). A specific seven amino acid sequence in between 6xHis CA and EF2-GGC peptide encodes for a TEV cleavage site (Figure 1.7). This permits the use of an enzyme, the TEV protease, for location specific peptide bond cleavage, separating the EF2-GGC peptide from the 6xHis CA protein. Following this cleavage step additional rounds of protein purification are required in order to obtain the newly liberated peptide in a highly pure, monomeric preparation for coupling to functionalised NP scaffolds in affinity resin manufacture.

1.5 Peptide Bond Cleavage

Following protein purification, peptide bond cleavage to remove affinity tags and/or to remove fusion protein partners is quite common. Some tags are feared to have adverse effects on protein structure and function and so are not desirable for downstream analytical studies of the protein itself (Waugh, 2005). In particular for the production of bio therapeutic proteins destined to become drugs for human consumption, the removal of affinity tags is essential before approval of its use as a medicinal agent. Cleavage to separate fusion proteins from the protein of interest may also be desired. In this study, 6xHis CA was incorporated as a fusion protein partner to the EF2-GGC peptide to enhance expression and solubility (Figure 1.6). Following expression and purification of this fusion protein, 6xHis CA is no longer needed and so the peptide bonds holding the EF2-GGC peptide and 6xHis CA together can be cleaved to release the product of interest, the EF2-GGC peptide. The separation of affinity tags and/or partner fusion proteins can be achieved by chemical or enzymatic peptide bond cleavage (Wood, 2014).

There are a number of different chemicals known to catalyse location specific cleavage of peptide bonds such as cyanogen bromide (CNBr), formic acid, hydroxylamine and 2-nitro-5-thiocyanobenzoic acid (NTCB) to name a few.
CNBr may be one of the more widely known chemicals for this use. A methionine (Met) residue is prerequisite for CNBr cleavage in a reaction where the nucleophilic sulfur group of methionine attacks the electrophilic carbon of CNBr (Gross & Witkop, 1962). This results in the replacement of bromide in CNBr to sulfur to form an iminolactone intermediate and the oxidised methionine converts to methionine sulfoxide. Water and acid, e.g. hydrochloric acid, are then required to complete peptide bond cleavage reaction as the iminolactone is hydrolysed in what is also known as a "Schiff base hydrolysis" reaction. This forms a homoserine lactone molecule and the fully cleaved protein. The cleaved protein now has a new C-terminus amino acid. It should be noted that CNBr cleavage may not take place when the Met residue is followed by either a serine or threonine amino acid as these can compromise water/acid iminolactone hydrolysis because they can react with it instead to form the homoserine end product without peptide bond cleavage.
Formic acid can cleave between aspartic acid and proline residues. Hydroxylamine cleaves between asparagine and glycine amino acids and NTCB at cysteine residues. The variety of cleavable residues is good; however, chemical cleavage can be quite a dangerous method as many of these chemicals are known to be volatile and toxic. One should carefully study the material safety data sheets before considering use of the

chemical. For example the MSDS for CNBr describes the potential hazards and risks associated with this chemical in detail. CNBr is in fact fatal if swallowed/ingested, can cause severe burns and is a volatile substance. Deactivation of this chemical is also quite dangerous as it generates great exothermic energy which may be explosive. Hydroxylamine has considerable toxic effects too. Potential routes of exposure to hydroxylamine are through the skin, eyes and mucous membrane absorption, ingestion and inhalation which may cause serious adverse effects including irritation, toxicity, and mutagenicity. For these reasons the use of chemicals for peptide bond cleavage is not the first choice and the alternative approach of enzymatic cleavage has become more popular.

Enzymatic peptide bond cleavage via proteases is not only more common but is a much safer option. As proteases are enzymes they have specific active sites that only catalyse cleavage at recognised sites. This is of the upmost importance in order to avoid random peptide bond cleavage which could potentially break important structural bonds of the protein itself. Some common enzymes that are employed for peptide bond cleavage include enterokinase, factor Xa, SUMO protease, TEV protease and thrombin (Terpe, 2003) (Table 1.2).

Peptide bond cleavage of a tag located at the N-terminal of the recombinant protein is often carried out using enterokinase or factor Xa. This is good as it returns the recombinant protein native N-terminal state, leaving no additional residues behind. However, these proteases have been known to cleave in other non-specific locations. Also as factor Xa contains disulphide bonds, its use is limited in conditions containing reducing agents (Young *et al.,* 2012). SUMO protease has been reported to work quite well in a broad range of conditions from 4 - 37°C and pH 5.5 - 10.5 enabling a broader range of proteins to be cleaved using this protease. Thrombin is a particularly good protease to use if the recombinant protein is present in strong detergents. This is notable for membrane protein purification as they often require the use of detergents to disrupt and help pull out the protein from the plasma membrane wall and therefore will be contained in a formulation of strong detergents post purification. The protease employed in this study for enzymatic peptide bond cleavage was the TEV protease. This enzyme is required to separate the fusion protein 6xHis CA from the EF2-GGC peptide.

1.5.1 TEV Protease

The TEV Protease is a cysteine protease, the name, referring to the catalytic triad of the enzyme's active site involving a cysteine residue, an aspartic acid and a histidine. TEV protease originates from the Tobacco Etch Virus (TEV) (Fang *et al.,* 2007). This 27 kDa enzyme recognises a specific seven amino acid sequence in which it catalyses peptide bond cleavage specifically between Glutamic and Glycine/Serine residues of this sequence (Figure 1.7, Glutamic and Glycine/Serine amino acids highlighted in red) (Blommel & Fox, 2007).

$$E - N - L - Y - F - Q \downarrow G/S$$

Figure 1.7. TEV Protease recognition site for peptide bond hydrolysis

This simple cleavage system is advantageous for the production of the EF2-GGC peptide as by incorporating this seven amino acid sequence in between the 6xHis CA and the EF2-GGC peptide amino acid sequences of the fusion protein, it provides a specific location for the cleavage and thus release of the EF2-GGC peptide of interest from its fusion protein partner, 6xHis CA (Figure 1.8).
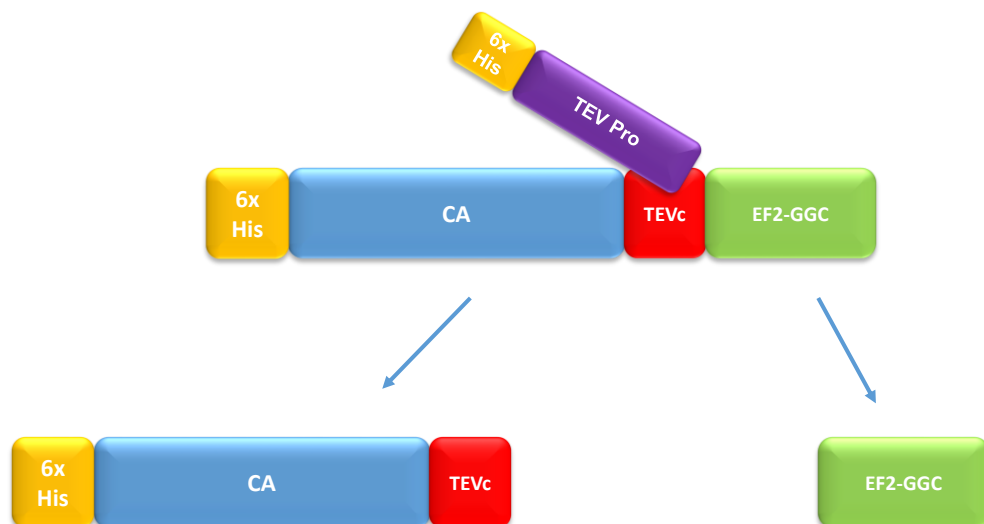
Figure 1.8. TEV protease cleavage of the EF2-GGC peptide fusion protein

The TEV protease used in this study is a mutated version of the wildtype. Early studies of this enzyme found auto inactivation of its catalytic function due to self-cleavage occurring as a common problem (Zdanov *et al.,* 2003; Kapust *et al.,* 2001). Also codon bias and low solubility were issues in the production of this enzyme. An amino acid substitution of valine in place of serine at position 219 (S219V mutation) of this enzyme's amino acid sequence not only eliminated the auto inactivation dilemma but increased the enzymes activity two fold (Wang, 2013; Tropea *et al.,* 2009). The enzyme also demonstrated low solubility, making purification of soluble, biologically active enzyme difficult. To overcome this the enzyme is expressed with MBP as a fusion protein as MBP imparts high expression and greater solubility properties to the protease.

This mutated TEV protease is available commercially from a number of different biotechnology companies such as Sigma Aldrich and Life Technologies. However in house production of this enzyme has been well documented in the literature and is substantially cheaper than purchasing the industrial manufactured enzyme and so was carried out as part of this study.

TEV protease demonstrates a number of advantages over other proteases. For example Factor Xa, enterokinase and thrombin have previously been observed to cleave in sites other that the designed target (Tropea *et al.,* 2009). This problem is not seen with TEV because it will only recognise and cleave the specific seven amino acid sequence presented in figure 1.8. Positioning of this seven amino acid TEV cleavage site is an important aspect to consider. If placed just before the N-terminal of the recombinant protein or in this study EF2-GGC peptide, following cleavage only a single amino acid glycine or serine will remain at the N-terminus. If placed at the C-terminal, following cleavage a six amino acid sequence will be left behind at the end of the protein sequence. In the case of the EF2-GGC fusion protein, cleavage results in a 6 amino acid sequence remaining on the c-terminal of 6xHis CA and one amino acid on the n-terminal of the EF2-GGC peptide. These remaining residues are not a concern as the 6xHis CA is not required any further following cleavage and one extra amino acid on the EF2-GGC peptide is not expected to have any adverse effects on peptide functionality. However, a point to note is that this minor issue may be avoided if the amino acid left behind was part of the actual native protein sequence, i.e. if glycine/serine is actually originally part of the n-terminal protein sequence.

1.6 EF2-GGC Peptide for a Miniaturised Protein Purification System

Once the EF2-GGC peptide of interest is cleaved and free from its fusion partner, 6xHisCA, it can be purified so as to achieve a suitable formulation for the next steps in building a miniaturised EF2 affinity resin; conjugation of the peptide to a functionalized NP surface.

As this study aims to build an affinity protein purification system for small scale protein preparations, the choice of material for the base of the resin is key to achieve this miniaturised scale. A solid inert support that provided a means for optimal directional display of the EF2-GGC ligand is required for the protein purification system to perform at its best. Inert resins for affinity chromatography exist in many different forms and sizes, however nano-sized supports, 100 nm nanoparticles (NPs) in particular were of interest for this study as it allowed the system to reach a nano-scale size. This smaller sized resin provided a number of other advantages as well. For the same volume of resin bed, much more 100 nm NPs would be present in comparison to another resin with a larger diameter e.g. 45 - 165 µm size of the sepharose CL-6B resin in which standard Ni-NTA resin by Qiagen is supplied. In this way the presence of more NPs equates to a much greater number of affinity ligands available for the tagged recombinant protein to bind and therefore the potential to purify larger quantities of protein using less material. Their small size also allows them to be handled in a wider variety of vehicles e.g. micro-centrifuge tubes or columns.

There are nanoparticle protein purification systems available currently with magnetic nanoparticles for this use appearing most often in the literature headlines (Kim *et al.* 2007; Fischer *et al.* 2013).Yet, the majority of work carried out with nanoparticles seem to be focused on drug delivery and cancer therapy development (Saptarshi *et al.* 2013). This leaves the area of nano-scale protein purification slightly under explored, even when the potential for many successful applications is quite significant. For example it could aid in high through-put identification of optimum protein purification conditions, enabling much better screening capabilities of a range of different proteins. Combining the use of less material in a reduced timeframe is most definitely a top area of interest in both academic and industrial settings where advances are still desired.

1.7 Aims

The aims of this study were to (Figure 1.9):

1) Develop a standard operating protocol for stepwise production of the EF2-GGC peptide from a bacterially expressed recombinant fusion protein

2) Couple the peptide to a functionalized nanoparticle scaffold to build a highly specific EF2-NP affinity resin and test the EF2-NP resin for resin functionality, i.e. capture and purification of EF1 tagged recombinant proteins from complex mixtures

EF2-GGC
Peptide
Production

Peptide
Conjuagtion to
NP Resin

Purification of
EF1 tagged
Proteins
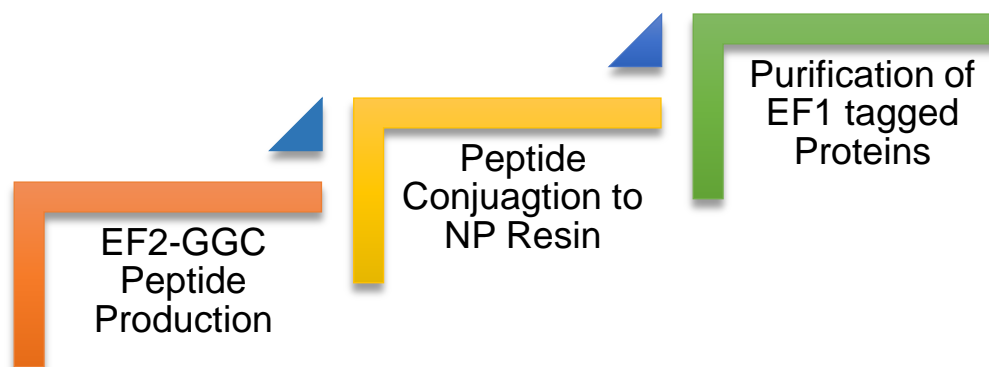
Figure 1.9. Schematic image of the aims of this study

Chapter 2

Materials & Methods

2.1 Materials

Table 2.1. Addresses of Suppliers

Supplier	Address
Addgene	75 Sidney Street, Suite 550A, Cambridge, MA 02139, USA
BD Bacto	1 Becton Drive, Franklin Lakes, NJ 07417-1880, USA
Bio-Rad Laboratories Ltd.	1000 Alfred Nobel Drive, Hercules, CA 94547, USA
DNA 2.0	Menlo Park, CA, USA
Formedium	Unit 1B, Hunstanton Commercial Park, Hunstanton, Norfolk PE365JQ, UK
GE Healthcare Life Sciences	Little Chalfont, Buckinghamshire, UK
Invitrogen/ Life Technologies	5791 Van Allen Way, Carlsbad, CA, USA
Millipore Corporation	290 Concord Road, Billerica, MA 01821-7037, USA
National Diagnostics	6407 Idlewild Road, Suite 211, Charlotte, NC 28212, USA
New England Biolabs	240 County Road, Ipswich, MA 01938-2723, USA
Novagen	Darmstadt, Germany
Qiagen	Strasse 1, 40724 Hilden, Germany
Sigma	Sigma-Aldrich Chemie Gmbh, P.O. 1120, 89552 Steinheim, Germany
Spectrum Labs	Rancho Dominguez, CA, USA
ThermoFisher Scientific	81 Wyman Street, Waltham, MA USA 02451

Table 2.2. List of Chemicals

Chemical	Source
30 % Bis-Acrylamide	National Diagnostics
Agar	BD Bacto
Glycine	Fisher Scientific
Ni-NTA Agarose	Qiagen
Overnight Express™ Instant TB Medium	Novagen
E.coli BL21 DES (Rosetta™) Competent Cells	Novagen
E.coli DH5α™ Competent Cells	ThermoFisher Scientific
EF2-GGC Recombinant Fusion Protein Plasmid	DNA 2.0
See Blue™ Plus Two Prestained Protein Ladder	Invitrogen
AcTEV Protease	Life Technologies
TEV Protease S219V Mutant (pRK793) *E.coli* Rosetta™ (DES)	Addgene
Tryptone	ForMedium
GeneJet Plasmid Midiprep Kit	ThermoFisher Scientific
Pierce™ BCA Protein Assay Kit	ThermoFisher Scientific

All other chemicals were purchased from Sigma Aldrich

2.2 Methods – EF2-GGC Peptide Production

2.2.1 Plasmid DNA Retrieval

Plasmid DNA of the EF2-GGC recombinant fusion protein plasmid (Figure 5.1, appendix I for plasmid DNA map), synthesised by DNA 2.0 (Menlo Park, California, U.S.A) was obtained from the supplied GFC filter paper disc. 100 µl of 10 mM Tris-HCl pH7.5 buffer was applied directly to the centre of the filter paper disc and incubated for 2 minutes at room temperature. A sterile syringe needle was used to puncture a small hole in the bottom of a sterile 0.5 ml micro centrifuge tube. Following the 2 minute incubation time the soaked filter paper was placed inside the punctured 0.5 ml micro centrifuge tube which was then placed into a fresh 1.5 ml micro centrifuge tube (with no puncture) and centrifuged at full speed for 1 minute in a table-top centrifuge at room temperature. The liquid that passed through into the 1.5 ml micro centrifuge tube contained the EF2 recombinant fusion protein plasmid DNA and was measured using the NanoDrop (Thermo Scientific) to verify plasmid DNA concentration of 20 ng/µl. All plasmid DNA were stored in -20°C.

2.2.2 Transformation of Bacterial Host Cells

Two different strains of competent *E.coli*, BL21 DE3 (Rosetta™) (Novagen) and DH5α™ (ThermoFisher Scientific), were transformed with the EF2 recombinant fusion protein plasmid DNA by heat shock. Plasmid DNA (10 ng) was incubated with 20 µl of re-suspended competent bacterial cells in a 1.5 ml micro centrifuge tube on ice for 5 minutes and then placed in a 42°C water bath for exactly 45 seconds to complete the heat shock step. The micro centrifuge tube was immediately transferred onto ice for a 2 minute incubation after which 80 µl of Super Optimal Broth with Catabolite repression media (SOC) (2 % w/v tryptone, 0.5 % w/v yeast extract, 10mM NaCl, 2.5 mM KCl, 10 mM MgCl$_2$, 10 mM MgSO$_4$, 20 mM glucose) was added. This (now 100 µl total volume) mix was incubated at 37°C, 250 rpm for 1 hour and then spread on a dual-selective Tryptone Yeast Extract (TYE) (1 % w/v tryptone, 0.5 % w/v yeast extract, 1.4 mM NaCl, 1.5 % w/v bacto agar) agar plate. Selective by means of supplementation with antibiotics; chloramphenicol (34 µg/ml) and ampicillin (100µg/ml) were required to select growth of *E.coli*, Rosetta and ampicillin only was required for *E.coli* DH5α. The TYE agar plates were incubated overnight (12-18 hours) at 37°C to allow for growth of successfully transformed single colonies.

2.2.3 Bacterial Starter Culture and Glycerol Stock Preparation

Under a flame, one isolated colony from each TYE plate was picked to inoculate a starter culture consisting of 10 ml 2X Tryptone Yeast extract (2X TY) (1.6 % w/v tryptone, 1 % w/v yeast extract, 85 mM NaCl) media supplemented with 1% glucose and the same antibiotics as described in 2.2.2, for each bacterial strain. The starter cultures were incubated at 37°C, 250 rpm for 16 h.

A glycerol stock of the successfully transformed bacteria, of both *E.coli* Rosetta and DH5α strains were prepared under a flame by adding 750 µl of a sterile 50 % glycerol solution to 750 µl of starter culture bacterial cells in a 2 ml micro centrifuge tube with a screw cap. This was mixed gently but thoroughly by pipetting and placed in -80°C freezer for long term storage.

2.2.4 Plasmid DNA Stock Preparation

To prepare long term EF2 recombinant fusion protein plasmid DNA stocks, a culture of the *E.coli* DH5α transformed cells were grown for plasmid DNA extraction. 100 ml of 2X TY media with ampicillin was inoculated with 1 ml of the starter culture. The culture was grown at 37°C, 250 rpm for 16h. The cells were harvested via centrifugation at 5,000 rpm, 4°C for 20 minutes. The supernatant containing only the media was discarded and the pellet of bacterial cells stored at -20°C until required for plasmid DNA extraction. Plasmid DNA extraction was carried out via Thermo Scientific GeneJET Plasmid Midiprep Kit protocol. All plasmid DNA stocks were measured using the NanoDrop to determine plasmid DNA concentration and stored at -20°C.

2.2.5 Fusion Protein Expression

For EF2-GGC recombinant fusion protein expression, the *E.coli* Rosetta cells were cultured first in a pilot scale preparation of 100 ml total culture volume to investigate protein expression levels before up-scaling to 2 l cultures. Protein expression was investigated by both auto and IPTG induction methods.

2.2.5.1 Auto-induction - Pilot Scale

Autoinduction was achieved using Overnight Express™ Instant TB Media (Novagen, Darmstadt, Germany). For a 100 ml expression culture, 6 g of Overnight Express™ and 1 ml of 100% glycerol was dissolved fully in 100 ml dH$_2$0 and microwaved on high power until bubbles appeared. The medium was sterilized through microwave irradiation by continuing to microwave for another 15-30 seconds after the bubbles appeared, with care taken to ensure the media did not boil over. The medium was left to cool before ampicillin (100 µg/ml) and chloramphenicol (34 µg/ml) antibiotics were

added and the culture inoculated with a 1:200 dilution (e.g. 0.5 ml in 100 ml) of the *E.coli* Rosetta starter culture. The bacterial cell culture was grown in a sterile conical flask, 5 times the size of the culture volume, at 37°C, 250 rpm, for 24 hours. After 4, 8, 12 and 24 h 1 ml of the culture was taken and saved for Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE), section 2.2.5.5, and stained by Coomassie blue, section 2.2.5.6, to investigate protein expression at that time point. The remaining cells of the expression culture were harvested via centrifugation at 5,000 rpm, 4°C for 20 minutes. The supernatant containing only the media was discarded and the pellet of bacterial cells stored at -20°C.

2.2.5.2 Isopropyl β-D-1-thiogalactopyranoside (IPTG) induction - Pilot Scale

IPTG induction was achieved via the addition of IPTG to the expression culture. 100 ml of 2X TY media supplemented with ampicillin (100µg/ml) and chloramphenicol (34µg/ml) antibiotics was inoculated a 1:20 dilution of the *E.coli* Rosetta starter culture (e.g. 5 ml starter culture into 100 ml media) and grown in a sterile conical flask, 5 times the size of the culture volume, at 37°C, 250 rpm, until the optical density at 600 nm (OD_{600}) of the culture reached an OD_{600} of ~ 0.6 - 0.8. When the culture reached this 0.6 $O.D_{600}$, IPTG was added to a final concentration of 1 mM. The culture was grown at 37°C, 250 rpm, for 6 h following addition of IPTG. After every hour, 1 ml of the culture was taken and saved for SDS-PAGE, section 2.2.5.5, and stained by Coomassie blue, section 2.2.5.6, to investigate protein expression at that time point. The remaining cells of the expression culture were harvested via centrifugation at 5,000 rpm, 4°C for 20 minutes. The supernatant containing only the media was discarded and the pellet of bacterial cells stored at -20°C.

2.2.5.3 Protein Expression Time-point Normalisation

The $O.D_{600}$ of each 1ml time-point, was measured via simple reads on a Cary 50 conc UV-Visible spectrophotometer (Varian). Each time-point was diluted with fresh media down to an OD_{600} of 0.6 in order to normalise the cell densities of all samples. All time-points were centrifuged at 10,000 x g, 4°C for 15 minutes to pellet the bacterial cells. The media supernatant was discarded and the pellet re-suspended in 50 µl 4X Lameilli Sample Buffer (LSB) (250 mM Tris-HCl pH 6.8, 10 % w/v SDS, 0.5 % w/v bromophenol blue, 25 % w/v glycerol, 6% v/v β-mercaptoethanol) and boiled at 99°C for 5 mins. An equal volume of each sample was loaded into a 10% SDS-PAGE and run as described in section 2.2.5.5 and Coomassie stained as per section 2.2.5.6.

2.2.5.4 Auto Induction - Scale Up

Following pilot scale protein expression investigations, autoinduction was chosen for protein expression scale up from 100 ml to a 2 litre culture. This method was carried out as described in section 2.2.5.1, with two changes: 1) the culture inoculum was changed to a 1:20 dilution, i.e. 100 ml starter culture into 1,900 ml Overnight Express™ Instant TB Media. 2) The culture was grown for total of 12 hours.

2.2.5.5 SDS and Native PAGE

Protein samples for SDS-PAGE were prepared by mixing the desired amount (µg) of protein sample in 4X LSB and boiled at 99°C for 5 mins. Polyacrylamide gels were cast in a cassette of glass plates, 10.0 x 8.0 cm, 1.5 mm width (Bio-Rad Laboratories Ltd.). 10, 12 or 16 % resolving gel were prepared with 5 % stacking gel (for gel recipes see table 2.1) and a 10 or 15 well comb depending on the number of samples to be run. All gels were run in Mini-PROTEAN® Tetra Vertical Electrophoresis Cell using PowerPac basic power supply (Bio-Rad Laboratories Ltd.). 10 and 12 % SDS-PAGE were run at 80 V for 15 mins, or until the protein samples reached the stacking/resolving gel interface, at which point the current was increased to 120 V for a further 60 mins or until the dye front has reached the end/just run off the gel. 16 % SDS-PAGE were run at 80 V for 120 mins, or until the dye front reached 1 cm up from the bottom of the gel. All standard SDS-PAGE were run in tris-glycine running buffers (25 mM Tris, 186 mM glycine, 0.1 % w/v SDS). Native PAGE was carried out using the same method as standard SDS-PAGE however this protocol omitted β-mercaptoethanol and, boiling of the samples, and SDS from all aspects, i.e. no SDS was permitted in the sample preparation, gel recipe and running buffer.

Table 2.3. Standard SDS/Native PAGE Recipes

Ingredient (ml)	Resolving Gel (%)			Stacking Gel (%)
	10	12	16	5
dH$_2$0	4.0	3.3	1.6	2.1
30 % Bis-acrylamide	3.3	4.0	4.27	0.5
1.5 M Tris pH 8.8	2.5	2.5	2	-
1 M Tris pH 6.8	-	-	-	0.38
10 % SDS	0.1	0.1	0.08	0.03
10 % APS	0.1	0.1	0.08	0.03
TEMED	0.004	0.004	0.008	0.003
Final Volume	**10**	**10**	**8**	**3**

2.2.5.6 PAGE Staining

SDS/Native PAGE were stained by Coomassie blue or silver staining. Coomassie blue staining was performed by washing the gel for 1 hr on a rocker at room temperature in a standard Coomassie blue stain (10 % acetic acid, 50 % methanol, 40 % dH$_2$0, 1 % w/v Coomassie blue reagent). The gel was then washed in de-stain solution (70 % dH$_2$0, 10 % acetic acid and 20 % methanol) until a satisfactory amount of Coomassie blue stain was removed from the gel and protein bands are clearly visible. Silver staining of SDS/Native PAGE required fixation for 30 mins in fixation solution (10 % acetic acid, 40 % ethanol, 50 % dH$_2$0), sensitization for 30 mins in sensitization solution (30 % ethanol, 52 mM sodium acetate, 0.2 % sodium thiosulfate, *0.5 % glutaraldehyde), three 5 minute washes in dH$_2$0, 20 min wash in 0.25 % w/v silver nitrate solution, two 1 minute washes in dH$_2$0, washing in developing solution (15 mM sodium carbonate, *0.08 % formaldehyde) until the protein bands appear and the desired intensity was reached, and 10 mM EDTA stopping solution for 10 mins to stop

the reaction. * = Glutaraldehyde and formaldehyde were added immediately before use. All gels were imaged by Cannon 210 scanner.

2.2.6 Fusion Protein Purification

Both native cell lysis and Ni-NTA IMAC were carried out on ice or at 4°C at all times to minimise any possible protein degradation. Samples of each fraction collected from native cell lysis and IMAC were run in 10 % SDS-PAGE, as described in 2.2.5.5 and stained by Coomassie blue, section 2.2.5.6.

2.2.6.1 Native Cell Lysis

The *E.coli* Rosetta bacterial cell pellets were thawed on ice for no longer than 15 minutes and the weight of each pellet recorded. Pre-cooled (4°C) lysis buffer (50 mM sodium phosphate, 300 mM sodium chloride and 10 mM imidazole, pH 8.0), was used to re-suspend and lysis the bacterial cell pellets, 5ml buffer per 1 gram pellet weight, until a homogenous solution was achieved. To this lysozyme (Sigma) was added, 1 mg per every 1 ml cell lysate and incubated on ice for thirty minutes. Following this incubation the cell lysate was sonicated using Soniprep probe on ice for six cycles of 10 second pulses at 200-300 W with a 10 second cooling period between each pulse, i.e. 10 seconds on and 10 seconds off. A universal nuclease (ThermoFisher Scientific), 1,250 units, was added to this sonicated lysate. To clear the lysate of cellular debris it was centrifuged at 4°C, 10, 000 x g for 30 minutes. A sample of both the supernatant and pellet debris was taken for SDS-PAGE analysis. The supernatant containing soluble protein was taken for IMAC and the pellet discarded.

2.2.6.2 Native IMAC

Ni-NTA resin (Qiagen) was equilibrated in pre-cooled lysis buffer prior to incubation with the cell lysate. For pilot scale, e.g. 100 ml original culture volume, 1 ml of a 50% Ni-NTA slurry (or 0.5 ml bead bed) was required. For larger cultures e.g. 1 L original culture volume, 6 ml bead bed minimum was required. The equilibrated beads were incubated with the cleared cell lysate, end-over mixing or on a roller at 4°C for 60 minutes minimum for pilot scale cultures or overnight for larger cultures i.e. 1 L expression cultures cell lysates. Following the incubation period the Ni-NTA resin-cell lysate mixture was transferred to a sterile 20 ml capacity empty column and the flow-through (FT) collected. The Ni-NTA resin bed was washed twice with 5 column volumes of wash buffer each time (50 mM sodium phosphate, 300 mM sodium chloride and 20 mM imidazole, pH 8.0). The protein was eluted four times with 1 column volume

elution buffer each time (50 mM sodium phosphate, 300 mM sodium chloride and 250 mM imidazole, pH 8.0). A sample from the FT, wash and elution fractions were taken for SDS-PAGE analysis. All elutions 1-4 were collected and pooled for immediate dialysis overnight.

2.2.6.3 Dialysis

In two, 2 litre capacity sterile glass bottles, 4 litres of 10 mM HEPES, 300 mM NaCl, was prepared and pre-cooled to 4°C before adjusting the pH to 7.4**.** Spectra/Por 3 Dialysis Tubing, 3.5 kDa molecular weight cut off (MWCO), 45 mm Flat-width, 15 meters/roll (50 ft.), (Spectrum Labs) was soaked in dH_20 for 30 mins minimum and rinsed thoroughly before use**.** One end of the tubing was folded over once and clipped ensuring no gaps remained**.** The pooled elution fraction was pipetted into the dialysis membrane bag, any air bubbles were pushed out and the open end of the tubing carefully folded over and clipped closed. A buoy was placed onto one of the clips to aid in floating the dialysis bag**.** The dialysis tubing containing the sample was placed in a 2 l beaker of pre-cooled dialysis buffer on a slow stirring magnetic stirrer at 4°C for 4 h**.** The dialysis buffer was changed after 4 h to a fresh batch and left overnight.

2.2.6.4 Protein Quantification

Protein quantification was determined using the bicinchoninic acid (BCA) assay, UV-Vis 280 or Nanodrop, or a combination of the three methods depending on sample preparation and volume. A BCA assay was carried out using the BCA protein assay reagent kit (Pierce) with bovine serum albumin (BSA) as the reference standard to generate a standard curve of 0.025 – 2 mg/ml. Protein samples were diluted 1:2, 1:5 or 1:10 with the same buffer the protein was prepared in prior to assay in order to fit within the BSA standard curve. All samples were carried out in triplicate. The assay was performed in a 96 well plate and absorbance of each sample read at 565 nm using Spectra Max 190 microplate reader (Molecular Devices). Protein concentration was determined against the standard curve and the average taken of each triplicate recorded as the final protein concentration. UV-Vis was carried out using a 1 ml capacity quartz cuvette (Hellma Analytics). Absorbance was measured from 800 to 200 nm by Cary 50 spectrophotometer. The absorbance at 280 nm was taken and the Beer Lambert Law ($A = ε \times l \times c$, where A = absorbance at 280 nm, $ε$ = extinction coefficient, l = path length, 1 cm, c = unknown concentration) applied to calculate total protein concentration. The Nanodrop ND-1000 spectrophotometer (Thermo Scientifc) was employed for protein quantification of fractions that were too low in volume for UV-Vis or BCA assay. The Nanodrop was blanked against the appropriate buffer after which 2

µl of sample was loaded and the absorbance at 280 nm was measured. The Beer Lambert Law was applied, similar to UV-Vis calculations. Protein purity was determined by densitometry analysis using ImageJ program of protein bands following SDS-PAGE.

2.2.7 Fusion Protein TEV Protease Cleavage

Cleavage of the IMAC purified and dialysed EF2-GGC recombinant fusion protein via TEV protease was initially carried out via a commercial enzyme, AcTEV Protease (Life Technologies) before an in-house TEV protease was produced. This reaction was carried out in 50 mM Tris-HCl, 0.5 mM EDTA and 1 mM Dithiothreitol (DTT), pH 8.0, incubation at 30°C for 8 h. To determine optimum enzyme to protein ratio the concentration of fusion protein was varied from 50, 100 and 200 µg to a fixed amount of AcTEV protease of 10 µg. Negative Control reaction: Fusion Protein in reaction buffer without any enzyme. An aliquot of each fraction was saved for SDS-PAGE, Coomassie blue staining and protein quantification as described in sections 2.2.5.5, 2.2.5.6 and 2.2.6.4 respectively.

2.2.8 Spin Concentrators

Pierce protein spin concentrators PES 10 kDa MWCO, 0.5 ml (Thermo Scientific) were employed as a method of size exclusion for the purification of the EF2-GGC peptide. The spin concentrator polyethersulfone (PES) membrane was rinsed in dH_20 to remove any residual glycerine or sodium azide before applying the sample. This was done by placing 500 µl dH_20 into the sample chamber, which was placed in the provided collection tube and capped before centrifugation at 15, 000 x g until <400 µl had passed into the collection tube. The filtrate and retentate were discarded and 500 µl of TEV cleaved fusion protein was placed into the sample chamber. The chamber containing the sample was placed into a clean collection tube and capped before centrifugation at 15, 000 x g for ~30 mins. A second sample of TEV cleaved fusion protein was incubated in 1 mM DTT for 1 hr at room temperature before spin concentration. An aliquot of each fraction was saved for SDS-PAGE, Coomassie blue staining and protein quantification as described in sections 2.2.5.5, 2.2.5.6 and 2.2.6.4 respectively.

2.2.9 EF2-Agarose Protein Purification

EF2-agrose resin, 2 ml resin bed, was equilibrated in lysis buffer (10 mM HEPES, 300 mM NaCl, 2 mM $CaCl_2$, pH 7.4). TEV cleaved fusion protein was incubated with the

equilibrated resin on a roller for 90 mins at 4°C. The FT was collected and the resin bed washed twice with 5 column volumes of wash buffer each time (10 mM HEPES, 300 mM NaCl, 0.5 mM $CaCl_2$, pH 7.4). The protein was eluted 3 times with one column volume elution buffer each time (10 mM HEPES, 300 mM NaCl, 10 mM EDTA, pH 7.4). An aliquot of each fraction was saved for SDS-PAGE, Coomassie blue staining and protein quantification as described in sections 2.2.5.5, 2.2.5.6 and 2.2.6.4 respectively.

2.2.10 Second Round IMAC

A second round of IMAC is required to purify the EF2-GGC peptide from the 6xHis CA and 6xHis TEV Protease following TEV protease cleavage. The TEV cleaved fusion protein was incubated with pre-equilibrated Ni-NTA resin, end-over-mixing at 4°C for 2 h. The FT was collected and the resin bed washed twice with 5 column volumes of wash buffer (50 mM sodium phosphate, 300 mM sodium chloride and 20 mM imidazole, pH 8.0) each time. The protein was eluted 3 times with one column volume elution buffer each time (50 mM sodium phosphate, 300 mM sodium chloride and 250 mM imidazole, pH 8.0). An aliquot of each fraction was saved for SDS-PAGE, Coomassie blue and silver staining and protein quantification as described in sections 2.2.5.5, 2.2.5.6 and 2.2.6.4 respectively. Only the fractions suspected to contain the EF2-GGC peptide, as determined by SDS-PAGE and protein quantification were taken and dialysed as described in 2.2.6.3 into a low sodium chloride buffer: 10 mM HEPES, 20 mM NaCl, 1 mM EDTA, 0.1 mM DTT, pH 7.4 for next steps.

2.2.11 Ion Exchange Chromatography

DEAE resin was equilibrated in a low sodium chloride buffer (10 mM HEPES, 20 mM NaCl, 1 mM EDTA, 0.1 mM DTT, pH 7.4). The equilibrated beads were incubated with the dialysed protein sample from second round Ni-NTA at 4°C, on a roller for 2 h. Following the incubation period the FT was collected and the resin bed washed twice with 5 column volumes of wash buffer each time (10 mM HEPES, 40 mM NaCl, 1 mM EDTA, 0.1 mM DTT, pH 7.4). The protein was eluted up to 7 times with 1 column volume elution buffer each time (10 mM HEPES, 500 mM NaCl, 1 mM EDTA, 0.1 mM DTT, pH 7.4), and whole elution fractions collected and saved. An aliquot of each fraction was saved for SDS and Native PAGE, silver staining and protein quantification as described in sections 2.2.5.5, 2.2.5.6 and 2.2.6.4 respectively.

2.2.12 Size Exclusion Chromatography

SEC was performed using two different types of resin, the first being Sephadex G50 (Sigma). Dry sephadex G50 media was soaked in running buffer (10 mM HEPES, 300 mM NaCl, pH 7.4) and packed into a 5 ml column, three quarters full, for SEC. This was performed on the bench top at room temperature with gravity flow delivering protein elution. The protein sample was loaded onto the column and allowed to soak completely into the resin before the bead bed was topped up with running buffer. A total of twelve 0.5 ml fractions were collected and an aliquot of each saved for SDS-PAGE, Coomassie blue staining and protein quantification as described in sections 2.2.5.5, 2.2.5.6 and 2.2.6.4 respectively.

The second type of resin employed for SEC was a pre-packed Superdex 75 10/300 GL column (GE Healthcare Life Sciences, Little Chalfont Buckinghamshire, UK). SEC with this resin was performed on ÄKTA Explorer instrumentation under the control of UNICORN 5.0 software (Amersham Biosciences) with a 1 ml injection loop. All SEC experiments were performed under manual control. Once the SEC column and injection loop were both connected to the system both pumps A and B and the injection loop were washed with running buffer to equilibrate. The flow rate was set to 0.5 ml/min and buffer was allowed to run through the column for 48 mins minimum to fully equilibrate.

A standard of 0.3 mg/ml ubiquitin, 8.5 kDa, (Sigma) was injected prior to each sample injection as a method to test the column and equipment performance and integrity and to serve as a good estimate as to what elution volume the EF2-GGC peptide may elute at as they are near in molecular weights.

Following sample injection, upon appearance of the first peak on the UNICORN chromatogram, 0.5 ml fractions were collected until the UV-Vis returned to a stable baseline and all of the protein had eluted from the SEC column. An aliquot of each fraction was saved for SDS-PAGE, silver staining and protein quantification as described in sections 2.2.5.5, 2.2.5.6 and 2.2.6.4 respectively.

Following the above protocol, three separate experiments were performed investigating the EF2-GGC peptide via SEC.

The first experiment investigated EF2-GGC peptide following elution from IEX. All samples were taken from IEX elution fractions. The first fraction, 1.6 mg/ml, had no changes/additions to buffer composition following elution from IEX. A second IEX elution fraction, 1.7mg/ml, was treated with 1 mM tris(2-carboxyethyl)phosphine (TCEP) incubation for 1 hr at room temperature prior to SEC, and a third IEX elution

fraction was treated with 10 mM TCEP incubation for 1 hr at room temperature prior to SEC. The running buffer, 10 mM HEPES, 150 mM NaCl, 1 mM EDTA, pH 7.4, was prepared, filter sterilized via a stericup-GP,0.22 µm, PES, 500 ml filter unit (MerckMillipore) under vacuum and degassed under vacuum for 15 mins while on a magnetic stirrer before addition of TCEP to a final concentration of 0.1 mM.

The second experiment investigated EF2-GGC peptide in the absence of sodium chloride. This investigation was performed on fractions collected from the first SEC experiment. Fractions 2 – 7, collected from all three samples of the first experiment were pooled and concentrated via Amicon® ultra-15 centrifugal filter units, 3.5 kDa MWCO (MerckMillipore). The spin concentrator's ultracel regenerated cellulose membrane was first rinsed in dH_2O to remove any residual glycerine or sodium azide before applying the sample. This was done by placing 5 ml dH_2O into the sample chamber, which was sitting in the provided collection tube and capped before centrifugation in a swinging bucket rotor at 4, 000 x g until for 10 mins. The dH_2O filtrate and retentate were discarded and the pooled protein sample was placed into the sample chamber. The sample was subjected to centrifugation in a swinging bucket rotor at 4, 000 x g until for 30 mins or until the sample was concentrated down to a final volume of 1 ml. The centrifugal filter membrane was rinsed out with 400 µl fresh buffer and this was added to the 1 ml concentrated protein.

This fraction was then exchanged into a sodium chloride-free buffer; 10 mM HEPES, 1 mM EDTA, 0.1 mM TCEP, pH 7.4 (this buffer was prepared as described in the first SEC experiment) using a smaller volume (total volume 2 ml) Amicon® ultra-2 centrifugal filter units 3.5 kDa MWCO (MerckMillipore). The 1.4 ml protein sample was concentrated via centrifugation, as previously described, down to 1 ml, to which an equal volume of sodium chloride-free buffer was added. This now 2 ml fraction was concentrated down again to a total volume of 1 ml and fresh buffer was added. These steps were repeated a total of three times to ensure complete buffer exchange. Again, following concentration the centrifugal filter membrane was rinsed out with 400 µl fresh buffer which this was added to the 1 ml concentrated protein to give a final volume of 1.4 ml. Protein quantification was determined via Nanodrop.

TCEP was added to the sample to a final concentration of 10 mM. Following a 1 hr incubation end-over-mixing at 4°C, 200 µl (0.142mg) was taken for SEC. 800 µl of fresh, sodium chloride-free running buffer was added to the 200 µl sample to bring the total volume up to 1 ml, as required to fill the injection loop. Another 200 µl (0.142mg) was taken for SEC after a 3 and 18 hr incubation.

The third and final SEC experiment investigated the EF2-GGC peptide conformation following a 3 hr incubation in the presence of 6 M GuHCl and 10 mm DTT at room temperature prior to SEC.

2.3 Methods – TEV Protease Production

2.3.1 Bacterial Starter Culture and Glycerol Stock Preparation

An inoculum loop was sterilized by washing in ethanol and passing through a flame. The loop was allowed to cool beside the flame before taking an inoculum of TEV protease *E.coli*, BL21 DE3 (Rosetta™) cells supplied from Addgene. This inoculum was streaked out for growth of single colonies on a dual selective (ampicillin and chloramphenicol antibiotics) TYE agar plate and incubated overnight (<16 hours) at 37°C.

Under a flame, one isolated colony from the TYE plate was picked to inoculate a starter culture; 10 ml of 2X TY media supplemented with 1% glucose and ampicillin and chloramphenicol. The starter culture was incubated at 37°C, 250 rpm for 16 h.

A glycerol stock of the TEV Protease bacterial strain was prepared as described earlier in 2.2.3.

2.3.2 TEV Protease Protein Expression

Three separate investigations into TEV protease expression was attempted (A, B and C), each on a pilot scale preparation of 100 ml total culture volumes. All expression cultures contained both ampicillin and chloramphenicol antibiotics. Standard SDS-PAGE, 12% gels, were run of the time points collected from each culture and Coomassie blue stained as per section 2.2.5.5 and 2.2.5.6

A. Protein expression was investigated initially by IPTG induction. This method had two changes to the protocol described in section 2.2.5.2. First, following the addition of IPTG, the temperature of the culture was reduced down to 30°C for the duration of growth and second, the culture was grown for a total of 8 h.

B. IPTG induction was trialled again with one more change to the protocol and auto induction was also included in the investigation. IPTG induction culture was grown for a total of 20h rather than 8h. Auto induction was performed with three changes to the protocol described in section 2.2.5.1. The culture was inoculated with a 1:20 dilution of the starter culture, grown at 30°C and for a total duration of 20 h.

C. A final investigation into TEV protease protein expression was performed via IPTG induction in new media adapted from Blommel & Fox, (2007): Terrific Broth (TB) (1.2 % peptone, 2.4 % yeast extract, 72 mM K_2HPO_4, 17 mM KH_2PO_4, 4 % glycerol), supplemented with 0.375 % aspartic acid and 2 mM $MgSO_4$. The IPTG induction was performed as described in section 2.2.5.2, i.e. the culture was inoculated with a 1:20 dilution of the starter culture, grown at 37°C until an OD_{600} of 0.6 was reached and IPTG was added to a final concentration of 1 mM. The culture was grown at 37°C, 250

rpm for 20 h. After 4 and 16 h, 1 ml of the culture was taken and saved for SDS-PAGE to investigate protein expression at that time point. The remaining cells of the culture were harvested via centrifugation at 5,000 rpm, 4°C for 20 minutes. The supernatant containing only the media was discarded and the pellet of bacterial cells stored at -20°C.

2.3.3 TEV Protease Protein Purification

Both native cell lysis and Ni-NTA IMAC were carried out as described in section 2.2.6. Changes to the method to note; no universal nucleases was added during cell lysis, dialysis buffer was 10 mM HEPES, 150 mM KCL, pH 7.4.

2.3.4 TEV Protease Functionality Test

The purified TEV protease was incubated with the substrate EF2-GGC recombinant fusion protein in order to firstly confirm that the enzyme is functionality active and secondly determine optimal cleavage conditions specific to the EF2-GGC recombinant fusion protein. All cleavage reactions were carried in 25 mM Tris-HCl, 0.1 mM EDTA, 0.5 mM DTT at pH 8.0, overnight at 4°C, at an enzyme to protein ratio of 1:50. The commercial enzyme, AcTEV protease acted as a positive control. Two negative controls were included, 1) the enzyme without any substrate and 2) the substrate without any enzyme, both in reaction conditions as outlined above. Three different total protein concentrations were investigated with the aim to identify if increased protein concentration lead to increased protein precipitation during cleavage. A total protein concentration of 2, 1 and 0.33 mg/ml reactions were investigated.

A second functionality test was carried out to investigate the effect of temperature on enzymatic activity. Two sets of reactions were carried out, one at 30°C and the other at 4°C. In each set a fixed amount of EF2-GGC recombinant fusion protein, 75 μg, was incubated with increasing concentrations of TEV protease; 0.5, 1, 2, 5 and 10 μg. An aliquot of each fraction was saved for SDS-PAGE and Coomassie blue staining as described in sections 2.2.5.5 and 2.2.5.6 respectively.

2.4 Methods – EF2-NP Functionality Investigation

Synthesis and coupling of purified EF2-GGC peptide was performed by Dr. Eugene Mahon, University College Dublin. 100 nm rhodamine doped silica cores were synthesized and amine functionalized before a bifunctional PEG_8-mal layer was covalently coupled. Thiol coupling of the EF2-GGC peptide to the malamide group of the PEG_8-mal layer resulted in irreversible coupling of the EF2-GGC peptide to the NP surface in a directional manner to create the EF2-NP affinity resin.

2.4.1 EF2-NP Purification

From a 1 mg/ml stock of EF2-NPs, 25 μl was taken and equilibrated in 10 mM HEPES, pH 7.4. This equilibrated resin was incubated with 100 μl of a native bacterial cell lysate containing EF1 tagged green fluorescent protein (EF1-GFP), for 12 mins at room temperature (native cell lysis of the EF1-GFP bacterial cell pellets was carried out according to section 2.2.6.1, with one change, a different lysis buffer of 10 mM HEPES, 300 mM NaCl, 2 mM $CaCl_2$ pH 7.4 was used). The EF2-NP cell lysate mixture was washed twice with 100 μl of lysis buffer each time. The captured protein was eluted twice with 100 μl of elution buffer each time (10 mM HEPES, 150 mM NaCl, 10 mM EDTA, pH 7.4). An aliquot of each fraction was saved for GFP fluorescence analysis, Native PAGE and silver staining as described in sections 2.4.2, 2.2.5.5 and 2.2.5.6 respectively.

2.4.2 GFP Fluorescence Analysis

GFP fluorescence in all fractions collected from EF2-NP purification was measured by excitation at 410 nm and absorbance at 510 nm in a 96 well plate by Spectra Max 190 plate reader. An image of the plate following this was captured by Fujifilm LAS-3,000 Intelligent Dark Box under blue light. This instrument was also used to capture an image of the samples following native PAGE.

Chapter 3:
Results

3.1 Results - EF2-GGC Peptide Production

A schematic overview of the experimental work flow to produce the EF2-GGC peptide is presented in figure 3.1. Results from each step of this optimized production line are shown in this section.
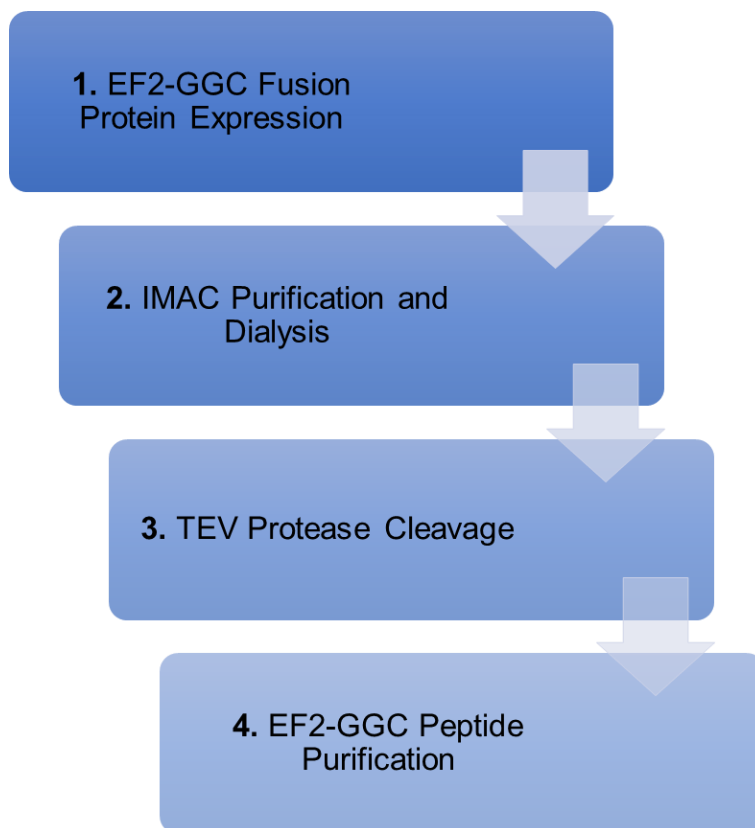

Figure 3.1. EF2-GGC Peptide Production Workflow Overview

3.1.1 Fusion Protein Expression

EF2-GGC recombinant fusion protein over-expression was explored by two induction methods: autoinduction and IPTG induction in figure 3.3, A) and B). Pilot scale culture volumes of 100 ml were investigated at fixed time points, before scaling up to 2 litre expression cultures. The apparent molecular weight of the EF2-GGC recombinant fusion protein is indicated by the red arrow at 36.446 kDa.

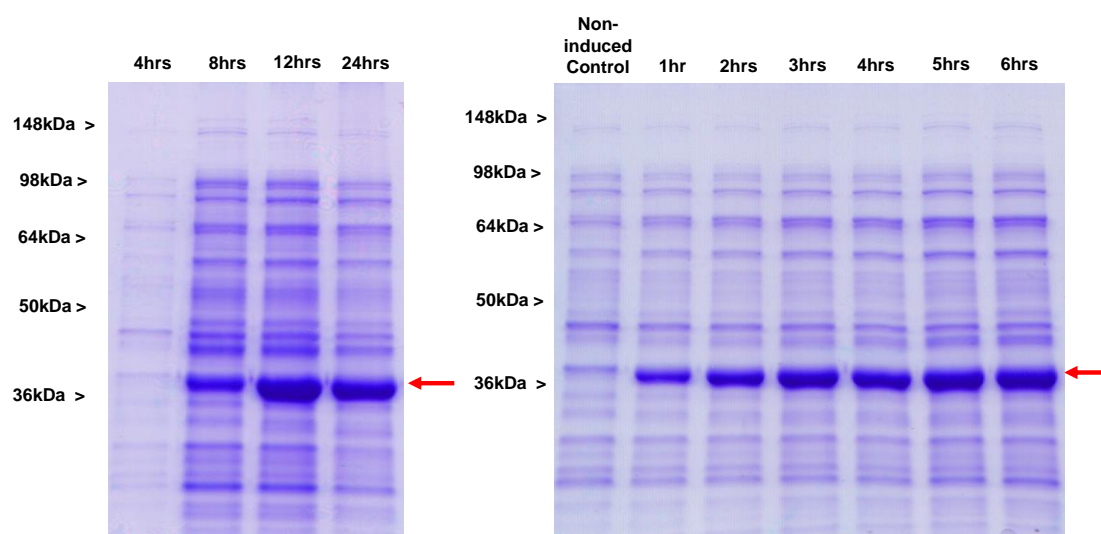**A) Auto Induction**

B) IPTG Induction

Figure 3.2. EF2-GGC Recombinant Fusion Protein Expression

Time points collected from pilot expression cultures of the EF2-GGC recombinant fusion protein (red arrow) were separated on a 10 % SDS-PAGE with Coomassie blue staining. A) Autoinduction culture with time points collected at 4 h intervals for a total of 24 h. B) IPTG-induction culture with time points collected every hour for a total of 6 h.

EF2-GGC recombinant fusion protein over-expression was successful in all attempts. The protein band seen just above the 36 kDa marker (red arrow) of figure 3.2 became more intense as the time of growth continued indicating this particular protein had increased expression over time. This along with the correct size positioning of the protein band suggests that this was correct over expression of the EF2-GGC recombinant fusion protein. Optimum protein expression levels can be seen at 12 h of growth for auto-induction and 6 h growth for IPTG induction, with similar band intensities between them.

3.1.2 Fusion Protein Purification – IMAC

Bacterial cells from 12 h autoindcution, containing the over-expressed EF2-GGC recombinant fusion protein were lysed in a pre-cooled native lysis buffer and Ni-NTA agarose was introduced for IMAC purification of the over expressed EF2-GGC fusion protein. From figure 3.2, 12 h autoindcution appeared to have the optimum level of target protein expression and so this expression culture condition and time point was chosen for protein purification.
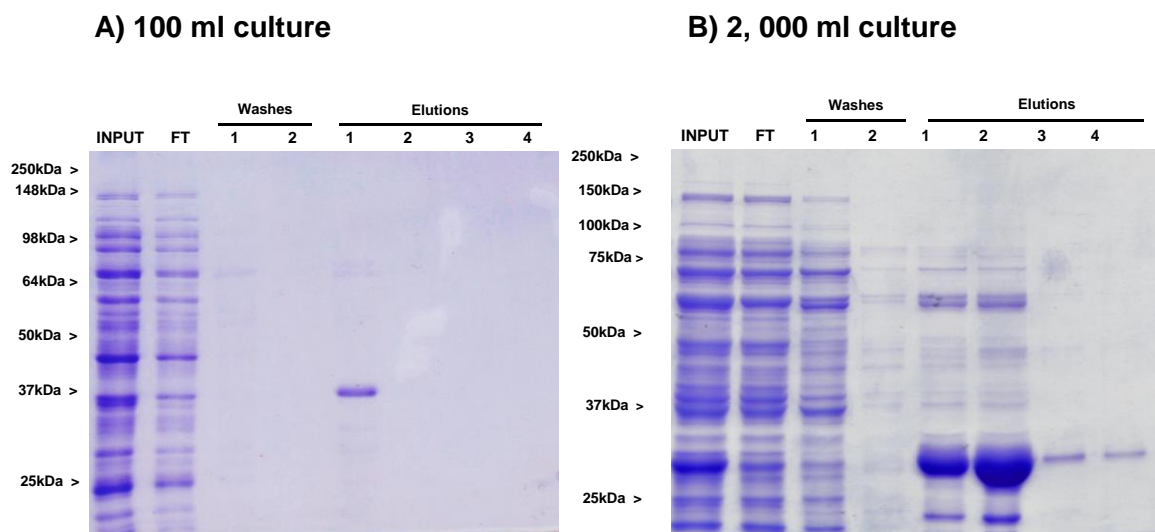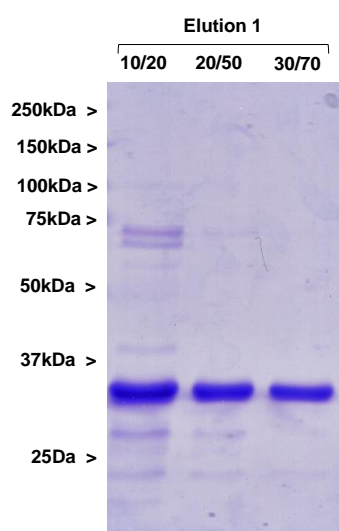
A) 100 ml culture   **B) 2, 000 ml culture**

Figure 3.3. Native IMAC purification of the EF2-GGC recombinant fusion protein. Fractions were separated on a 10 % SDS-PAGE with Coomassie blue staining. Input = cell lysate, FT = flow through, Washes and elutions are bracketed and numbered. (A). IMAC of a pilot scale expression culture (B). IMAC of a large scale expression culture volume (both originate from autoinduction cultures)

The first lane labelled "input" of figure 3.3, the protein band present just below the 37 kDa marker indicates the over expressed EF2-GGC fusion protein was present in the cell lysate. A band of the same molecular weight in the FT and wash lanes was not seen in image (A), indicating the resin had captured all fusion protein in the cell lysate and none was washed away. However when compared to image (B), there was a band at the same molecular weight (36.4 kDa) seen in the FT and wash lanes, however was much less intense indicating that despite the majority of the EF2-GGC fusion protein binding to the resin, some had not or had weakly bound and had been washed away. A band in both images (A) and (B), just under 37 kDa can be seen in the elution lanes

indicating native Ni-NTA purification was successful at purifying the EF2-GGC fusion protein from bacterial cell lysate. Particularly intense in elution lanes 1 and 2 of image (B), revealing that the displacement of the EF2-GGC recombinant fusion protein was quite immediate. Protein quantification was performed after dialysis because the high imidazole concentration of the elution buffer were incompatible with BCA assay reagents. On average a total protein yield of 49.5 mg/L was returned (n = 2) (Table 3.2).

3.1.2.1 IMAC Optimization

Ni-NTA IMAC was optimized with the objective to obtain the maximum yield of target protein with highest purity. This was achieved through varying the concentration in the lysis and wash buffers.

Elution 1	10/20	20/50	30/70
Total Protein Conc. (mg/ml)	1.09	0.48	0.39
Densitometry (%)	1.0	0.71	0.64

Table 3.1. Densitometry Analysis of EF2-GGC fusion protein bands in figure 3.5.

Figure 3.4. Optimization of Ni-NTA IMAC.

Elution one fractions from three different IMAC purifications of the EF2-GGC recombinant fusion protein, varying concentrations of imidazole in the lysis and wash buffers were separated on a 10 % SDS-PAGE with Coomassie blue staining. 10/20 = 10 mM imidazole in the lysis buffer and 20 mM imidazole in the wash buffer, 20/50 = 20 mM imidazole in the lysis and 50 mM imidazole in the wash buffer, 30/70 = as previous, 30 mM lysis/70 mM wash.

Increasing the concentration of imidazole in the lysis and wash buffers lead to less non-specific protein binding to the Ni-NTA resin and gave greater purity of the eluted protein. However this also decreased the yield of target protein retrieved as determined by a combination of protein quantification by a BCA assay and densitometry analysis of the EF2-GGC fusion protein bands in figure 3.4. Table 3.1 summarises the densitometry analysis providing the percentage band intensity of the EF2-GGC recombinant protein bands from all three samples, versus the concentration of total protein in each elution fraction.

3.1.3 Fusion Protein AcTEV Protease Cleavage

The commercial enzyme, AcTEV protease, was incubated with semi-purified EF2-GGC recombinant fusion protein for enzymatic cleavage at a specific location to achieve separation of 6xHis CA from the peptide of interest, EF2-GGC.
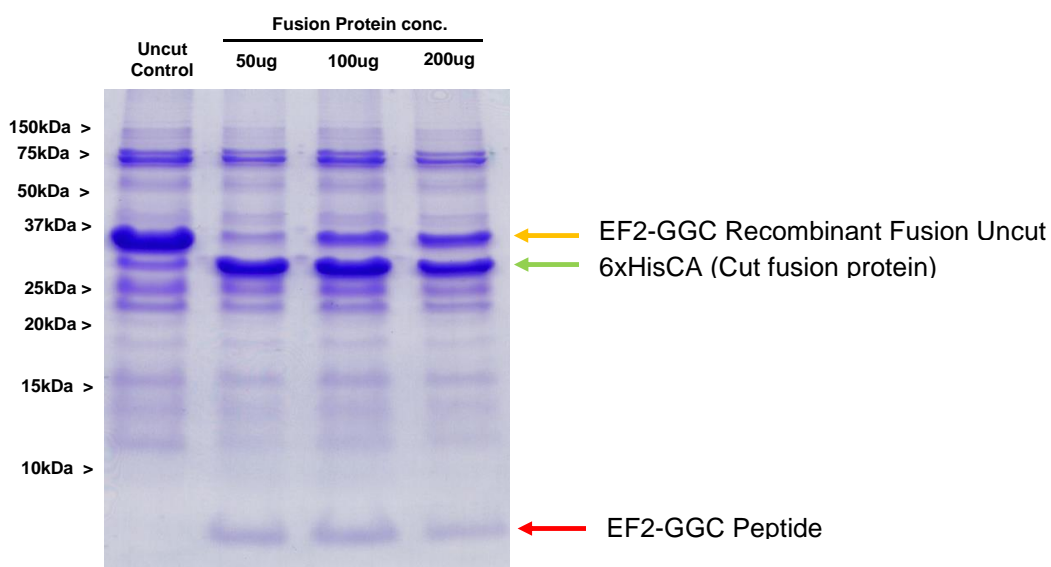
Figure 3.5. TEV protease cleavage of the EF2-GGC recombinant fusion protein. Fractions were separated on a 16 % SDS-PAGE with Coomassie blue staining. The first lane contains the negative control, uncut EF2-GGC recombinant fusion protein. The following three lanes contain increasing concentrations of fusion protein (bracketed), 50, 100 and 200 μg, to a fixed 10 units of commercial AcTEV protease and show protein bands for the uncut EF2-GGC recombinant fusion protein (yellow arrow), cut EF2-GGC recombinant fusion protein, therefore just 6xHis CA (green arrow) and the released EF2-GGC peptide (red arrow).

AcTEV protease was successful in catalysing the cleavage of the EF2-GGC fusion protein as concluded by the presence of the protein band under the 10 kDa marker (red arrow) indicating this was the EF2-GGC peptide at ~4.38 kDa (Figure 3.5). The cut fusion protein having lost 4.38 kDa was now smaller in molecular weight, 32.06 kDa, and was present just above the 25 kDa marker (green arrow). Remaining uncut EF2-GGC recombinant fusion protein can be seen at 32 kDa (orange arrow).

In this experiment increasing concentrations of fusion protein, 50, 100 and 200 µg, to a fixed 10 units of commercial AcTEV protease enzyme was investigated to obtain the optimum fusion protein to enzyme ratio for cleavage conditions. The first lane contains the uncut control (or negative control), fusion protein with no enzyme, thus the protein remains at the molecular weight of the complete recombinant fusion protein; 36.4 kDa as it is not cleaved. In the following lanes, fusion protein with enzyme, all lanes showed successful cleavage of the fusion protein. Lane two can be concluded to have the best cleavage with minimal fusion protein remaining at ~ 36.4 kDa as the majority had been cleaved by the enzyme. Lanes three and four required more than 10 units of AcTEV protease in order to effectively cleave all of the fusion protein.

Production of the TEV protease and optimization of cleavage conditions was investigated further in section 3.2.

3.1.4 EF2-GGC Peptide Purification

A number of different methods were investigated for their ability to purify the EF2-GGC peptide from 6xHis CA and TEV protease following cleavage. A simple size exclusion approach was tested first in the form of spin concentrators. Another method tested was purification by an EF2-agarose affinity resin whereby homo-dimerization of the EF2 hands were considered as a method of purification.

A second round of Ni-NTA chromatography was the obvious choice as an indirect purification method as it can bind both his tagged proteins, 6xHis CA and 6xHis TEV protease, separating them from the untagged EF2-GGC peptide. Another route to purify the peptide was to employ IEX. This method of protein purification was chosen because the isoelectric point (pI) of the EF2-GGC peptide was estimated to be 4.21 by the bioinformatics tool ExPASY and so in buffers with a pH above this, the peptide would exhibit a net negative charge (Table 3.2). This characteristic was exploited using a weak anion exchanger, diethylaminoethanol (DEAE), to which it could be assumed

that the EF2-GGC peptide would demonstrate tight binding, separating the peptide from all other positively or neutral charged proteins i.e. 6xHis CA pI of 7.00 (Table 3.2) that would not bind to the resin.

And finally, SEC was employed initially as a polishing step to provide high grade purity of EF2-GGC peptide. 6xHis CA, and 6xHis TEV protease are 32 and 27 kDa respectively and the EF2-GGC peptide is 4.38 kDa (Table 1.3). This significant difference in size allows for good separation of the protein mixture if any remaining 6xHis CA or TEV protease remained.

Table 3.2 Overview of the EF2-GGC Recombinant Protein characteristics, before and after TEV protease cleavage

Protein	EF2-GGC Recombinant Fusion Protein	6xHis CA	EF2-GGC Peptide	6xHis TEV Protease
No. of Amino Acids	329	289	40	249
Molecular Weight (kDa)	36.446	32.081	4.382	28.617
pI	6.27	7.00	4.21	9.95
Extinction Co-efficient	52,035	46,410	5,500	32,220

Despite the abundant types of protein purification techniques available to achieve purification of the EF2-GGC peptide, one concern was the stability of the peptide following cleavage. 6xHis CA has acted as a stabilizing, solubility enhancing partner protein and so separation from this protein by TEV cleavage was a concern as it was unknown as to how this recombinantly produced peptide will formulate without the stability of 6xHis CA. Consequently investigations into EF2-GGC peptide formulation was an objective before the next steps in building the nanoscale EF2 affinity resin could be carried out, thus so along with purification, investigations the effect of different buffer compositions upon peptide conformation was achieved through SEC.

3.1.4.1 Spin Concentrators

Pierce protein spin concentrators PES 10 kDa MWCO, 0.5 ml were investigated as an alternative method of size exclusion to purify the EF2-GGC peptide from the larger molecular weight protein contaminants.
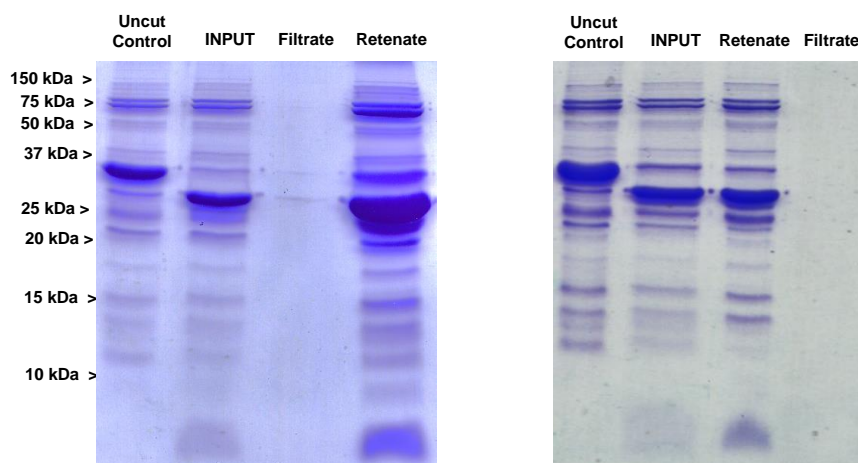

Figure 3.6. Purification of the EF2-GGC peptide via Spin Concentrators.
Fractions were separated on a 16 % SDS-PAGE with Coomassie blue staining. (A) Untreated (B) Sample treated with 1 mM DTT for 1 hr at room temp, prior to spin column purification

The first lane in each image, (A) and (B) of figure 3.6 shows the EF2-GGC recombinant fusion protein before TEV protease cleavage. The second lane labelled shows the EF2-GGC recombinant fusion protein after TEV protease cleavage and was the sample put into the spin concentrator for separation, i.e. labelled "input". No protein was found to have passed through the 10 kDa MWCO spin concentrators as no protein was found in the filtrate fractions. The sample that was left behind, (i.e. that did not pass through the membrane) labelled "retentate" was a thus a concentrated fraction of the original input sample. Incubation of the sample with 1 mM DTT prior to spin concentration did not have any effect on protein separation.

3.1.4.2 EF2-Agarose Purification

Another method investigated for EF2-GGC peptide purification was the EF1-EF2 protein purification system itself. EF2-agarose resin was incubated with the TEV protease cleaved fusion protein and purification carried out as described in section 2.2.9.

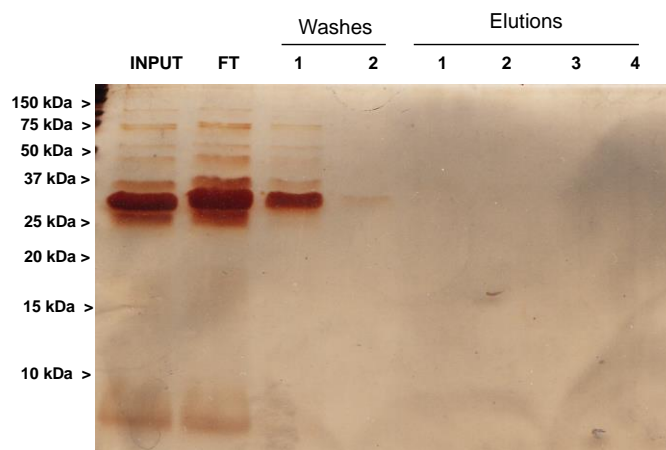**Figure 3.7. Purification of the EF2-GGC peptide via EF2-agarose resin.**
Fractions separated on a 16 % SDS-PAGE with silver staining. INPUT = cell lysate, FT = flow through, Washes and elutions are bracketed and numbered.

This method failed to purify the EF2-GGC peptide as no protein appeared to have bound to the EF2-agarose resin as all protein eluted in the FT and wash fractions (Figure 3.7).

3.1.4.3 Second Round IMAC

A second round of IMAC was investigated as a method to remove 6xHis CA and 6xHis TEV protease to purify the EF2-GGC peptide.
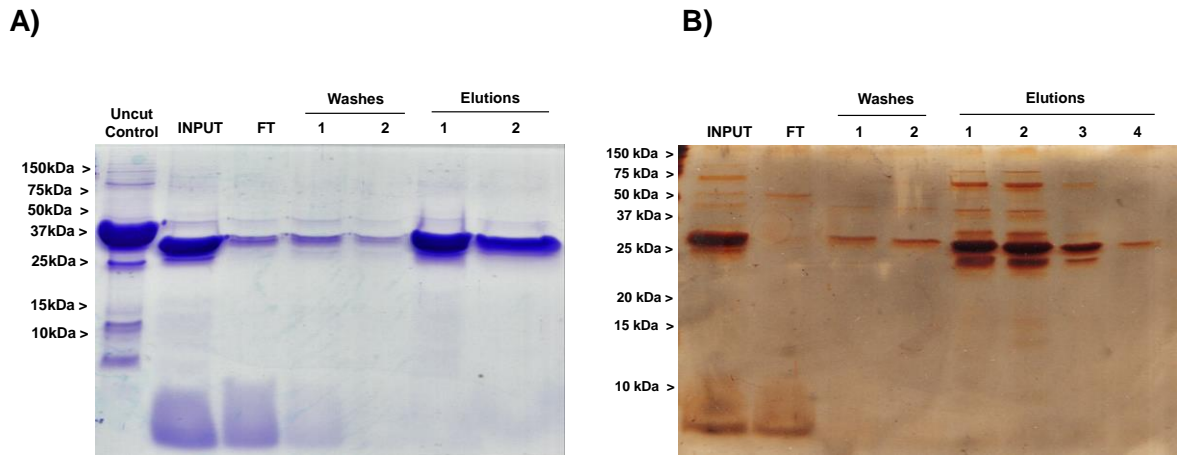
A)   **B)**

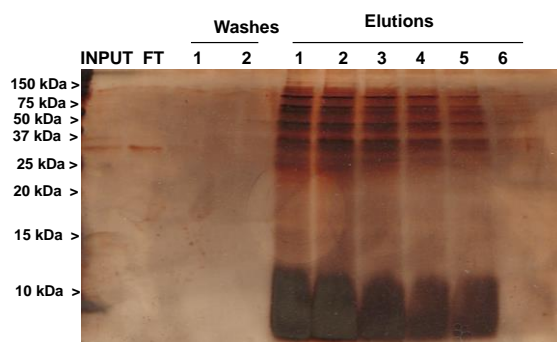**Figure 3.8. Second round of IMAC for purification of the EF2-GGC peptide.**
Fractions separated on a 16 % SDS-PAGE. A) Coomassie blue staining. B) Silver staining. INPUT = cell lysate, FT = flow through, Washes and elutions are bracketed and numbered.

The EF2-GGC peptide had no affinity tag and thus did not bind to the Ni-NTA agarose resin as proven by the protein band below the 10 kDa marker in both images A) and B) of figure 3.8, present in the FT and wash fractions. As CA and TEV protease both have 6xHis tags they did bind to the resin and were seen in the elution lanes. Some 6xHis CA and possibly TEV protease upper bands appeared with the EF2-GGC peptide in the FT and wash lanes indicating that these proteins were not bound by the resin. Despite this the majority is captured by the Ni-NTA resin as they are represented by particularly intense protein bands in the elution lanes. In this way the majority of 6xHis CA and TEV protease was separated from the EF2-GGC peptide but because a small amount of 6xHis CA (and possibly TEV protease) was still present in the FT and wash 1 lanes containing the EF2-GGC peptide, other purification steps were necessary to remove these contaminants.

3.1.4.4 Ion Exchange Chromatography

To ensure complete removal of the 6xHis CA and possible TEV protease upper bands and obtain a pure EF2-GGC peptide preparation another purification step using IEX was investigated.

A) SDS-PAGE   **B) Native PAGE**
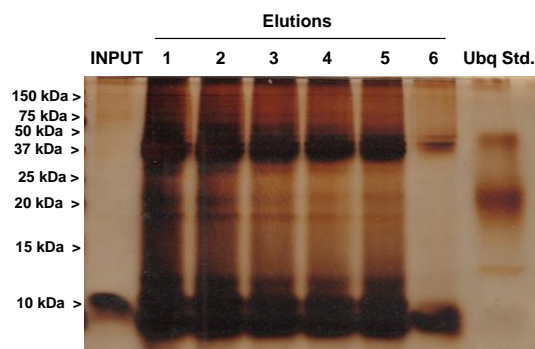

Figure 3.9. IEX for purification of the EF2-GGC peptide.
Fractions were separated on 16 % polyacrylamide gels. A) SDS-PAGE B) Native PAGE (no SDS present). Both gels were subjected to silver staining. INPUT = cell lysate, FT = flow through, Washes and elutions are bracketed and numbered.

The positively charged DEAE resin of IEX was successful in the capture and elution of the negatively charged EF2-GGC peptide as shown by the presence of the EF2-GGC peptide band in all elution lanes (Figure 3.9). However there were also higher molecular weight bands in all elution lanes which may possibly represent aggregated EF2-GGC peptide or may still represent the presence of some higher mw contaminating proteins, possibly 6xHis CA. Native PAGE suggests that those upper bands were aggregated EF2-GGC peptide as the higher molecular weight bands in image B) were greatly reduced comparing to the same mw region in image A), suggesting that the denaturing sample preparation of image A) fractions reduced peptide aggregation. A yield of 3.7 mg/L of EF2-GGC peptide was obtained from IEX, determined by UV-Vis (n=2) (Table 3.3).

Table 3.3 Summary of total protein yields following each step in EF2-GGC peptide production

Purification Process	Total Purified Protein Yield (mg/L)		
	n = 1	n = 2	Average
IMAC I	52.25	46.85	49.55
IMAC II	14.22	11.3	12.76
% Return from IMAC II	27.2	24.12	25.66
IEX	4.25	3.32	3.79
Theoretical EF2-GGC Peptide Yield	6.3	5.64	5.97
Actual % EF2 peptide obtained	67.46	58.86	63.16

Protein quantification at this point revealed the amount of pure EF2-GGC peptide in total that was retrieved from this protocol. On average (n=2) 3.79 mg/L of EF2-GGC peptide was obtained using this protocol (Table 3.3). When compared to table 1.2, this result is moderately better than most reported there. The peptide was 8.3 times smaller than the complete fusion protein and an estimation as to what the theoretical peptide yield would be was calculated (Table 3.3). Calculations found that on average (n = 2) 63.16 % of the theoretical yield of EF2-GGC peptide was recovered using this protocol. Further optimizations would be necessary to achieve nearer to a 100 % return.

3.1.4.5 Size Exclusion Chromatography

SEC was investigated as an alternative method to achieve purification of the EF2-GGC peptide. Two different resins were investigated for SEC. Sephadex G50 was tested first on a TEV cleaved EF2-GGC fusion protein.
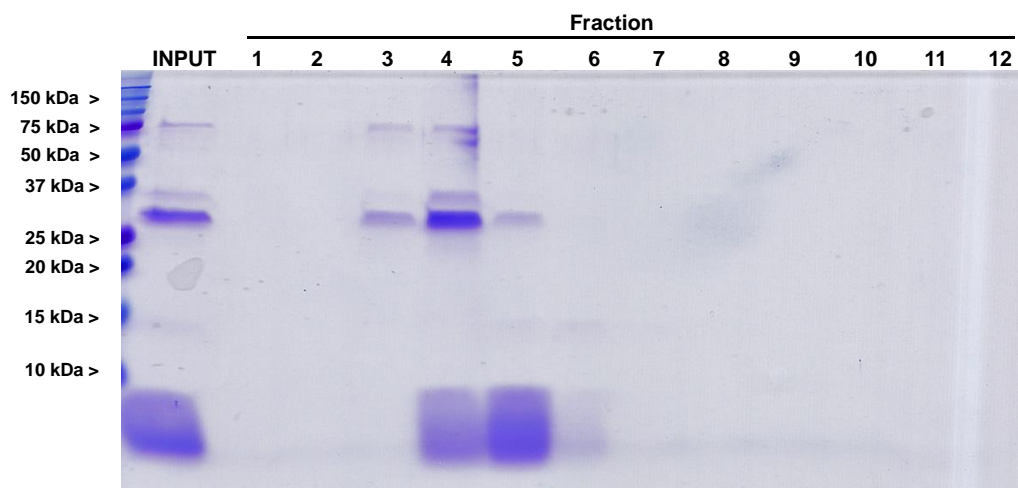
Figure 3.10. SEC via Sephadex G50 for purification of the EF2-GGC peptide.
The fractions collected, numbered and bracketed, were separated on a 16 % SDS-PAGE and Coomassie blue stained.

Fractions 3, 4, 5 and 6 were the only fractions to contain protein in figure 3.10. SEC here was semi-successful with some separation between the higher mw proteins from the EF2-GGC peptide. However as fraction 4 was quite similar to the input fraction this indicates that the protein sample was not separated to a satisfactory degree and thus this method did not prove to be useful for size exclusion purification of the EF2-GGC peptide.

SEC was attempted again with a different resin, namely Superdex 75 in combination with FPLC. This resin has a separation range for proteins between 3,000 and 70,000 Daltons and was employed as a final polishing step to ensure purification of the EF2-GGC peptide following second round IMAC and IEX steps and to aid in investigations into variations in the peptide buffer conditions.

A) i
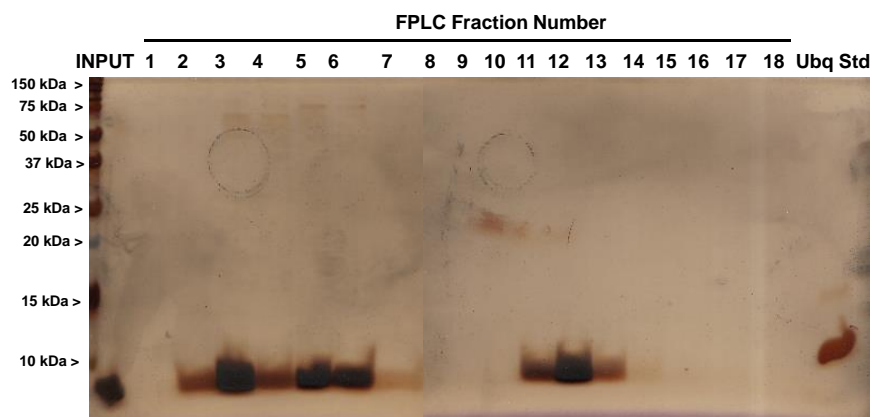
A) ii
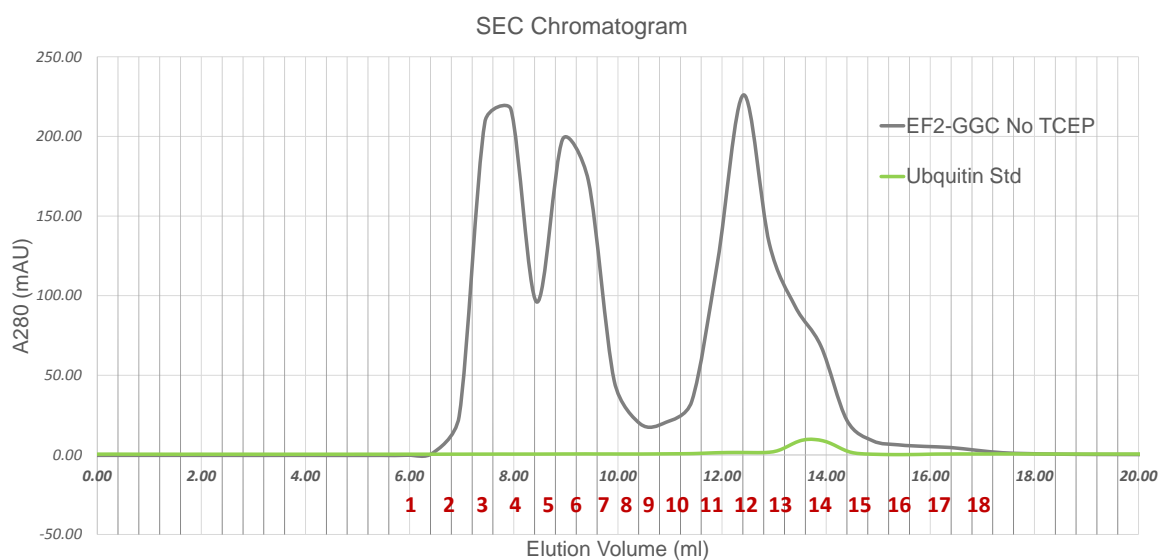
Figure 3.11. SEC via FPLC for purification of the EF2-GGC peptide – TCEP Treatment

A) EF2-GGC, No TCEP i) 16 % SDS-PAGE with silver staining of fractions collected following elution from SEC column. ii) SEC chromatogram, UV-Vis A280 (mAU) vs Elution volume (ml), fraction numbers in red under x-axis.
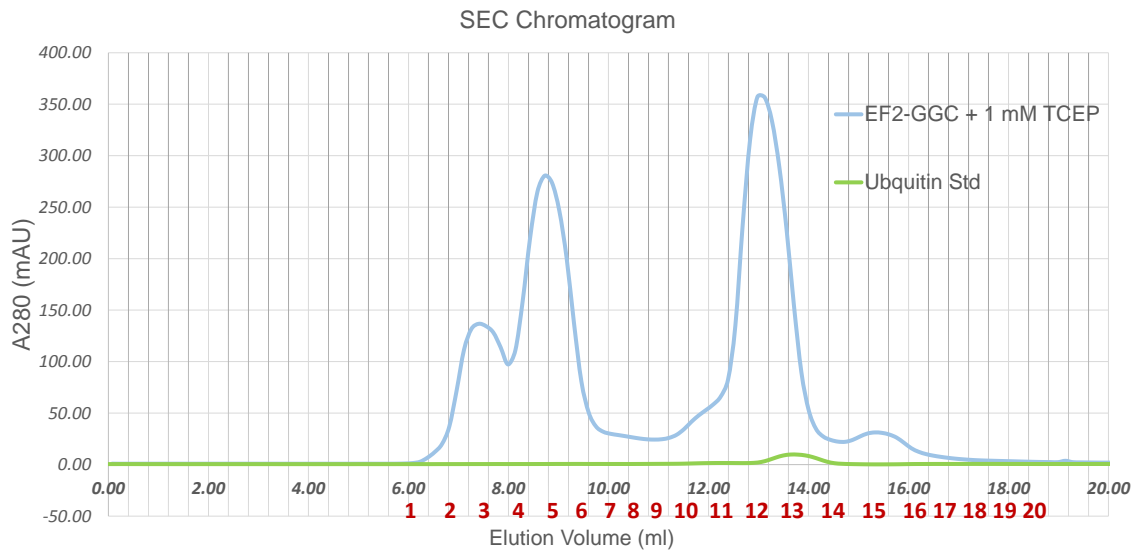
B) i
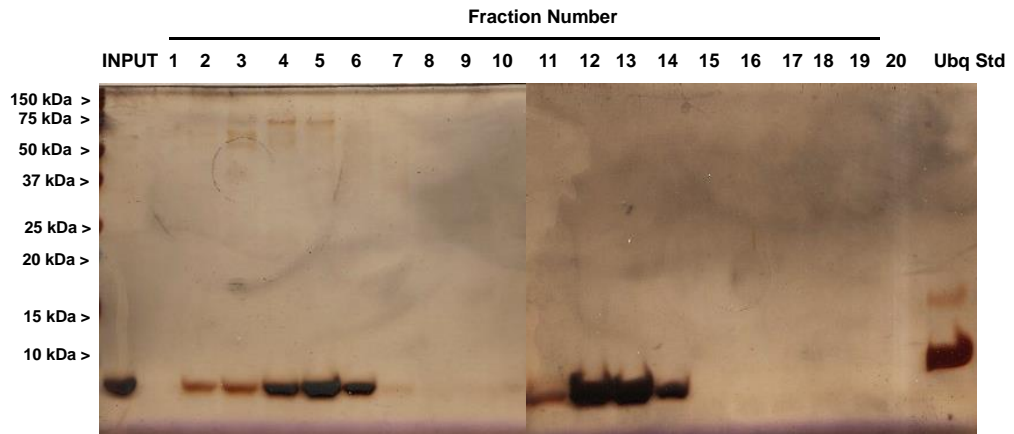
B) ii

Figure 3.11. SEC via FPLC for purification of the EF2-GGC peptide – TCEP Treatment

B) EF2-GGC + 1 mM TCEP treatment prior to SEC. i) 16 % SDS-PAGE with silver staining of fractions collected following elution from SEC column. ii) SEC chromatogram, UV-Vis A280 (mAU) vs Elution volume (ml), fraction numbers in red under x-axis.
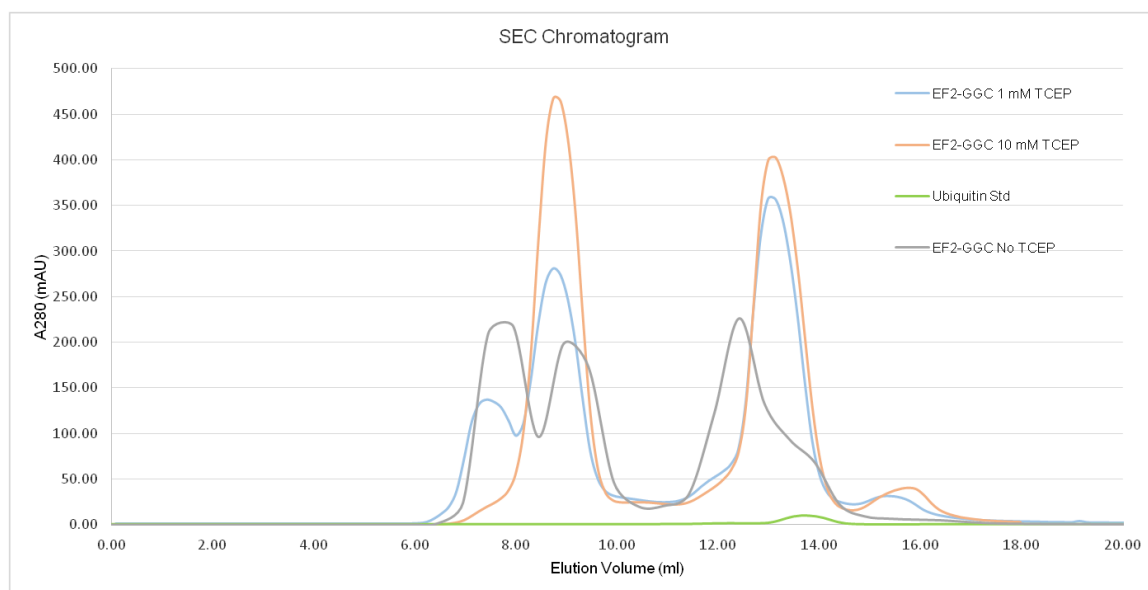
C)

Figure 3.11. SEC via FPLC for purification of the EF2-GGC peptide – TCEP Treatment

C) Over-laid SEC chromatogram presenting 4 samples, UV-Vis A280 (mAU) vs Elution volume (ml).

Elution fractions from IEX, containing EF2-GGC peptide were resolved by SEC and the results presented in figure 3.11. Figure 3.11 A) i and B) i show fractions collected upon elution of the protein from the SEC column separated on 16 % standard SDS-PAGE and silver stained. All fractions collected are numbered and bracketed under the label "Fractions". Figure 3.11 A) ii and B) ii are SEC chromatograms of the eluted protein sample corresponding to A) i and B) i respectively. Figure3.11C) is a SEC chromatogram where chromatograms of three different samples were overlaid into the one graph for direct comparison of elution volumes.

Interpretations of the results presented in figure 3.11 indicate that there was a high degree of aggregation of the EF2-GGC peptide. The peptide elutes from the SEC column over a broad range of fractions indicating that there were different molecular weight proteins present in the sample, which therefore could suggest a high degree of different molecular weight aggregates, or perhaps the presence of higher molecular weight contaminants. With a molecular weight of 4.382 kDa the peptide was expected to elute from the column around ~ 16 ml, after the ubiquitin (8.5 kDa) standard peak at

~14 ml, however this did not occur and peaks in all chromatograms can be seen much earlier around 6.5 ml. The SEC chromatogram analysis in combination with SDS-PAGE of the fractions collected indicate that there was EF2-GGC peptide present in those early peak fractions along with some higher molecular weight contaminants as there were faint protein bands around the 50 and 75 kDa marker (Figure 3.11 A) i and B) i). However the majority of the protein in those early fractions was the EF2-GGC peptide suggesting that it was peptide aggregation that was responsible for the peaks seen at 6.5 ml and onwards elution volume and not the presence of higher molecular weight proteins.

Following interpretation of Figure 3.11A) i and ii, TCEP was added in an effort to reduce/eliminate this high aggregation of the EF2-GGC peptide. Incubation of the EF2-GGC peptide in 1 mM TCEP for 1 h at room temperature prior to sample injection did reduce aggregation of the EF2-GGC peptide. This was proven in the chromatograms as the first peak in image A) ii had reduced in height when comparing to the same peak in B) ii (see C) for overlapped chromatograms). Also it can be seen in the SDS-PAGE analysis where the protein bands for the EF2-GGC peptide were less intense in fractions 2-7 of B) i when compared to the same fractions in A) i. A third sample was investigated where the amount of TCEP was increased to 10 mM for 1 h incubation at room temperature prior to SEC. This did reduce peptide aggregation, proven by C), the peaks indicating protein elution had shifted to the right at a later elution volume and thus a lower molecular weight protein. Nevertheless this step did not eliminate peptide aggregation completely.

Another attempt to reverse aggregation of the EF2-GGC peptide was to remove sodium chloride from the buffer composition as this was thought to be promoting aggregation.
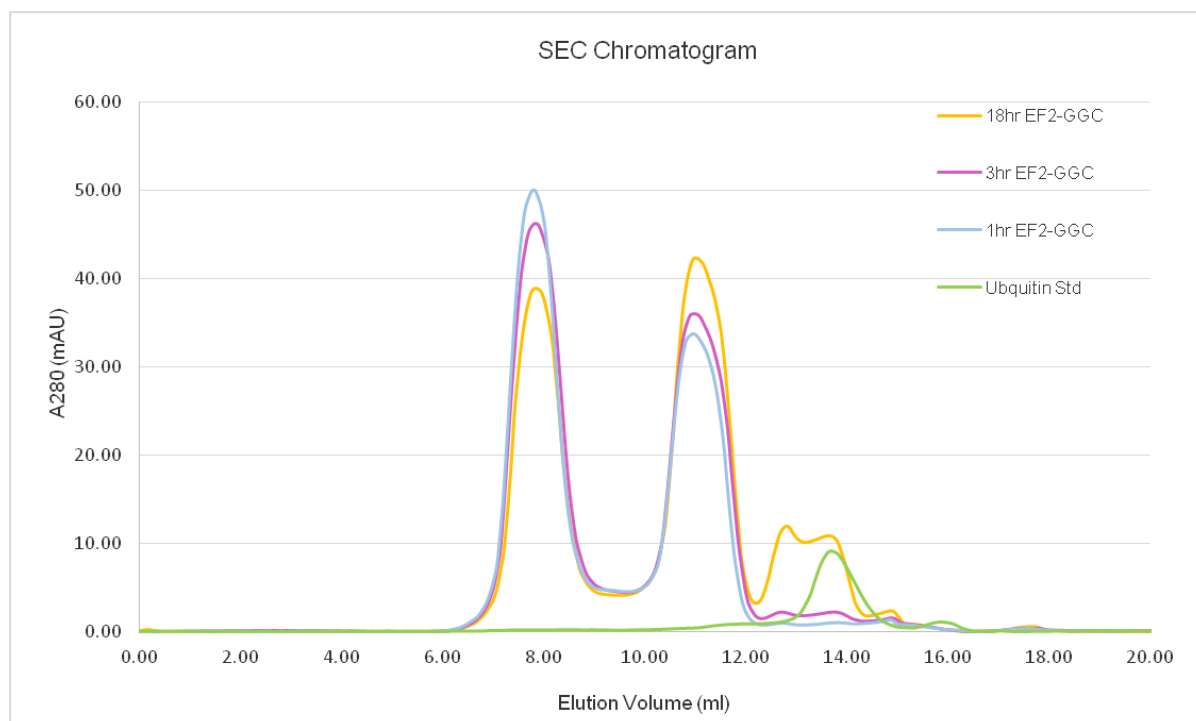
Figure 3.12. SEC via FPLC for purification of the EF2-GGC peptide – No Sodium Chloride Treatment.
Over-laid SEC chromatogram presenting 4 samples, UV-Vis A280 (mAU) vs Elution volume (ml).

The EF2-GGC peptide was exchanged into a sodium chloride-free buffer (10 mM HEPES, 1 mM EDTA, 0.1 mM TCEP pH 7.4) and incubated in this for 1, 3 and 18 h prior to SEC via FPLC. Again the removal of sodium chloride did help to reduce aggregation of the EF2-GGC peptide, but did not eliminate peptide aggregation completely (Figure 3.12). The longer the incubation in a sodium chloride-free buffer the less aggregation of the EF2-GGC peptide as observed in the 18 hr sample as it had a lower peak height comparing to the 1 and 3 h samples at 8 ml elution volume. The 18 h sample increased in peak height at 13 ml. This suggests that the aggregated peptide in the 8 ml peak (first peak) had reduced in aggregation and was seen to elute from the column in a later elution volume of 13 ml indicating a lower molecular weight (Figure 3.12).

To fully denature the EF2-GGC peptide and in that way eliminate peptide aggregation it was incubated in 6 M guanidine hydrochloric acid (GuHCl) and 10 mM DTT for 3 h at room temperature prior to SEC via FPLC.

(A)
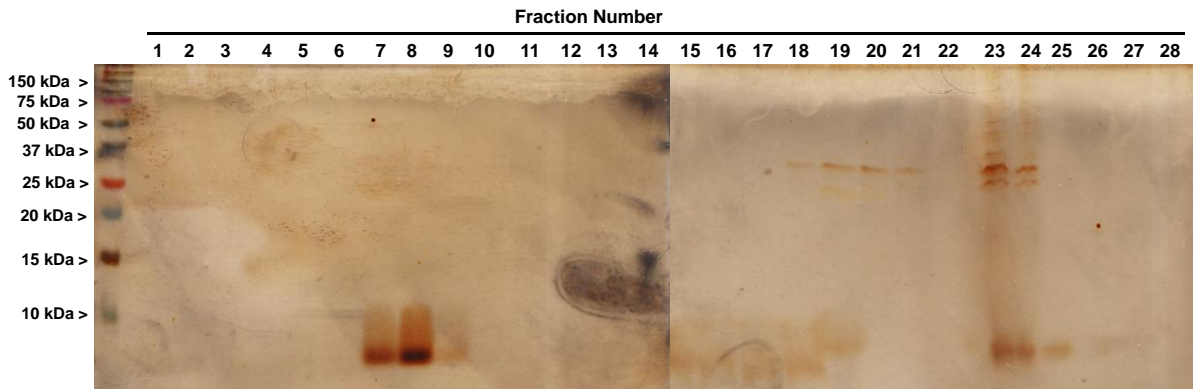
(B)
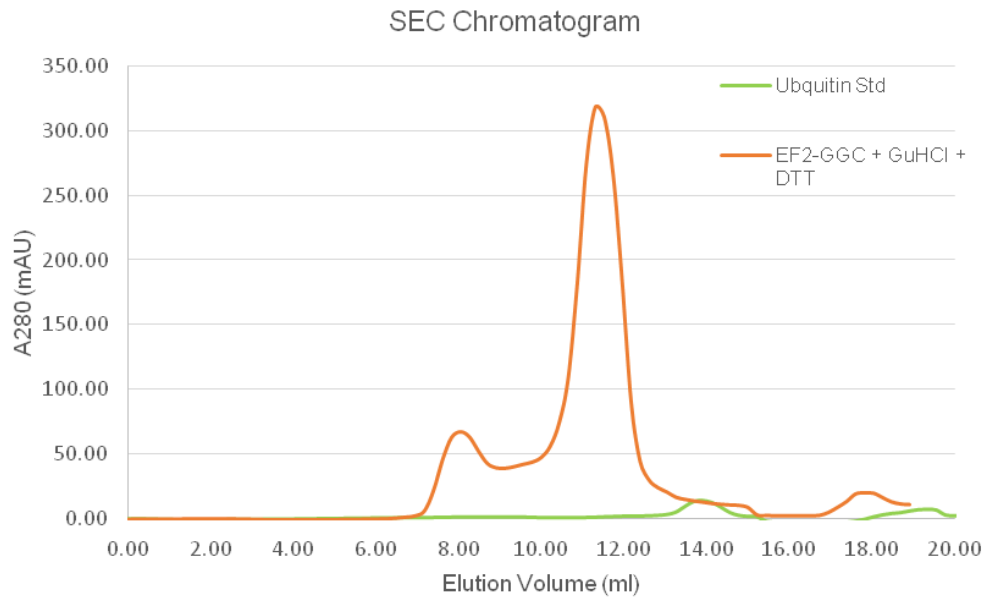
Figure 3.13. SEC via FPLC for purification of the EF2-GGC peptide – 6M GuHCl and 10 mM DTT Treatment.

(A) 16 % SDS-PAGE with silver staining of fractions collected following elution from SEC column, (B) SEC chromatogram, UV-Vis A280 (mAU) vs Elution volume (ml).

Incubation in 6 M GuHCl and 10 mM DTT did reduce peptide aggregation as shown by the large peak in the chromatogram of figure 3.13, appearing at a later elution volume of 11.5/12 ml in comparison to previous SEC experiments shown in figures 3.11 and 3.12. This later elution volume indicates smaller protein and thus possibly less aggregation of the peptide. Despite this later elution volume this method still failed to produce non-aggregated peptide at an expected elution volume of 16 ml or later.

3.2 Results - TEV Protease Production; Commercial versus In-house

An alternative to purchasing the commercial AcTEV protease was to express and produce the enzyme in house.

3.2.1 TEV Protease Protein Expression

TEV protease expression was screened at a pilot scale in three separate investigations as outlined in section 2.3.2. The protein was translated as a fusion protein to MBP with a TEV cleavage sequence in between. In this way, due to the presence of this protease *in vivo*, it catalysed the cleavage of this fusion protein and separated the 6xHis TEV protease from the MBP during culture incubation. Therefore two over-expressed protein bands may be interpreted from the SDS-PAGE analysis, one for MBP at 42.5 kDa and another for 6xHis TEV protease at 27.5 kDa (Figure 3.14).
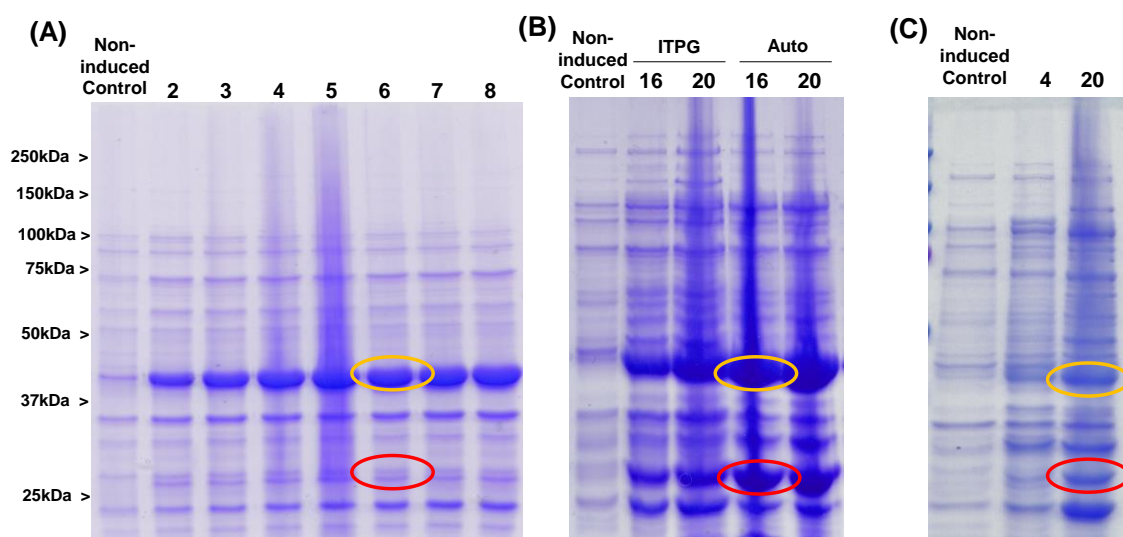
Figure 3.14. 6xHis TEV Protease Protein Expression

A-C samples were separated on a 12 % SDS-PAGE with Coomassie blue staining, numbers above lanes indicate number of hours at which the sample was taken, MBP (orange circle) and 6xHis TEV protease (red circle) (A) IPTG Induction with time points collected every hour for a total of 8 h (B) Revised IPTG-Induction and new auto-induction culture with time points collected after 16 and 20 h only (C) IPTG induction using new supplemented TB media.

Initial 6xHis TEV protease expression culture represented in Figure 3.14A did not show any noteworthy over expression of the 6xHis TEV protease at any time point. The band just above the 37 kDa marker, MBP, orange circle, did show a slight increase in intensity as the time of culture incubation continued indicating slight protein expression over time. This suggests that the TEV protease must also have been expressed however no particularly intense band representing 6xHis TEV protease was observed. Protein expression results presented in Figure 3.14B showed clear over expression of the target protein, 6xHis TEV protease (red circle). MBP was also expressed to an equal degree as expected (orange circle). Autoinduction, 20 h culture was concluded to be the optimum method for 6xHis TEV protease over expression due to the high intensity of this protein band in comparison to all other images in figure 3.14. Figure 3.14C also shows good over expression of 6xHis TEV protease after 20 h culture incubation.

3.2.2 TEV Protease Protein Purification – IMAC

Results from IMAC purification of the over expressed 6xHis TEV protease protein as described in section 2.3 are presented in this section. This procedure was not carried out for the culture shown in Figure 3.14A due to the low protein expression levels interpreted from the SDS-PAGE. However IMAC purification was carried on the cultures presented in Figure 3.14B and C.

The total protein yield obtained from the dialysed elution fractions following IMAC were determined via a BCA assay. IMAC purification of the IPTG culture presented in Figure 3.14B, returned a 0.16 mg/ml total protein (4 ml total volume = 0.65 mg/4 ml). IMAC purification of the auto induction culture also presented in figure 3.14B returned a 0.135 mg/ml (4 ml total volume = 0.542 mg/4 ml). The IMAC purification fractions of the culture presented in Figure 3.14C is shown below in Figure 3.15.
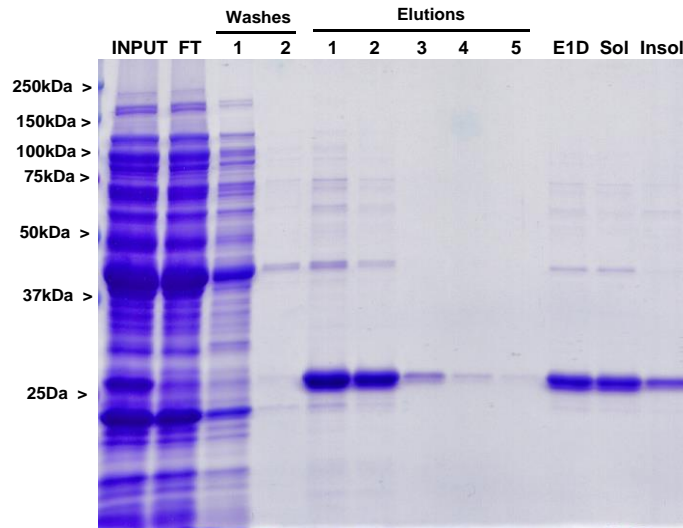
Figure 3.15. Native IMAC purification fractions of 6xHis TEV protease protein.
Fractions were separated on a 12% SDS-PAGE with Coomassie blue staining. Input =
cell lysate, FT = flow through, Washes and elutions are bracketed and numbered, E1D
= Elutions 1-4 following dialysis, Sol = soluble fraction following dialysis, Insol =
insoluble fraction following dialysis, i.e. precipitated protein

IMAC purification of 6xHis TEV protease was successful. In the first lane labelled
"INPUT", of figure 3.15, a band just above the 25 kDa marker was present indicating
the over expressed 6xHis TEV protease protein is present in the cell lysate. A band of
the same molecular weight was faintly visible in the FT and wash lanes indicating that
the resin did not capture all 6xHis TEV protease present in the cell lysate and therefore
more resin was needed. Despite this, it can be assumed the majority of the protein was
captured by the resin as these band were much less intense than the band in the
INPUT lane. Weakly bound protein was washed away in the wash lanes 1 and 2. A
strong band above the 25 kDa was seen in all elution lanes, particularly intense in
elution lanes 1 and 2, indicating native Ni-NTA purification was successful at purifying
the 6xHis TEV protease from the bacterial cell lysate. Elution fractions 1-4 were pooled
and dialysed however this purified protein did precipitate during overnight dialysis. The
soluble and insoluble fractions are visible in the last two lanes of this gel image. It can
be interpreted that less than 50 % of the total protein had precipitated overnight
suggesting that the dialysis buffer was not suitable for this protein. The total protein
yield determined via a BCA assay following dialysis returned a 0.22 mg/ml total protein
(47.5 ml total volume = 10.45 mg/47.5 ml). It can be assumed that there was more
protein present before precipitation occurred.

3.2.3 TEV Protease Functionality Test

The over expressed and purified 6xHis TEV protease was investigated for enzyme activity against the substrate, purified EF2-GGC recombinant fusion protein. The commercial enzyme, AcTEV protease acted as a positive control for this experiment and two negative controls were included, 1) the enzyme without any substrate and 2) the substrate without any enzyme, both in reaction conditions.
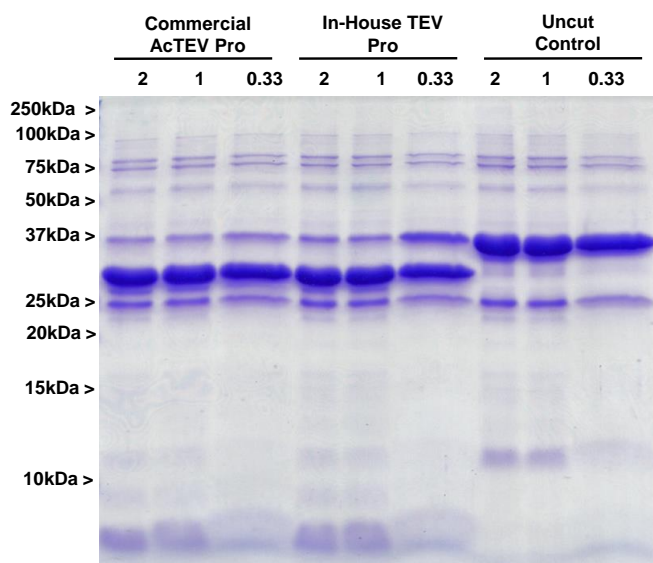
Figure 3.16. Commercial AcTEV protease versus In-house produced 6xHis TEV protease Activity Assay. Fractions separated on a 16% SDS-PAGE with Coomassie blue staining**.** Numbers under the brackets indicate overall protein reaction concentrations as 2 mg/ml, 1 mg/ml and 0.33 mg/ml within each sample set.

Peptide bond cleavage of the substrate, the EF2-GGC recombinant fusion protein was successful using both the in-house produced 6xHis TEV protease and the commercial AcTEV protease (figure 3.16). The ratio of protein to enzyme, in all reactions was 50:1 (µg). The numbers 2, 1 and 0.33 refer to the overall concentration of the reaction, i.e. a 2 mg/ml solution was cleaved (i.e. 200 µg fusion protein to 4 µg enzyme in a 100 µl total reaction volume = 2 µg/µl or 2 mg/ml), a 1 mg/ml and 0.33 mg/ml solution. The total protein concentration was varied in order to investigate if a higher overall concentration of protein would lead to increase protein precipitation. No precipitation was observed in any reaction.

The positive control for this experiment was in the first three lanes, bracketed under the label "Commercial AcTEV Pro". This enzyme catalysed peptide bond cleavage of the

36.44 kDa EF2-GGC fusion protein which resulted in the release of the EF2-GGC peptide at 4.382 kDa and the remaining 6xHis CA fusion protein reduced to 32 kDa. The negative control was in the last three lanes, bracketed under the label "Uncut Control". The samples in these three lanes had no enzyme and thus no peptide bond cleavage occurred, therefore the EF2-GGC recombinant protein remained at 36.44 kDa. No additional proteins bands at 32 kDa or 4.382 kDa were present in these lanes indicating that no other protein of that size was present in these samples and therefore the bands in the positive control and sample lanes must represent the cut fusion protein and released EF2-GGC peptide.

The test lanes were the middle three lanes, bracketed under the label "In-house TEV Pro". These lanes show similar protein band profiles to the positive control in that the EF2-GGC recombinant fusion protein was cut by the in-house produced TEV protease and produced two new protein bands at 32 kDa for the 6xHis CA protein and at 4.382 kDa for the EF2-GGC peptide. Therefore the in-house produced 6xHis TEV protease was functionally active and could cleave the substrate in reaction conditions. However, as there was a band for uncut EF2-GGC fusion protein remaining this suggests that more 6xHis TEV protease would be required to efficiently cleave all of the substrate.

A direct comparison of the in-house produced 6xHis TEV protease enzymatic activity at 30°C and 4°C was investigated and results shown in figure 3.17.
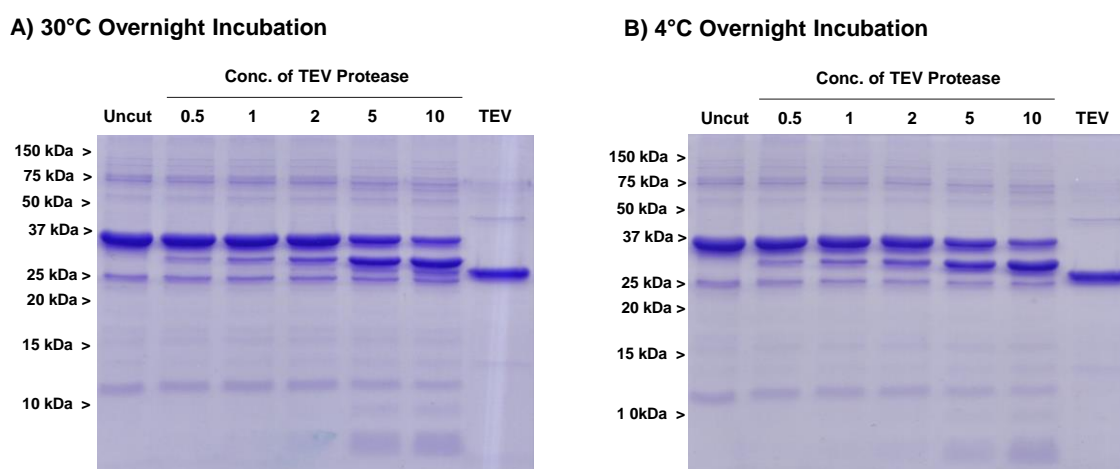
Figure 3.17. In-house produced 6xHis TEV protease activity at 30°C and 4°C.
Fractions were separated on a 16% SDS-PAGE with Coomassie blue staining. The concentration of enzyme is labelled and represented by the numbers under the bracket from 0.5 up to 10 μg. Two negative controls 1) Fusion protein substrate alone in the first lane, labelled "uncut", 2) In-house 6xHis TEV protease (10 μg loaded) alone, labelled "TEV".

A fixed amount of EF2-GGC recombinant fusion protein, 75 μg, was incubated with increasing concentrations of 6xHis TEV protease; 0.5, 1, 2, 5 and 10 μg (Figure 3.17). The first and last lanes in both images (A) and (B) of figure 3.17 represent the negative controls. The first lane was uncut EF2-GGC recombinant protein, whereby this sample was incubated in reaction conditions with no enzyme. The last lane was in-house produced 6xHis TEV protease incubated in reaction conditions with no substrate. Both controls show no presence of any protein bands in the same position/Mw as 6xHis CA or EF2-GGC peptide. This therefore confirms that these bands are only a result of correct enzymatic cleavage of the fusion protein. It also demonstrates that overnight cleavage at 4°C was suitable for peptide bond cleavage and did not significantly reduce the enzymes catalytic cleavage efficiency/activity.

3.3 Results - EF2-NP Functionality

Despite heavy aggregation of the EF2-GGC peptide identified through SEC, an aliquot was taken and investigated for coupling to functionalized NPs. These NPs, now displaying the EF2-GGC peptide were tested for the ability of the peptide to capture and purify a test protein, EF1-GFP, from a native bacterial cell lysate. An EF2-GGC peptide that was chemically synthesised via a commercial company was employed as a positive control and as a sample for direct comparison to the in-house produced peptide in this experiment. Another aim of this experiment was to identify if the GFP would remain functional following native purification.

A)
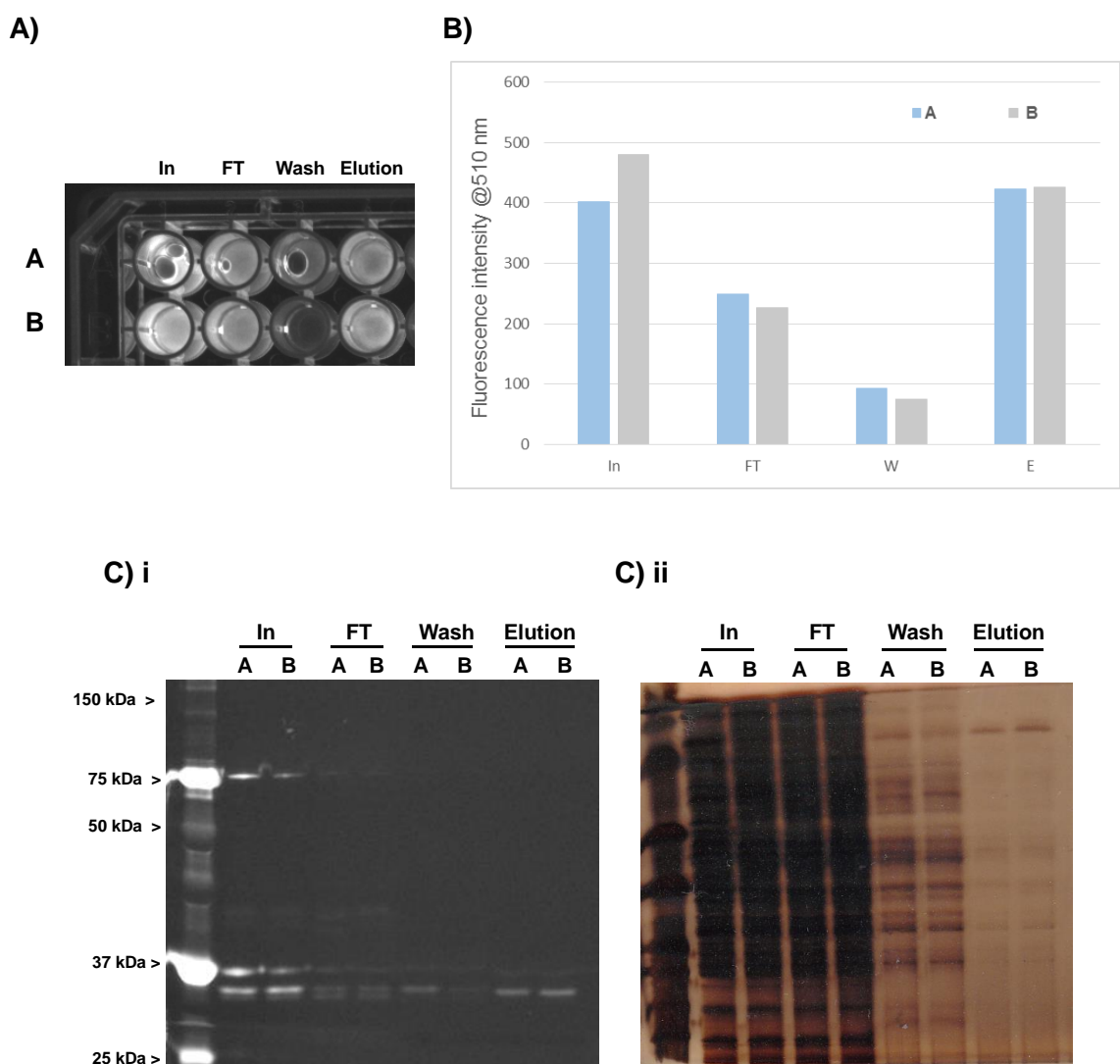

B)

C) i

C) ii

Figure 3.18. Selective capture and release of target GFP-EF1 by EF2-NPs
A) Fractions collected from EF1-EF2 purification, In = INPUT, FT= Flow through, wash and elution fractions, A = in-house produced EF2 peptide coupled NPs, B =

68

commercial EF2 peptide coupled NPs, measuring GFP fluorescence (excitation 485, emission 510 nm) in 96 well plate. B) Bar chart representing GFP fluorescence in image A. C) Native PAGE of fractions in image A, C) i image of gel under UV light, C) ii image of gel following silver staining.

Both EF2-NPs affinity resins were successful in the capture and purification of EF1-GFP directly from native bacterial cell lysate as GFP was seen in both elution lanes in figure 3.18 C) i. The in-house produced peptide demonstrated similar capture and elution functionality to the commercial peptide EF2-NPs, suggesting that the in-house peptide was equal in performance to the commercial peptide.

Following purification from native bacterial lysate by EF2-NPs, EF1-GFP still retained all functionality demonstrating that neither the EF1 tag nor the purification procedure denatured the protein in any way. This can be concluded as the protein in all fractions, representing each stage of the purification protocol, exhibited fluorescence, figure 3.18 A. The fluorescence readings were depicted in a bar chart in Figure 3.18B. The highest fluorescence levels were in the input samples. Sample A has a lower reading of fluorescence compared to sample B however this was most likely due to the bubble that was present in the well that can be seen in A), which would hinder correct reading of the sample. The same reasoning could be applied for FT and wash fractions of sample A (the in-house produced EF2-GGC peptide).

These same samples following electrophoresis on a 12 % Native PAGE were imaged to identify the presence of native GFP, C) i. The same gel following imaging was silver stained, C) ii. From the elution bands of C) ii, it was hard to conclude whether any GFP has been purified however when compared to C) i, it was clear that there is GFP present in low amounts and therefore the purification system worked. As expected the Input fractions of C) i show the largest amount of GFP fluorescence. The FT fractions show similar levels of GFP fluorescence indicating that some GFP did not bind to the resin. The wash fraction for sample A shows some EF1-GFP had washed off the resin in contrast to sample B, where no EF1-GFP was seen suggesting that this resin held the captured EF1-GFP better than the resin in sample A. However both elution fractions show equal amounts of EF1-GFP released.

Chapter 4

Discussion of Results and Conclusion of Study

4.1 Discussion – EF2-GGC Peptide Production

4.1.1 Fusion Protein Expression

As described in section 1.4.1, to produce the EF2-GGC peptide it was expressed in the form of a recombinant fusion protein. Recombinant protein expression was controlled in the bacterial host through the inducible lac operon.

Pilot scale protein expression screening (Figure 3.2) revealed that the EF2-GGC recombinant fusion protein demonstrated good expression and solubility in both auto and IPTG induction methods. Optimum protein expression was observed at 12 h for autoinduction and 6 h for IPTG-induction. For up-scaled protein expression cultures 12 h autoinduction was herein employed because of the protocol's ease of use. As described in the methods section 2.2.5.2, IPTG induction required culture OD monitoring and the addition of IPTG when the OD was acceptable. These steps were time consuming and introduced unnecessary risks for culture contamination. Therefore despite similar levels of protein expression observed, for experimental simplicity 12 h autoinduction was chosen as the method to continue with for up-scaled expression of this protein.

Protein expression studies for both autoinduction and IPTG induction were not investigated past 24 h and 6 h time points. These time points were chosen upon reference and comparison to numerous published methods and with regards to autoinduction, reference to the manufacturer guidelines (Berrow *et al.* 2006; Novagen 2009). Other institutions including New England BioLabs (https://www.neb.com/protocols/1/01/01/protein-expression-using-bl21de3-c2527) and EMBL (http://www.embl.de/pepcore/pepcore_services/protein_expression/ecoli/) report and recommend protocols for IPTG induction for *E.coli* BL21 DE3 cells specifically, with up to only 5 h for protein expression at 37°C. In this study target protein expression in a standardised IPTG induction system appeared to plateau at 5 and 6 h culture growth (figure 3.2B) and so it was suspected that target protein expression began to saturate at this time. Therefore to avoid progression of the expression culture into the cell death phase which may result in loss of the target protein to inclusion bodies, the expression culture was stopped at 6 h. Similar reasoning applied for autoinduction in that the media manufacturer Novagen recommends for the cell culture growth to be incubated until saturation which for 37°C *E.coli* cultures usually occurs around 8-10 h for harvesting the cells. Despite all lanes of figure 3.2A having equal sample loading it may be possible that some sample from the 24 h lane might have been pulled out with

movement from the running buffer in the electrophoresis gel tank and hence why this lane may appear to have less sample loaded.

Both systems were investigated at 37°C as this temperature is known for optimal growth of *E.coli,* however over-expression of a particular protein may be optimal at a lower temperature such as 25-30°C (Gadgil *et al.,* 2005). In this case future work could focus on screening of how different culture temperatures would affect both cell growth and protein expression (Sivashanmugam *et al.,* 2009). Due to the short culture incubation of the IPTG induction protocol, time points to investigate protein expression were taken every hour. Autoinduction was run for 12 h and so time points were spaced out every 4 h for the first 12 h and then a final time point taken at 24 h. Due to timing it was not possible to take a time point at 16 and 20 h and so a second expression culture should be inoculated at a different time of day in order to obtain these time points.

There were problems with the consistency of the protein yield return from each expression culture following IMAC purification and this was thought to be due to the expression culture conditions. Microwave irradiation was used for preparation of the autoinduction media as an alternative method over conventional autoclaving programmes for media sterilization (Border & Rice-Spearman, 1999). However it was noted that improper or insufficient microwaving by the user led to contamination of the bacterial cell culture, observed by an increased weight of the cell pellet following bacterial cell harvest but decreased the yield of protein retrieved following IMAC. This suggests the majority of the bacterial cells in the pellet may have been dead or contaminated during or before growth, hence lower protein yields. Another factor may have been the expression culture inoculum. The OD_{600} of the starter culture inoculum was not measured prior to culture inoculation and therefore it was impossible to know if the same density of cells were introduced into the expression culture each time. Overall culture pH, plasmid loss and limited availability of dissolved oxygen are other factors that potentially could have contributed to protein yield variability (Sivashanmugam *et al.,* 2009). To address these problems in future, the starter culture inoculum should be standardised to a fixed OD_{600} reading prior culture inoculation therefore all inoculums would be equal, reducing the risk for culture growth variability.

4.1.2 Fusion Protein Purification – IMAC

IMAC protein purification via Ni-NTA agarose was successful in purifying the 6xHis tagged EF2-GGC recombinant fusion protein from native bacterial cell lysate confirmed by a band at the expected mw for this protein of ~ 36.44 kDa present in the elution lanes of both images in figure 3.3.

Optimization of this protocol in terms of protein yield and purity of the eluted protein was investigated through optimization of buffer composition and concluding from the results figure 3.4 and table 3.1, a 10/20 mM imidazole concentration in the lysis/wash buffers respectively was chosen to obtain good yields with slight compromising on the purity.

To ensure experimental validation was not at risk at any point in this particular procedure the Ni-NTA resin used for each purification was washed following each use in 0.5 M NaOH to remove any possible bacterial and/or endotoxin contamination of the resin (see appendix III for general Ni-NTA washing protocol, adapted from the QIAexpressionist™, 5th Ed.). The resin was stored in 20 % ethanol at 4°C until required for re-use. Re-use of the resin was only performed on identical proteins, i.e. if a batch of resin was used for fusion protein purification, it would not be then used for TEV protease purification. After a maximum of 5 protein purifications the resin was stripped of its nickel and re-charged according to Qiagen protocols (Appendix III).

Following purification it was necessary to dialyse the eluted protein immediately to remove high concentrations of imidazole salts because precipitation, also known as "salting out" of the protein occurred rapidly and has been seen in many other cases (Berg *et al.,* 2002). It should be noted the yield of protein returned did vary considerably however this was suspected to be as a result of expression culture variability and not due to any aspect of the IMAC protocol, therefore all aspects of IMAC were maintained in all future purifications.

4.1.3 Fusion Protein TEV Protease Cleavage

Both the AcTEV and in house produced TEV protease cleavage were successful in cleaving the EF2-GGC recombinant fusion protein. The ratio of enzyme to protein was optimized to achieve efficient cleavage with minimal consumption of the enzyme. Figure 3.5, section 3.1.3 concluded 10 units of commercial AcTEV enzyme to every 50 µg fusion protein and figure 3.16, section 3.2.3 concluded 1 µg of in house TEV protease to every 50 µg fusion protein.

Following in house TEV protease IMAC, the protease precipitated out of solution significantly during overnight dialysis and this was suspected to be due to the formation of disulphide bonds, because this enzyme is a cysteine protease and in a concentrated sample these proteases may form disulphide bridges and perpetuate precipitation (Tropea *et al.* 2009). The dialysis buffer should therefore be amended to include 5 mM DTT minimum to prevent future protein precipitation. As the yield of protease from this batch was sufficient for the foreseeable future cleavage reactions, no further expression and purification of 6xHis TEV protease was performed and therefore no further optimization was carried out. This enzyme was employed in all future cleavage reactions for EF2-GGC peptide production.

4.1.4 EF2-GGC Peptide Purification

At this point in the EF2-GGC peptide production line, the EF2-GGC peptide was required in a pure monomeric formulation suitable for affinity resin manufacture.

Initial EF2-GGC peptide purification investigations involving the use of 10 kDa MWCO spin concentrators were not successful even with the addition of DTT, as figure 3.6 (section 3.1.4.1) verifies no protein had passed through the spin concentrator membrane. This could suggest aggregation of the EF2-GGC peptide.

Previous work published by Julenius *et al.*, (2002) which reported the EF2-GGC peptide has a tendency to homodimerize prompted an experiment into purification by homodimerization using an EF2-agarose affinity resin. However results here also failed to produce purified peptide (figure 3.7). Despite this method failing to purify the peptide, it did indicate that the EF2-agarose resin is a very specific resin and will not bind non-specific proteins. I concluded from these results that these methods were fruitless and another method for peptide purification was investigated, namely IMAC.

4.1.4.1 Second Round IMAC and IEX

A second round of Ni-NTA IMAC proved to be somewhat effective in purifying the EF2-GGC peptide indirectly shown by semi-pure preps in the FT and wash fractions (Figure 3.8). However despite the addition of more Ni-NTA agarose to address the remaining fusion protein contaminates, it did not provide any significant improvement to peptide purity.

IEX, following second round IMAC, was indeed more effective in purifying the EF2-GGC peptide. Figure 3.9A, where gel samples had been fully reduced, had greatly diminished higher mw bands (37-75 kDa) in comparison to figure 3.9B, containing non-reduced samples. This suggested that these higher mw proteins were perhaps aggregated peptide and when subjected to a denaturing preparation, the peptide separated into its monomeric form, reducing its overall mw.

To fully address this reoccurring problem of peptide aggregation the buffer composition of the EF2-GGC peptide needed to be re-evaluated. SEC was therefore employed to identify and compare the effects of different conditions on the state of aggregation of the peptide.

4.1.4.2 SEC

SEC using Sephadex G50 resin, figure 3.10, section 3.1.5.5 showed promising results to be successful in purifying the EF2-GGC peptide directly following TEV cleavage. Continuing on from these results an automated approach to SEC was taken with a different resin namely Superdex 75 aiming to achieve greater protein separation.

Figure 3.11 confirmed the suspected aggregation of the EF2-GGC peptide with protein eluting over a broad range of fractions. An attempt to reverse peptide aggregation by treating EF2-GGC peptide samples with 1 and 10 mM TCEP for 1 h incubation at room temperature prior to SEC failed to deliver complete separation. TCEP, as a reducing agent, was employed to break disulphide bonds that may (were suspected to) have formed due to oxidation of thiol side chains of the terminal cysteine residue of the EF2-GGC peptide (Figure 1.4), causing or promoting peptide aggregation (Burns *et al.* 1991; Liu *et al.,* 2010). TCEP acts by donating an electron to the oxidised functional group, hence why it is known as a reducing agent. The end result is irreversible giving it an advantage over using other reducing agents such as DTT or β-mercaptoethanol. This may be because there was such a high degree of aggregation the TCEP couldn't act on buried disulphide bridges and hence eliminate peptide aggregation completely.

A second SEC experiment involved removal of sodium chloride from the EF2-GGC peptide buffer formulation (Figure 3.12). This was performed due to a previous study, Julenius *et al.*, 2002, finding that the presence of salt increased EF2 dimerization. Comparing to figure 3.11A) ii, the peaks of the samples presented in figure 3.12 appeared at a later elution volume, again indicating a lower molecular weight protein, and therefore perhaps a lower degree of peptide aggregation. The longer the

incubation in a salt free buffer the greater the reduction in peptide aggregation however yet no clear monomeric peptide was observed. IEX elution in high sodium chloride of 500 mM may have contributed to peptide aggregation and so in future work this elution step should be changed to elute via lowering the pH or removed this step altogether.

Again, the final SEC investigation to reverse peptide aggregation by fully denaturing in 6 M GuHCl and 10 mM DTT prior to SEC failed, figure3.13. As this experiment and numerous previous efforts were unsuccessful to produce monomeric EF2-GGC peptide, this study as a whole to produce an EF2-GGC peptide suitable for affinity resin manufacture was ineffective and therefore I can conclude that this approach should not be continued. Nonetheless, in the face of heavy EF2-GGC peptide aggregation, an aliquot was taken for investigation into how an in-house produced peptide might perform on a NP surface for the capture and purification of EF1 tagged proteins.

4.2 Discussion – EF2-GGC NP Functionality

EF2-GGC peptide functionality following conjugation to functionalized NPs was successful as figure 3.18 revealed the peptide did bind and purify EF1-tagged GFP from a native bacterial cell lysate.
Therefore, in conclusion, despite aggregation of the EF2-GGC peptide, this experiment demonstrated that this in-house produced peptide was functional when coupled to a NP scaffold in the capture and elution of EF1 tagged GFP from native bacterial lysates. This experiment also provided proof of concept that binding between the EF1 and EF2 peptides is a highly specific, tight interaction that can be exploited as a promising protein purification tool. Optimization of this protein purification system would benefit from efficient, high grade EF2 peptide manufacture. Upon production of a suitable EF2-GGC peptide for resin manufacture, the way in which the EF2-GGC peptide is displayed could be optimized, e.g. linker lengths between the EF2 ligand and NP surface could be varied, or the method of ligand coupling, e.g. amine versus thiol coupling could be investigated in order to fully optimize affinity resin performance.

4.3 Conclusions

The aims of this project were to produce a recombinant, 40 amino acid helix-loop-helix peptide (EF2-GGC) and test its ability to function as the affinity ligand on a nanoscale affinity scaffold. I can conclude that while both of these aims were achieved, this

protocol was ultimately deemed ineffective due to the high degree of aggregation of the EF2-GGC peptide rendering it unsuitable for sustainable affinity resin manufacture.

Interpretation of the results suggest that it was upon TEV protease cleavage of the EF2-GGC recombinant fusion protein that peptide aggregation began. This therefore implies that the EF2-GGC peptide alone was highly unstable, not a desirable trait to have in an affinity ligand. Optimization of this cleavage step would be most definitely required in order to reduce/remove peptide aggregation before continuing in resin manufacture. This alongside reports from Li, (2011) who suggest that optimization in the fusion protein cleavage efficiency contributed to a greater yield return of peptide. Peptide aggregation may be due to a number of factors mediated by pH, ionic strength and concentration. One approach to reduce protein aggregation in particular was discussed by Maggio (2008) in "*Novel Excipients Prevent Aggregation in Manufacturing and Formulation of Protein and Peptide Therapeutics*" for Bioprocess Technical, whereby trifluoroethanol (TFE) is discussed and highlighted in how it helps to stabilize alpha helices. This would be of particular interest to aid with the EF2-GGC peptide stability. Other options to reduce aggregation include lyophilisation in that it avoids risk of oxidation, however it may only delay the issue as upon reconstitution precipitation and aggregation are commonly encountered (Zhang MZ *et al.,* 1995).

4.4 Future Objectives

The potential of the EF1-EF2 system as a powerful protein purification tool is very much real, however the need for efficient EF2 peptide production is a major bottleneck in the progression of this technique.

For future work to produce monomeric EF2-GGC peptide for affinity resin manufacture, investigations must be made into the characteristics and properties of the peptide from an amino acid level in order to decipher the requirements for a stable, functional EF2 peptide ligand. For example, for this study a tryptophan residue was introduced into the EF2-GGC peptide sequence. This was required for means of protein quantification however, tryptophan is known to have a hydrophobic side chain and so the addition of this amino acid may have increased the overall hydrophobicity of the EF2-GGC peptide and therefore promoted aggregation. Therefore to avoid this occurring future work should focus on the evaluation of the EF2-GGC peptide amino acid composition and extending from those results expression and purification of the peptide may be undertaken.

Bibliography

Andersson, L. et al., 2000. Large-Scale Synthesis of Peptides. *Biopolymers (Peptide Science)*, 55, pp.227–250.

Belkacemi, L. et al., 2002. Calcium-Binding Proteins Distribution and Implication in Mammalian Placenta. , 19(1), pp.57–64.

Berg, J.M., Tymoczko, J.L. & Stryer, L., 2002. The Purification of Proteins Is an Essential First Step in Understanding Their Function. In *Biochemistry*. New York: W H Freeman. Available at: http://www.ncbi.nlm.nih.gov/books/NBK22410/.

Berrow, N.S. et al., 2006. Recombinant protein expression and solubility screening in Escherichia coli: A comparative study. *Acta Crystallographica Section D: Biological Crystallography*, 62(10), pp.1218–1226.

Block, H. et al., 2009. Chapter 27 Immobilized-Metal Affinity Chromatography (IMAC). A Review. *Methods in Enzymology*, 463(C), pp.439–473.

Blommel, P.G. & Fox, B.G., 2007. A Combined Approach to Improving Large-Scale Production of Tobacco Etch Virus Protease. *Protein expression and purification*, 55(1), pp.53–68.

Bommarius, B. et al., 2010. Cost-effective expression and purification of antimicrobial and host defense peptides in Escherichia coli. *Peptides*, 31(11), pp.1957–65. Available at: http://dx.doi.org/10.1016/j.peptides.2010.08.008.

Bornhorst, J.A. & Falke, J.J., 2000. Purification of Proteins Using Polyhistidine Affinity Tags. *Methods in Enzymology*, (326), pp.245–254.

Bray, B.L., 2003. Large-scale manufacture of peptide therapeutics by chemical synthesis. *Nature Reviews, Drug Discovery*, 2(July), pp.587–593.

Burns, J.A. et al., 1991. Selective reduction of disulfides by tris(2-carboxyethyl)phosphine. *The Journal of Organic Chemistry*, 56(8), pp.2648–2650. Available at: http://pubs.acs.org/doi/abs/10.1021/jo00008a014.

Dell'Orco, D. et al., 2005. Electrostatic contributions to the kinetics and thermodynamics of protein assembly. *Biophysical journal*, 88(3), pp.1991–2002. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1305251&tool=pmcentr ez&rendertype=abstract [Accessed February 13, 2015].

Fang, L. et al., 2007. An improved strategy for high-level production of TEV protease in

Escherichia coli and its purification and characterization. *Protein Expression and Purification*, 51, pp.102–109.

Fischer, I. et al., 2013. Continuous protein purification using functionalized magnetic nanoparticles in aqueous micellar two-phase systems. *Journal of Chromatography A*, 1305, pp.7–16. Available at: http://dx.doi.org/10.1016/j.chroma.2013.06.011.

Furman, J.L., Chiu, M. & Hunter, M.J., 2015. Early Engineering Approaches to Improve Peptide Developability and Manufacturability. *The AAPS Journal*, 17(1), pp.111–120.

Gadgil, M., Kapur, V. & Hu, W.-S., 2005. Transcriptional response of Escherichia coli to temperature shift. *Biotechnology progress*, 21(3), pp.689–99. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15932244.

Gross, E. & Witkop, B., 1962. Nonenzymatic cleavage of peptide bonds: the methionine residues in bovine pancreatic ribonuclease. *The Journal of biological chemistry*, 237(6), pp.1856–60. Available at: http://www.ncbi.nlm.nih.gov/pubmed/13902203.

Islam, T. et al., 2014. A novel strategy for the purification of a recombinant protein using ceramic fluorapatite-binding peptides as affinity tags. *Journal of chromatography. A*, 1339, pp.26–33. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24661869.

Jad, Y.E. et al., 2015. Peptide synthesis beyond DMF : THF and ACN as excellent and friendlier alternatives †‡. *Royal Society of Chemistry*, (13), pp.2393–2398.

Julenius, K. et al., 2002a. Coupling of Ligand Binding and Dimerization of Helix-Loop-Helix Peptides : Spectroscopic and Sedimentation Analyses of Calbindin D 9k EF-Hands. *Proteins: Structure, Function and Genetics*, (47), pp.323–333.

Julenius, K. et al., 2002b. Coupling of ligand binding and dimerization of helix-loop-helix peptides: spectroscopic and sedimentation analyses of calbindin D9k EF-hands. *Proteins,* 47(3), pp.323–33. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11948786 [Accessed October 31, 2014].

Kapust, R.B. et al., 2001. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein engineering*, 14(12), pp.993–1000.

Kapust, R.B. & Waugh, D.S., 1999. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein science : a publication of the Protein Society*, 8(8), pp.1668–1674.

Kim, J.S. et al., 2007. Highly-Efficient Purification of Native Polyhistidine-tagged Proteins by Multivalent NTA-modified Magnetic Nanoparticles. *Bioconjugate Chemistry*, 18(2), pp.333–341.

Krishnamurthy, V.M. et al., 2008. Carbonic anhydrase as a model for biophysical and physical-organic studies of proteins and protein-ligand binding. *Chemical reviews*, 108(3), pp.946–1051. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18335973.

Kuo, W.H.K. & Chase, H. a., 2011. Exploiting the interactions between poly-histidine fusion tags and immobilized metal ions. *Biotechnology Letters*, 33(6), pp.1075–1084.

Kyle, S., James, K.A.R. & Mcpherson, M.J., 2012. Recombinant production of the therapeutic peptide lunasin. *Microbial Cell Factories*, 11(28), pp.1–8.

Li, Y., 2011. Recombinant production of antimicrobial peptides in Escherichia coli : A review. *Protein Expression and Purification*, 80(2), pp.260–267. Available at: http://dx.doi.org/10.1016/j.pep.2011.08.001.

Lichty, J.J. et al., 2005. Comparison of affinity tags for protein purification. *Protein Expression and Purification*, 41(1), pp.98–105.

Lindman, S. et al., 2009. Green fluorescence induced by EF-hand assembly in a split GFP system. *Protein science : a publication of the Protein Society*, 18(6), pp.1221–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2774432&tool=pmcentr ez&rendertype=abstract [Accessed February 13, 2015].

Lindner, P. et al., 1997. Specific detection of His-tagged proteins with recombinant anti-His tag scfv-phosphatase or scFv-phage fusions. *BioTechniques*, 22(1), pp.140–149.

Linse, S. et al., 1987. Structure-function relationships in EF-hand calcium-binding proteins. Protein engineering and biophysical studies of calbindin D9k. *Biochemistry*, pp.6723–6735. Available at: http://pubs.acs.org/doi/abs/10.1021/bi00395a023 [Accessed October 31, 2014].

Liu, P. et al., 2010. A Tris (2-Carboxyethyl) Phosphine (TCEP) Related Cleavage on Cysteine-Containing Proteins. *Journal of the American Society for Mass Spectrometry*, 21(5), pp.837–844. Available at: http://dx.doi.org/10.1016/j.jasms.2010.01.016.

Maggio, E.T.T., 2008. Novel Excipients Prevent Aggregation in Manufacturing and

Formulation of Protein and Peptide Therapeutics. *BioProcess International*, 6, pp.58–65. Available at: http://www.aegisthera.com/pdf/BioProcess_ProTek.pdf.

Merrifield, R.B., 1963. Solid Phase Peptide Synthesis. I. The Synthesis of. *Journal of the American Chemical Society*, 85(14), p.2149.

Montalbetti, C.A.G.N. et al., 2005. Amide bond formation and peptide coupling. *Tetrahedron*, 61(740), pp.10827–10852.

Novagen, 2009. Overnight Express Instant LB and TB Media User Protocol. , pp.1–6.

O'Connell, D. et al., 2012. Affinity Tag Patent Doc WO2013102684.pdf. , pp.1–47. Available at: http://www.google.com/patents/WO2013102684A1?cl=en.

Pipkorn, R. et al., 2014. A Peptide & Peptide Nucleic Acid Synthesis Technology for Transporter Molecules and Theranostics – The SPPS. *International Journal of Medical Sciences*, 11(7), pp.697–706.

Prasad, K.V.S.R.G., Bharathi, K. & Haseena Banu, B., 2011. Applications of peptide coupling reagents - An update. *International Journal of Pharmaceutical Sciences Review and Research*, 8(1), pp.108–119.

Rigaut, G. et al., 1999. A Generic Protein Purification Method For Protien Complex Characterization and Proteome Exploration. *Nature biotechnology*, 17(October), pp.1030–1032. Available at: http://www.nature.com/nbt/journal/v17/n10/abs/nbt1099_1030.html.

Rodríguez, V., Asenjo, J.A. & Andrews, B.A., 2014. Design and implementation of a high yield production system for recombinant expression of peptides. *Microbial Cell Factories*, 13(65), pp.1–10.

Rudolph, R. & Lilie, H., 1996. In vitro Folding of Inclusion Body Proteins. *FASEB Journal*, 10, pp.49–56.

Saptarshi, S.R., Duschl, A. & Lopata, A.L., 2013. Interaction of nanoparticles with proteins : relation to bio-reactivity of the nanoparticle. *Journal of Nanobiotechnology*, 26(11), pp.1–12.

Schmidt, T.G.M. & Skerra, A., 2007. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nature protocols*, 2(6), pp.1528–1535.

Sheehan, J.C. & Hess, G.P., 1995. Communications to the Editor; A New Method of Forming Peptide Bonds. *Communications to the Editor*, pp.1067–1068.

Sivashanmugam, A. et al., 2009. Practical protocols for production of very high yields

of recombinant proteins using Escherichia coli. *Protein science : a publication of the Protein Society*, 18(5), pp.936–48. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2771296&tool=pmcentr ez&rendertype=abstract.

Skerra, A. & Schmidt, T.G.M., 1999. Applications of a peptide ligand for streptavidin: The Strep-tag. *Biomolecular Engineering*, 16(1-4), pp.79–86.

Terpe, K., 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied microbiology and biotechnology*, 60(5), pp.523–533.

Thundimadathil, J., 2013. Challenges in the large scale production of peptides. *Speciality Chemicals Magazine*, (May), pp.26–28.

Tropea, J.E., Cherry, S. & Waugh, D.S., 2009. High Throughput Protein Expression and Purification S. A. Doyle, ed. , 498(1), pp.297–307. Available at: http://link.springer.com/10.1007/978-1-59745-196-3 [Accessed July 10, 2014].

Wang, Y. et al, 2013. Insight into the structural stability of wild type and mutants of the tobacco etch virus protease with molecular dynamics simulations. *Journal of Molecular Modelling*, 19(11), pp.4865–4875.

Waugh, D.S., 2005. Making the most of affinity tags. *Trends in Biotechnology*, 23(6), pp.316–320.

Wood, D.W., 2014. New trends and affinity tag designs for recombinant protein purification. *Current opinion in structural biology*, 26, pp.54–61. Available at: http://www.sciencedirect.com/science/article/pii/S0959440X14000463 [Accessed July 12, 2015].

Young, C.L., Britton, Z.T. & Robinson, A.S., 2012. Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. *Biotechnology Journal*, 7(5), pp.620–634.

Zdanov, A.S. et al., 2003. Tobacco Etch Virus Protease: Crystal Structure of the Active Enzyme and Its Inactive Mutant. *Russian Journal of Bioorganic Chemistry*, 29(5), pp.415–417.

Zheng, C.F. et al., 1997. A new expression vector for high level protein production, one step purification and direct isotopic labeling of calmodulin- binding peptide fusion proteins. *Gene*, 186(1), pp.55–60.
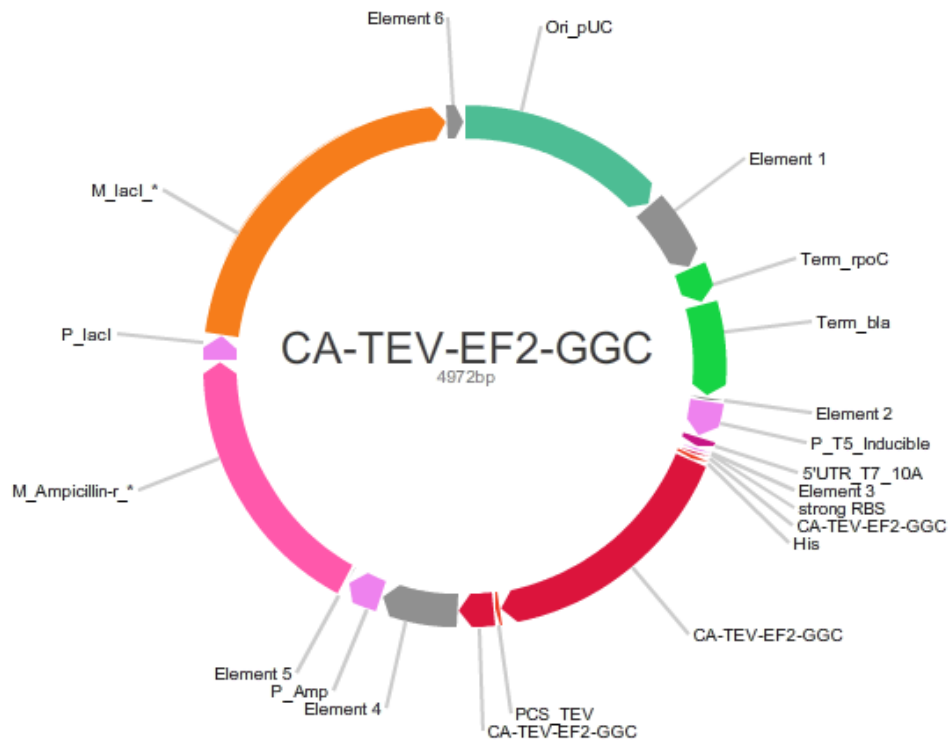
Figure 5.1 EF2-GGC Recombinant Fusion Protein Plasmid Map

Appendix II

EF2-GGC Recombinant Fusion Protein Amino Acid Sequence

MRGSHHHHHHGMASMTGGQQMGRASPDWGYDDKNGPEQWSKLYPIANGNNQSP
VDIKTSETKHDTSLKPISVSYNPATAKEIINVGHSFHVNFEDNDNRSVLKGGPFSDSYR
LFQFHFHWGSTNEHGSEHTVDGVKYSAELHVAHWNSAKYSSLAEAASKADGLAVIG
VLMKVGEANPKLQKVLDALQAIKTKGKRAPFTNFDPSTLLPSSLDFWTYPGSLTHPPL
YESVTWIICKESISVSSEQLAQFRSLLSNVEGDNAVPMQHNNRPTQPLKGRTVRASFE
NLYFQGLKGPSTLDELFEELDKDGDGQVSFEEWQVLVKKISQGGC

Number of amino acids: 329

Molecular weight: 36446.5

Theoretical pI: 6.27

Extinction coefficient (Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water): 52035

EF2-GGC Peptide Amino Acid Sequence

GLKGPSTLDELFEELDKDGDGQVSFEEWQVLVKKISQGGC

Number of amino acids: 40

Molecular weight: 4382.8

Theoretical pI: 4.21

Extinction coefficient (Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water): 5500

6xHis CA (+ TEV cleavage site) Amino Acid Sequence

MRGSHHHHHHGMASMTGGQQMGRASPDWGYDDKNGPEQWSKLYPIANGNNQSP
VDIKTSETKHDTSLKPISVSYNPATAKEIINVGHSFHVNFEDNDNRSVLKGGPFSDSYR
LFQFHFHWGSTNEHGSEHTVDGVKYSAELHVAHWNSAKYSSLAEAASKADGLAVIG
VLMKVGEANPKLQKVLDALQAIKTKGKRAPFTNFDPSTLLPSSLDFWTYPGSLTHPPL
YESVTWIICKESISVSSEQLAQFRSLLSNVEGDNAVPMQHNNRPTQPLKGRTVRASFE
NLYFQ

Number of amino acids: 289

Molecular weight: 32081.7

Theoretical pI: 7.00

Ext. coefficient   46410

Appendix III

General Ni-NTA Resin Washing

Note: The re-use of Ni-NTA resins should only be performed with identical recombinant proteins and be limited to a maximum of 5 runs.

Ni-NTA resin should be washed following each use for desorption of contaminants:

1. Wash the resin with 15 bed volumes of 0.5M NaOH (allow for a contact time of 30mins minimum) (resin will turn a rusty orange colour)
2. Remove the NaOH solution by washing resin with 10 bed volumes of distilled H_20 (colour of resin should return back to pale blue)
3. For storage wash resin with 2 bed volumes of 20-30% ethanol, re-suspend beads in the ethanol solution and store at 2-8°C
4. If using resin immediately, wash/equilibrate with 10 bed volumes of the appropriate lysis buffer

Ni-NTA Resin Stripping and Regeneration

For complete metal elimination and re-charging of the resin:

1. Wash the column with 2 volumes of Regeneration Buffer (6 M GuHCl, 0.2 M acetic acid).
2. Wash the column with 5 volumes of H2O.
3. Wash the column with 3 volumes of 2% SDS.
4. Wash the column with 1 volume of 25% EtOH.
5. Wash the column with 1 volume of 50% EtOH.
6. Wash the column with 1 volume of 75% EtOH.
7. Wash the column with 5 volumes of 100% EtOH.
8. Wash the column with 1 volume of 75% EtOH.
9. Wash the column with 1 volume of 50% EtOH.
10. Wash the column with 1 volume of 25% EtOH.
11. Wash the column with 1 volume of H2O.
12. Wash the column with 5 volumes of 100 mM EDTA, pH 8.0.
13. Wash the column with H2O.
14. Recharge the column with 2 volumes of 100 mM NiSO4.
15. Wash the column with 2 volumes of H2O.
16. Wash the column with 2 volumes of Regeneration Buffer.

Equilibrate with 2 volumes of a suitable buffer (e.g. native lysis buffer) if using immediately or 20-30 % EtOH for storage.