

Biophysical characterisation of human eukaryotic elongation factor 1 Beta and its interaction with human eukaryotic elongation factor 1 Gamma

Nnenna Elebo

Supervised by: Dr Ikechukwu A. Achilonu

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

July, 2017

Declaration

I, **Elebo Nnenna C, (500688)**, am a student registered for the degree of MSc in Biochemistry in the academic year 2015/2016.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged.
- I have not submitted this work before for any other degree or examination at this or any other University.
- The information used in the Dissertation **HAS NOT** been obtained by me while employed by, or working under the aegis of, any person or organisation other than the University.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

signature



4th day of July, 2017

Dedication

I dedicate this work to the people who have always encouraged, supported and prayed for me. My amazing husband Benjamin for your unconditional love, my daughters Somma, Muna and Unoaku, my sisters Grace and Joy, my brothers Wilson, Max, Okey and Chibu and finally my lovely mom Hannah.

Abstract

Eukaryotic protein synthesis occurs in three phases: initiation, elongation and termination. The elongation phase is mediated by elongation factors. Elongation factors are divided into elongation factor 1 (eEF1) and elongation factor 2 (eEF2). Elongation factor 1 complex are proteins that mediates the extension of growing polypeptide chains by adding one amino acid residue at a time. The eEF-1 complex comprises of four subunits, eEF1 α , eEF1 β , eEF1 γ and eEF1 δ . The β -subunit of elongation factor 1 complex (eEF1) plays a central role in the elongation step of eukaryotic protein biosynthesis, which essentially involves interaction with the α -subunits (eEF1 α) and γ -subunits (eEF1 γ). To biophysically characterise heEF1 β , three *E. coli* expression vector systems was constructed for recombinant expression of the full length (FL-heEF1 β), amino terminus (NT-heEF1 β) and the carboxyl terminus (CT-heEF1 β) regions of the protein. NT-heEF1 β was created from the FL-heEF1 β by site-directed mutagenesis using mutagenic forward and reverse primers. The results suggest that heEF1 β is predominantly alpha-helical and possesses an accessible hydrophobic cavity in the CT-heEF1 β . Both FL-heEF1 β and NT-heEF1 β forms dimers of size 62 kDa and 30 kDa, respectively, but the CT-heEF1 β is monomeric. FL-heEF1 β interacts with the N-terminus GST-like domain of heEF1 γ (NT-heEF1 γ) to form a 195 kDa complex, or a 230 kDa complex in the presence of oxidised glutathione. On the other hand, NT-heEF1 β forms a 170 kDa complex with NT-heEF1 γ and a high molecular weight aggregate of size greater than 670 kDa. This study affirms that the interaction between heEF1 β and heEF1 γ subunits occurs at the N-terminus regions of both proteins, also the N-terminus region of heEF1 β is responsible for its dimerisation and the C-terminus region of heEF1 β controls the formation of an ordered eEF1 β - γ oligomer, a structure that may be essential in the elongation step of eukaryotic protein biosynthesis.

Acknowledgement

Very special thanks to the following people:

Dr Ikechukwu A. Achilonu for his unending support, guidance, patience and motivation throughout the course of my study. I would not have been able to achieve this without your inspiration and mentorship, you are truly Godsent and for this am very grateful.

Prof Heini Dirr for your guidance and advice which have been so helpful, am so privileged and honoured to have worked in your lab.

The members of the Protein Structure-Function Research Unit, University of the Witwatersrand for all your assistance and friendship.

I would also like to acknowledge the University of the Witwatersrand and the National Research Foundation for financial assistance.

Table of contents

Declaration	ii
Abstract	iv
Acknowledgements.....	v
List of abbreviations	viii
List of Figures	x
List of Tables	xi
CHAPTER 1 Introduction and literature review	1
1.1 The translational elongation factor complex	1
1.2 Roles of elongator factor complex in protein biosynthesis	1
1.3 Human eukaryotic elongation factor 1 gamma.....	2
1.3.1 Ligandin functions of human eukaryotic elongation factor 1 gamma	5
1.4 Human eukaryotic elongation factor 1 beta.....	5
1.5 Protein-protein interactions	6
1.6 Protein-protein interactions between the subunits of eEF-1 complex.....	10
1.7 Aims and objectives.....	11
CHAPTER 2 Materials and methods.....	12
2.1 Materials	12
2.2 Methods	12
2.2.1 Construction of expression vectors.....	12
2.2.2 Plasmid extraction and sequencing.....	13
2.2.3 Site Directed Mutagenesis	14
2.2.4 Transformation and overexpression	16
2.2.5 Purification	18
2.2.5.1 Immobilised Metal Affinity Chromatography.....	18
2.2.5.2 Ion exchange chromatography.....	19

2.2.5.3	Glutathione-Agarose affinity chromatography	19
2.2.6	Sodium dodecyl sulfate-polyacrylamide gel	20
2.2.6.1	Tricine-based sodium dodecyl sulfate-polyacrylamide gel	20
2.2.6.2	Glycine-based sodium dodecyl sulfate-polyacrylamide gel	21
2.2.7	Liquid chromatography-Mass spectrometry/mass spectrometry	21
2.2.8.	Protein quantification.....	22
2.2.8.1	Estimation of the molecular weight of the proteins.....	22
2.2.8.2	Determination of protein concentration.....	22
2.2.9	Determination of protein quality using Absorbance Spectrometry	23
2.2.10	Structural characterisation	24
2.2.10.1	Far-UV circular dichroism.....	24
2.2.10.2	Fluorescence spectroscopy	25
2.2.11	Size exclusion-high performance liquid chromatography	27
2.2.12	Protein-protein and protein-ligand interaction.....	27
CHAPTER 3 Results.....		29
3.1	Introduction.....	29
3.2	Vector sequencing	29
3.3	Protein expression and purification	29
3.3.1	FL-heEF1 β	31
3.3.2	NT-heEF1 β	31
3.3.3	CT-heEF1 β	31
3.3.4	NT-heEF1 γ and CT-heEF1 γ	36
3.4	Quantitative and qualitative evaluation of purified proteins	36
3.5	Structural characterisation	42
3.5.1	Secondary structure characterisation by far-UV CD	42
3.5.2	Tertiary structure characterisation by tryptophan ANS fluorescence.....	44
3.5.3	Quaternary structure characterisation by SE-HPLC.....	47
3.6	Functional characterisation by protein-protein interactions	51
CHAPTER 4 Discussion and conclusion.....		54
4.1	Discussion.....	54
4.2	Conclusion and prospective study	58
References.....		60

List of Abbreviations

ϵ	molar extinction coefficient for absorption
$[\theta]$	mean residue ellipticity
A ₂₈₀	absorbance at 280nm
ANS	8-anilino-1-naphthalene sulfonate
BSA	Bovine serum albumin
°C	degree Celsius
CD	circular dichroism
CT	carboxyl terminal
DTT	dithiothretiol
EDTA	ethylenediaminetetra-acetic acid
FL	full length
G-site	glutathione binding site
GDP	guanidine diphosphate
GSH	reduced glutathione
GSSH	oxidised glutathione
GST	glutathione-S-transferase
GTP	guanidine triphosphate
heEF1 β	human eukaryotic elongation factor 1Beta
heEF1 γ	human eukaryotic elongation factor 1Gamma
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β -D-1-thiogalactopyranoside
kDa	kilo Dalton
<i>l</i>	pathlength
LB	lysogeny broth
LC-MS/MS	Liquid chromatography-mass spectrometry/ mass spectrometry
MCS	multiple cloning site
mM	milli molar
<i>N</i>	native state
NT	amino terminal
PCR	polymerase chain reaction
PDB	Protein Databank
PPI	protein-protein interactions

RNA	Ribonucleic acids
SDS-PAGE	Sodium dodecylsulfate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SE-HPLC	Size exclusion-high performance liquid chromatography
SOC	Super optimal broth
UV	ultraviolet

The IUPAC-IUBMB one- and three-letter abbreviations for the 20 standard amino acids were used

List of Figures and Table

	Page number
Figure 1.1: Translation elongation complex	3
Figure 1.2: Ribbon representation of the three dimensional structure of heEF1 γ	4
Figure 2.1: Schematic representation of constructs used in this study	15
Figure 2.2: Amino acid composition of FL-heEF1 β	15
Figure 2.3: Site directed mutagenesis	17
Figure 3.1: Production of FL-heEF1 β protein	30
Figure 3.2: IMAC purification profile of FL-heEF1 β	32
Figure 3.3: IMAC purification of FL-heEF1 β	33
Figure 3.4: DEAE-Agarose chromatography purification of FL-heEF1 β	33
Figure 3.5: IMAC purification of NT-heEF1 β	34
Figure 3.6: DEAE-Agarose chromatography purification of NT-heEF1 β	34
Figure 3.7: Lysate and pure samples of heEF1 β proteins	35
Figure 3.8: Purification of NT-heEF1 γ	38
Figure 3.9: Purification of CT-heEF1 γ	38
Figure 3.10: Graph of molecular weight of FL-heEF1 β protein against distance	39
Figure 3.11: Graph of absorbance against dilution factor	40
Figure 3.12: Absorbance spectrometry of FL-heEF1 β	41
Figure 3.13: Mass spectrometry results	41

Figure 3.14: Far-UV CD spectra.....	43
Figure 3.15: Extrinsic ANS Fluorescence emission spectra.....	45
Figure 3.16: Fluorescence signal of FL-heEF1 β	46
Figure 3.17: Gel filtration standards resolved by SE-HPLC	48
Figure 3.18: Quaternary structure characterisation of heEF1 β	49
Figure 3.19: Quaternary structure characterisation of heEF1 γ	50
Figure 3.20: Protein-protein interaction by SE-HPLC	52
Figure 3.21: Protein-protein interaction (NT-heEF1 β and NT-heEF1 γ).....	53

List of Tables

Table 1: Non-Standard materials and suppliers	11
Table 2: Summary of the secondary structure content	43

CHAPTER 1

Introduction and literature review

1.1 The translational elongation factor complex

In the cell, proteins are required to achieve various tasks. In protein biosynthesis, biological cells generate new proteins. Protein synthesis in living cells occurs as a result of the translation of the genetic information encoded in the messenger RNA (mRNA) into a sequence of amino acids in the polypeptide chain. Eukaryotic protein biosynthesis involves three separate stages. They are initiation, elongation and termination (Chi *et al.*, 1992). Eukaryotic elongation factors (eEF) play important roles in attaining precision during the translation process and are conserved throughout evolution (Olareswaju *et al.*, 2004). In both prokaryotes and eukaryotes the small ribosome subunit aids the binding of the mRNA while the large ribosome subunit aids in peptide formation (Kozak, 1999). Translation process in prokaryotes is much faster than in eukaryotes. The elongation factors in prokaryotes are namely: EF-Tu and EF-Ts while in eukaryotes they are eukaryotic elongation factors 1 and 2 (eEF1 and eEF2). The eukaryotic elongation factor 1 (eEF1) is further divided into four subunits namely eEF1 α , eEF1 β , eEF1 γ and eEF1 δ (Chi *et al.*, 1992; Le Sourd *et al.*, 2006).

1.2 Roles of elongator factor complex in protein biosynthesis

Eukaryotic elongation factor 1 complex (eEF1) is made up of proteins responsible for extending a polypeptide chain through the addition of amino acid, one residue at a time. Eukaryotic elongation factor 1 is made up of two entities namely: G-protein (eEF1 α) and nucleotide exchange factor (eEF1 β , eEF1 γ and eEF1 δ). The G-protein plays an important role in selection of the amino acids and also the transferring of the amino acid to the acceptor site (A-site) of the ribosome thereby forming a complex (Le Sourd *et al.*, 2006; Corbi *et al.*, 2010). The nucleotide exchange factor (eEF1 β , eEF1 γ and eEF1 δ) is required to regenerate eEF1 α from the complex (eEF1 α -GDP) to an active complex form (eEF1 α -GTP) (Ito *et al.*, 2004; Corbi *et al.*, 2010). The aminoacyl-transferRNA (aa-tRNA) stimulates eEF1 α to convert GTP to GDP by detaching GDP-bound eEF1 α from the ribosome. Thereby leaving only the aminoacyl-tRNA (aa-tRNA)

attached to the acceptor site (Olaewaju *et al.*, 2004). The eukaryotic elongation factor 1 gamma (eEF1 γ) appears to be associated with β and δ subunits, and stimulates eEF1 β in initiating the exchange of GDP to GTP on the subunit (Le Sourd *et al.*, 2006) as shown in Figure 1.

1.3 Human eukaryotic elongation factor one gamma

The human eukaryotic elongation factor 1 gamma (heEF1 γ) is about 47-52 kDa (Le Sourd *et al.*, 2006). It was initially characterised in invertebrate *Artemia salina* (Maessen *et al.*, 1987; Gillen *et al.*, 2008) and the human sequence has been published (Kumabe *et al.*, 1992; Sanders *et al.*, 1992). It is made up of two domains connected by amino acids of approximately 60 residues, which are rich in lysine (Jeppesen *et al.*, 2003). These two domains are (i) an amino terminal (NT) glutathione-S-transferase (GST) like domain and (ii) a protease-resistant carboxyl terminal (CT) domain. The GST-like domain is approximately 25 kDa and it is made up of α -helical and β -strands on the thioredoxin sub-domain (Vanwetswinkel *et al.*, 2003b) (Figure 1.2a). The binding sites for eEF1 β and eEF1 δ are both on this domain. This domain has been shown to be phosphorylated at threonine 46 and threonine 223 (Le Sourd *et al.*, 2006). The protease resistant carboxyl terminal contains a high number of aromatic amino acids and about 20% of all residues form antiparallel β -sheets enclosed by helices (Vanwetswinkel *et al.*, 2003a). The available structural information on heEF1 γ is the nuclear magnetic resonance spectroscopy (NMR) structure of the carboxyl terminal domain (PDB1D:1PBU) (Vanwetswinkel *et al.*, 2003b) (Figure 1.2b).

Although, heEF1 γ appears dispensable for translation, its absence does not seem to affect the rate of translational elongation but it has been found to have other roles such as: the GST domain of the heEF1 γ has been shown to aid in detecting oxidative stress (Jeppesen *et al.*, 2003; Olaewaju *et al.*, 2004). Over expression of heEF1 γ occurs in several tumours and cancer, influencing tumour aggressiveness (Mimori *et al.*, 1996; Mathur *et al.*, 1998; Al-Maghrebi *et al.*, 2005). The heEF1 γ has affinity for membrane and cytoskeleton elements, and helps to anchor the other subunits of the eEF1 complex to the cytoskeleton (Kim *et al.*, 2007). It is also a positive regulator of NF- κ B signalling pathway (Seth *et al.*, 2005; Liu *et al.*, 2014).

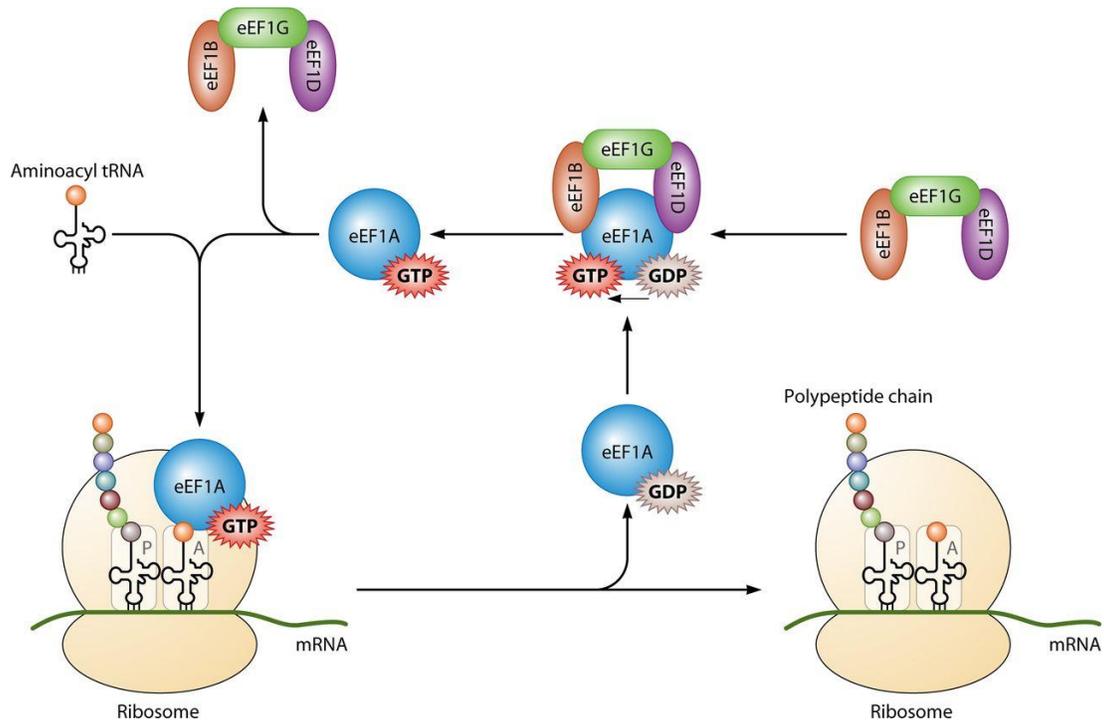


Figure 1.1: Translation elongational complex: GTP forms complex with the eukaryotic elongation factor 1A (eEF1A) and transports an aa-tRNA to the A site of the ribosome. GTP is broken down to GDP if codon-anticodon recognition takes place and eEF1A-GDP will be released. The eEF1A-GTP complex which is the active complex is regenerated by the exchange of GDP to GTP (Li *et al.*, 2013).

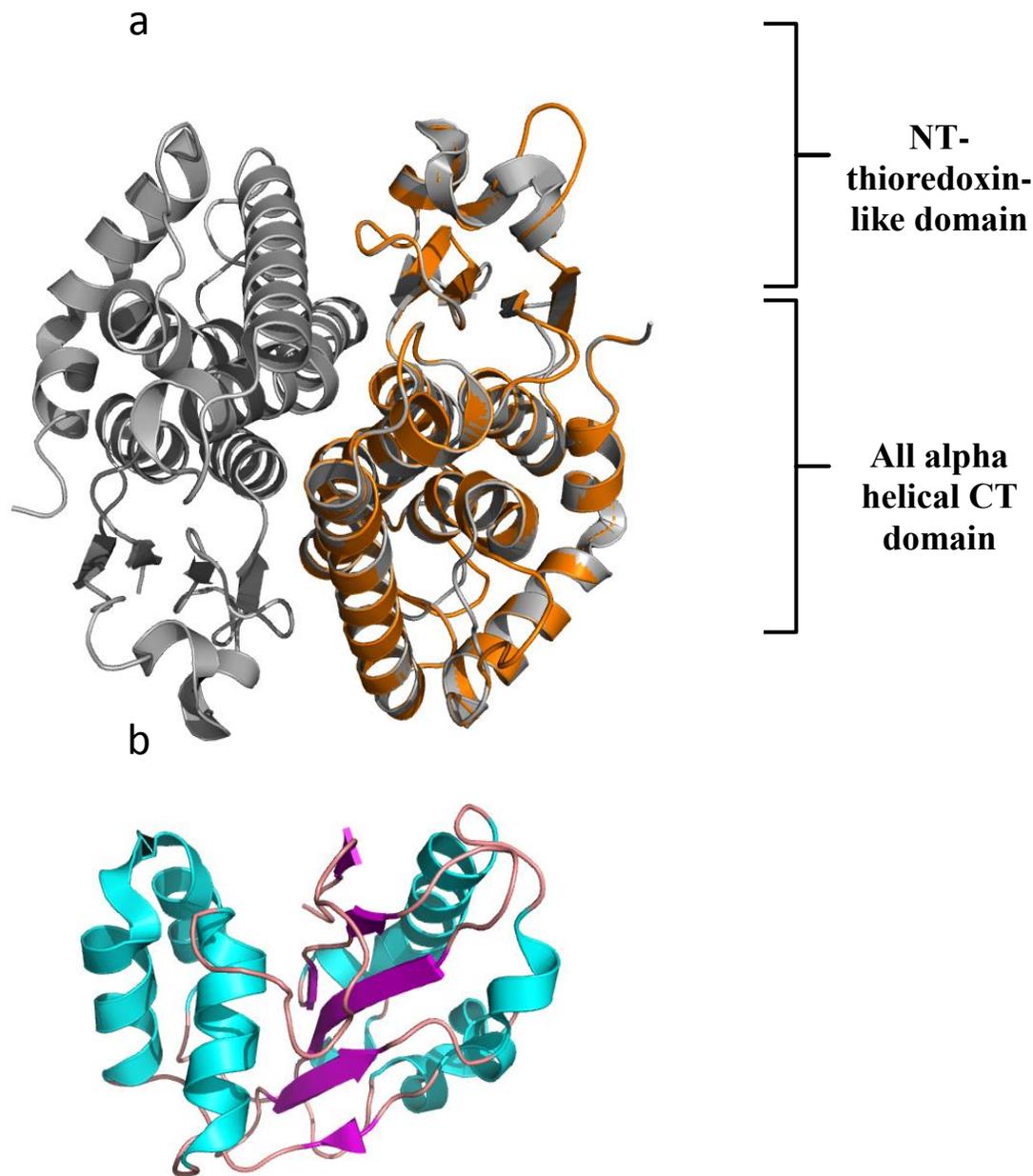


Figure 1.2: Ribbon representation of the three dimensional structure of heEF1 γ . (a) Superimposition of the homology model of the NT-GST like domain of heEF1 γ (orange) into the modelling template (PDB ID: 4ECJ; GST from *Pseudomonas aeruginosa* in complex with glutathione). The homology model was generated through the Swiss Model structural prediction server (Guex and Peitsch, 1997) by Dr Ikehukwu Achilonu (b) Solution structure (PDB ID: 1PBU) of the C-terminus domain of the heEF1 γ subunit (Vanwetswinkel *et al.*, 2003a).

1.3.1 Ligandin functions of human eukaryotic elongation factor 1 gamma

The GST-like NT domain of the heEF1 γ has a reduced glutathione (GSH) binding site but cannot be compared to the other mammalian classes of GST (Achilonu *et al.*, 2014) and this may be due to the heEF1 γ having a G-site different from the typical GSTs and this is evident in the ability of heEF1 γ to bind oxidised glutathione (GSSG) more firmly than reduced glutathione (GSH) (Tshabalala *et al.*, 2016). The preference of GSSG over GSH could be important in certain diseases such as multiple sclerosis, Parkinson's, Alzheimer's and tumorigenesis, because studies have shown that decreased molar ratio of GSH/GSSG is related to oxidative stress seen in these diseases (Jones *et al.*, 2000). Physiologically GSH (about 10 mM) is more abundant than GSSG (0.1 mM). This could be the reason why heEF1 γ is relatively high in cancer cells compared to other eEF1 subunits (Ernst *et al.*, 1978).

1.4 The human eukaryotic elongation factor 1 beta

Human eukaryotic elongation factor one beta (heEF1 β) has 225 amino acids and its monomeric weight is about 26-30 kDa (Pérez *et al.*, 1998). The heEF1 β has two domains which are: amino terminal (NT) domain and the carboxyl terminal (CT) domain connected to each other by a section of acidic amino acids residues (van Damme *et al.*, 1990). The NT domain is highly homologous with a GST-like CT domain, while the CT domain of the heEF1 β contains about 100 amino acids and has the nucleotide exchange activity. The amino acid sequence of the rabbit's eEF1 has 98% homology with human eEF1 except for isoleucine 72, glycine 43, and arginine 78 present in the rabbit being replaced while alanine replaced valine 156 (Chen and Traugh, 1995). Studies have shown that both eEF1 β and eEF1 δ are homologous because they contain nucleotide exchange activity and they share over 81% sequence similarity from the acid-rich region downstream (van Damme *et al.*, 1990).

The heEF1 β catalyses the GDP/GTP exchange activity on eEF1 α and is also very essential in the regeneration of eEF1 α (Chen and Traugh, 1995) which completes one elongation cycle. It also plays an important part in the oxidative stress response pathway (Olareswaju *et al.*, 2004). The eukaryotic elongation factor 1 beta (eEF1 β) has been shown to be a marker for detecting cellular senescence (Byun *et al.*, 2009). It is also important in the formation of high molecular weight

eEF1 complexes by providing binding sites for both α and γ subunits. (van Damme *et al.*, 1990; van Damme *et al.*, 1992).

1.5 Protein-protein interaction

Specific complimentary recognition of two or more peptides to form a stable structure is termed protein-protein interactions (Werther and Seitz, 2008). Protein-protein interactions (PPIs) lead to the formation of dimeric or multimeric proteins. Multimeric proteins are found in the cytosol, cell membrane and cell organelles (Hardy *et al.*, 1988). Many proteins self-associate to form homodimers. Homodimerisation can occur between monomers in solution with or without intervention from promoters (Nussinov *et al.*, 1998). Most homodimers exist in their dimeric state and it is very difficult to separate them without denaturing their individual monomeric structures. PPIs are usually complex and their stability is attributed to and regulated by some environmental conditions such as changes in pH, temperature, ionic strength, and covalent modifications such as phosphorylation (Markus and Benezra, 1999). PPIs can be classified based on stability and mechanism of the protein-protein complex.

PPIs occur between hetero-oligomer (non-identical) and homo-oligomer (identical) peptide chains. Homologous protein oligomers can be arranged in an heterologous or isologous manner with structural symmetry (Goodsell and Olson, 2000). Heterologous interaction involves the use of different interfaces which can lead to unending aggregation because it does not have a closed symmetry, while isologous interaction uses the same surface on both monomers (Nooren and Thornton, 2003).

Obligate PPIs involves the use of promoters that are unstable structures when they are on their own and an example of such promoters is Arc repressor dimer which is very essential for DNA binding (Jones and Thornton, 1996). Non-obligate PPIs involves the use of promoters that can exist on their own and their components are stable independently. Also each interacting pair of proteins has their own unique complex interface. Examples of non-obligate PPIs include HYHEL-5 with lysozyme which is an antibody-protein complex and enzyme-inhibitor complex trypsin found in bitter gourd (Jones and Thornton, 1996; Archakov *et al.*, 2003).

PPIs are classified as permanent or transient based on their life span and nature of interaction. Permanent complexes are stable with protein-protein interfaces that are closely packed together and have fewer intersubunit hydrogen bonds (Jones and Thornton, 1996). Their surface properties are very close to that of the protein core because they are extension of the protein folding. Permanent complexes exhibit the highest complementarity while transient or temporary complexes have the lowest complementarity (Jones and Thornton, 1996; Tsai *et al.*, 1997). Non-obligate interactions can be permanent or transient while obligate interactions are permanent (Nooren and Thornton, 2003).

In multimeric proteins, PPIs are essential for normal functioning of cells (Teichmann, 2002) which includes: transportation of cholesterol and lipids among certain cells in the body which is achieved by the interaction between the plasma protein apolipoprotein E and low density lipoprotein receptors (Mahley, 1988), catalysing metabolic reactions and changing specificity of the protein (Peng *et al.*, 2016). PPIs form the basis of the quaternary structure (Jones and Thornton, 1995) and changes in the quaternary state of the protein can lead to a biological function or activity (Nooren and Thornton, 2003). PPIs are important biological regulators as seen in during the association of polypeptides with each other or with nucleic acids and phospholipids (Pawson and Nash, 2003). They are also very essential in designing drugs, optimisation of drug therapies already in use (Archakov *et al.*, 2003; Wendt, 2012) as well as cancer therapeutic strategy (Peng *et al.*, 2016). PPIs facilitate biochemical functions such as enzyme cooperativity and signal transduction (Jones and Thornton, 1995; Jones and Thornton, 1996). In order to understand the dynamics and stability of proteins, PPIs in cells are very essential in discovering the structure and functions of many unknown proteins (Figeys, 2002). Some novel proteins can be assessed by describing their localisation in the cell (Teichmann, 2002) and certain functions could be allotted to the protein based on the known functions of their interacting partners (Schwikowski *et al.*, 2000). PPIs can be driven by polar interactions (hydrogen bonds and van der waals²), electrostatic interactions (salt bridges) and hydrophobicity among others.

Hydrophobicity is one of the major driving forces in the stabilisation of protein folding and in PPIs. Hydrophobic interactions describes the free energy gained when non-polar residues of proteins interact in polar environment (Kauzmann, 1959), and can be known as solvent entropy.

Hydrophobic effect on protein structure was first identified by Kauzmann (1959) and he proposed that in aqueous environment, proteins tend to bury the non-polar amino acid residues while orienting those with polar or charged side chains to interact with the solvent molecules. Entropy lost by protein molecules while forming complexes was compensated by the entropy gained by water as a result of the accessible protein surface area (Chothia and Janin, 1975). This process is entropically driven at room temperature because the addition of non-polar molecules to water disrupts the hydrogen bonded structure of water hence the water molecules arrange themselves so that they can have greater contacts with themselves and lesser contacts with the non-polar substance (Geiger *et al.*, 1979; Stillinger, 1980). Averaged values of the contact surface hydrophobicity represent the mean hydrophobicity value of the protein core and its surface (Janin *et al.*, 1988). Hydrophobic areas in the contact interface are arranged as patches and the proteins associate through the hydrophobic patches on their surfaces (Jones and Thornton, 1996). The number and size of these patches vary. Hydrophobic interactions are greater in permanent complexes and lesser in temporary complexes. Some of the examples of PPIs driven by hydrophobicity are dimeric porcine insulin, $\alpha\beta$ dimer of horse oxy-haemoglobin and bovine trypsin-pancreatic inhibitor complex (Chothia and Janin, 1975).

PPIs can also be driven by hydrogen bonds. Hydrogen bonding is a non-covalent interaction which involves the sharing of hydrogen atom between hydrogen bond donor group such as hydroxyl group (-OH) or amino group (NH_2) and hydrogen bond acceptor groups such as oxygen atom or nitrogen atom which are the peptide backbone groups and the polar amino side chains in proteins (Sticke *et al.*, 1992). Hydrogen bond is an intrinsic component of PPI and has been suggested to bring about specificity in PPI (Fersht, 1987). Backbone hydrogen bonds are mainly local whereas at the domain interfaces most hydrogen bond contributions are due to polar residues since these contacts are mostly non-local. The strength of the hydrogen bond will depend on the relative angles and distances. Hydrogen bonds involving the main chain atoms determine the stability of the protein's secondary structures (Creighton, 1991). In α -helix the backbone is tightly wound around a central axis with a full turn for every 3.6 amino acid residues. For every amino acid residue forming the helix except for the end residues, a hydrogen bond is formed between the hydrogen attached to the nitrogen of the backbone and the carbonyl oxygen of the amino acid residue four positions along the chain whereas in β -pleated sheets

hydrogen bonding occurs between the nitrogen atoms and carbonyl oxygen atoms of the backbone. Intra-molecular hydrogen bonds form to replace the fluctuating intra and inter molecular bonds that form in the unfolded protein so that the native state is favoured enthalpically in the folded conformations. Though hydrogen bonds are not the main folding force they are also important in maintaining protein stability (Dill, 1990). Hydrogen bonds are observed at protein interfaces at an average of about 10 bonds per interface or one bond per each 100-200 Å² (Jones and Thornton, 1996). Examples of PPIs driven by hydrogen bonds are human immunodeficiency virus (HIV) protease (Navia *et al.*, 1989) and subtilisin inhibitor homodimer (Mitsui *et al.*, 1979).

Salt bridges also known as specific charge contact are electrostatic interactions formed between acidic amino acids which are negatively charged such as aspartic or glutamic acid and basic amino acids that are positively charged such as arginine, lysine and histidine. Based on the geometry, location in the protein, whether they are hydrogen bonded or not, the energy contribution of salt bridges varies between 5-15 kcal mol⁻¹ (20-60 kJ mol⁻¹) per ion pair. Due to the high amount of energy required for the transfer of charged ions from a polar to a non-polar environment, the amount of ion pairs at the dimer interfaces and domains of the proteins are low. This energy is known as Born energy and it is about 80 kJ mol⁻¹ (Dill, 1990). Salt bridges are thereby responsible for correct packing and binding specificities in protein interiors, domains and dimer interfaces. Salt bridges are usually stabilising although they stabilise proteins only under favourable packing conditions in a non-polar environment (Kumar and Nussinov, 1999).

Association of proteins is mainly due to complementarity in structure and also the co-operation of some weak forces such as van der waals' interaction. Although van der waals' interactions are less energetic, they are more numerous than hydrogen bonds. The overall contribution of van der waals' in PPIs is very small, but has been found to be essential in determining which protein recognise another protein (Chothia and Janin, 1975).

Proteins forming PPI must maintain a stable conformational surface to enable recognition and binding interaction. Most proteins are specific in their choice of binding partners and examples include hormone-receptor and antibody-antigen complex, while some are multispecific having more than one binding partners examples include regulatory pathways such as RhoGAP which is an intracellular cell signalling network (Nooren and Thornton, 2003). Complementarity of shape

and chemistry which determines the free energy of binding brings about specificity. Non-native PPI partners will not form PPI. The propensity of PPI is an index of native and active protein performing a biological function.

1.6 Protein-protein interaction between the subunits of eEF1 complex

Complex formation between the four subunits of eEF1 increases the activity of the molecule (Motorin *et al.*, 1991). The CT-domain of eEF1 α interacts with the CT-domain of eEF1 β and this interaction is of great importance because eEF1 β triggers the exchange of GDP to GTP and protects eEF1 α against tryptic cleavage (van Damme *et al.*, 1992). The CT-domain of eEF1 δ interacts with the CT-domain of eEF1 α (Janssen *et al.*, 1994). Furthermore eEF1 δ interacts with eEF1 γ and the complex eEF1 $\beta\gamma$ but not with eEF1 β alone, this could be due to the conformational changes in both secondary and tertiary structure of the complex, thereby creating a binding site for eEF1 δ (Janssen *et al.*, 1994). Recombinant eEF1 $\beta\gamma\delta$ stimulate the activity of EF1 α by up to 10-fold, indicating that it is a functional complex that produces a greater level of stimulation than both eEF1 β and eEF1 $\beta\gamma$ (Sheu and Traugh, 1997). In order to acquire more detailed information on how the complexes formed by the subunits of eEF1 are structurally and functionally related to each other as well as the role of eEF1 β in the complex eEF1 $\alpha\beta\gamma\delta$ during protein synthesis, there is need for a more advanced study of the β - γ subunits interactions. Detailed analysis is required to assess comprehensively how proteins interact and exchange information.

The β and γ subunits of eEF1 complex form a single functional unit, which can only be dissociated with the use of denaturants (Chen and Traugh, 1995). The β subunit binds to the NT-domain of the γ subunit (van Damme *et al.*, 1990). Although, studies have shown that the nucleotide exchange activity resides in the β subunit and not γ subunit (Janssen and Moller, 1988), the rate of GDP exchange increased in $\alpha\beta\gamma$ complex when compared to $\alpha\beta$ subunit (van Damme *et al.*, 1990). The amino terminal of the complex eEF1 $\beta\gamma$ is the binding site for eEF1 δ subunit (Janssen *et al.*, 1994) since eEF1 δ subunit has never been reported to interact with eEF1 β alone. The level of expression of eEF1 $\beta\gamma$ increased in certain tumour cells (Veremieva *et al.*, 2014) indicating independent effect of this complex in some tumours and hence the importance for extensive study of this protein-protein interaction.

1.7 Aim and objectives

The β -subunit of elongation factor 1 complex (eEF1) plays an essential role in the elongation step in eukaryotic protein biosynthesis, which essentially involves interaction with the α - and γ -subunits. The functional rationale for these complexes (EF1 $\beta\gamma$ and EF1 $\beta\alpha$) are not fully understood, hence the need for further studies on the protein-protein interactions. This work was aimed to biophysically characterise heEF1 β by constructing three *E. coli* expression vector systems for recombinant expression of the full length (FL-heEF1 β), N-terminus (NT-heEF1 β) and the C-terminus (CT-heEF1 β) regions of the protein and to qualitatively assess its interaction with heEF1 γ in the presence oxidised glutathione ligand (GSSG). These aims are to be achieved with the following objectives:

- Confirm the identity of the gene containing FL-heEF1 β by sequencing the plasmid.
- Create NT-heEF1 β fragment from FL-heEF1 β using site directed mutagenesis.
- Express and purify FL-heEF1 β , NT-heEF1 β and CT-heEF1 β to homogeneity.
- Secondary characterisation of FL-heEF1 β , NT-heEF1 β and CT-heEF1 β using far-UV circular dichroism (CD).
- To assess the hydrophobic binding pockets of heEF1 β using extrinsic-ANS binding fluorescence assay.
- Quaternary structure characterisation of FL-heEF1 β , NT-heEF1 β and CT-heEF1 β using size exclusion high pressure/performance liquid chromatography (SE-HPLC).
- Functional characterisation by protein-protein interactions of FL-heEF1 β , NT-heEF1 β and CT-heEF1 β with heEF1 γ using size exclusion high performance liquid chromatography (SE-HPLC) in the presence and absence of oxidised glutathione ligands (GSSG).

CHAPTER 2

Materials and methods

2.1 Materials

The synthesised codon harmonised gene encoding full length heEF1 β which were cloned using NdeI and Bam HI restriction sites into pET-28a plasmid to create pTFL-heEF1 β were kindly provided by Dr Ikechukwu Achilonu. All other reagents used were of analytical grade.

Table 1: Non-standard materials and suppliers. These are some of the chemicals and materials used during this research study and companies where they were purchased.

Materials	Source	Location
Mutagenesis primers	Inqaba Biotech	Pretoria, South Africa
GeneJet™ plasmid miniprep kit	Fermentas	Ontario, Canada
Quikchange II XL Site directed mutagenesis kit	Stratagene	La Jolla, Ca, USA
SDS Molecular weight markers	Fermentas	Ontario, Canada
Isopropyl β -D-1-thiogalactopyranoside (IPTG)	Sigma-Aldrich	St. Louis, MO, USA
Dithiothreitol (DTT)	Melford Laboratories	Suffolk, UK
Coomassie Brilliant Blue-G250	Sigma-Aldrich	St Louis, MO,USA
BL21 codon plus competent cells	New England Biolabs	Ontario, Canada
IMAC columns (nickel and cobalt)	GE Healthcare	Buckinghamshire, UK
Antibiotics (ampicilin and chloramphenicol)	Roche Diagnostics	Manheim, Germany
Yarra™ 3u SEC-2000, LC Column	Phenomenex	Torrence, CA, USA

2.2 Methods

2.2.1 Construction of expression vectors

The gene sequence encoding heEF1 β open reading frame (ORF) with the accession number CAG33106.1 was retrieved from GenBank and was codon harmonised to enable expression in *Escherichia coli* (*E.coli*). Codon harmonisation is an algorithm developed from the relationship between the secondary protein structure and codon usage frequencies in a heterologous expression in order to improve expression (Angov *et al.*, 2008). The harmonised sequence encoding FL-heEF1 β and CT-heEF1 β were synthesised and cloned into pET-28a and pET-11a vector by GenScript Corporation (NJ, USA) for the expression of FL-heEF1 β and CT-heEF1 β

respectively. Both the CT-heEF1 β and FL-heEF1 β sequences have NT-hexahistidine tag incorporated into them to enable purification using immobilised metal affinity chromatography. The synthesized gene was cloned using the NdeI and BamHI restriction sites into the plasmids to create pTFL-heEF1 β (Achilonu *et al.*, 2014). Six proteins construct which are: FL-heEF1 γ , NT-heEF1 γ , CT-heEF1 γ , FL-heEF1 β , NT-heEF1 β and CT-heEF1 β were used in this study (Figure 2.1). The FL-heEF1 γ , NT-heEF1 γ and CT-heEF1 γ protein plasmids were provided from previous research (Achilonu *et al.*, 2014).

2.2.2 Plasmid extraction and sequencing

The pET-28a is a bacterial expression vector with T7 lac promoter, thrombin cleavage site and an N-terminal histagged sequence. The histag allows for effective detection and purification of the protein. The pET-28a vector encodes a gene for kanamycin resistance by while the pET-11a plasmid encodes a gene for ampicillin resistance by NdeI/BamHI, thus selecting only the cells containing the plasmids. The plasmids (pTFL-heEF1 β) were extracted from the cells using GeneJetTM plasmid miniprep kit based on alkaline lysis (Bimboim and Doly, 1979) in accordance with the protocol detailed as follows: 1 ml cell culture was centrifuged at 27000 $\times g$ twice, to harvest the cells after which 250 μ l of resuspension solution was added to the cells to maintain optimal pH and chelate divalent cations because it contains EDTA which inhibit enzymes from cleaving the plasmid DNA. A pipette was used to resuspend the cells by gentle aspiration. The cells were then lysed under alkaline conditions by the adding 250 μ l of lysis solution containing SDS (sodium dodecylsulfate) and NaOH until the solution becomes viscous and slightly clear. Thereafter, 350 μ l of the neutralisation solution containing sodium acetate was added and mixed thoroughly by inverting the tube four to six times to neutralise the reaction and to precipitate out the proteins and larger genomic DNA. The cell solution was centrifuged at 27000 $\times g$ for 5 min to pellet cell debris and chromosomal DNA, after which the supernatant was decanted into a GeneJet spin column which contains a silica-based membrane that binds the plasmid DNA and then centrifuged at 27000 $\times g$ for 1 min, the column was placed back in the same collection tube after the flow-through was discarded. The plasmid DNA bound to the column was washed twice with 500 μ l washing solution centrifuged at 27000 $\times g$ for 1 min and the column placed back in the same collection tube after discarding the flow through. The plasmid solution was centrifuged again at 27000 $\times g$ for 1 min to remove residual wash solution and ethanol in the DNA plasmid. The GeneJet spin column was then transferred into a fresh Eppendorf tube. The plasmid DNA

was eluted with elution buffer and then incubated at room temperature for 2 min. The plasmid DNA solution was then centrifuged at $27000\times g$ for 2 min and stored at -20°C . The purified plasmid was sent to Inqaba Biotech (Pretoria, South-Africa) for sequencing. The chromatogram received from Inqaba Biotech contained gene sequence which was translated using the online server ExPASy translate (Artimo *et al.*, 2012). The translated amino acid sequences (Figure 2.2) were then compared with the sequence in the database using the basic alignment search tool (BLAST) and to confirm the proteins identity (Altschul *et al.*, 1990).

2.2.3 Site Directed Mutagenesis

Site directed mutagenesis involves the extensive analysis of gene structure and function (Shenoy and Visweswariah, 2003). The amino terminus (NT) fragment of both proteins heEF1 β and heEF1 γ were created out of their respective full length (FL-heEF1 β and FL-heEF1 γ) by site directed mutagenesis (Figure 2.3) using mutagenic primers (reverse and forward). For heEF1 β , a stop codon was encoded at Lys 79 (AAA \rightarrow TAA) which is a single nucleotide change. The length of plasmid and each protein (FL-heEF1 β and FL-heEF1 γ) is approximately 6000 bases and the time of cycle is about 3 h. The reaction mixtures in a total volume of 50 μl consisted of:

- I- **HeEF1 β** : 2.5 μl FL- heEF1 β -pET 28 plasmid, 1.5 μl (125 ng) forward primer, 1.5 μl (125 ng) reverse primer, 5.0 μl reaction buffer, 1.0 μl dNTP mix, 1.5 μl Quiksolution reagent and 37 μl sterile MilliQ H₂O.
- II- **HeEF1 γ** : 1.0 μl FL- heEF1 γ -pET 11 plasmid, 2.0 μl (125 ng) reverse primer, 2.0 μl (125 ng) forward primer, 5 μl Reaction buffer, 1.0 μl dNTP mix, 1.5 μl Quiksolution reagent and 37.5 μl sterile MilliQ H₂O.

The Biorad MycyclerTM was used and the cycling parameters used were as follows: two amplification cycles of 2 min at 95°C as the starting temperature followed by 18 cycles consisting of denaturation step at 95°C for 20 sec, an annealing step at 60°C for 10 sec which is a temperature suitable for the primers, then an extension step at 68°C for 220 min. Finally one cycle of 68°C for 5 sec, this is the holding temperature. 2 μl of *Dpn I* was added to the reaction mixture and incubated at 37°C for 5 min to cleave methylated DNA (Braman *et al.*, 1996). The reaction products were then used to transform *E.coli* JM109 competent cells as follows:

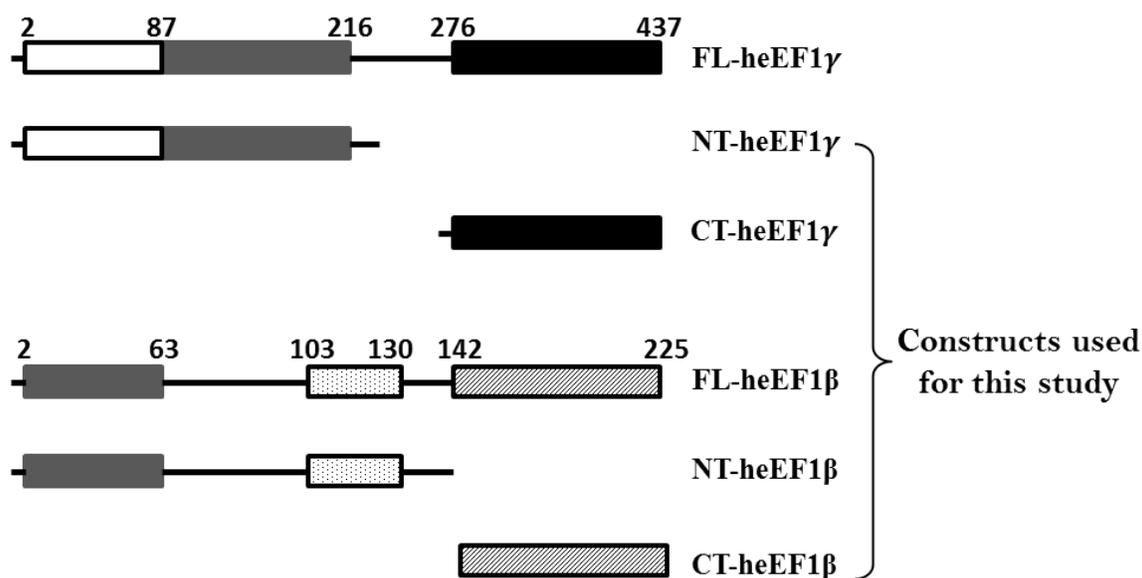


Figure 2.1: Schematic representation of the constructs used in this study. For the heEF1 γ protein three constructs used are: CT- heEF1 γ and NT-heEF1 γ were used. NT-heEF1 γ was created through site directed mutagenesis by encoding a stop codon between the NT-domain and CT-domain of FL-heEF1 γ . While for the heEF1 β protein the constructs were: FL-heEF1 β , CT-heEF1 β and NT-heEF1 β . NT-heEF1 β created by site directed mutagenesis which involved encoding a stop codon on Lys 79 of the FL-heEF1 β .

MGSSHHHHHSSGLVPRGSHMGFGDLKSPAGLQVLNDYLADKS
YIEGYVPSQADVAVFEAVSSPPADLCHALRWYNHIKSYEKEK
ASLPGVKKALGKYGPADVEDTTGSGATDSKDDDDIDLFGSDDE
EESEEAKRLREERLAQYESKKAKKPALVAKSSILLDVKPWDDE
TDMAKLEECVRSIQADGLVWGSSKLVVPGYGIKKLQIQCVVED
DKVGTDMLEEQITAFEDYVQSMDVAAFNKI

Figure 2.2: Amino acid composition of FL-heEF1 β . The NT-domain is in red with an hexahistidine tag and the underlined sequence represents the CT-GST like region of the NT-domain. The CT- heEF1 β is indicated in black.

2 μl of the JM109 cells was thawed on ice, 1 μl of the NT-heEF1 β mixture was added to it and then left on ice for 30 min to stabilise the lipid membranes of the cells. The cells were heat-shocked for 45 sec at 42°C on a heating block to alter the state of the fluid membrane by increasing its permeability and allowing the DNA to enter the cell. It was immediately put back on ice to cool down for 5 min. Thereafter, 750 μl of SOC media [2% (w/v) tryptone, 0.5% (w/v) yeast extract, 250 mM KCl, 2 M MgCl₂, 1 M glucose] was added to the transformed cells to provide nutrients and allow the cells to grow. The cells were then incubated at 37°C for 1 h with shaking at 230 rpm and then plated onto lysogeny broth (LB) agar plates [1% (w/v) tryptone, 1.5% (w/v) agar, 0.5% (w/v) yeast extract, 1% (w/v) NaCl] containing 30 $\mu\text{g}/\text{ml}$ kanamycin sulfate and subsequently incubated overnight at 37°C. Transformants were selected at random and the plasmid DNA was extracted from the JM109 cells using the molecular biology ThermoScientific kit protocol as described above (section 2.2.2).

2.2.4 Transformation and over expression

Recombinant proteins (FL-heEF1 β , CT-heEF1 β , NT-heEF1 β , CT-heEF1 γ and NT-heEF1 γ) were used to transform *E. coli* BL21 Codon Plus (Agilent) expression cells. The competent cells were transformed using the method described by Chung and colleagues (1989). BL21 codon plus competent cells were thawed on ice for 5-10 min after which 2 μl of pT-heEF1 β plasmid was added. The reaction mixture was then incubated on ice for 30 min, heat shocked at 42°C for 45 secs on a heating block and immediately transferred on ice for 5 min. The cells were grown by adding 750 μl of SOC media [2% (w/v) tryptone, 0.5% (w/v) yeast extract, 250 mM KCl, 2 M MgCl₂, 1M glucose] to the reaction mixture followed by incubation at 37°C using a shaker incubator at 250 rpm agitation for 1 h. The cells were then placed on LB- agar plates [1% (w/v) tryptone, 1.5% (w/v) agar, 0.5% (w/v) yeast extract, 1% (w/v) NaCl] containing the antibiotics 100 $\mu\text{g}\cdot\text{mL}^{-1}$ kanamycin and 30 $\mu\text{g}\cdot\text{mL}^{-1}$ chloramphenicol for heEF1 β and 100 $\mu\text{g}\cdot\text{mL}^{-1}$ ampicillin and 30 $\mu\text{g}\cdot\text{mL}^{-1}$ chloramphenicol for heEF1 γ and then incubated at 37°C overnight (~ 16 h).

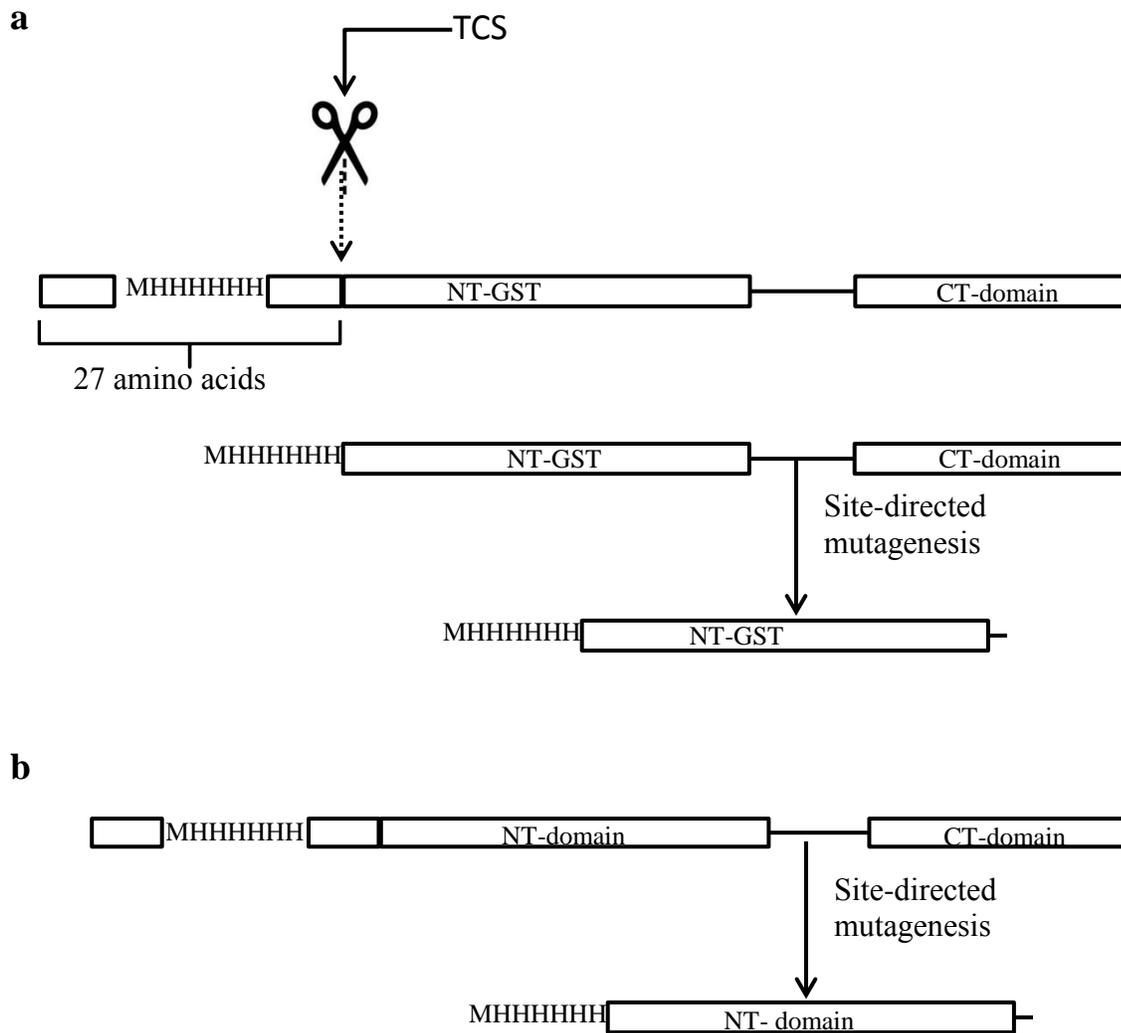


Figure 2.3: Site directed mutagenesis using mutagenic primers (reverse and forward). (a) Schematic representation of site directed mutagenesis used to create NT- heEF1 γ from FL- heEF1 γ which is an NT-GST-like domain (b) Schematic representation of site directed mutagenesis used to create NT- heEF1 β from FL- heEF1 β by encoding a stop codon at Lys79 (AAA→TAA single nucleotide change).

The transformed cells were then picked from the colonies observed on the LB-agar plates and added to freshly prepared sterile 2×YT media [1% (w/v) yeast extract, 1.6% (w/v) tryptone, 0.5% (w/v) NaCl] supplemented with 100 µg.mL⁻¹ kanamycin and 30 µg.mL⁻¹ chloramphenicol and incubated at 37°C overnight with 250 rpm agitation. A 50-fold dilution was used to inoculate fresh sterile 2×YT media supplemented with 100 µg.mL⁻¹ kanamycin and 30 µg.mL⁻¹ chloramphenicol, 10 µl antifoam 204 was added and then incubated at 37°C, 250 rpm agitation till an OD₆₀₀ ~0.5 was reached. Cell culture was then chilled on ice for 10 min. Cold induction was done using 0.5 mM IPTG for expression of the proteins (FL-heEF1β, CT-heEF1β, NT-heEF1β, CT-heEF1γ and NT-heEF1γ). The cells were grown for a further 6 h, incubated at 30°C with shaking at 250 rpm to achieve optimum protein expression. Aliquots of 1000 µl cell culture collected at 0, 2 h, 4 h, 6 h and 16 h post-induction was analysed using tricine-SDS PAGE. Cells were pelleted by centrifugation at 5000×g for 25 min. Harvested cells were resuspended with buffer A [50 mM Tris-HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1 M NaCl, 40 mM imidazole pH 7.4] and then stored at -20°C.

2.2.5 Purification

2.2.5.1 Immobilised metal affinity chromatography

The frozen lysed cell suspensions were thawed at 37°C and then thawed on ice by sonication for five cycles of 30 sec bursts, using a power output of 12 with a Sonicator Ultrasonic Processor (Misonix Incorporated). The lysed cells were then centrifuged at 25000×g for 30 min at 4°C to pellet the insoluble fraction. Decanted supernatant (soluble fraction) was subsequently loaded to a 5 ml nickel resin column for FL-heEF1β, CT-heEF1β, NT-heEF1β and CT-heEF1γ and cobalt resin for NT-heEF1γ which has been pre-equilibrated with buffer A using the ÄKTA FPLC (fast protein liquid chromatography) purification system (GE Healthcare) coupled to a computer with prime view 1.0 software. The column was then washed with 10 column volumes of the buffer A followed by 10 column volumes of buffer B [50 mM Tris-HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1 M NaCl, 40 mM imidazole, 1% (v/v) Triton X-100, pH 7.4] and finally washed with 10 column volumes buffer A to remove excess Triton X-100. The bound proteins were eluted off the column using the eluting buffer C [50 mM Tris-HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1 M NaCl, 250 mM imidazole pH 7.4]. Eluted proteins were collected in fractions and tested for

presence of protein using Bradford reagent. Tricine SDS-PAGE was used to analyse the purity of the protein and the concentration assessed by Beer-Lambert's law.

2.2.5.2 Ion exchange chromatography

Ion exchange chromatography involves the interaction between the charged molecules in the mobile phase and the oppositely charged groups attached to the stationary phase. Diethylaminoethyl are positively charged ion exchange matrices used as anion exchangers because they bind to proteins that have an overall negative charge. Proteins are released from the resin by increasing the salt concentration of the buffer or changing the pH of the solution. IMAC purified fractions of FL-heEF1 β and NT-heEF1 β proteins were pulled together and dialysed (16 h, 4°C) against buffer D [50 mM Tris-HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1 mM EDTA, 1 mM EGTA, 0.1 mM TCEP pH 7.4] and then further purified by loading onto a 20 ml DEAE column that is connected to an Äkta Prime FPLC system, which has been pre-equilibrated with buffer D. The proteins were eluted using the DEAE elution buffer E [50 mM Tris-HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1 mM EDTA, 1 mM EGTA, 0.1 mM TCEP, 300 mM NaCl, pH 7.4]. Eluted protein fractions were collected and assessed for purity using SDS-PAGE gel.

2.2.5.3 Glutathione-Agarose affinity chromatography

Glutathione (GSH)-Agarose affinity chromatography is used for non-denaturing and highly selective purification of proteins containing glutathione such as glutathione peroxidase and GST. Due to the GST-like amino terminal domain of heEF1 γ , both NT-heEF1 γ and FL-heEF1 γ have high affinity for glutathione. The agarose beads were washed thoroughly with 10 column volumes of buffer F [50 mM Tris-HCl, 0.02% (w/v) NaN₃, 1 M NaCl, pH 7.4] to equilibrate the column. The IMAC purified fractions NT-heEF1 γ protein collected were pulled together and dialysed (16 h, 4°C) against buffer G [50 mM Tris-HCl, 0.02% (w/v) NaN₃, 1 M NaCl, 2 mM DTT pH 7.4] to remove the excess imidazole. The dialysed proteins were then passed through the GSH-agarose column and then washed with 10 column volumes of buffer H [50 mM Tris-HCl, 0.02% (w/v) NaN₃, 1 M NaCl, 2 mM DTT, 1% (v/v) Triton x-100 pH 7.4] followed by 10 column volumes of buffer F to remove excess Triton x-100 detergent. Bound proteins were eluted with buffer I [10 mM Glycine-NaOH, 0.02% (w/v) NaN₃ pH 10]. Eluted protein fractions

were collected and their pH immediately adjusted to ~7.5 by adding 1 M Tris-HCl pH 7.4 at 25% (v/v). The purity of the proteins collected was assessed using tricine SDS-PAGE.

2.2.6 Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a technique used to assess the expression and purity of the protein. Both glycine SDS-PAGE and tricine SDS-PAGE could be used based on the size of the protein, but in this study both were used based on the availability in the laboratory.

2.2.6.1 Tricine-based sodium dodecyl sulfate polyacrylamide gel electrophoresis

The size, solubility and purity analysis of the proteins were assessed using the tricine SDS-PAGE according to the method described by Herman Schägger. The principle of this technique is similar to SDS-PAGE except that tricine is used instead of glycine as the trailing ion and it is mainly used in the separation of small proteins and peptides smaller than 30 kDa (Schägger, 2006). Protein samples were prepared in 2:1 dilution with loading buffer [0.5 mM Tris-HCl, 20% (v/v) glycerol, 10% (w/v) SDS, 100 mM β -mercapto ethanol, 0.05% bromophenol blue, pH 6.8] and boiled at 100°C for 5 min to ensure complete denaturation before loading onto gel. The separating gel consists of 0.6% (w/v) acrylamide/bisacrylamide, 1% (w/v) gel buffer (3 \times), 0.3% (v/v) glycerol, 1.5% (w/v) ammonium persulfate, 0.005% (w/v) TEMED and the stacking gel consists of 0.1% (w/v) acrylamide/bisacrylamide, 0.3% gel buffer (3 \times), 0.009% (w/v) ammonium persulfate, 0.005% (w/v) TEMED. The anode buffer used consist of 1 M Tris-HCl, pH 8.9 and the cathode buffer consist of 1 M Tris, 1 M Tricine, 1% SDS, pH 8.25. 20 μ l of the protein sample was loaded into the SDS-PAGE wells and electrophoresed at 160 volts for 90 min using a PowerPacTM Basic Bio-Rad electrophoresis system. The molecular weight marker used contained a mixture of seven proteins: lysozyme (14.4 kDa), β -lactoglobulin (18.4 kDa), restriction endonuclease Bsp98I (25 kDa), lactate dehydrogenase (35 kDa), ovalbumin (45 kDa), bovine serum albumin (66.2 kDa), and β -galactosidase (116 kDa). The gels were stained in a staining solution [0.25% (w/v) Coomassie Blue R250, 10% (v/v) glacial acetic acid and 45% (v/v) methanol] for 3-4 h and then destained in 50% (v/v) methanol, 40% (v/v) water and 10% (v/v) glacial acetic acid until the background was clear.

2.2.6.2 Glycine-based sodium dodecyl sulphate-polyacrylamide gel electrophoresis

Sodium dodecyl sulphate (SDS) is an anionic detergent which gives uniform negative charge to proteins when bound to them (Pitt-Rivers and Impiombato, 1968). SDS also denatures the proteins into their individual polypeptide units. The expressed proteins and purified protein fractions collected were subjected to discontinuous SDS-PAGE according to the Laemmli method (Laemmli, 1970). The discontinuous gel system consisted of a 12% separating gels [12% (w/v) acrylamide, 1.35% (w/v) bisacrylamide, 0.25 M Tris-Cl pH 8.8, 0.1% (w/v) SDS, 0.05% (v/v) ammonium persulfate and 0.1% (v/v) TEMED] and 4% acrylamide stacking gels [4% (w/v) acrylamide, 0.36% (w/v) bis-acrylamide, 0.05 M Tris-HCl pH 6.8, 0.01% (w/v) SDS, 0.005% (v/v) ammonium persulfate and 0.2% (v/v) TEMED]. The loading buffer [0.5 mM Tris-HCl, 20% (v/v) glycerol, 10% (w/v) SDS, 100 mM β -mercapto ethanol, 0.05% bromophenol blue, pH 6.8] was mixed with the protein in the ratio 1:2 and then boiled at 100°C for 5 min before loading onto the gels. The gels were electrophoresed using the electrode buffer (anode and cathode) [250 mM Tris-HCl, 192 mM glycine, 0.1% (w/v) SDS (pH 8.3)] at 160 V for approximately 90 min using a PowerPacTM Basic Bio-Rad electrophoresis system. The molecular weight markers used are the same as in section 2.2.6.1. The gels were then stained with Coomassie Brilliant Blue R-250 staining solution same as in section 2.2.6.1 for 3-4 h, followed by destaining with a 1:5:4 (acetic acid : methanol : water) solution overnight.

2.2.7 Liquid Chromatography-Mass Spectrometry/ Mass Spectrometry

Mass Spectrometry (MS) can be used to identify proteins through peptide mass fingerprinting (PMF) and sequence-specific peptide fragmentation. Mass spectrometry (MS) involves separation of samples according to their mass-to-charge (m/z) ratio.

Liquid chromatography (LC) or gel electrophoresis are the standard approach to protein identification (Thiede *et al.*, 2005). LC involves two phases namely a mobile phase and a stationary phase which is attached to a bed with the samples to be separated distributed selectively between both phases. It is a process whereby sample particles undergoes sorption and desorption on the stationary phase. Larger particles elute before the smaller ones.

Liquid chromatography-Mass spectrometry (LC-MS) involves the combination of two techniques which are the separation technique of liquid chromatography (HPLC) and mass

analysis abilities of mass spectrometry. The use of this combined technique is essential because it is sensitive, accurate and can tolerate various levels of contaminants (Kaufmann, 1995). The 12% SDS-PAGE gel of pure FL-heEF1 β after electrophoresis was sent to CSIR (Council for Scientific and Industrial Research Pretoria, South Africa). An in-gel trypsin digest and LC-MS/MS was carried out in order to determine the protein contents of the bands. The bands containing the protein to be analysed was extracted from the gel and destained using 50 mM ammonium bicarbonate (NH₄HCO₃)/50% methanol, after which 55 mM iodoacetamide in 25 mM NH₄HCO₃ was used to carry out alkylation which was then exposed to in-gel digest using trypsin at 37°C overnight. The digest was then resuspended in 50% acetonitrile/5% formic acid solution which was analysed by rapid separation liquid chromatography using a Dionex Ultimate 3000 RSLC system attached to a QSTAR ELITE mass spectrometer. Electrospray ionisation was used to ionise the peptide and the QSTAR ELITE mass spectrometer was used to measure the mass of the ions. Protein pilot using the Paragon search engine (AB Sciex) (Shilov *et al.*, 2007) was used to match the obtained MS/MS spectra with proteins in a UniSwiss database supplied. Proteins with percentage confidence above 95% were reported for the bands analysed.

2.2.8 Protein quantification

2.2.8.1 Estimation of the molecular weight of the proteins

The molecular weights of the proteins were derived from the equation of the line, from the graph of Log of molecular weights of the standards used in the gels against the distance travelled (cm). The distance travelled by the protein is measured from the electrophoretogram and substituted in the equation of the line to derive the molecular weight of the protein or can be interpolated from the graph.

2.2.8.2 Determination of protein concentration

The concentrations of the heEF1 β proteins (FL-heEF1 β , CT-heEF1 β and NT- heEF1 β) as well as heEF1 γ proteins (CT-heEF1 γ and NT-heEF1 γ) were determined using a Jasco V-630 spectrophotometer. The Beer-Lambert law was applied as follows:

$$A = \epsilon cl \quad (2.1)$$

$A = \epsilon cl$ where A is the absorbance, ϵ is the molar extinction coefficient ($M^{-1}cm^{-1}$) at a given wavelength (usually 280 nm for protein detection), c is the molar concentration and l is the pathlength (cm).

$$Conc (mg/ml) = \frac{Slope \times M_r}{\epsilon} \quad (2.2)$$

Slope is gotten from the equation of the line in the graph of absorbance versus wavelength. The molar extinction coefficients was derived theoretically from ExPASy ProtParam (Gasteiger *et al.*, 2005). For heEF1 β proteins: $2 \times 29,910 M^{-1}cm^{-1}$ for FL-heEF1 β because it is a dimer, $48,930 M^{-1}cm^{-1}$ for CT-heEF1 β and $15930 M^{-1}cm^{-1}$ for NT-heEF1 β while for heEF1 γ proteins: $86860 M^{-1}cm^{-1}$ for FL-heEF1 γ , $37930 M^{-1}cm^{-1}$ for NT-heEF1 γ and $48930 M^{-1}cm^{-1}$ for CT-heEF1 γ .

Absorbance at 280 nm was measured for a serially diluted solution of the protein solution using doubling dilution method and the concentration determined by fitting a linear regression to six points of absorbance versus wavelength. The buffer readings were subtracted from the absorbance readings used for the concentration determination. The determined interference at 340 nm was subtracted from the measured absorbance at 280 nm to correct the effects of light scattering. Substitution of the corrected absorbance, the extinction coefficient, and the pathlength (1 cm) into the Beer-Lambert law (above) yields the molar concentration (M) of soluble protein.

2.2.9 Determination of protein quality using absorbance spectrometry

The absorbance of the proteins was monitored at wavelength 260 nm–340 nm to check for DNA contamination and protein aggregation. 5 μM of pure protein in 10 mM Tris-HCl pH 7.4 was assessed using the Jasco V-630 spectrophotometer (Analytical Solutions). The data was collected at 20°C and are an average of three accumulations. All the data collected were buffer corrected by subtracting the data of the blank solution from the protein data. A graph of absorbance against wavelength was plotted.

2.2.10 Structural characterisation

2.2.10.1 Far-UV circular dichroism

Circular dichroism (CD) spectroscopy is used to analyse the secondary structure of the protein by using optically active molecules to measure the difference in absorbance of left and right circularly polarised light (Creighton, 1997). CD spectroscopy provides information about the environments of the chromophores of the aromatic amino acids as well as the contributions from disulfide bonds (Kelly and Price, 1997). Aromatic amino acids and disulfide groups have characteristic absorption bands in the near-UV range (250-300 nm).

The protein backbone absorbs strongly in the far-UV region (170-250 nm) which leads to characteristic secondary structure spectra (Woody, 1995). This wavelength range thereby gives a good indication of the secondary structural content of proteins such as α -helices and β -sheets. Proteins that have high α -helical content display characteristic minima at 208 and 222 nm and a strong peak at 190 nm whereas β -sheets give one trough near 217 nm and a peak in the 195-200 nm range (Woody, 1995).

Secondary structure content of FL-heEF1 β , CT-heEF1 β and NT-heEF1 β were assessed using far-UV CD. Measurements were performed using Jasco J1500 spectropolarimeter (Jasco, Tokyo, Japan) at 20°C in the wavelength range of 190–250 nm. Stock solutions of the proteins FL-heEF1 β , CT-heEF1 β and NT-heEF1 β were dialysed against buffer J (0.1 M Tris-HCl pH 7.4) which was filtered using a 0.2 μ m filter to remove particles which can affect the polarisation of light. The far-UV CD spectra of the proteins were collected five times using a 2 mm pathlength quartz cuvette with a data pitch of 0.2 nm, scanning speed of 100 nm.min⁻¹, band width of 0.5 nm and 1 sec response time. All spectra were buffer corrected by subtracting the spectra of the blank solution (0.1 M Tris-HCl pH 7.4) from the protein spectra. The raw CD data was converted to mean residue ellipticity $[\theta]$ using the following formula.

$$[\theta]_{MRE} = \frac{100\theta}{cnl} \quad (2.3)$$

Where θ is the CD milli degree, c is the protein concentration, n is the number of residues and l is the pathlength in cm. The quantity of secondary structure of the proteins was estimated using Dichroweb algorithm. The raw far-UV CD data was submitted to the server and deconvoluted using the CONTILL algorithm implemented in the Dichroweb server (Whitmore and Wallace, 2004).

2.2.10.2 Fluorescence spectroscopy

The naturally occurring fluorophores in proteins are tryptophan, tyrosine and phenylalanine. Fluorescence occurs when a molecule is excited from the ground state and returns to the ground state as emission of light at a longer wavelength (Lakowicz, 1983). Fluorescence can be used to analyse the tertiary structure of proteins particularly changes in the local environment of innate fluorophores (Joseph and Lakowicz, 1999). Phenylalanine has a small quantum yield therefore its emission is not noticeable while that of tyrosine is slightly higher than that of phenylalanine, Trp residues are roughly five times more sensitive than Tyr, mainly because Trp has a molar extinction coefficient of $5.5 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$ at 280 nm which is greater than the extinction coefficient for Tyr at 274 nm (Eftink, 1995). The indole ring of tryptophan is highly sensitive to solvent polarity (Lakowicz and Masters, 2008). Hence, fluorescence in the near-UV range is particularly sensitive to the environment of tryptophan residues. It depends on how many tryptophan residues and the region of the protein (buried or on the surface) where they are located (Creighton, 1997). A blue-shifted spectrum indicates that the tryptophan residue is buried in an environment which is nonpolar, while a red shifted spectrum indicates that the tryptophan residue is exposed in an environment which is polar. The tryptophan's indole ring is highly sensitive to solvent polarity (Joseph and Lakowicz, 1999). Emission spectra of this residue reflect the polarity of its surrounding environment.

ANS-based extrinsic fluorescence

ANS (8-anilino-1-naphthalene sulfonate) is an hydrophobic dye used as an extrinsic fluorescence probe (Engelhard and Evans, 1995). The hydrophobic pockets in the proteins were assessed by extrinsic ANS fluorescence. ANS fluorescence is quenched in an aqueous or polar environment, but when it binds to a hydrophobic surface there will be an increase in the fluorescence quantum

yield and the maximum emission wavelength becomes blue-shifted (Gasyimov and Glasgow, 2007). Free ANS excited at 390 nm emits at 540 nm. However, when ANS binds to exposed hydrophobic sites on a protein, the emission is lowered to around 470 nm. The change in emission wavelength depends on the quantum yield of ANS and the hydrophobicity of the ANS binding sites available on the protein (Stryer, 1965).

A stock solution of ANS was prepared in buffer K [50 mM Tris-HCl pH 7.4, 0.5 M NaCl and 0.02% (w/v) NaN₃] away from light in accordance with the general procedure for spectrophotometrically determining the concentration of ANS using extinction coefficient (ϵ) of 5000 M⁻¹cm⁻¹ at 350 nm. Protein (5 μ M) was incubated with 200 μ M of freshly prepared ANS away from light for 60 min, to achieve equilibrium. A series of blanks were generated, each containing 200 μ M ANS and treated in similar manner as the ANS-Protein samples. The samples were analysed using Jasco FP-6300 spectrofluorimeter with a 10 mm pathlength cuvette and 200 nm.min⁻¹ scan speed. Samples were excited at 380 nm with a slit width of 5 nm and emission spectra (average of 5 scans) were recorded from 390 to 600 nm. Spectra were produced from an average of three accumulations. The spectra were recorded at 20°C, buffer corrected, and are an average of three accumulations at a scan speed of 200 nm.min⁻¹.

ANS-protein binding curve was determined using varying concentrations of ANS (0-200 μ M) titrated against fixed concentration (5 μ M) of each protein construct. The fluorescence emission intensity at 465 nm were extracted and plotted against the ANS concentration and a single site ligand-binding curve was fitted based on the equation:

$$F_{cor} = \frac{F_{max} [Protein]}{K_d + [protein]} \quad (2.4)$$

Where F_{cor} is the corrected fluorescence, F_{max} is the max fluorescence and K_d is the dissociation constant. Data fitting was done using Sigma plot v 13.0.

2.2.11 Size exclusion-high performance liquid chromatography

Size exclusion chromatography (SEC) is a technique involving the separation of molecules based on their sizes in solution. Larger particles are eluted faster. Size exclusion high performance liquid chromatography (SE-HPLC) has more advantages such as reusable column without repacking and regeneration, increased speed of analysis and good resolution among others (Tayyab *et al.*, 1991).

The dynamic volume and quaternary structure of the proteins (FL-heEF1 β , CT-heEF1 β , NT-heEF1 β , FL-heEF1 γ , CT-heEF1 γ and NT-heEF1 γ) were assessed using analytical SE-HPLC. The procedure was carried out on a LC Phenomenex HPLC column along with a Guard cartridge column. The column attached to a Shimadzu Prominence HPLC system (SPD20A) at a flow rate of 0.2-0.3 ml.min⁻¹ was pre-equilibrated with buffer L [50 mM Tris- HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1 M NaCl, pH 6.8] which has been filtered and degassed. After equilibration the standard gel filtration marker was injected to calibrate the column. The standard gel filtration molecular weight marker containing the following proteins was used: thyroglobulin (670 kDa), γ -globulin (154 kDa), ovalbumin (44 kDa), myoglobin (17 kDa) and vitamin B₁₂ (1.35 kDa). The quaternary structure characteristics were determined by injecting 20 μ l of each protein (5 μ M) onto the column and eluted isocratically with the buffer at a flow rate of 0.2-0.3 ml.min⁻¹ for FL-heEF1 β , CT-heEF1 β , NT-heEF1 β , NT-heEF1 γ and CT-heEF1 γ . The log of the molecular weight of the proteins was interpolated from the graph of molecular weights against retention time of the standards.

2.2.12 Protein-protein and protein-ligand interaction

Oxidised glutathione (GSSG) have been shown to be related to cellular oxidative stress as seen in some diseases such as tumorigenesis, multiple sclerosis (Jones *et al.*, 2000). The column was equilibrated with the buffer M [50 mM Tris- HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1M NaCl, 5 mM GSSG pH 6.8]. Each of the proteins (FL-heEF1 β , CT-heEF1 β , NT- heEF1 β , CT-heEF1 γ and NT- heEF1 γ) was incubated with 20 mM GSSG for 30 min at 20°C and then loaded onto the column. Protein-protein interaction between heEF1 β and heEF1 γ was also assessed. Equimolar amounts of FL-heEF1 β and CT-heEF1 γ , CT-heEF1 β and NT-heEF1 γ , NT- heEF1 β and CT-heEF1 γ , NT- heEF1 β and NT-heEF1 γ , FL-heEF1 β and NT-heEF1 γ were prepared and

incubated for 30 min at 20°C and then injected into the SE-HPLC column in the presence and absence of oxidised glutathione ligand.

CHAPTER 3

Results

3.1 Introduction

The biochemical and biophysical characterisation of target proteins are important elements in improving the success rates of their structural studies. Biophysical characterisation of heEF1 β involved the construction of three *E. coli* expression vector systems for recombinant expression of the FL-, NT- and the CT- regions of the protein using pET-28a for FL and NT and pET-11a plasmid for CT. The proteins were purified to homogeneity to prevent interference by contaminants. Quantitative and qualitative analysis of the protein was then carried out before the structural characterisations of the proteins.

3.2 Vector sequencing

The chromatograms received from Inqaba Biotec (Pretoria, South Africa) contained gene sequences, which were translated using the online server ExPASy Translate (Artimo *et al.*, 2012). The translated amino acid sequences were then compared with the sequence in the database using the basic alignment search tool (BLAST). The proteins sequences of both FL-heEF1 β and FL-heEF1 γ have very close sequence identity with that found in the database.

3.3 Protein expression and purification

The conditions for producing the maximum amount of soluble protein for all the proteins were found to be cold induction with a final concentration of 0.5 mM IPTG and expression at 30°C for 6 h post-induction (Figure 3.1). The soluble protein was partitioned into *E. coli* inclusion bodies on inducing at 37°C. Most of the recombinant proteins were found in the soluble cell lysate. The controls are cells cultured without IPTG and they did not show any detectable over-expressed protein.

All the proteins (FL-heEF1 β , NT-heEF1 β , CT-heEF1 β , NT-heEF1 γ and CT-heEF1 γ) were purified by IMAC and as shown on the SDS-PAGE electrophoretogram. IMAC purification of

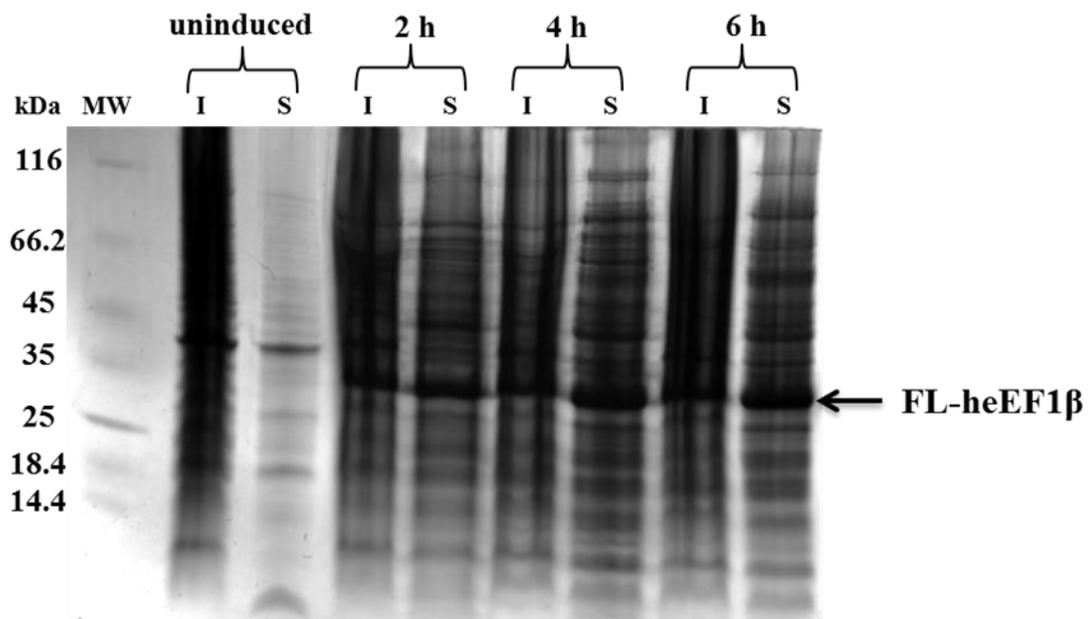


Figure 3.1: Tricine SDS-PAGE electrophoretogram showing production of FL-heEF1 β . Induction for optimal protein production, when exposed to 0.5 mM IPTG concentration for varying time periods (2 h, 4 h and 6 h) at 37°C. I represents insoluble and S represents soluble. The arrow points at the protein band between 25-35 kDa corresponding to the theoretical weight.

the protein was performed using HisTrap columns. The hexahistidine-tagged fusion protein binds to the nickel or cobalt ions which are covalently attached to the matrix and dilution was monitored at 280 nm (Figure 3.2). The hexahistidine-fusion protein has stronger affinity for the nickel or cobalt ions and outcompetes the low concentration of imidazole and protein contaminants. The proteins are eluted with a high concentration of imidazole (350 mM).

3.3.1 FL-heEF1 β

The FL-heEF1 β have more than one band as seen in the IMAC SDS-PAGE electrophoretogram, (Figure 3.3) which indicates the presence of contaminants hence the need for other purification methods. DEAE-agarose ion exchange chromatography was shown to be effective in purifying FL-heEF1 β proteins to homogeneity as seen in Figure 3.4.

3.3.2 NT-heEF1 β

The NT-heEF1 β protein also required further purification as observed from the IMAC SDS-PAGE electrophoretogram (Figure 3.5) there was more than one band. DEAE-agarose ion exchange chromatography was also used for second step purification (Figure 3.6).

3.3.3 CT-heEF1 β

The CT-heEF1 β protein required only one purification step which was IMAC (Figure 3.7). Only one band was seen in the electrophoretogram indicating that the protein is pure.

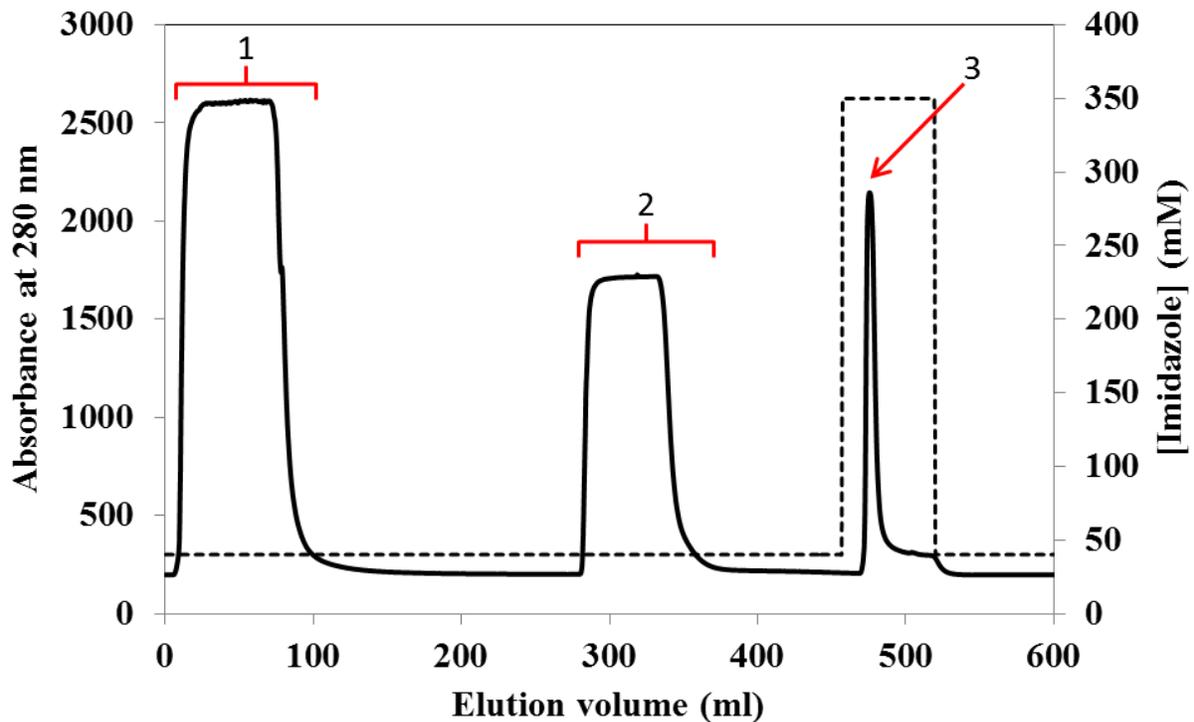


Figure 3.2: IMAC purification profile of FL-heEF1 β . The solid line represents the absorbance at 280 nm while the dashed lines represent the imidazole concentration. Peak 1 is the protein injection, peak 2 is the Triton X-100 wash and peak 3 is the one-step imidazole elution. The soluble fraction subsequent to cell lysis was loaded onto a HisTrap nickel column pre-equilibrated with 50 mM Tris-HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1M NaCl, 40 mM imidazole pH 7.4. The histagged proteins bind to the column while the non-specific proteins pass through as their binding affinity is relatively weak as seen in peak 1. Absorbance increases again during the wash because of the detergent Triton X-100 present in the washing buffer, which is used to remove non-specific bound proteins or proteins involved in hydrophobic interactions with the column as seen in peak 2. Afterwards the equilibration buffer was used to remove excess Triton X-100 and the absorbance decreases again until it returned to baseline. A high concentration of imidazole (350 mM) displaces the proteins from binding to the column. The proteins elutes in a sharp peak as seen in peak 3.

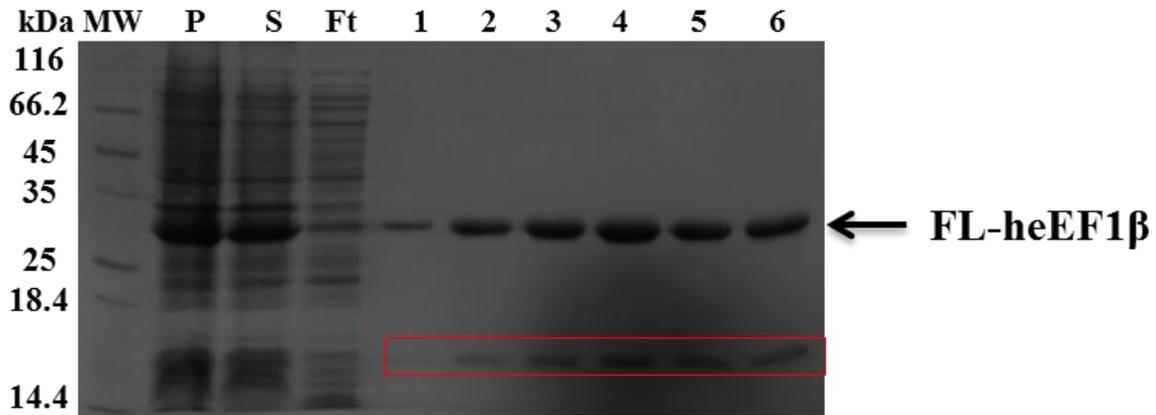


Figure 3.3: Tricine SDS-PAGE electrophoretogram showing the IMAC purified samples of FL-heEF1 β , where MW represents the molecular weight marker, P represents pellet, S represents the supernatant, Ft represents the flow through and 1-6 represents the eluted fractions of the protein collected at increased imidazole concentration (300 mM). The red outlined bands are impurities.

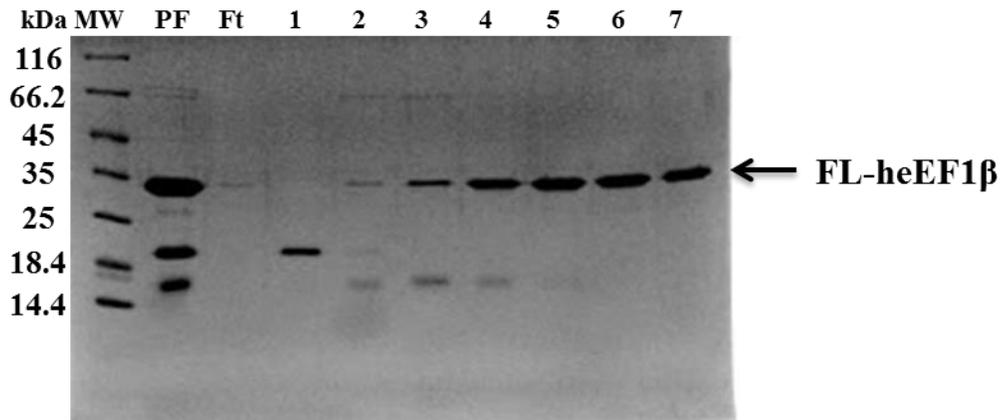


Figure 3.4: 12% Glycine SDS-PAGE electrophoretogram showing the DEAE-agarose ion exchange chromatography purified samples of FL-heEF1 β . IMAC purified samples were pooled together and dialysed (16 h, 4°C) against the dialysis buffer (50 mM Tris-HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1 mM EDTA, 1 mM EGTA, 500 μ l TCEP pH 7.4). Purified samples of FL-heEF1 β were eluted with 0.4 M NaCl. Where Mw is the molecular weight markers, PF represents the pooled fractions of IMAC purified FL-heEF1 β samples, Ft is the flow through while 1-7 represents the eluted protein fractions; fractions 1-4 still have some impurities while fractions 5-7 are pure samples of FL-heEF1 β .

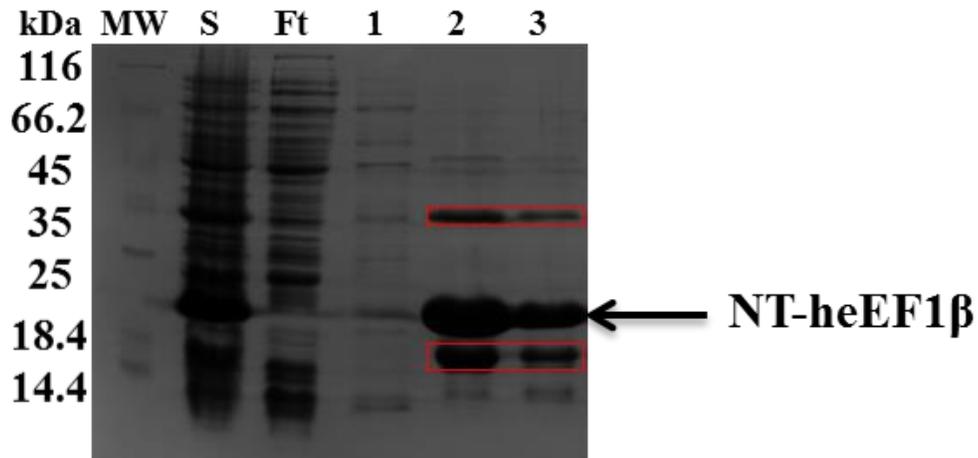


Figure 3.5: Tricine SDS-PAGE electrophoretogram used to assess the IMAC purified samples of NT-heEF1 β where MW represents the molecular weight marker, S represents the supernatant, Ft represents the flow through and 1-3 represents the eluted fractions of the protein collected at increased imidazole concentration (300 mM). The red highlighted bands and other smaller bands are contaminants, hence the need for further purification step.

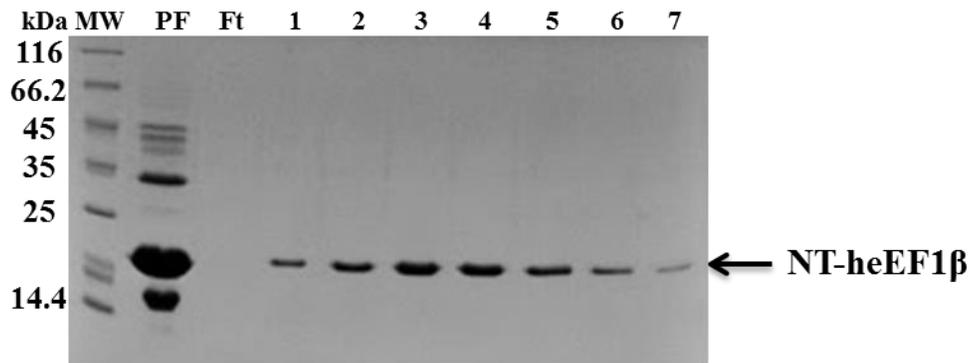


Figure 3.6: 12% Glycine SDS-PAGE electrophoretogram showing the DEAE-agarose ion exchange chromatography purified NT-heEF1 β samples. IMAC purified samples were pooled together and dialysed (16 h, 4°C) against the dialysis buffer (50 mM Tris-HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1 mM EDTA, 1 mM EGTA, 500 μ l TCEP pH 7.4). PF represents the IMAC purified samples pooled together and dialysed. FT is the flow through; it has no band indicating that the protein binds completely to DEAE-agarose. 1-7 represent the purified samples of NT-heEF1 β eluted with 0.4 M NaCl. From the gel, only one band is observed for all the fractions collected indicating that the protein is very pure.

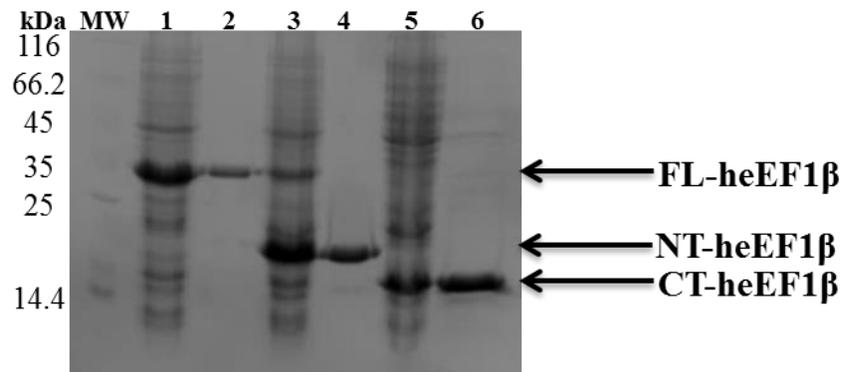


Figure 3.7: 12% Glycine SDS-PAGE electrophoretogram showing the heEF1 β proteins used in this study: FL-heEF1 β , NT-heEF1 β and CT- heEF1 β . MW is the molecular weight marker. Lane 1, lane 3 and lane 5 are the lysate samples of FL-heEF1 β , NT-heEF1 β and CT- heEF1 β respectively while lane 2, lane 4 and lane 6 are the purified samples of FL-heEF1 β , NT-heEF1 β and CT- heEF1 β respectively.

3.3.4 NT-heEF1 γ and CT-heEF1 γ

The NT-heEF1 γ was purified by IMAC after which the eluted protein fractions were pooled together and further purified by GSH-Agarose chromatography (Figure 3.8) while the CT-heEF1 γ protein was purified to homogeneity by IMAC alone (Figure 3.9).

3.4 Quantitative and qualitative evaluation of the purified proteins

A graph of the Log of the molecular weight of the standards against distance migrated on the gel (cm) in Figure 3.4 was plotted, and used to determine the molecular weight of the protein (FL-heEF1 β) as seen in Figure 3.10. The equation of the line from the graph;

$$y = -0.1219x + 2.1796 \quad (3.1)$$

where x is the distance (cm) migrated by the protein on the gel and y is the Log of molecular weight of the protein. From the electrophoretogram (Figure 3.4) the distance migrated by the protein was substituted into equation 3.1. Therefore the molecular weight of the protein was computed to be approximately 28 kDa.

About 7.0 g of wet cells was obtained in 1 liter culture for FL-heEF1 β , while 10.4 g and 9.5 g of wet cells were obtained for CT-heEF1 β and NT-heEF1 β respectively. FL-heEF1 β had a reasonable level of soluble expression which yielded approximately 28.5 mg of protein while CT-heEF1 β and NT-heEF1 β yielded approximately 71 mg and 65 mg of protein respectively. Protein concentration was determined by serial dilution. A graph of absorbance (A_{280} - A_{340}) against dilution factor was plotted for the FL-heEF1 β protein. Slope was derived from the equation of the line (Figure 3.11). Equation of the line was derived to be:

$$y = 4.4663x + 0.0019 \quad (3.2)$$

Slope was derived to be 4.4663 and substituted in equation 3.2 above

$$\text{Conc (mg/ml)} = \frac{4.4663 \times 67000}{2 \times 29910} = 5.0024 \text{ mg/ml}$$

$$\text{Conc (mol/L)} = \frac{\text{reacting mass}}{\text{molar mass}} = \frac{5.0024}{67000} = 0.000075 \text{ mol/L} \sim 75 \mu\text{M}$$

Therefore concentration of FL-heEF1 β was computed to be 75 μM .

The protein quality was determined using absorbance spectrometry, which is a graph of absorbance against wavelengths 260-340 nm to ensure that there was no DNA contamination or protein aggregation (Figure 3.12). Following protein expression, purification, quantity and quality evaluation, the identity of the proteins was confirmed by peptide sequencing mass spectrometry in CSIR (Pretoria, South Africa) to be heEF1 β and heEF1 γ (Figure 3.13). The results showed that all the peptides were identified with >95% confidence.

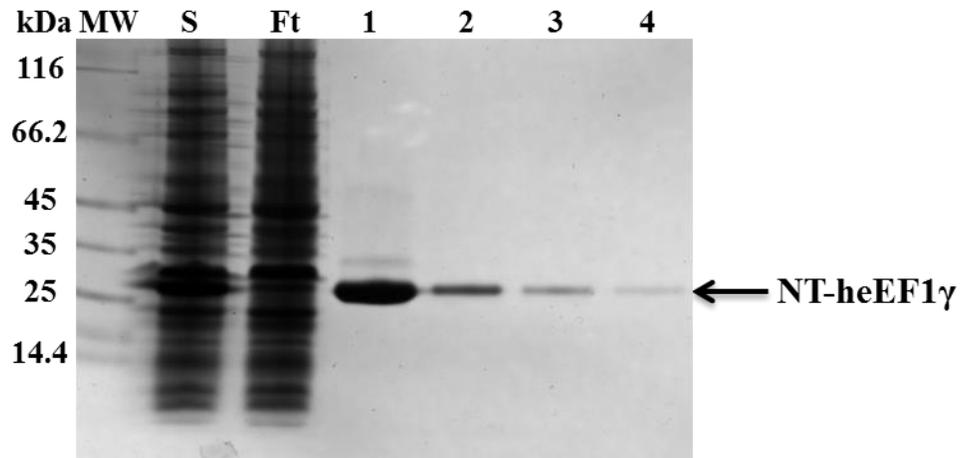


Figure 3.8: Tricine SDS-PAGE electrophoretogram showing the eluted pure fractions of NT-heEF1 γ . S represents the supernatant, FT represents the flow through and 1,2,3 and 4 represents the eluted pure fractions of NT-heEF1 γ . The arrow points to the band of the protein which falls between 25 and 35 kDa.

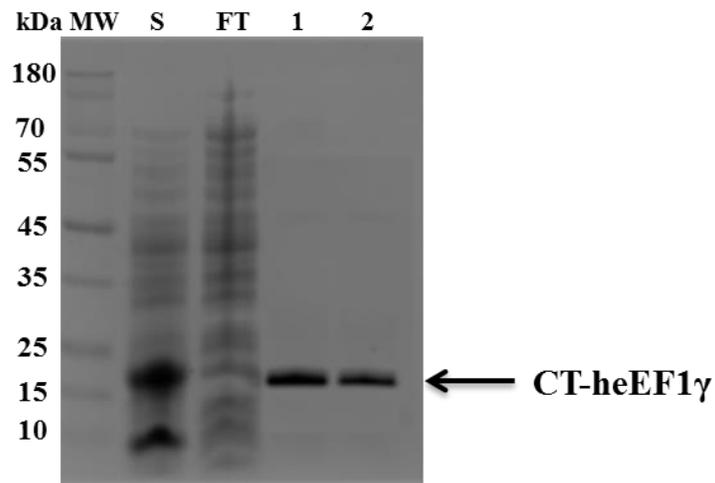


Figure 3.9: Tricine SDS-PAGE electrophoretogram showing the eluted pure fractions of CT-heEF1 γ . S represents the supernatant, FT represents the flow through and 1 and 2 represent the eluted pure fractions of CT-heEF1 γ . The arrow points to the protein band.

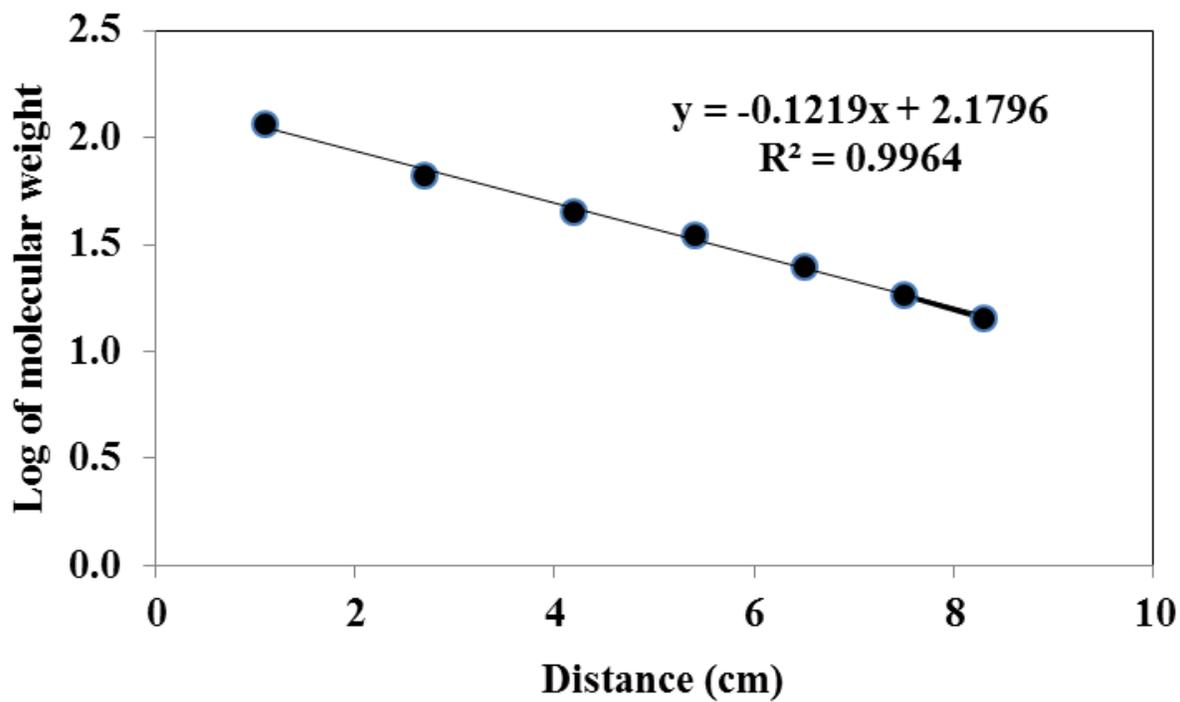


Figure 3.10: Graph of Log of molecular weight against distance migrated by the protein on the gel in Figure 3.4. The equation of the line was used to compute the molecular weight of the protein, by substituting the value of the distance migrated on the gel (cm) into the equation.

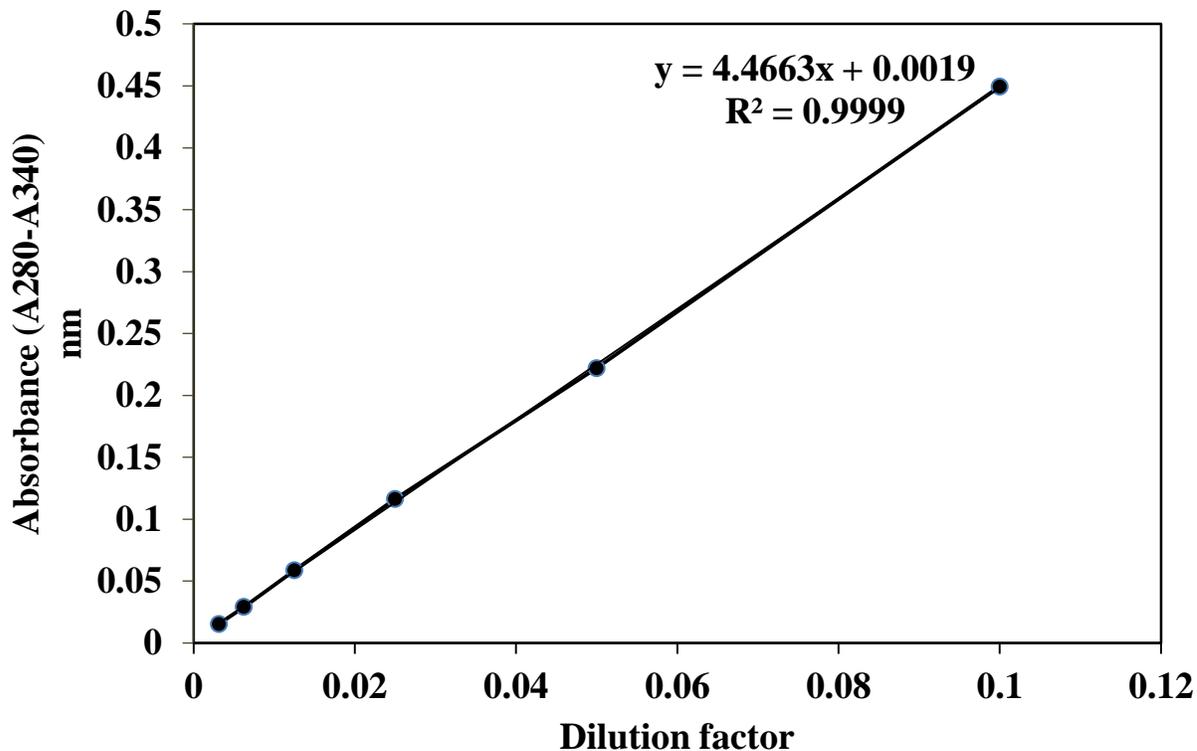


Figure 3.11: A graph of absorbance against dilution factor. The value of R^2 was derived to be 0.9999~1 and the slope which is 4.4663 was used to determine the concentration of the protein FL-heEF1 β . Six dilutions of the proteins was prepared by adding 100 μ l of the protein to 900 μ l of buffer and mixed thoroughly. 500 μ l of the sample was taken and added to 500 μ l of buffer and was done serially in five Eppendorf™ tubes. Absorbance at 340 nm was subtracted from the absorbance at 280 nm to remove the interferences by noise or light.

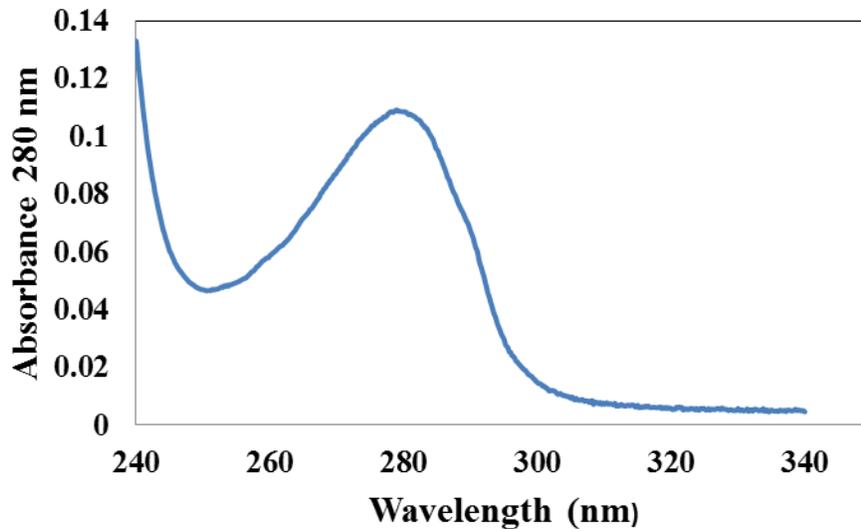


Figure 3.12: A graph of absorbance at 280 nm against wavelength for FL-heEF1 β protein. No DNA contamination was observed because there was no peak formed at 260 nm which is the wavelength at which DNA absorbs UV light. There was no protein aggregation because there was no peak formed at 340 nm.

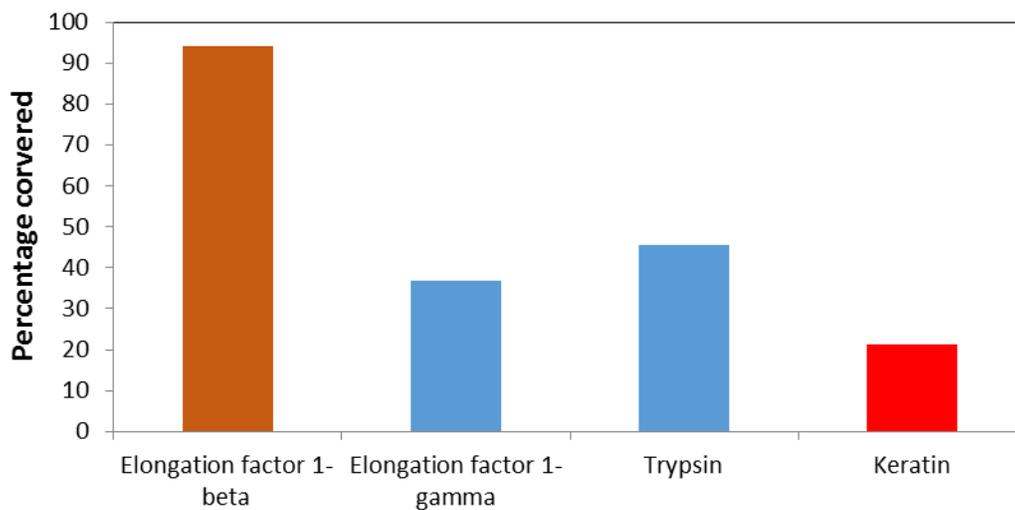


Figure 3.13: A chart showing the mass spectrometry results of the samples sent to CSIR Pretoria. More than 40% of the peptides are heEF1 γ and more than 90% are heEF1 β . All peptides were identified with >95% confidence. The keratin impurities could be from the skin while the trypsin was used to breakdown the proteins to smaller peptides.

3.5 Structural characterisation

Far-UV circular dichroism spectroscopy was used to analyse the secondary structure of the protein. Fluorescence spectroscopy was used to provide information on the tertiary structure by looking at changes to the local environment of tryptophan residues. Analytical size exclusion high pressure liquid chromatography (SE-HPLC) was used to determine the quaternary structure of the protein and to assess the protein-protein interactions.

3.5.1 Secondary Structure characterisation by far-UV CD

Information on the secondary structure of the FL-heEF1 β , NT-heEF1 β and CT-heEF1 β was obtained using far-UV CD at pH 7.4 and 20°C. Far-UV CD spectra of the proteins denatured in 8 M urea were also collected to make certain that the spectra for the native protein are distinguished. The far-UV CD spectra for all the three proteins were recorded over a far-UV CD wavelength range of 190 nm to 250 nm. The results from the Figure 3.14 indicate that the heEF1 β protein is rich in alpha helices because of the negative peak at 208 nm. The FL-heEF1 β and NT-heEF1 β display minima at 208 and 222 nm, and a peak at 190 and 195 nm respectively, which is typical of proteins with predominant alpha helical content. The spectra data CT-heEF1 β has only one negative peak at 208 nm and a peak at 190 nm, which indicates a mix of alpha helix and β sheets. The CD spectra data were deconvoluted using the CONTILL algorithm in Dichroweb as seen in Table 2. The data suggest that FL-heEF1 β and NT-heEF1 β are predominantly α -helical, which corresponds to the spectra profile. The data for CT-heEF1 β is not conclusive because the rmsd is above 0.1 (Table 2).

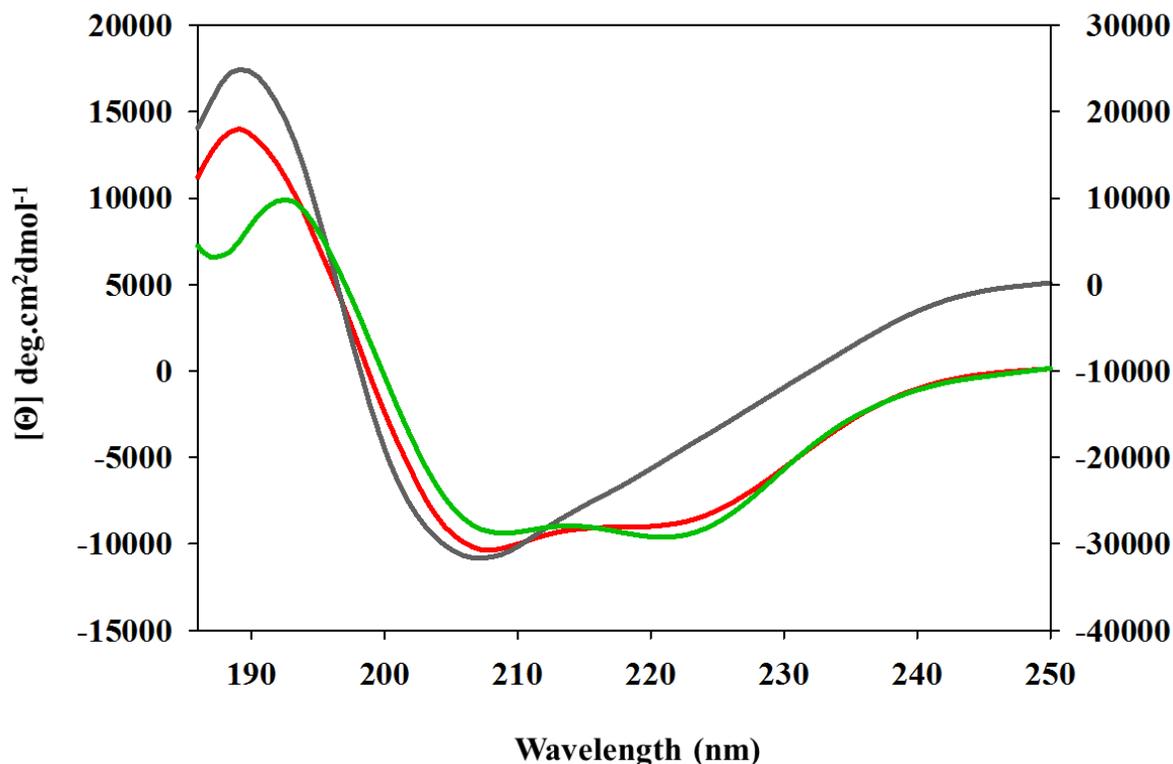


Figure 3.14: Far-UV CD spectra of FL-heEF1 β (red), NT-heEF1 β (green) and CT-heEF1 β (grey). Spectra were collected using samples of 2 μ M protein in 0.1 mM Tris-HCl pH 7.4. The similar spectra (red and green) indicate that both proteins have similar secondary structure which is predominantly α -helix.

Table 2: Summary of the secondary structure content calculated using the CONTILL algorithm implemented in the Dichroweb server. Normalised root-mean-square deviation (NRMSD) indicates the goodness of fit of calculated data to experimental data, with <0.1 acceptable as good fit.

Protein construct	Secondary structure content (%)				rmsd
	α helix	β strand	β turns	Unordered	
FL-heEF1 β	56.1	21.0	23.0	7.9	0.04
NT-heEF1 β	52.4	9.2	36.3	0.2	0.02
CT-heEF1 β	38.5	27.9	14.9	17.2	0.12

3.5.2 Tertiary structure characterisation by extrinsic tryptophan ANS fluorescence

ANS (8-anilino-1-naphthalene sulfonate) is an amphipathic dye used as an extrinsic fluorescence probe because it binds to accessible hydrophobic pockets in a protein (Gasymov and Glasgow, 2007). Increase in quantum yield and simultaneous blue shift in the maximum emission wavelength (λ_{max}) of ANS fluorescence from 510 nm (free ANS in a polar environment) to 480 nm indicates that the protein binds to ANS. The result from Figure 3.15 shows that both FL-heEF1 β and CT-heEF1 β have accessible binding pockets because of the increase in quantum yield and change in the λ_{max} of the fluorescence spectra from 510 to 480 nm. There was a slight increase in quantum yield for NT-heEF1 β , but it does not have any binding pocket accessible to ANS because the λ_{max} remained at 510 nm. The apparent increase in quantum yield could be due to the interaction between the dye and the cluster of highly charged surface available amino acids in the protein. To further quantify heEF1 β -ANS binding, a fixed amount of heEF1 β was titrated with increasing amounts of ANS and the resulting curve was fitted using a single or multiple site binding models implemented on Sigma v 12 (Figure 3.16). The results show that the binding of ANS to the accessible hydrophobic pocket in heEF1 β follows a single dose dependent hyperbolic profile. ANS concentration tends to saturate at concentration $>200 \mu\text{M}$. The fit statistics satisfies a single site binding, yielding a K_d of $\sim 70.5 \mu\text{M}$ of ANS.

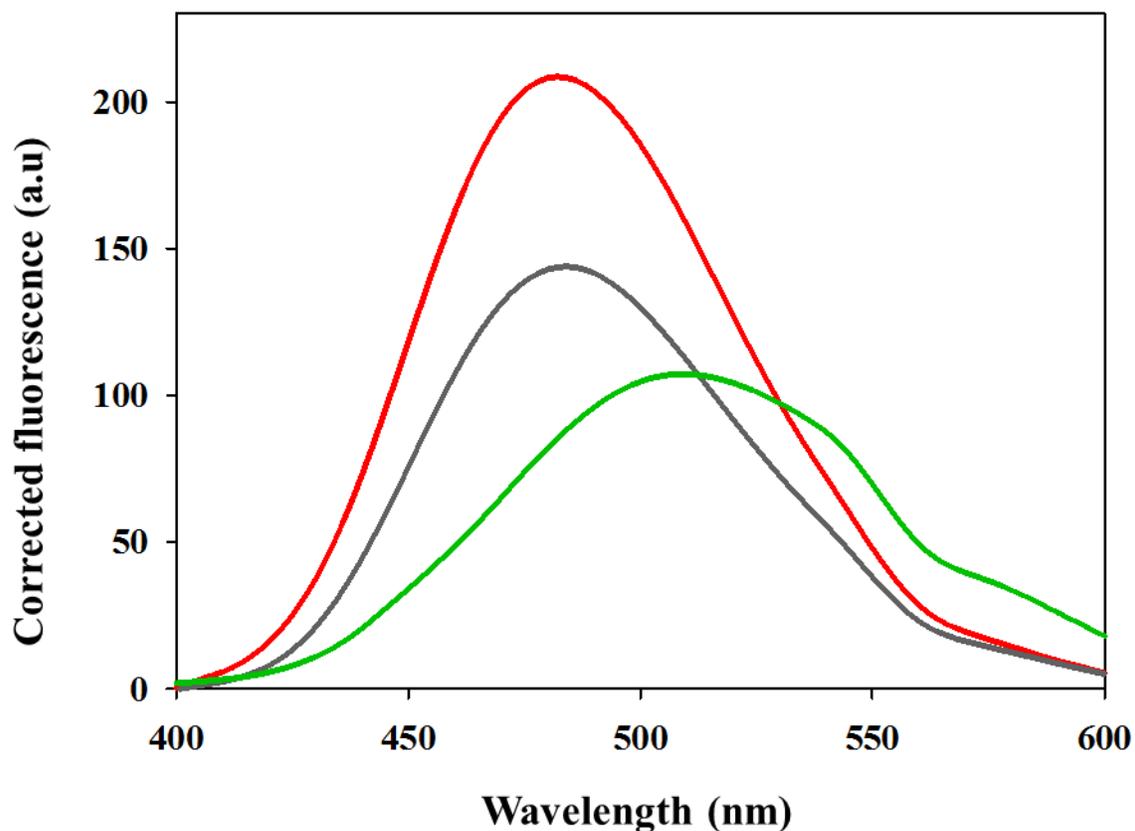


Figure 3.15: Extrinsic ANS Fluorescence emission spectra. Spectra of 200 μM ANS bound to 2 μM FL-heEF1 β (red), CT-heEF1 β (grey) and NT-heEF1 β (green). ANS was selectively excited at 390 nm and the spectra recorded over the 400-600 nm range using an excitation and emission bandwidth of 5 nm. Each spectrum is the average of three accumulations of three replicate samples in 50 mM Tris-HCl pH 7.4 containing 0.5 M NaCl and 0.02 % (w/v) NaN_3 . The spectra of protein bound to ANS were corrected for the fluorescence contribution from free unbound ANS.

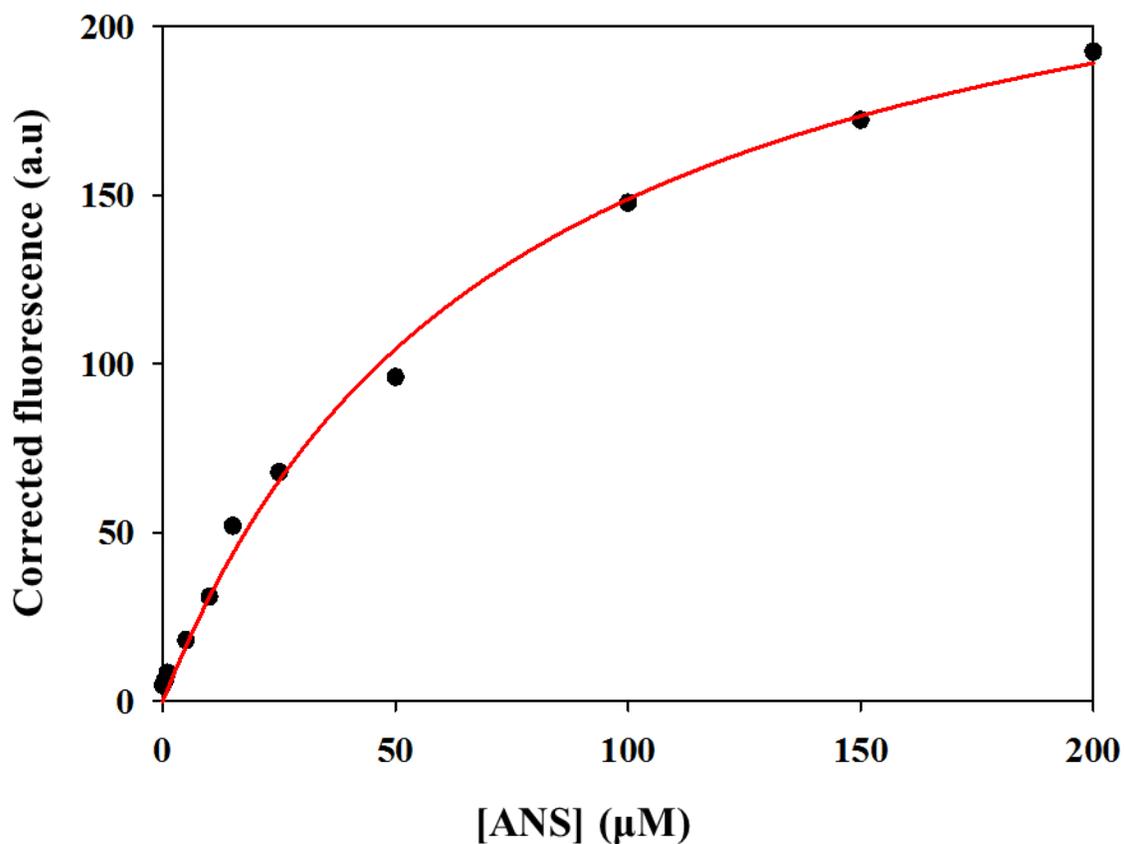


Figure 3.16: Fluorescence signal of FL-heEF1 β as a function of ANS concentration. 2 μM FL-heEF1 β was titrated with increasing ANS concentration excited at 295 nm and emission at 480-510 nm. Each data point represents an average of three replicate experiments. The data was fitted to a hyperbolic function using Sigma Plot v 12. K_d for ANS binding to the FL-heEF1 β is 70.5 μM

3.5.3. Quaternary structure characterisation by SE-HPLC

Analytical size exclusion high performance liquid chromatography (SE-HPLC) was used to determine the oligomeric state of heEF1 β . The gel filtration standards were used to calibrate the column and a graph of the Log of molecular weights against retention time was plotted (Figure 3.17). From the individual chromatogram of the proteins as seen in Figure 3.18 the retention times were used to interpolate the molecular weight of the protein or by using the equation of line. The FL-heEF1 β eluted with a double peak at 62 kDa and 25 kDa corresponding to a possible homodimer and monomer, with the dimer being the predominant quaternary structure at the condition used for the analysis. CT-heEF1 β eluted at a single peak at 14 kDa indicating a monomeric state. NT-heEF1 β which is approximately 15 kDa in its monomeric form, eluted at a peak approximately 30 kDa which is indicative of a homodimer as seen in Figure 3.18b. From the overall results it indicates that there was no high order oligomeric states of heEF1 β constructs and that the observed masses are comparable to the theoretically predicted monomeric masses of each construct. The NT-heEF1 γ eluted at a peak approximately 23 kDa which is a monomer and as a dimer at 46 kDa in the presence of GSSG. CT-heEF1 γ eluted at 19 kDa which is predominantly monomeric (Figure 3.19).

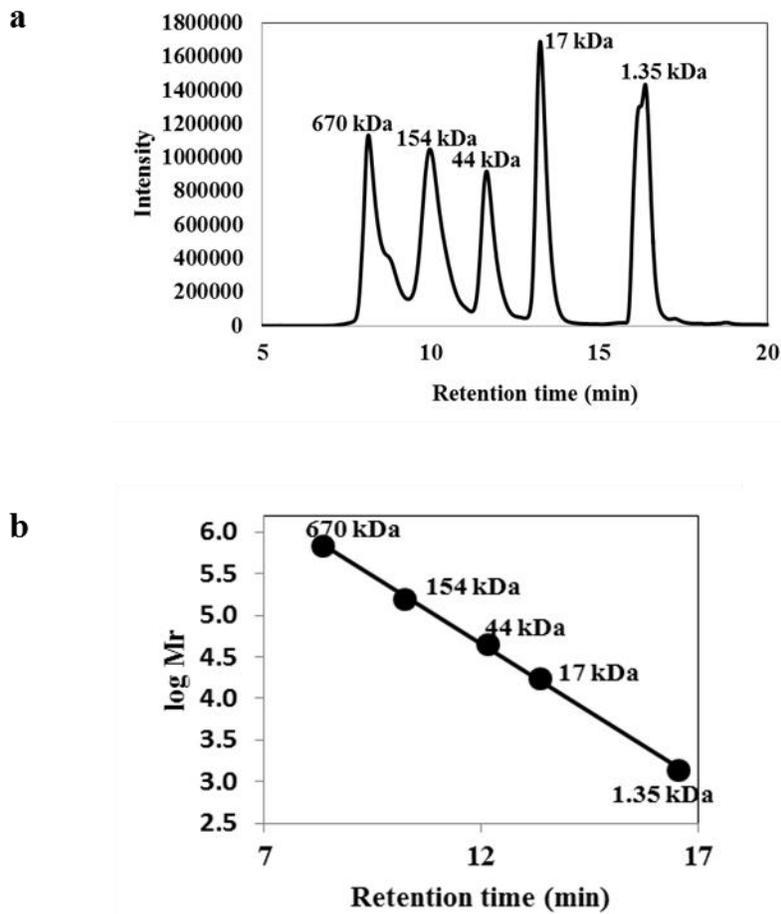


Figure 3.17: (a) Chromatogram of the gel filtration standards resolved by SE-HPLC at a flow rate of 0.2-0.3 ml.min⁻¹ at 20 °C. The standards were used to calibrate the column. Standards with larger weights are eluted first. (b) Graph of Log of molecular weight of gel filtration standards against retention time (min) was used to determine the molecular weight of the desired protein. 670 kDa-Thyroglobulin (bovine), 154 kDa- γ -globulin, 44 kDa-Ovalbumin, 17 kDa-Myoglobin, 1.35 kDa-Vitamin B₁₂. The molecular weight of the protein can be interpolated from the line graph, if the retention time is known.

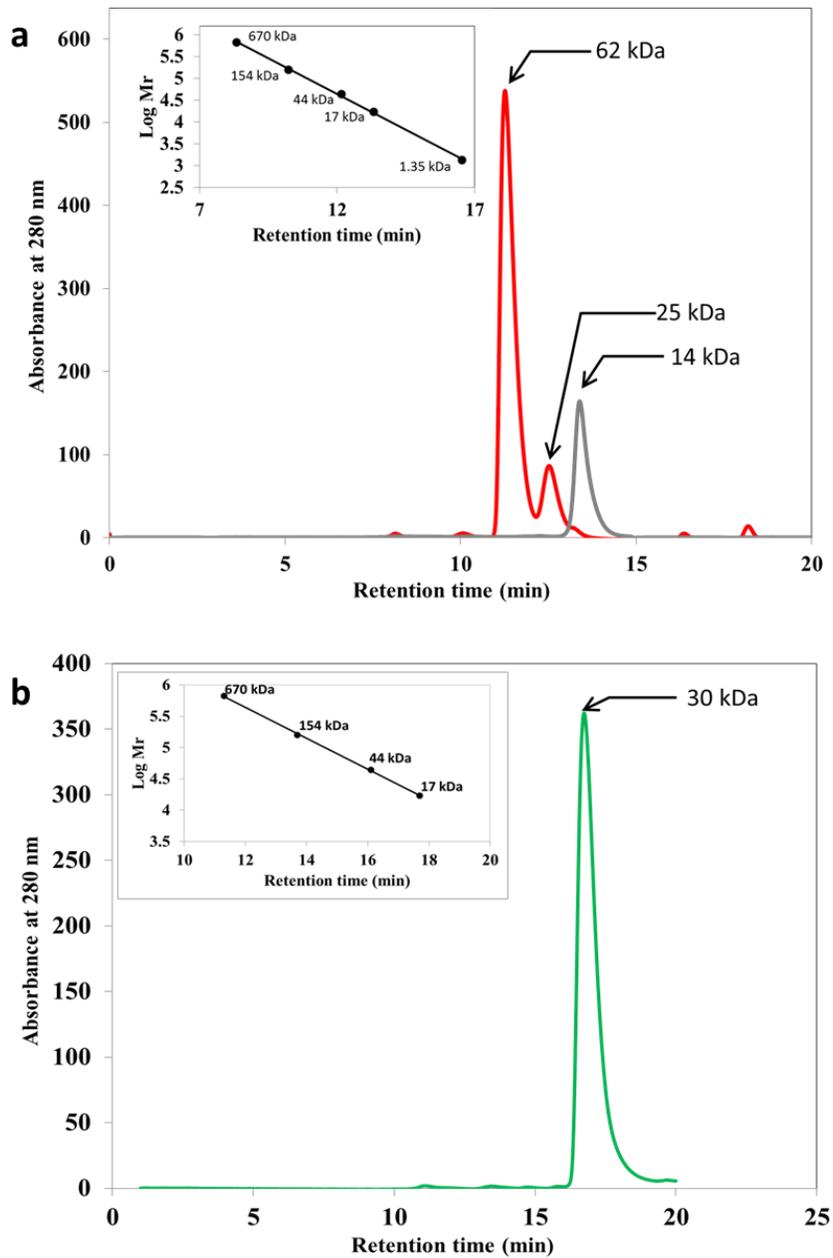


Figure 3.18: Chromatograms showing the quaternary structure characterisation of heEF1 β resolved by SE-HPLC. (a) Chromatogram of FL-heEF1 β (red) and CT-heEF1 β (grey) showing that FL-heEF1 β is dimeric at 62 kDa and monomeric at 25 kDa which implies that it is predominantly dimeric, while CT-heEF1 β is monomeric at 14 kDa. (b) Chromatogram of NT-heEF1 β showing that the protein is dimeric.

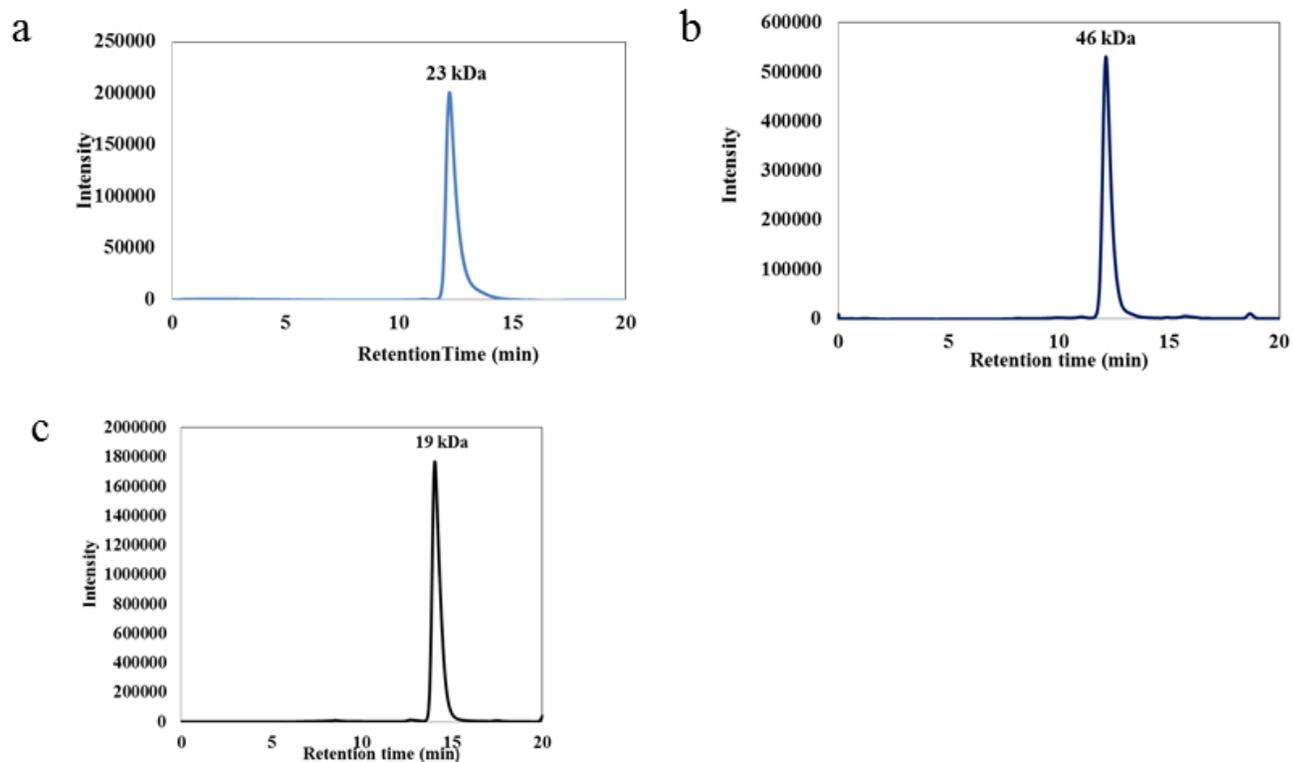


Figure 3.19: Individual chromatograms of the heEF1 γ proteins using SE-HPLC. (a) Chromatogram of NT-heEF1 γ , the protein is monomeric at 23 kDa (b) Chromatogram of NT-heEF1 γ in the presence of GSSG ligand. The protein is dimeric at 46 kDa (c) Chromatogram of CT-heEF1 γ which is also predominantly monomeric at 19 kDa.

3.6 Functional characterisation by protein-protein interaction

Analytical SE-HPLC was also used to characterise the protein-protein interaction between heEF1 β and heEF1 γ . The heEF1 β protein does not interact with 19 kDa CT-domain of heEF1 γ because the peaks in the chromatogram corresponds to the dimeric form of heEF1 β (64 kDa), the monomeric form of heEF1 β (23 kDa) and the monomeric form of CT-heEF1 γ (19 kDa) as seen in Figure 3.20. There is no interaction between CT-heEF1 β and NT-heEF1 γ because the peaks of the mixture corresponds with the individual peaks for the monomeric NT-heEF1 γ (23 kDa) and CT-heEF1 β (14 kDa) as observed in Figure 3.20b. The heEF1 β interacts with NT-heEF1 γ and this interaction takes place at the N-terminus domain of both proteins. The complex between dimeric heEF1 β and monomeric NT-heEF1 γ is approximately 195 kDa, which is ~44% larger than the predicted mass of 110 kDa, assuming FL-heEF1 β -NT-heEF1 γ complex assumes $\gamma:\beta\beta:\gamma$ (monomeric NT-heEF1 γ and dimeric FL- heEF1 β) conformation. The interaction between FL-heEF1 β and NT-heEF1 γ is not affected by the dimerisation of NT-heEF1 γ (46 kDa) in the presence of oxidised glutathione (GSSG) as seen in Figure 3.20d. However the sizes of the complex increases to 230 kDa which is ~36% larger than the predicted size of 156 kDa if the complex between FL-heEF1 β and NT-heEF1 γ assumes $\gamma\gamma:\beta\beta:\gamma\gamma$ (dimeric FL- heEF1 β and NT-heEF1 γ) conformation. No higher oligomeric state was observed in the interaction between FL-heEF1 β and NT-heEF1 γ .

When NT-heEF1 β was incubated with NT-heEF1 γ for complex formation, the mixture resolved in three distinct peaks which are 23 kDa, 129 kDa and >670 kDa corresponding to the monomeric NT-heEF1 γ , a possible $\gamma:\beta\beta:\gamma$ (monomeric NT-heEF1 γ and dimeric NT-heEF1 β) conformation and high order oligomers, respectively as seen in Figure 3.21. The 129 kDa peak is 32% greater than the expected 76 kDa size for $\gamma:\beta\beta:\gamma$ conformation.

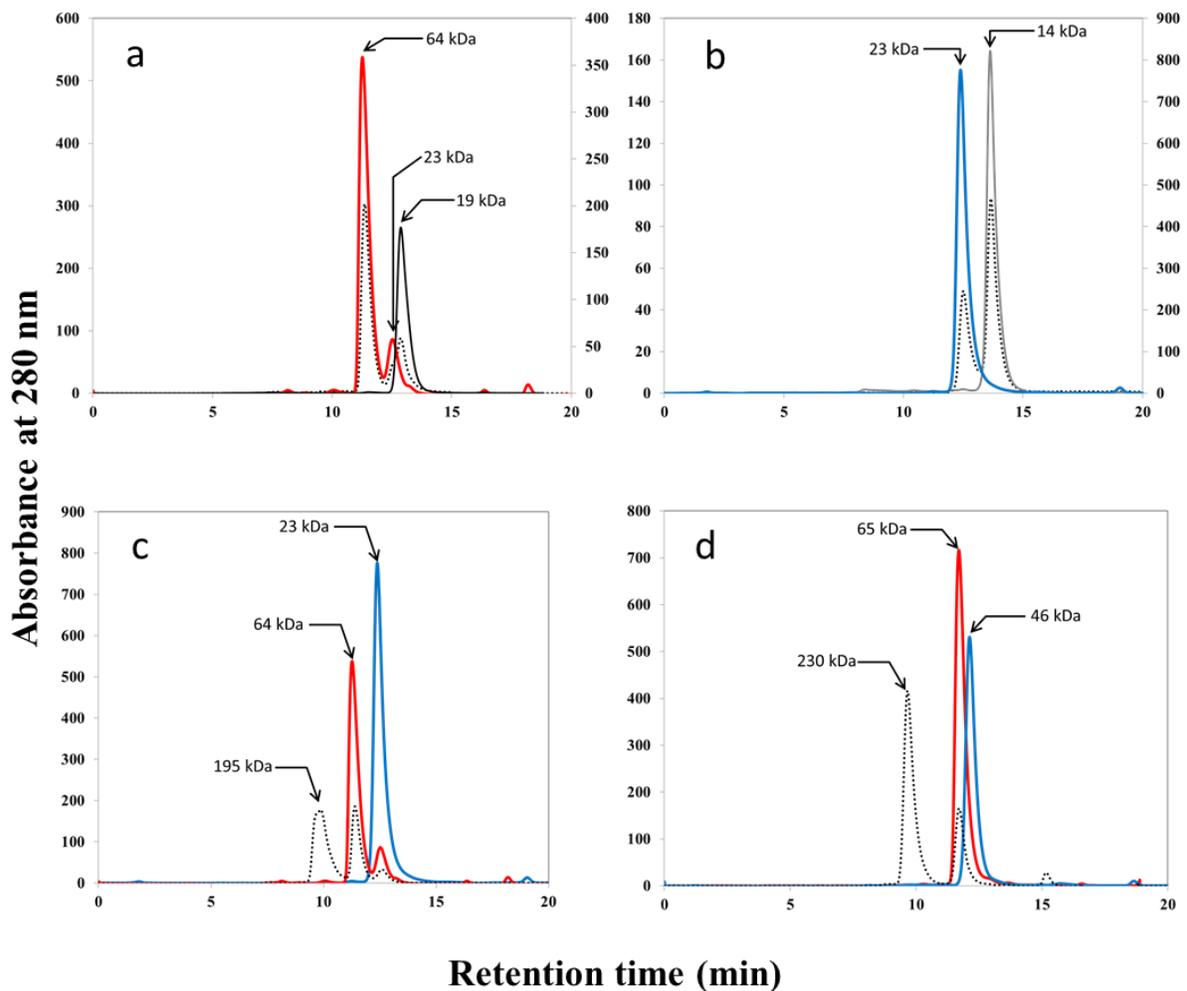


Figure 3.20: Protein-protein interaction chromatograms resolved by SE-HPLC. (a) FL-heEF1 β (red) and CT-heEF1 γ (black) and the dotted line indicating no interaction because the peaks correspond to the individual proteins (FL-heEF1 β and CT-heEF1 γ). (b) NT-heEF1 γ (blue) and CT-heEF1 β (grey) and the dotted line indicating no interaction because the peaks correspond to the individual proteins (NT-heEF1 γ and CT-heEF1 β). (c) FL-heEF1 β and NT-heEF1 γ interaction and the dotted line indicating the interaction resolved at 195 kDa, while the other peaks correspond to the individual protein. (d) FL-heEF1 β and NT-heEF1 γ interaction in the presence of GSSG. The dotted line indicates the interaction resolved at 230 kDa while the other peak corresponds to the FL-heEF1 β .

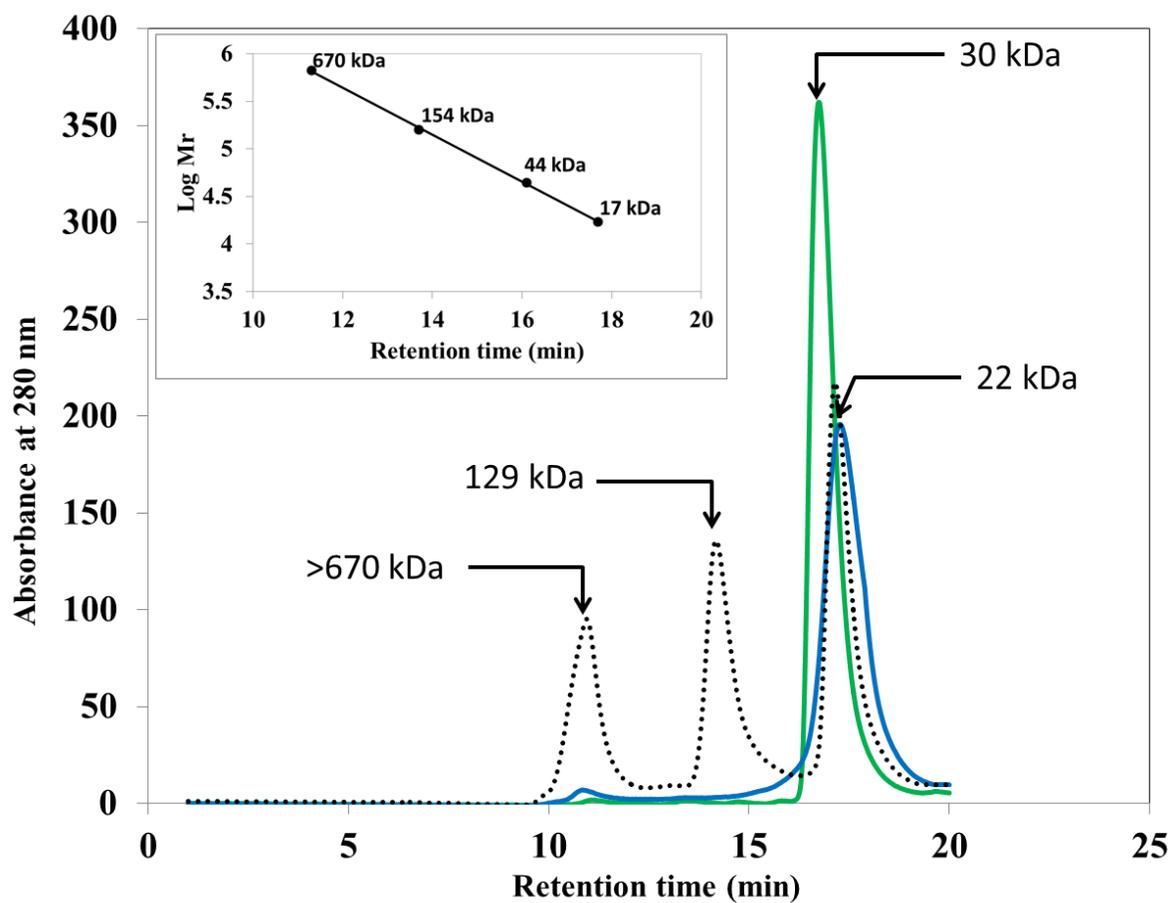


Figure 3.21: Protein-protein interaction (NT-heEF1 β and NT-heEF1 γ) resolved by SE-HPLC. The dotted lines represent the interaction profile of both proteins which resolved in three different peaks: 22 kDa (monomeric NT-heEF1 γ), 129 kDa (monomeric NT-heEF1 γ and dimeric NT-heEF1 β) and high order oligomer > 670 kDa. The green line is the NT-heEF1 β which is dimeric while the blue line is the NT-heEF1 γ respectively.

CHAPTER 4

Discussion and conclusion

4.1 Discussion

Codon harmonisation of the gene encoding the proteins was very vital for the expression of soluble proteins. Heterologous protein expression could lead to the formation of insoluble aggregates and low expression of soluble proteins due to differences in identical codon usage between the expression and natural host. The plasmids pET-28a and pET-11a were used to clone the genes containing the proteins in *E. coli* JM109 cells and are ampicillin and kanamycin resistant respectively. JM109 cells are not expression cells hence the need to extract the plasmid DNA from it. They both have N-terminal histagged thrombin cleavage which had no effect on the expression and purification of the proteins. The thrombin cleavage site on the NT-domain was to ensure that there were minimal changes in the amino acid composition of the protein after purification. In this study the thrombin was not cleaved and it had no effect on the structure and function of the proteins. Site-directed mutagenesis was used to create the NT-heEF1 β because it was cheaper purchasing the plasmid. A stop codon was encoded at lys 79 because it is easier to create a stop codon at the lysine residue by changing AAA \rightarrow TAA.

Expression of pET-28 plasmid vector which is kanamycin resistant with FL-heEF1 β and NT-heEF1 β and pET-11a vector which is ampicillin resistant with CT-heEF1 β insert in *E. coli* BL21 (DE3) yielded reasonable levels of proteins which could be due to the codon harmonisation of the gene and using certain expression conditions such as cold induction, 0.5 mM IPTG and expression at 30°C for 6 h post-induction. Protein aggregation was not encountered and minimal amount of protein was found in the insoluble lysate.

Isolation and purification of proteins is crucial and central to structure-function studies such as enzymology, protein-protein or protein-ligand interactions. A low concentration of imidazole was used in the equilibration buffer [50 mM Tris-HCl, 0.02% (w/v) NaN₃, 10% (v/v) glycerol, 1 M NaCl, 40 mM imidazole, pH 7.4] to get rid of non-specific proteins binding to the column since imidazole has a higher affinity for the nickel ions than the protein. Cellular DNA fragments

that may have bound to the protein during sonication were removed by high concentration of salt (1 M NaCl) in the equilibration buffer. Salt wash is of great importance because the sodium and chloride ions interact with the exposed charged side chains of the protein thereby preventing it from interacting with the negatively-charged DNA backbone which then passes through the column. Small amount of the protein was found in the insoluble fraction and could be due to the formation of inclusion bodies as a result of incorrect folding of the proteins. Two step purification systems was carried out because of the extra bands seen on the 12% glycine SDS-PAGE electrophoretogram after IMAC purification which could be as a result of impurities.

For FL-heEF1 β and NT-heEF1 β the second purification step was DEAE-agarose ion exchange chromatography which is an anion exchanger. The heEF1 β protein has a pI of 4.88, which means that the protein is negatively charged and hence can bind to the positively charged DEAE resin. The pure protein was eluted with high salt (NaOH) concentration. Non-specific proteins and impurities that did not bind to the resin passed through the column. The second purification step for NT-heEF1 γ is GSH-agarose affinity chromatography because it has a GST-like NT-domain which has affinity for glutathione. Bound proteins were eluted with 10 M glycine-NaOH instead of using glutathione because of the presence of glutathione in the background. Eluted protein fractions were collected and their pH immediately adjusted to ~7.5 by adding 1 M Tris-HCl pH 7.4 at 25% (v/v) to prevent the loss in the secondary and tertiary characteristics of the protein due to increase in pH.

Protein quantification was evaluated by determining the concentration of the protein using serial dilution. Large quantities of the protein was obtained for structural studies and could be due to the over expression of the soluble proteins.

Confirming the integrity of the protein at the secondary, tertiary and quaternary structural levels provide valuable information required for further protein structure and function studies. The far-UV CD data confirmed that heEF1 β is predominantly α -helical and the NT-heEF1 β construct used in this study is also predominantly α -helical; and putting into context the structure of the C-terminus domain, the entire heEF1 β should be rich in α -helices and random coils. When heEF1 β is divided into an NT- and a CT- domain, the secondary structure content of both domains altered and does not sum up to the individual secondary structure in the full length polypeptide.

The reduction in the α -helical content of the NT-heEF1 β could be as a result of formation of β -strands and β -turns secondary structures (Table 2). Also, the acid rich region in the NT-heEF1 β construct could also have become more unordered in the absence of the CT-domain of the protein. The reason why the truncation of heEF1 β resulted in the formation of more unordered region and its implication with regards to protein-protein interaction with the γ subunit of eEF1 complex have not been ascertained.

The tertiary structure of heEF1 β was examined with the aim of identifying possible accessible hydrophobic pockets in the protein. This may provide us with an insight into the type of forces responsible for the formation of heEF1 β γ complex. ANS was used as an extrinsic fluorescence probe. The presence of hydrophobic cavities accessible to ANS may indicate that the interaction between β and the γ subunit of the eEF1 β complex could be chiefly driven by hydrophobic interaction (Slavik *et al.*, 1982; Gasymov and Glasgow, 2007), although it depends on which region of heEF1 β . The absence of ANS cavities may indicate that ionic or electrostatic forces may be the mediating force behind such protein-protein interaction. There was interaction between heEF1 β and ANS as seen in Figure 3.15 above, which indicates accessible hydrophobic clefts in heEF1 β . These hydrophobic clefts appear to be in the CT-domain where heEF1 β interacts with the GTPase α -subunit (van Damme *et al.*, 1992). The binding of ANS to the hydrophobic pocket in the C-terminus follows a single site binding model as seen in Figure 3.16, which suggests that ANS interacts with a single site in heEF1 β . The absence of any hydrophobic cavity accessible to ANS in the N-terminus region of heEF1 β could indicate that electrostatic forces may be involved in the formation of heEF1 β γ complex. This could mean that the interaction could be at the acid-rich cluster.

The quaternary structure determination using analytical SE-HPLC was done at pH 6.8 because the matrix of the column disintegrates at pH above 7.0. The resolution of the column is adequate enough to enable precise estimation of the oligomeric state of a protein between 1.5 kDa and 670 kDa. The CT-heEF1 β elutes as a monomer with a tumbling volume consistent with the theoretically predicted size, while the FL-heEF1 β is dimeric with an apparent molecular weight of ~62 kDa. This indicates that the NT-domain of heEF1 β could be responsible for the dimerisation of the protein. In the absence of any high order oligomeric states of heEF1 β , the

GST-like fold found in the NT-domain of heEF1 β could be the initiator of dimer formation in heEF1 β . This region is not found in heEF1 δ , which resolves as a high order oligomer with molecular weight >670 kDa. Moreover, the absence of the CT-domain of heEF1 β did not prevent the dimerisation of NT-heEF1 β , neither did the secondary structure modifications observed in NT-heEF1 β prevent its dimerisation. Structural prediction of heEF1 β suggests that part of the NT-domain shares similar fold with the α -helical CT-domain of GSTs which could be responsible for dimerisation of the protein. NT-heEF1 γ undergoes ligand-induced dimerization in the presence of oxidised glutathione. This ligand-induced dimerization did not affect the tendency of NT-heEF1 γ to form a complex with NT-heEF1 β . Efforts to replicate this data using reduced glutathione did not yield conclusive results.

The interaction between the β and the γ -subunits was examined using analytical SE-HPLC, which has a very high resolving power. Although it has been postulated that the interaction between heEF1 β and heEF1 γ occur at the NT-domains of both proteins (van Damme *et al.*, 1990), it was still necessary to rule out the possibility that the CT-domain of heEF1 β may be involved in β - γ interaction. The results show that the interaction indeed occurs at the NT-domains of both proteins, which is a further proof that both proteins are active.

Previous studies have shown that β and δ forms stable complex with γ subunit, and these complexes ($\beta\gamma$ and $\delta\gamma$) stimulate nucleotide exchange activity by the α -subunit (van Damme *et al.*, 1992). The β and δ -subunits share over 81% sequence similarity from the acid-rich region downstream (van Damme *et al.*, 1990). Thus, the α -subunit is able to form transient interaction with β and δ -subunits, which enhances nucleotide exchange activity. If the δ and β subunits forms stable complex with the γ subunit (Sheu and Traugh, 1997), it means that the only possible interacting region is the acid rich cluster, which is common in both subunits. Furthermore, steric hindrance will not permit the formation of $\gamma\beta\alpha$ or $\gamma\delta\alpha$ complexes because of the proximity between the acid rich region and the C-terminus α -binding domain (in β and δ subunits). From the results so far, there may be one binding site per protein per β and γ subunit as suggested by (Sheu and Traugh, 1997). There was no high order oligomer when NT-heEF1 γ interacted with FL-heEF1 β , and the conformations $\gamma:\beta\beta:\gamma$ and $\gamma\gamma:\beta\beta:\gamma\gamma$ may be probable. The expected molecular weights of the complex will not always correspond to the observed molecular weights

because of the orientations of the complex and how it affects the hydrodynamic volumes. There is a possibility of the formation of high order complex with $\gamma\gamma:\beta\beta:\gamma\gamma$ conformation ($[\dots\gamma\gamma:\beta\beta:\gamma\gamma:\beta\beta:\gamma\gamma:\beta\beta:\gamma\gamma:\beta\beta:\gamma\gamma\dots]_n$) if one binding site per molecule is probable. But the absence of higher molecular weight complex indicates that there is a possibility that the interaction is controlled to prevent the formation of higher molecular complexes when the β and the γ subunits interact. In the absence of the CT-domains in the β -subunit, there was a formation of high order molecular complex (>670 kDa) when NT-heEF1 β interacts with heEF1 γ as seen in Figure 3.22. The actual molecular basis of this high order complex formation in the absence of residues 140-225 in NT-heEF1 β could not be ascertained at this point in time. However, based on the experimental conditions used in this study, the CT-domain may be a role player in the interaction between the β and the γ -subunits of EF1 complex.

The heEF1 β protein was characterised with respect to its structure and interaction with heEF1 γ by dissecting heEF1 β into the NT-domain and the α -binding catalytic CT-domain. There is an unordered region in central region of the protein that contains a cluster of acidic amino acids residues.

4.2 Conclusion and prospective study

The heEF1 β was biophysically characterised with regards to its secondary, tertiary, quaternary and protein-protein interaction with the N-terminus GST-like domain of heEF1 γ . This study reveals that heEF1 β is predominantly α -helical with an ANS-accessible hydrophobic cavity on the CT- α -subunit binding domain of the protein. This suggests that the interaction between heEF1 β and heEF1 γ may be driven by electrostatic forces that could be contributed by the cluster of acidic amino acid residues in the central region of the polypeptide chain. The heEF1 β protein exists predominantly as a dimer of approximately 62 kDa and the NT-domain of heEF1 β is responsible for the dimerisation of the protein. The NT-heEF1 β interacts with the NT-domain of heEF1 γ in a possible 1:2 ratio (dimeric β : monomeric γ or dimeric β : dimeric γ) without formation of high order molecular complexes. The absence of residue 140-225 in heEF1 β resulted in high ordered complex with NT-heEF1 γ , suggesting that the CT-domain of heEF1 β may be the modulator for the formation of highly ordered $\beta:\gamma$ complex. This study could serve as an informative reference for understanding the molecular basis of protein-protein interaction between the β and γ -subunits of eEF1 complex in eukaryotes. The β - γ complex is very essential

in the nucleotide exchange activity of eEF1 and could be used extensively for commercial purpose in the mass production of proteins. The β - γ complex has been seen to increase in certain tumour cells and hence could be manipulated for pharmaceutical purposes or drug targets in order to understand their influence on tumour cells in cancer research studies. Further research such as quantitative study of the interaction between the β and γ subunits of the eEF1 using surface plasmon resonance (SPR) and phosphorylation studies of the β and γ subunits of eEF1 and its effects on their interaction, could improve our knowledge on this PPI.

REFERENCES

1. **Achilonu, I., Siganunu, T. P. and Dirr, H. W.** (2014) Purification and characterisation of recombinant human eukaryotic elongation factor-1 gamma, *Protein expression and purification* 99: 70-77.
2. **Al-Maghrebi, M., Anim, J. T. and Olalu, A. A.** (2005) Up-regulation of eukaryotic elongation factor-1 subunits in breast carcinoma, *Anticancer research* 25(3C): 2573-2577.
3. **Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J.** (1990) Basic local alignment search tool, *Journal of molecular biology* 215(3): 403-410.
4. **Angov, E., Hillier, C. J., Kincaid, R. L. and Lyon, J. A.** (2008) Heterologous protein expression is enhanced by harmonising the codon usage frequencies of the target gene with those of the expression host, *Public library of science (PLoS) One* 3(5): e2189.
5. **Archakov, A. I., Govorun, V. M., Dubanov, A. V., Ivanov, Y. D., Veselovsky, A. V., Lewi, P. and Janssen, P.** (2003) Protein-protein interactions as a target for drugs in proteomics, *Proteomics* 3(4): 380-391.
6. **Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E. et al.** (2012) ExPASy: SIB bioinformatics resource portal, *Nucleic acids research* 40(Web Server issue): W597-603.
7. **Bimboim, H. and Doly, J.** (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA, *Nucleic acids research* 7(6): 1513-1523.
8. **Braman, J., Papworth, C. and Greener, A.** (1996) Site-directed mutagenesis using double-stranded plasmid DNA templates, *In Vitro Mutagenesis Protocols*: 31-44.
9. **Byun, H. O., Han, N. K., Lee, H. J., Kim, K. B., Ko, Y. G., Yoon, G., Lee, Y. S., Hong, S. I. and Lee, J. S.** (2009) Cathepsin D and eukaryotic translation elongation factor-1 as promising markers of cellular senescence, *Cancer Research* 69(11): 4638-4647.
10. **Chen, C.-J. and Traugh, J. A.** (1995) Expression of recombinant elongation factor 1 beta from rabbit in *Escherichia coli*. Phosphorylation by casein kinase II, *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 1264(3): 303-311.
11. **Chi, K., Jones, D. V. and Frazier, M. L.** (1992) Expression of an elongation factor-1c related sequence in adenocarcinomas of the colon, *Gastroenterology* 103: 98-102.
12. **Chothia, C. and Janin, J.** (1975) Principles of protein-protein recognition, *Nature* 256(5520): 705-708.
13. **Corbi, N., Batassa, E. M., Pisani, C., Onori, A., Di Certo, M. G., Strimpakos, G., Fanciulli, M., Mattei, E. and Passananti, C.** (2010) The eEF1 γ subunit contacts RNA polymerase II and binds vimentin promoter region, *Public library of science (PLoS) One* 5(12): e14481.
14. **Creighton, T. E.** (1991) Stability of folded conformations: Current opinion in structural biology 1991, 1: 5-16, *Current Opinion in Structural Biology* 1(1): 5-16.

15. **Creighton, T. E.** (1997) *Protein structure: a practical approach*: Oxford university press.
16. **Dill, K. A.** (1990) Dominant forces in protein folding, *Biochemistry* 29(31): 7133-7155.
17. **Eftink, M. R.** (1995) [22] Use of multiple spectroscopic methods to monitor equilibrium unfolding of proteins, *Methods in enzymology* 259: 487-512.
18. **Engelhard, M. and Evans, P. A.** (1995) Kinetics of interaction of partially folded proteins with a hydrophobic dye: evidence that molten globule character is maximal in early folding intermediates, *Protein Science* 4(8): 1553-1562.
19. **Ernst, V., Levin, D. H. and London, I. M.** (1978) Inhibition of protein synthesis initiation by oxidized glutathione: activation of a protein kinase that phosphorylates the α subunit of eukaryotic initiation factor 2, *Proceedings of the National Academy of Sciences* 75(9): 4110-4114.
20. **Fersht, A. R.** (1987) The hydrogen bond in molecular recognition, *Trends in Biochemical Sciences* 12: 301-304.
21. **Figeys, D.** (2002) Functional proteomics: mapping protein-protein interactions and pathways, *Current opinion in molecular therapeutics* 4(3): 210-215.
22. **Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D. and Bairoch, A.** (2005) *Protein identification and analysis tools on the ExPASy server*: Springer.
23. **Gasymov, O. K. and Glasgow, B. J.** (2007) ANS fluorescence: potential to augment the identification of the external binding sites of proteins, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1774(3): 403-411.
24. **Geiger, A., Rahman, A. and Stillinger, F.** (1979) Molecular dynamics study of the hydration of Lennard-Jones solutes, *The Journal of Chemical Physics* 70(1): 263-276.
25. **Gillen, C. M., Gao, Y., Niehaus-Sauter, M. M., Wylde, M. R. and Wheatly, M. G.** (2008) Elongation factor 1B γ (eEF1B γ) expression during the molting cycle and cold acclimation in the crayfish *Procambarus clarkii*, *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* 150(2): 170-176.
26. **Goodsell, D. S. and Olson, A. J.** (2000) Structural symmetry and protein function, *Annual review of biophysics and biomolecular structure* 29(1): 105-153.
27. **Guex, N. and Peitsch, M. C.** (1997) Swiss-model and the Swiss-Pdb Viewer: an environment for comparative protein modeling, *electrophoresis* 18(15): 2714-2723.
28. **Hardy, S., Holmgren, J., Johansson, S., Sanchez, J. and Hirst, T. R.** (1988) Coordinated assembly of multisubunit proteins: oligomerisation of bacterial enterotoxins in vivo and in vitro, *Proceedings of the National Academy of Sciences* 85(19): 7109-7113.
29. **Ito, T., Marintchev, A. and Wagner, G.** (2004) Solution structure of human initiation factor eIF2 α reveals homology to the elongation factor eEF1B, *Structure* 12(9): 1693-1704.
30. **Janin, J., Miller, S. and Chothia, C.** (1988) Surface, subunit interfaces and interior of oligomeric proteins, *Journal of molecular biology* 204(1): 155-164.
31. **Janssen, G. and Moller, W.** (1988) Elongation factor-1 $\beta\gamma$ from *Artemia*, *European Journal of Biochemistry* 171(1-2): 119-128.
32. **Janssen, G., Van Damme, H., Kriek, J., Amons, R. and Möller, W.** (1994) The subunit structure of elongation factor-1 from *Artemia*. Why two alpha-chains in this complex?, *Journal of Biological Chemistry* 269(50): 31410-31417.

33. **Jeppesen, M. G., Ortiz, P., Shepard, W., Kinzy, T. G., Nyborg, J. and Andersen, G. R.** (2003) The crystal structure of the glutathione S-transferase-like domain of elongation factor-1 β from *Saccharomyces cerevisiae*, *Journal of Biological Chemistry* 278(47): 47190-47198.
34. **Jones, S., Marin, A. and Thornton, J. M.** (2000) Protein domain interfaces: characterisation and comparison with oligomeric protein interfaces, *Protein Engineering* 13(2): 77-82.
35. **Jones, S. and Thornton, J. M.** (1995) Protein-protein interactions: a review of protein dimer structures, *Progress in biophysics and molecular biology* 63(1): 31-65.
36. **Jones, S. and Thornton, J. M.** (1996) Principles of protein-protein interactions, *Proceedings of the National Academy of Sciences* 93(1): 13-20.
37. **Joseph, R. L. and Lakowicz, R.** (1999) Principles of fluorescence spectroscopy: Kluwer Academic/Plenum Publishers, New York.
38. **Kaufmann, R.** (1995) Matrix-assisted laser desorption ionization (MALDI) mass spectrometry: a novel analytical tool in molecular biology and biotechnology, *Journal of biotechnology* 41(2): 155-175.
39. **Kauzmann, W.** (1959) Some factors in the interpretation of protein denaturation, *Advances in protein chemistry* 14: 1-63.
40. **Kelly, S. M. and Price, N. C.** (1997) The application of circular dichroism to studies of protein folding and unfolding, *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* 1338(2): 161-185.
41. **Kim, S., Kellner, J., Lee, C. H. and Coulombe, P. A.** (2007) Interaction between the keratin cytoskeleton and eEF1 β affects protein synthesis in epithelial cells, *Nature structural & molecular biology* 14(10).
42. **Kozak, M.** (1999) Initiation of translation in prokaryotes and eukaryotes, *Gene* 234(2): 187-208.
43. **Kumabe, T., Sohma, Y. and Yamamoto, T.** (1992) Human cDNAs encoding elongation factor-1 gamma and the ribosomal protein L19, *Nucleic acids research* 20(10): 2598.
44. **Kumar, S. and Nussinov, R.** (1999) Salt bridge stability in monomeric proteins, *Journal of molecular biology* 293(5): 1241-1255.
45. **Lakowicz, J. R.** (1983) Principles of fluorescence spectroscopy, *Plenum Press, (New York) Chapter 5*: 111-150.
46. **Laemmli, U. K.** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227: 680-685.
47. **Lakowicz, J. R. and Masters, B. R.** (2008) Principles of fluorescence spectroscopy, *Journal of Biomedical Optics* 13(2): 9901.
48. **Le Sourd, F., Boulben, S., Le Bouffant, R., Cormier, P., Morales, J., Belle, R. and Mulner-Lorillon, O.** (2006) eEF1B: At the dawn of the 21st century, *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 1759(1): 13-31.
49. **Li, D., Wei, T., Abbott, C. M. and Harrich, D.** (2013) The unexpected roles of eukaryotic translation elongation factors in RNA virus replication and pathogenesis, *Microbiology and Molecular Biology Reviews* 77(2): 253-266.
50. **Liu, D., Sheng, C., Gao, S., Jiang, W., Li, J., Yao, C., Chen, H., Wu, J., Chen, S. and Huang, W.** (2014) eEF1 β is a positive regulator of NF- κ B signaling pathway, *Biochemical and biophysical research communications* 446(2): 523-528.

51. **Maessen, G., Amons, R., Zeelen, J. and Möller, W.** (1987) Primary structure of elongation factor-1 γ from *Artemia*, *Federation of European biochemical societies (FEBS) letters* 223(1): 181-186.
52. **Mahley, R. W.** (1988) Apolipoprotein E: cholesterol transport protein with expanding role in cell biology, *Science* 240(4852): 622.
53. **Markus, M. and Benezra, R.** (1999) Two isoforms of protein disulfide isomerase alter the dimerisation status of E2A proteins by a redox mechanism, *Journal of Biological Chemistry* 274(2): 1040-1049.
54. **Mathur, S., Cleary, K. R., Inamdar, N., Kim, Y. H., Steck, P. and Frazier, M. L.** (1998) Overexpression of elongation factor-1 γ protein in colorectal carcinoma, *Cancer* 82(5): 816-821.
55. **Mimori, K., Mori, M., Inoue, H., Ueo, H., Mafune, K., Akiyoshi, T. and Sugimachi, K.** (1996) Elongation factor-1 gamma mRNA expression in oesophageal carcinoma, *Gut* 38(1): 66-70.
56. **Mitsui, Y., Satow, Y., Watanabe, Y., Hirono, S. and Iitaka, Y.** (1979) Crystal structures of *Streptomyces subtilisin* inhibitor and its complex with subtilisin BPN', *Nature* 277: 447-452.
57. **Motorin, Y. A., Wolfson, A. D., Löhr, D., Orlovsky, A. F. and Gladilin, K. L.** (1991) Purification and properties of a high molecular mass complex between Val-tRNA synthetase and the heavy form of elongation factor-1 from mammalian cells, *European Journal of Biochemistry* 201(2): 325-331.
58. **Navia, M. A., Fitzgerald, P. M., McKeever, B. M., Leu, C. T., Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L. and Springer, J. P.** (1989) Three dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1, *Nature* 337(6208): 615-620.
59. **Nooren, I. M. and Thornton, J. M.** (2003) Structural characterisation and functional significance of transient protein–protein interactions, *Journal of molecular biology* 325(5): 991-1018.
60. **Nussinov, R., Xu, D. and Tsai, C. J.** (1998) Mechanism and evolution of protein dimerisation, *Protein Science* 7(3): 533-544.
61. **Olarewaju, O., Ortiz, P. A., Chowdhury, W. Q., Chatterjee, I. and Kinzy, T. G.** (2004) The translation elongation factor eEF1B plays a role in the oxidative stress response pathway, *RNA biology* 1(2): 89-94.
62. **Pawson, T. and Nash, P.** (2003) Assembly of cell regulatory systems through protein interaction domains, *Science* 300(5618): 445-452.
63. **Peng, X., Wang, J., Peng, W., Wu, F. X. and Pan, Y.** (2016) Protein–protein interactions: detection, reliability assessment and applications, *Briefings in Bioinformatics*: bbw066.
64. **Pérez, J. M., Kriek, J., Dijk, J., Canters, G. W. and Möller, W.** (1998) Expression, purification, and spectroscopic studies of the guanine nucleotide exchange domain of human elongation factor, EF-1 β , *Protein expression and purification* 13(2): 259-267.
65. **Pitt-Rivers, R. and Impiombato, F. A.** (1968) The binding of sodium dodecyl sulphate to various proteins, *Biochemical Journal* 109(5): 825-830.
66. **Sanders, J., Maassen, J. and Möller, W.** (1992) Elongation factor-1 messenger-RNA levels in cultured cells are high compared to tissue and are not drastically affected further by oncogenic transformation, *Nucleic acids research* 20(22): 5907-5910.

67. **Schägger, H.** (2006) Tricine–SDS–PAGE, *Nature Protocols* 1: 16-22.
68. **Schwikowski, B., Uetz, P. and Fields, S.** (2000) A network of protein–protein interactions in yeast, *Nature biotechnology* 18(12): 1257-1261.
69. **Seth, R. B., Sun, L., Ea, C. K. and Chen, Z. J.** (2005) Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF- κ B and IRF3, *Cell* 122(5): 669-682.
70. **Shenoy, A. R. and Visweswariah, S. S.** (2003) Site-directed mutagenesis using a single mutagenic oligonucleotide and DpnI digestion of template DNA, *Analytical Biochemistry* 319(2): 335-336.
71. **Sheu, G.-T. and Traugh, J. A.** (1997) Recombinant subunits of mammalian elongation factor-1 expressed in *Escherichia coli* subunit interactions, elongation activity and phosphorylation by protein kinase CKII, *Journal of Biological Chemistry* 272(52): 33290-33297.
72. **Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A., Tang, W. H., Keating, S. P., Hunter, C. L., Nuwaysir, L. M. and Schaeffer, D. A.** (2007) The paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra, *Molecular and Cellular Proteomics* 6(9): 1638-1655.
73. **Slavik, J., Horák, J., Říhová, L. and Kotyk, A.** (1982) Anilinonaphthalene sulfonate fluorescence and amino acid transport in yeast, *The Journal of membrane biology* 64(3): 175-179.
74. **Sticke, D. F., Presta, L. G., Dill, K. A. and Rose, G. D.** (1992) Hydrogen bonding in globular proteins, *Journal of molecular biology* 226(4): 1143-1159.
75. **Stillinger, F. H.** (1980) Water revisited, *Science* 209(4455): 451-457.
76. **Stryer, L.** (1965) The interaction of a naphthalene dye with apomyoglobin and apohemoglobin: a fluorescent probe of non-polar binding sites, *Journal of molecular biology* 13(2): 482-495.
77. **Tayyab, S., Qamar, S. and Islam, M.** (1991) Size exclusion chromatography and size exclusion HPLC of proteins, *Biochemical Education* 19(3): 149-152.
78. **Teichmann, S. A.** (2002) Principles of protein-protein interactions, *Bioinformatics* 18(suppl 2): S249-S249.
79. **Thiede, B., Höhenwarter, W., Krah, A., Mattow, J., Schmid, M., Schmidt, F. and Jungblut, P. R.** (2005) Peptide mass fingerprinting, *Methods* 35(3): 237-247.
80. **Tsai, C. J., Lin, S. L., Wolfson, H. J. and Nussinov, R.** (1997) Studies of protein-protein interfaces: a statistical analysis of the hydrophobic effect, *Protein Science* 6(1): 53-64.
81. **Tshabalala, T. N., Tomescu, M. S., Prior, A., Balakrishnan, V., Sayed, Y., Dirr, H. W. and Achilonu, I.** (2016) Energetics of glutathione binding to human eukaryotic elongation factor-1 gamma: Isothermal titration calorimetry and molecular dynamics studies, *The Protein Journal* 35(6): 448-458.
82. **van Damme, H., Amons, R., Karssies, R., Timmers, C., Janssen, G. and Möller, W.** (1990) Elongation factor-1 β of artemia: localization of functional sites and homology to elongation factor-1 δ , *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 1050(1): 241-247.

83. **van Damme, H. T., Amons, R. and Moller, W.** (1992) Identification of the sites in the eukaryotic elongation factor-1 alpha involved in the binding of elongation factor-1 beta and aminoacyl-tRNA, *European Journal of Biochemistry* 207(3): 1025-34.
84. **Vanwetswinkel, S., Kriek, J., Andersen, G. R., Dijk, J. and Siegal, G.** (2003a) Letter to the Editor: 1 H, 15 N and 13 C resonance assignments of the highly conserved 19 kDa C-terminal domain from human elongation factor-1B γ , *Journal of biomolecular NMR* 26(2): 189-190.
85. **Vanwetswinkel, S., Kriek, J., Andersen, G. R., Güntert, P., Dijk, J., Canters, G. W. and Siegal, G.** (2003b) Solution structure of the 162 residue C-terminal domain of human elongation factor-1B γ , *Journal of Biological Chemistry* 278(44): 43443-43451.
86. **Veremieva, M., Kapustian, L., Khoruzhenko, A., Zakharychev, V., Negrutskii, B. and El'skaya, A.** (2014) Independent overexpression of the subunits of translation elongation factor complex eEF1H in human lung cancer, *BioMed Central (BMC) Cancer* 14(1): 1.
87. **Wendt, M. D.** (2012) Protein-protein interactions as drug targets *Protein-Protein Interactions*: Springer.
88. **Werther, M. and Seitz, H.** (2008) *Protein-protein interaction*: Springer.
89. **Whitmore, L. and Wallace, B.** (2004) Dichroweb, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic acids research* 32(suppl 2): W668-W673.
90. **Woody, R. W.** (1995) [4] Circular dichroism, *Methods in enzymology* 246: 34-71.