CHARACTERIZATION OF A HUMAN INHIBITORY ANTIBODY FRAGMENT AGAINST TISSUE FACTOR

by

JAN-G VERMEULEN

This thesis is submitted in fulfilment of the requirements for the degree

M.Med.Sc (Human Molecular Biology)

In the Faculty of Health Sciences

Department of Haematology and Cell Biology

at the

University of the Free State

Bloemfontein

Republic Of South-Africa

November 2011

Supervisor: Professor S.M. Meiring

Co-Supervisior: Professor E van Heerden

ACKNOWLEDGEMENTS

Contents

and pET22b (+) in *E.coli* TOP10

List of Figures

List of Tables

List of Abbreviations

(IPTG)	Isopropyl-β-D-thiogalactoside
AP-1	Activator Protein-1
C-terminus	Carboxy terminus
DIC	Disseminated intravascular coagulation
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EGR-1	Early growth response protein 1
Fab	Fragment antigen-binding
GuHCl	Guanidine hydrochloride
HIT	Heparin induced thrombocytopenia
HIV-AIDS	Human immunodeficiency virus - Acquired immune deficiency syndrome
IgG	Immunoglobulin
IL-6	Interleukin 6
INR	International Normalisation Ratio
IMAC	Immobilized metal ion affinity chromatography
LPS	Lipopolysaccharide
MWCO	Molecular weight cut off
NF-$_K$B	Nuclear factor kappa-light-chain-enhancer of activated B cells
NHLS	National health laboratory services
PAR	Protease-activated receptor
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PPP	Platelet poor plasma
RNA	Ribonucleic acid
rNAPc2	Recombinant nematode anticoagulant protein c2
rTFPI	Recombinant tissue factor Pathway Inhibitor
scFv	Single chain variable antibody fragment
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SP1	Specificity Protein 1
TBS	Tris buffered saline
TF_{AS}	Alternatively spliced tissue factor
TF_{FL}	Full length tissue factor
TFI-scFv	Tissue factor inhibitor single chain variable fragment
TF_{MP}	Tissue factor baring microparticles
TNF-α	Tissue necrosis factor alpha
tRNA	Transfer ribonucleic acid
VE	Venous thromboembolism
VEGF	Vascular endothelial growth factor
V_H	Heavy chain variable region
V_L	Light chain variable regions

CHAPTER 1
General Introduction

1.1 Background

Tissue factor is a transmembrane glycoprotein that functions as the primary initiator of coagulation in response to mechanical or chemical damage of a blood vessle. It plays a key role in the pathology of thrombosis and thrombotic complications associated with cardiovascular disease (Mackman, 2006). These diseases are annually responsible for 17 million deaths worldwide, 80% of these deaths occur in middle- and low income countries such as South Africa (Callow, 2006). Thrombotic complications associated with non-thrombotic disorders and diseases such as obesity, diabetes, cancer and HIV-AIDS drastically increase mortality in these diseases (Nofer *et al.*, 2002; Kakkar *et al.*, 2003; Levine *et al.*, 2003; Granata *et al.*, 2004; Milson *et al.*, 2007; Pantanowitz and Dezube, 2010). Consequently, a significant challenge in the field of cardiovascular- and thrombotic research is the development of safe and effective anticoagulant agents with which to supplement and eventually replace current commercial anticoagulants.

The inhibition of tissue factor is a recent development in anticoagulant therapy that addresses many of the shortcomings of the current anticoagulants. The inhibition of the initiation step of coagulation prevents the formation of "new" thrombin and down regulates the positive feedback mechanism responsible for thrombin amplification (Golino, 2002). Tissue factor inhibition also limits the *vice versa* activation between coagulation and inflammation (Golino, 2002; Chu, 2005; Steffel *et al.*, 2006a). A further potential advantage is that tissue factor inhibitors may only act at the site of vascular injury while leaving the physiological haemostasis intact (Golino, 2002). Animal models have shown that the inhibition of tissue factor poses little risk of haemorrhage (Hanson and Sakariassen, 1998; Roque´ *et al.*, 2000; Carmeliet and Collen, 1998). It has also been shown that the inhibition of the tissue factor – factor VIIa complex is safer and thus results in less bleeding complications than inhibition of other coagulation factors such as thrombin and factor X (Miura *et al.*, 2006).

Furthermore, tissue factor plays an important role in the pathogenesis of arterial thrombosis, arthrosclerosis, obesity, type 2 diabetes and HIV-AIDS (Samad *et al.*, 2001; Diamant *et al.*, 2002; Kamihura *et al.*, 2005; Steffel *et al.*, 2006b; Zwicker *et al.*, 2007; Khunnawat *et al.*, 2008). Apart from this role, it also contributes to the metastasis and angiogenesis observed in certain malignancies (Hobbs *et al.*, 2007; Rak *et al.*, 2009). The interaction between thrombosis- and inflammation mechanisms is especially important as the physiological response of the one system is responsible for the activation of the other (Levi and van der Poll, 2005; von Känel *et al.*, 2008). It is this *vice versa* activation between these two physiological processes that enhances the pathological severity of thrombosis and other diseases exhibiting thrombotic complications (Chu, 2005). The inhibition of tissue factor is therefore of great importance as it may provide a novel antithrombotic therapy which addresses the complications of these diseases.

Recognising the potential of tissue factor inhibition as a novel approach to antithrombotic therapy, our laboratory utilized phage display technology to identify a 26 kDa single chain antibody fragment (scFv) which functionally inhibits human tissue factor (Meiring *et al.*, 2009). Preliminary *in vitro* studies have shown that this antibody fragment directed against tissue factor extended prothrombin times in a dose dependent manner. The purified JT C5 single chain variable fragment (scFv), hereon referred to as tissue factor inhibitor scFv (TFI-scFv) increased prothrombin time in a concentration-depen- dent manner with a two-fold prolongation at 1.5 mM (39 ng/ml) The antibody has also been shown to exhibit an inhibitory effect on *in vitro* thrombin generation. It prolonged the lag phase prior to thrombin generation burst as well as reduced the maximum amount of thrombin generated at peak (Meiring *et al.*, 2009). Additional characterisation of the scFv was hampered by low protein yields and financial complications associated with the initial purification process. The aim of this study is therefore to improve the production of this antibody fragment in order to determine the efficacy in animal models.

1.2 References

CALLOW, A.D. 2006. Cardiovascular disease 2005 - The global picture. *Vascular Pharmacology*, 45, 302-307.

CARMELIET, P. & COLLEN, D. 1998. Molecule in focus: Tissue factor. *The International Journal of Biochemistry & Cell Biology*, 30, 661-667.

CHU, A.J. 2005. Tissue factor mediates inflammation. *Archives of Biochemistry and Biophysics*, 440, 123-132.

DIAMANT, M., NIEUWLAND, R., PABLO, R.F., STURK, A., SMIT, J.W.A. & RADDER, J.K. 2002. Elevated Numbers of Tissue Factor Exposing Microparticles Correlates With Componets of the Metabolic Syndromes in Uncomplicated Type 2 Diabetes Mellitus. *Circulation*, 106, 2442-2447.

GOLINO, P. 2002. The inhibitors of the tissue factor:factor VII pathway. *Thrombosis Research*, 106, V257-V265.

GRANATA, G., IZZO, T., MICCO, P.D., BONAMASSA, B., CASTALDO, G., VIGGIANA, V.G., PICILLO, U., CASTALDO, G. & NIGLIO, A. 2004. Thromboembolic events and haematological diseases: a case of stroke as clinical onset of a paroxysmal nocturnal haemoglobinurai. *Thrombosis Journal*, 2, 10-14.

HANSON, S.R. & SAKARIASSEN, K.S. 1998. Blood flow and antithrombotic drug effects. *American Heart Journal*, 135, S132-145.

HOBBS, J., ZAKARIJA, A., CUNDIFF, D., DOLL, J., HYMEN, E., CORNWELL, M., CRAWFORD, S., LIU, N., SIGNAEVSKY, M. & SOFF, G. 2007. Alternative spliced human tissue factor promotes growth and angiogenesis in pancreatic cancer tumor model. *Thrombosis Research*, 120, S13-S21.

KAKKAR, A. K., LEVINE, M., PINEDO, H. M., WOLFF, R. & WONG, J. 2003. Venous Thrombosis in Cancer Patients: Insights from the FRONTLINE Survey. *The Oncologist*, 8, 381-388.

KAMIHURA, Y., WADA, H., NOBORI, T., KOBAYASHI, T., SASE, T., NISHIKAWA, M., ISHIKURA, K., YAMADA, N., ABE, Y., NISHIOKA, J., NAKANO, T. & SHIKU, H. 2005. Elevated levels of leukocyte tissue factor mRNA in patients with venous thromboembolism. *Thrombosis Research*, 116, 307-312.

KHUNNAWAT, C., MUKERJI, S., HAVLICHEK, D., TOUMA, R. & ABELA, G.S. 2008. Cardiovascular Manifestations in Human Immunodefeiciency Virus-Infected Patients. *American Journal of Cardiology*, 102, 635-642.

LEVI, M. & VAN DER POLL, T. 2005. Two-Way Interaction Between Inflammation and Coagulation. *Trends in Cardiovascular Medicine*, 15, 254-259.

LEVINE, M.N., LEE, A.Y. & KAKKAR, A.K. 2003. From Trousseau to targeted therapy: New insights and innovations in thrombosis and cancer. *Journal of Thrombosis and Haemostasis*, 1, 1456-1463.

MACKMAN, N. 2006. The role of tissue factor in hemostasis and thrombosis. *Blood Cells, Molecules, and Diseases*, 36, 104-107.

MEIRING, S.M., VERMEULEN, J. & BADENHORST, P.N. 2009. Development of an Inhibitory Antibody Fragment to Human Tissue Factor using Phage Display Technology. *Drug Development Research*, 70, 199-205.

MILSON, C., YU, J., MAY, L., MEEHAN, B., MAGNUS, N., AL-NEDAWI, K., LUYENDYK, J., WEITZ, J. I., KLEMENT, P., BROZE, G., MACKMAN, N. & RAK, J. 2007. The role of tumor- and host -related tissue factor pools in oncogene-driven tumor progression. *Trombosis Research*, 129, S82-S91.

MIURA, M., SEKI, N., KOIKE, T., ISHIHARA, T., NIIMI, T., HIRAYAMA, F., SHIGEBAGA, T., SAKAI-MORITANI, Y., KAWASAKI, T., SAKAMOTO, S., OKADA, M., OHTA, M. & TSUKAMOTO, S. 2006. Potent and selective TF/VIIa inhibitors containing a neutral P1 ligand. *Bioorganic & Medicinal Chemistry*, 14, 7688-7705.

NOFER, J., KEHREL, B., FOBKER, M., LEVKAU, B., ASSMANN, G. & ECKARDSTAIN, A.V. 2002. HDL and ateriosclerosis: behond reverse cholesteerol transport. *Atherosclerosis*, 161, 1-16.

PANTANOWITZ, L. & DEZUBE, B. 2010. Monocytes tied to HIV-associated thrombosis. *Blood*, 115, 156-157.

RAK, J., MILSOM, C., MAGNUS, N. & YU, J. 2009. Tissue factor in tumor progression. *Best Practice & Research Clinical Haematology*, 22, 71-83.

ROQUE´, M., REIS, E.D., FUSTER, V., PADUREAN, A., FALLON, J.T., TAUBMAN, M. B., CHESEBRO, J.H. & BADIMON, J.J. 2000. Inhibition of Tissue Factor Reduces Thrombus Formation and Intimal Hyperplasia After Porcine Coronary Angioplasty. *Journal of the American College of Cardiology*, 36, 2303–2310.

SAMAD, F., PANDEY, M. & LOSKUTOFF, D. 2001. Regulation of tissue factor gene expression in obesity. *Blood*, 98, 3353-3358.

STEFFEL, J., AKHMEDOV, A., FANHDRICH, C., RUSCHITZKA, F., LÜSCHER, T. T. & TANNER, F. 2006 a. Differential effect of celecoxib on tissue factor expression in human endothelial and vascular smooth muscle cells. *Biochemical and Biophysical Research Communications*, 349: 597-603.

STEFFEL, J., LÜSCHER, T.F. & TANNER, F.C. 2006 b. Tissue Factor in Cardiovascular Diseases: Molecular Mechanism and Clinical Implications. *Circulation*, 113, 722-731.

VON KÄNEL, R., NELESEN, G.Z., MILLS, P.J., ZIEGLER, M.G. & DIMSDALE, J.E. 2008. Relationship between Heart rate variability, interleukin-6, and solible tissue factor in healthy subjects. *Brain, Behavior and Immunity*, 22, 461-468.

ZWICKER, J.I., FURIE, B.C. & FURIE, B. 2007. Cancer-associated thrombosis. *Oncology/Hematology*, 62, 126-136.

CHAPTER 2
Literature Review

2.1 Tissue factor

Tissue factor (also referred to as coagulation factor III, thromboplastin or CD142) is the primary physiological initiator of coagulation in response to mechanical- or chemical damage of a blood vessel (Kirchhofer *et al.*, 1997; Butenas *et al.*, 2005; Steffel *et al.*, 2006b). It is constitutively expressed by a variety of extra-vascular tissues that are not in direct contact with circulating blood, particularly in organ capsules, cells of the epidermis, mucosal epithelium and the vascular adventitial layer (Carmeliet and Collen, 1998; Levi and van der Poll, 2005; Mackman, 1995). This specific distribution results in the formation of a protective "haemorrhage envelope" that surrounds arteries and organs and provides a haemostatic barrier necessary for immediate response to injury (Levi and van der Poll, 2005; Lösche, 2005; Hoffman, 2008).

Intravascular cells in contrast do not express tissue factor constitutively. Expression in these cells is induced in response to various stimuli such as pro-inflammatory cytokines (interleukin-1, tumour necrosis factor-α), mitogens (fibroblast growth factor, platelet-derived growth factor, epidermal growth factor, transforming growth factor β-1,vascular endothelial growth factor, insulin, angiotensin II, α-thrombin), hormones, endotoxins, virus infections, immune complexes, modified lipoproteins and biogenic amines (Carmeliet and Collen, 1998; Golino, 2002; Kamihura *et al.*, 2005; Peng *et al.*, 2007; Bugge *et al.*, 1996).

Apart from its fundamental role in haemostasis, tissue factor also functions in the induction of intracellular Ca^{2+} signals (Kirchhofer and Nemerson, 1996; Steffel *et al.*, 2006b). Tissue factor mediates the activation of protease-activated receptor (PAR)-2 signalling by forming a complex with factor VIIa and factor Xa (Belting *et al.*, 2004). PAR signalling functions in a variety of biological process, including angiogenesis, inflammation, metastasis and cell migration (Bognanov *et al.*, 2003; Krikun *et al.*, 2009).

Tissue factor plays a role in angiogenesis and vasculogenesis and is especially important as it is an "immediate early gene" necessary for embryonic survival (Beck and D'Amore, 1997; Watanabe *et al.*, 1999; Mackman, 2004). During early embryonic development it is expressed in a diverse range of embryonic cells when coagulation factors are minimally expressed and/or unable to cross the placenta barrier (Bugge *et al.*, 1996; Carmeliet and Collen, 1998). As the growing embryo demands greater blood supply from the yoke sac, tissue factor positive immature blood vessels acquire a muscular outer wall in order to accommodate the increase in blood pressure (Carmeliet and Collen, 1998; Cines *et al.*, 1998). Tissue factor deficient blood vessels lack this outer muscle wall and are unable to handle the higher blood pressures. This results in blood vessel rupture and ultimately embryo death due to massive haemorrhaging (Bugge *et al.*, 1996; Carmeliet *et al.*, 1996; Beck and D'Amore, 1997).

2.2 Tissue factor gene

The human tissue factor gene is located on chromosome one region 1p21-22 (Carmeliet *et al.*, 1996; Förster *et al.*, 2006). The promoter (See: Figure 2.1) is spanned by three sequences with a high frequency of Cytosine and Guanine (CpG islands) (Lin *et al.*, 1996). These islands can be methylated in order to regulate tissue factor gene expression in certain cell types (Norris *et al.*, 1991). The promoter region consist of five Specificity Protein 1 sites (SP1) of which three overlap with three Early growth response protein 1 sites (EGR-1), two Activator Protein-1 sites (AP-1) and a nuclear factor kappa-light-chain-enhancer of activated B cells (NF-$_K$B) (Cui *et al.*, 1995; Mackman, 1995). SP1 sites regulate basal expression, while AP-1 and NF-$_K$B mediate lipopolysaccharide (LPS) and cytokine induction in monocytes and endothelial cells (Förster *et al.*, 2006). The three overlapping SP1/EGR-1 sites mediate serum growth factor induction in epithelial cells (Mackman, 1995).

The tissue factor gene itself is organized into six exons and five introns and spans a total of 12,4 kbp (Förster *et al.*, 2006). Only two transcript variants of tissue factor have been identified, namely the primary transcript and the alternative transcript (See: Figure 2.2(A)) (Nemerson, 1988; Bognanov *et al.*, 2003; Steffel *et al.*, 2006b).
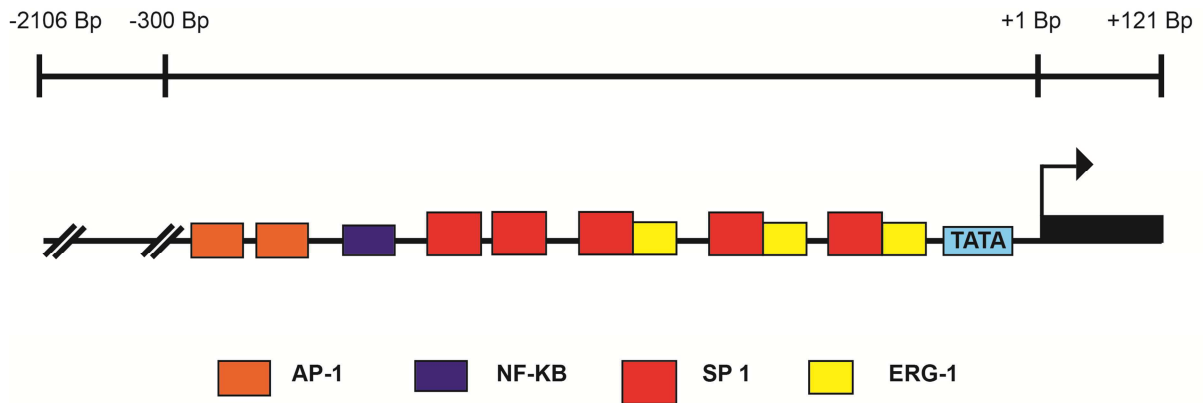
Figure 2.1: Tissue factor promoter structure

The human tissue factor promoter region spans a total 2.2 kbP (-2106 Bp to +121Bp), with the majority of transcription factor binding sites located between the -300 Bp to +1 Bp region. It consists of five SP1 (■) sites, three of which overlap with three EGR-1 (■) sites, two AP-1 (■) sites and a NF-$_K$B site (Blue). An arrow indicates the start of transcription site downstream of the TATA box (Light Blue).

Adapted from Mackman (1995) and Lin *et al.* (1996).

The primary transcript, also referred to as full length tissue factor (TF$_{FL}$), encodes for 295 amino acid residue (46 kDa) transmembrane glycoprotein (Kirchhofer *et al.*, 1997; Bognanov *et al.*, 2003; Butenas *et al.*, 2005; Steffel *et al.*, 2006b). Exon 1 encodes for a 32 amino acid signal peptide which is proteolytically cleaved during transport to the plasma membrane (Förster *et al.*, 2006). As a result, the mature protein consists of only 263 amino acid residues (Rønning *et al.*, 1996). Exons 2-5 encode the 219 amino acid extracellular domain which binds to coagulation factors VII and X and mediates the coagulant activity (Golino, 2002; Kamihura *et al.*, 2005; Lösche, 2005; Østerud *et al.*, 2007; Peng *et al.*, 2007).

Exon 6 encodes for the 23 amino acid transmembrane domain and 21 amino acid cytoplasmic tail (Carmeliet and Collen, 1998; Mackman, 2007). The transmembrane domain anchors tissue factor to the cell surface in order to achieve full pro-coagulant activity (Rauch *et al.*, 2006). The cytoplasmic domain regulates PAR 2 signal transduction during angiogenesis (O'Brien *et al.*, 1993; Hobbs *et al.*, 2007).

A) Tissue Factor Gene Structure

Exon 1
1-223

Exon 2
1392-1503

Exon 3
5594-5793

Exon 4
8491-8669

Exon 5
9277-9436

Exon 6
11157-12434
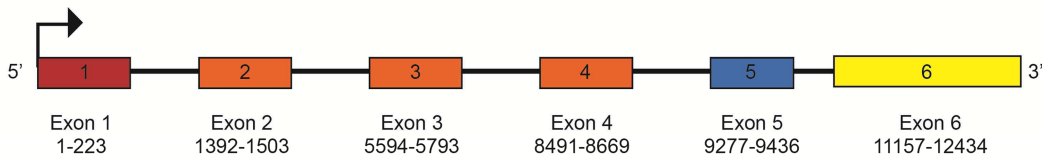
B) Exon composition of transcription variants

Primary Transcript

Membrane bound Tissue Facor
263 Amino Acids

Alternative Transcript

Alternative Tissue facor
206 Amino Acids
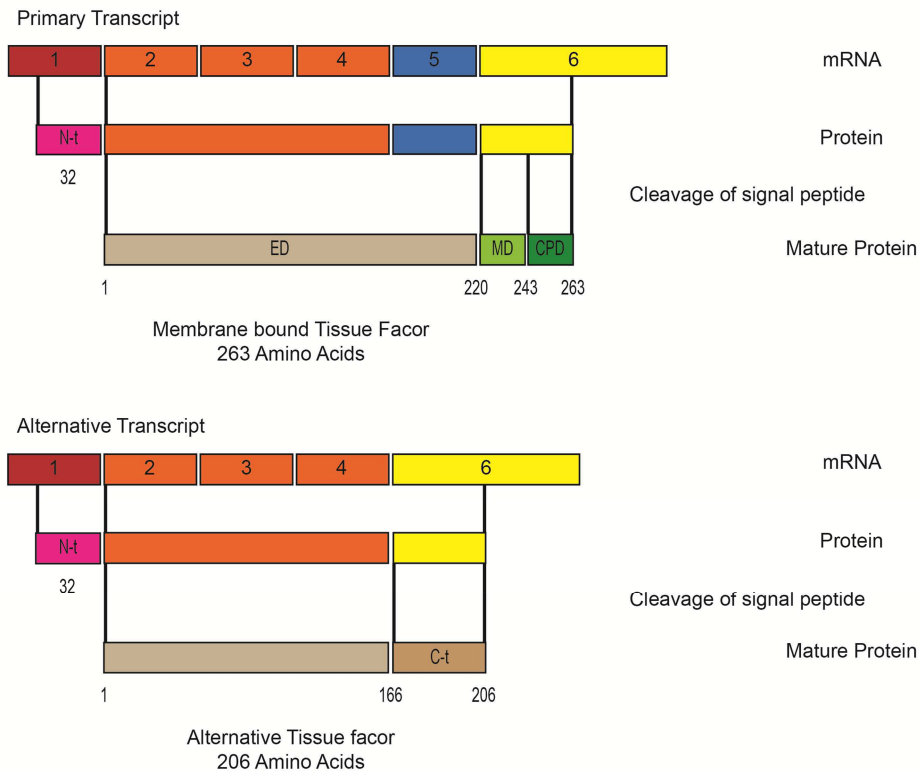
Figure 2.2 (A): Tissue factor gene structure

Structure of the human TF (*F3*) gene located on chromosome 1 p21-22. An arrow indicates the start of transcription site. Exons are numbered 1 to 6 and base pair lengths are indicated. Exon 1 (■) encodes for a 32 amino acid signal peptide. Exons 2-4 (■) and exon 5 (■) encode for the extracellular domain which binds to coagulation factors VII and X. Exon 6 (■) encodes for the transmembrane domain and cytoplasmic tail.

Figure 2.2 (B): Exon composition of transcript variants

Exons 1 to 4 are identical in both transcripts. The alternative transcript (Accession no: AF487337) compared to the primary transcript (Accession no: NM_001993), excludes exon 5 (■) due to the splicing of exon 4 directly in to exon 6. Both transcripts possess a 32 amino acid N-terminus (Nt- ■), which serves as a signal peptide in intercellular transport. The primary transcript consists of 3 domains; a 219 amino acid extracellular domain (ED- ■), a 23 amino acid membrane domain (MD- ■) and a 21 amino acid cytoplasmic domain (CDP- ■). The alternative transcript consists of 206 amino acids of which 1 to 166 are identical to the primary transcript and amino acid 167-206 form a unique C-terminus (Ct- ■).

Compiled from Carmeliet and Collen (1998); Rauch and Nemerson (2000); Bognanov *et al.* (2003); Förster *et al.* (2006); Rauch *et al.* (2006); Censarek *et al.* (2007); Hobbs *et al.* (2007) and Mackman (2007).

Although it was originally believed that TF_{FL} was mainly present on extra vascular cells, it has also been detected circulating freely in association with microparticles in healthy subjects (Censarek *et al.*, 2007). These tissue factor bearing microparticles (TF_{MP}) are sub-micrometric (0.1 - 1µm) membrane fragments that are constitutively released from the surface of cells (George, 2008). These membrane fragments lack a nucleolus and are incapable of protein synthesis (Lechner and Weltermann, 2008). Cell activation, apoptosis and shear stress up-regulates the formation of microparticles (Nomura *et al.*, 2008). TF_{MP} consists of an anionic phospholipid surface and as a result is highly pro-thrombotic (Biró *et al.*, 2003). TF_{MP} is believed to circulate in an "encrypted" form (Mezzano *et al.*, 2008). This "encrypted form" prevents the inappropriate activation of the coagulation cascade (Mackman, 2006). Although microparticles originate from various cell types the majority are derived from platelets, leucocytes, erythrocytes and endothelial cells (Biró *et al.*, 2003; George, 2008). The role of TF_{MP} in the coagulation is summarised in Figure 2.3.

The alternative transcript, also referred to as alternatively spliced tissue factor (TF_{AS}), is expressed as a 206 residue soluble protein (See: Figure 2.2 (B)) (Bognanov *et al.*, 2003; Butenas *et al.*, 2005; Vieira *et al.*, 2007; Mackman, 2008). During transcription exon 4 is spliced directly into exon 6 which results in an alteration of the reading frame (Brüggeman *et al.*, 2006). Amino acid residues 1-166 of TF_{AS} are identical to the main transcript, while residues 167-206 represent a unique Carboxy-terminus (C-terminus) (Mackman, 2004; Mackman, 2007). This alteration to the protein structure results in the loss of the transmembrane domain (Rauch *et al.*, 2006). As a result, TF_{AS} is unlikely to be an effective initiator of coagualation, since an assossiation with a membrane is required in order to achieve full pro-coagulant activity (Sturk-Maquelin *et al.*, 2003; Mackman and Morrissey, 2008). Some research groups proposed that TF_{AS} is inactive while circulating within blood and only becomes pro-coagulant when it co-localizes with platelets and is exposed to phospholipids (Bognanov *et al.*, 2003).
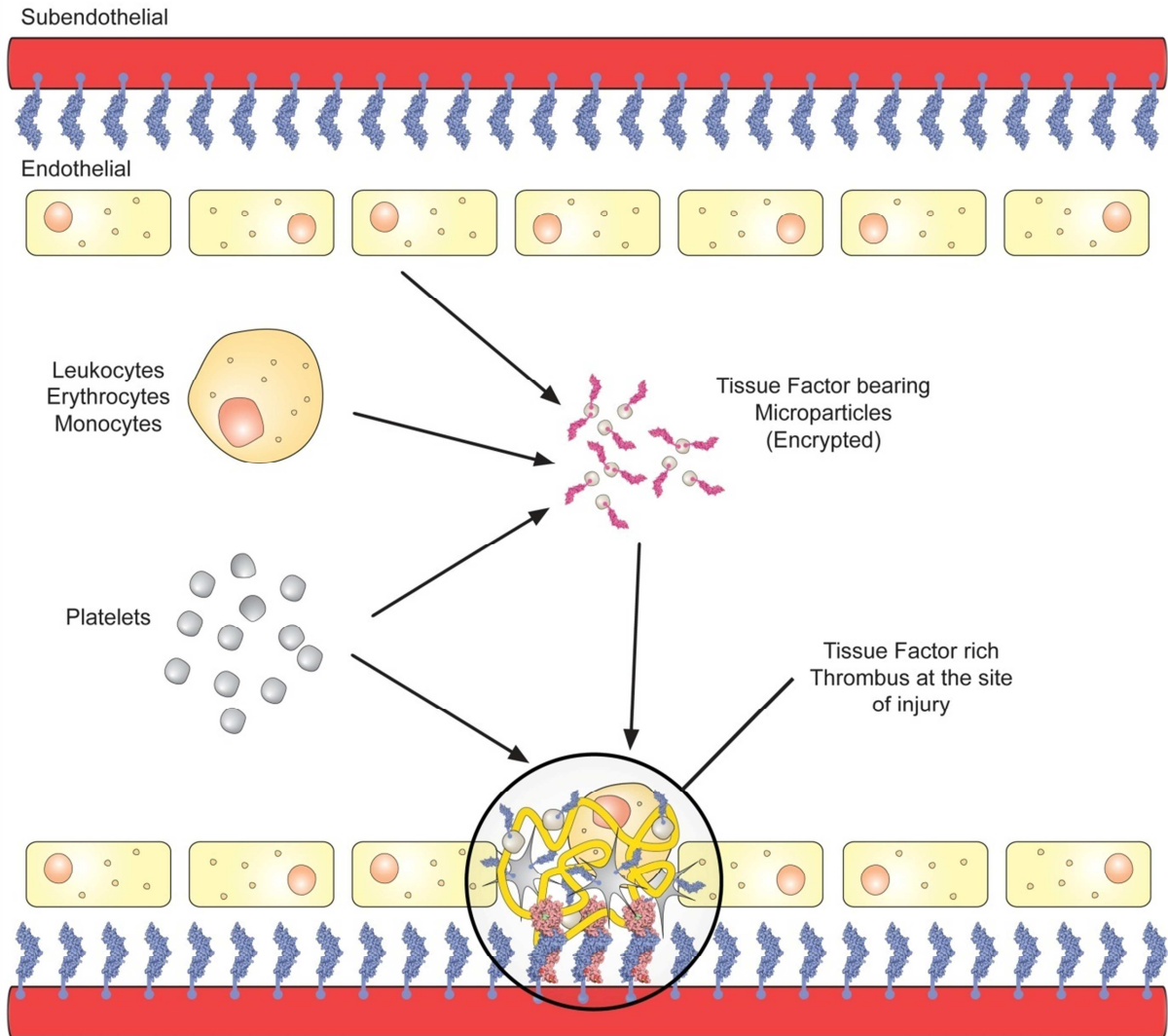
Figure 2.3: Tissue factor at the site of injury

Upon vessel damage, coagulation is initiated by constitutively expressed extra-vascular tissue factor. As the thrombus forms a tissue factor rich micro-environment, it is maintained by various sources of blood borne tissue factor. Platelets express TF_{FL} upon activation, inducing tissue factor expression in surrounding cells when incorporated into the thrombus and contribute to the pool of encrypted TF_{MP}. Monocytes, leukocytes and erythrocytes also contribute to the pool of encrypted TF_{MP}. The encrypted tissue factor becomes de-encrypted upon adhesion and fusion to active platelets. This result in a fibrin rich thrombus restricted to the site of vascular injury.

Compiled from Rauch and Nemerson (2000); Engelmann (2006); Mackman (2006); Butenas *et al*. (2008); Lechner and Weltermann (2008); Mezzano *et al*. (2008) and Nomura *et al*. (2008).

2.3 Tissue factor in coagulation

Under normal physiological conditions blood circulates as a liquid, however upon vascular damage, coagulation is rapidly initiated in order to prevent haemorrhaging (McVey, 1999; Norris, 2003). The current scheme of coagulation involves a series of cascade reactions which are divided into two phases; initiation and amplification (See: Figure 2.4) (Price *et al.*, 2004; Frédérick *et al.*, 2005).

Tissue factor initiates coagulation through its interaction with Factor VII (Dahlbäck, 2000; Price *et al.*, 2004). Factor VII is a vitamin K dependent plasma protein circulating freely in both active (1%, Factor VIIa) and the inactive (99%, Factor VII) forms within the blood (Hellstern *et al.*, 1997; Norris, 2003). Upon vessel damage exposed tissue factor binds to both forms of Factor VII and develops an active tissue factor-Factor VII$_a$ complex. This complex is the most potent activator of the coagulation cascade known (Norris, 2003). This active complex activates both Factor IX and Factor X via proteolytic cleavage. Activated Factor IX (Factor IX$_a$) in association with Factor VIII is responsible for the activation of additional Factor X (Dahlbäck, 2000; Butenas *et al.*, 2005). Factor X$_a$ combines with Factor V to form a prothrombinase complex. This complex is located on the surface of membranes and catalyzes the conversion of prothrombin to thrombin (Frédérick *et al.*, 2005; Vos, 2006). The initiation phase of coagulation is responsible for the generation of a minute amount of thrombin (< 4 % of total thrombin) (Nemerson, 1988; Dahlbäck, 2000; Butenas *et al.*, 2007).

The amplification of the thrombin occurs via a positive feedback mechanism where the initial amount of thrombin activates factors V, VIII and XI (Gresley and Agnelli, 2002; Butenas *et al.*, 2003). This results in the generation of large quantities of thrombin, *via* the same mechanism originally initiated by tissue factor (Meiring *et al.*, 2002; Butenas *et al.*, 2007). This large quantity of thrombin is required for the rapid formation of the fibrin thrombus clot through the activation of fibrinogen and factor XIII (Frédérick *et al.*, 2005). Activated fibrin assists in the activation of additional factor XIII which stabilizes the thrombus through the cross-linking of the fibrin fibres (Dahlbäck, 2000).
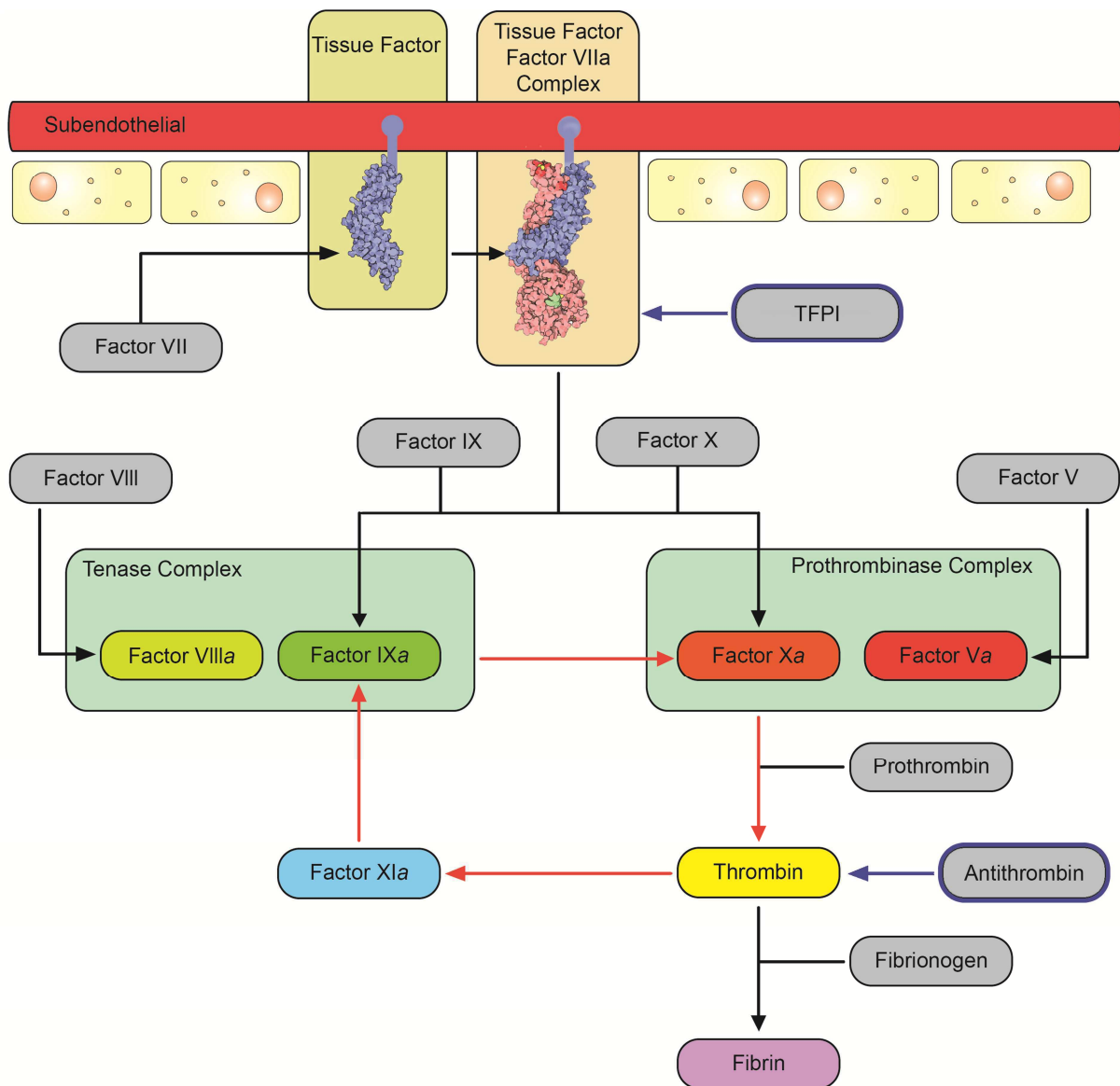
Figure 2.4: The extrinsic coagulation cascade

The interaction of coagulation factors within the tissue factor pathway of coagulation. The Cascade is initiated through the interaction of tissue factor and Factor VII. The thrombin initiated feedback mechanism is responsible for the generation of > 96 % of total thrombin is indicated by red arrows. The natural anti-coagulant pathway consisting of Tissue Factor pathway inhibitor (TFPI) and Antithrombin are outlined in blue

Derived from Frédérick *et al.* (2005) and Butenas *et al*. (2007).

Thrombin also activates and recruits additional tissue factor-bearing platelets into the thrombus in order to ensure a tissue factor and platelet rich environment that is restricted to the site of vascular injury (Bates and Weitz, 1998; Elkelboom *et al*., 2003; Hirsh, 2003).

The underlying physiological function of the tissue factor pathway is the localized generation of large quantities of thrombin in response to vascular trauma, in order to prevent bleeding (Fredenburg *et al.*, 2004). Although thrombin plays an essential role in coagulation, it is however important to remember that it is tissue factor that is responsible for the initiation of the coagulation cascade (Nemerson, 1988). Elevated levels of tissue factor expression have been implicated in the pathogenesis of various thrombotic disorders (Steffel *et al.*, 2006b). The next section describes the patho-physiology of tissue factor.

2.4 Pathlogical role of tissue factor

2.4.1 Insulin- and obesity associated thrombosis

An increase in tissue factor expression and activity has been implicated in the pathogenesis of thrombotic and cardiovascular anomalies often associated with metabolic disorders. High glucose concentrations in blood lead to a thrombin-induced expression of tissue factor in endothelial cells (Steffel *et al.*, 2006b). Hyperglycaemic conditions have also been associated with an increase in tissue factor expression in monocytes, elevated levels of tissue factor-bearing microparticles and an increase in activity of circulating tissue factor (Kern *et al.*, 2001; Diamant *et al.*, 2002; Lim *et al.*, 2004)*.* This aberrant expression, accompanied with increased tissue factor activity is believed to be responsible for high cardiovascular mortality and thrombotic complications observed in type 2 diabetes mellitus (Sommeijer *et al.*, 2006). Abnormal glucose homeostasis such as insulin resistance and hyper-insulinemia is quite often associated with obesity, a far more common metabolic disorder. Obesity affects more than 30% of the population in North America. It is characterized by the excessive increase of adipose tissue mass. Obesity leads to increased expression of tissue necrosis factor alpha (TNF-α) and interleukin 6 (IL-6) by adipose tissue, which in turn elevates the level of tissue factor expression in intravascular- and endothelial cells (Hotamisligil *et al.*, 1995; Kern *et al.*, 2001; Samad *et al.*, 2001). Obese patients suffering from hypercholesterolemia are at an even greater risk of developing thrombotic complications. High levels of low density lipoprotein and very low density lipoprotein have been shown to stimulate both the secretion of tissue factor as well as the activation of the tissue factor/Factor

VIIa complex (Nofer *et al.*, 2002; Llorente-Cortés *et al.*, 2004). This leads to an increase in tissue factor expression and enhances thrombotic tendency which contributes to cardiovascular complications (Samad *et al.*, 2001; Levi and van der Poll, 2005).

2.4.2 Tissue factor in atherosclerosis

Tissue factor contributes to the patho-physiology of atherosclerosis and acute coronary syndromes (unstable angina and myocardial infarction) that are associated with plaque rupture (Giesen *et al.*, 1999; Gresley and Agnelli, 2002; Marsik *et al.*, 2003; Viles-Gonzalez and Badimon, 2004).

Several factors contribute to the excessive expression of tissue factor during atherosclerosis. Inflammatory cytokines and interleukins are closely associated with atherosclerosis and are potent inducers of tissue factor expression in monocytes during the early stages of atherogenesis. During later stages excessive tissue factor expression is also detected in foam cells, endothelial cells and smooth muscle cells (Levi *et al.*, 2005; Lösche, 2005; Migdalski *et al.*, 2005; Steffel *et al.,* 2006b). Blood borne tissue factor levels are also increased due to the release of inflammatory cytokines (Morrissey, 2003; Steffel *et al.,* 2006b). Inflammation further enhances the thrombogenicity of atherosclerotic plaque through the down regulation of the physiological anticoagulant systems (Levi *et al.*, 2005). Very high levels of tissue factor are found within an atherosclerotic plaque. This is due to excessive tissue factor expression by macrophage derived foam cells, monocytes, vascular smooth muscle cells, endothelial cells as well as the presence of apoptotic microparticles of monocytic and lymphocytic origin. All of these cell types are abundant within the atherosclerotic plaque (Marsik *et al.*, 2003; Viles-Gonzalez and Badimon, 2004; Lösche, 2005; Migdalski *et al.*, 2005).

Once plaque rupture occurs the highly thrombotic core is exposed to the circulating clotting factors in the blood stream which lead to the rapid and excessive thrombin generation *via* tissue factor coagulation pathway (Morrissey, 2003; Viles-Gonzalez and Badimon, 2004; Migdalski *et al.*, 2005). The tissue factor-rich necrotic core is exposed to the circulatory coagulation factors and platelets resulting in the initiation

of coagulation (Rauch *et al.*, 2000; Lösche, 2005; Migdalski *et al.*, 2005). Platelets contribute to the thrombogenicity by inducing tissue factor expression on the surface of leukocytes and especially monocytes. The monocytes in turn transfer tissue factor-bearing particles back to platelets via CD15 interaction. These tissue factor-rich platelets bind to the sub-endothelial matrix of the ruptured atherosclerotic plaque. This provides a continuous supply of tissue factor to the already tissue factor-saturated necrotic core (Rauch *et al.*, 2000; Morrissey, 2003; Lösche, 2005). Activated platelets also contribute to the pool of blood borne tissue factor by shedding parts of their plasma membrane as tissue factor-bearing microparticles (Lösche, 2005). This excessive amount of tissue factor activity is responsible for the production of large quantities of thrombin which leads to arterial thrombosis commonly observed during atherosclerosis (Giesen *et al.*, 1999; Migdalski *et al.*, 2005). Tissue factor is also considered as the principle molecule linking fibrous plaque disruption to thrombo- and pulmonary embolism (Giesen *et al.*, 1999; Kamihura *et al.*, 2005).

2.4.3 Malignancy and thrombosis

Although the molecular mechanism still remains unclear, there is a well established association between certain malignancies and thrombosis (Levine *et al.*, 2003; Zwicker *et al.*, 2007). The presence of a malignant disease increases the risk of disseminated intravascular coagulation (DIC) and venous thromboembolism (VE) (Hembroug *et al.*, 2003; Nadir *et al.*, 2006; Tilly *et al.*, 2008). The risk of thrombotic complications is further increased with the administration of hormone replacement- and chemotherapies (Kakkar *et al.*, 2003; Zwicker *et al.*, 2007). Zwicker *et al.* (2007) reported that 25 % of cancer patients suffer from deep vein thrombosis postoperatively. The pathophysiology involves an intricate two-way interaction between the malignancy and the coagulation system (Hembroug *et al.*, 2003). The presence of malignant disease significantly enhances thrombotic tendency, which in turn sustains and even in some cases promotes tumour growth. Tissue factor is the main pro-coagulant responsible for the thrombotic tendencies observed during malignancies (Hembroug *et al.*, 2003; Nadir *et al.*, 2006; Zwicker *et al.*, 2007; Hoffman, 2008). A variety of malignant cells exhibit enhanced tissue factor activity through excessive expression of tissue factor on lipid rich surfaces (Levine *et al.*,

2003; Hembroug *et al.*, 2003). This partially explains why the risks of thrombotic complications are not uniform across all forms of tissue factor expressing malignant cells (Zwicker *et al.*, 2007).

Pulmonary embolism is common in ovarian, pancreatic, brain and bone cancers although not in head and neck cancers (Kakkar *et al.*, 2003; Zwicker *et al.*, 2007). Tissue factor is also constitutively expressed by human leukaemia- and lymphomas cells (Levine *et al.*, 2003; Nadir *et al.*, 2006). The expression of tissue factor in tumour cells induces the expression of vascular endothelial growth factor (VEGF), which in turn further enhances tissue factor expression (Levine *et al.*, 2003; Zwicker *et al.*, 2007; Hoffman, 2008). Many malignant cells elicit an inflammatory host response that stimulates tissue factor expression in monocytes via pro-inflammatory cytokines (Lee *et al.*, 2006; Zwicker *et al.*, 2007). Elevated levels of platelet- and tumour derived microparticles also contribute to thrombogenicity during cancer (Zwicker *et al.*, 2007; Tilly *et al.*, 2008). This aberrant expression coupled with an increase in blood borne tissue factor expression and enhanced thrombotic tendency contribute to pathological complications associated with malignancies.

2.5 Therapeutic Importance of Tissue Factor

Tissue factor provides an ideal target for the development of novel anticoagulants due to its fundamental role in thrombotic disorders (Frédérick *et al.*, 2005). Tissue factor targeting agents interfere with the very first step of coagulation, thereby suppressing all downstream effectors as well as preventing a multitude of other tissue factor-mediated effects (Golino, 2002; Steffel *et al.*, 2006a). There are several theoretical advantages to tissue factor inhibition. Firstly, the inhibition of the initiation step of coagulation would hamper the production of the initial amount of thrombin and thus prevent the positive feedback mechanism responsible for thrombin amplification. Secondly, inhibitors of tissue factor may not affect the integrity of physiological haemostasis as a whole, as they will only function at the site of vascular injury where tissue facor is exposed (Golino, 2002). Thirdly, the inhibition of tissue factor has reduced bleeding tendency in comparison to "first-generation" anticoagulant drugs (Frédérick *et al.*, 2005). Furthermore, it is believed that

relatively low concentrations of tissue factor inhibitor would be required to sufficiently suppress overactive coagulation.

2.6 Anticoagulants

Currently the most commonly prescribed antithrombotic drugs are warfarin, heparin and low molecular weight heparin. These drugs have been in use for over fifty years and have been invaluable in post surgery- and long term treatment of thrombosis (Ardit *et al.*, 2006; Copeland and Cheryl, 2009). Despite the popularity and effectiveness of these antithrombotic drugs, treatment is still far from ideal due to clinical as well as practical drawbacks (Keeling, 2006).

The antithrombotic effect of warfarin is mediated by vitamin K dependant coagulation factors (Prothrombin, factors VII, -IX and -X) and is therefore very sensitive to dietary fluctuations of vitamin K (Mueller, 2004; Seljeflot *et al.*, 2004). The efficacy of the drug is also influenced by various other factors such as; genetic polymorphisms that affect the rate at which the drug is metabolized, a slow onset of action (five hours to several days), a narrow therapeutic range as well as many drug to drug and drug to food interactions (Mueller, 2004; Ardit *et al.*, 2006; Copeland and Cheryl, 2009; Weitz, 2009). As a result, patients undergoing warfarin treatment require close dosage monitoring in order to maintain desired international normalised ratio (INR) (Francis, 2004). Although bleeding is the more severe complication associated with warfarin treatment, other adverse effects include skin and adipose tissue necrosis (Coumadin-induced skin necrosis), allergic or hypersensitivity reactions, gastrointestinal discomfort, oedema, and fetal abnormalities to name just a few (Mueller, 2004).

Unfractionated heparin acts as an antithrombotic by enhancing the ability of antithrombin to neutralize the serine active site of thrombin as well as factors IXa, -Xa, -XIa, -XIIa, -XIIIa and plasmin (Mueller, 2004; Franchini, 2005). Heparin however binds non-specifically to platelets, endothelial cells, macrophages as well as plasma proteins (platelet factor 4 and high molecular weight von Willebrand multimeres). This leads to a reduction in the amount of viable heparin available to interact with functional antithrombin (Mueller, 2004; Frédérick *et al.*, 2005). In addition, heparin as

well as the low molecular weight heparin-antithrombin complex is unable to target thrombin bound to fibrin, factor Xa bound to platelets or the prothrombinase complex (Weitz and Crowther, 2002). Some of the adverse effects of heparin treatment include bleeding, heparin induced thrombocytopenia (HIT), osteoporosis, renal dysfunction and hematoma or skin necrosis (Mueller, 2004; Franchini, 2005). As a result heparin treatment also requires careful monitoring to ensure correct dosage (van Veen *et al.*, 2008). This does not only increase the overall cost of treatment but also contributes to poor patient compliance (Hirsh, 2003; Ardit *et al.*, 2006).

Although these anticoagulants have been in use for many years and despite their relative effectiveness, there is still a great need for the development of safe and effective therapies for long term treatment of thrombosis. Due to the role of tissue factor as the primary initiator of the coagulation cascade, its inhibition provides a novel antithrombotic strategy which may address many of the complications associated with current anticoagulants.

2.7 Inhibition of tissue factor

Many different therapeutic approaches have been explored in order to develop antithrombotic agents to target tissue factor as well as the tissue factor/Factor VIIa complex. The following section is an overview of the different inhibition strategies that have been explored. The current therapeutic approaches to the inhibition of tissue factor are summarised in Figure 2.5.

2.7.1 Inhibition of tissue factor synthesis

The inhibition of tissue factor synthesis at DNA- and mRNA level has been increasingly explored in recent years. One novel mechanism for the inhibition of tissue factor is through the use of triplex forming oligonucleotides. These short DNA sequences bind to the major groove of the double stranded DNA via hydrogen bonding. The resulting tri-stranded DNA prevents the binding of transcription factors and the transcription of the gene. Tissue factor levels are reduced by targeting the major groove in the promoter region of the tissue factor gene (Qianning *et al.*, 2008). A similar approach is the use of hybridising antisense oligonucleotides that bind to

tissue factor mRNA and prevent gene translation (Stephens and Rivers, 1997; Steffel *et al.*, 2006a).

A hairpin ribozyme that destroys the tissue factor mRNA in vascular smooth muscle has also been developed (Steffel *et al.*, 2006b). An indirect method of inhibiting tissue factor synthesis is by the use of curcumin. It is a natural occurring pigment which suppresses the expression of TNF-α transcription factors. This relates to a reduction in tissue factor expression in endothelial cells during inflammation (Steffel *et al.*, 2006b). Although these inhibitors are relatively effective *in vitro*, further research is required to determine their effectiveness *in vivo*.

2.7.2 Direct inhibition of tissue factor action

Apart from a multitude of synthetic peptide tissue factor inhibitors, a large array of tissue factor targeting polyclonal, monoclonal and antibody fragments have also been developed. These antibody inhibitors vary in their size, binding site and ability to inhibit the function of tissue factor (Gresle and Agnelli, 2002; Jacquemin and Saint-Remy, 2004). Interestingly, the relative efficacy of the antibodies is related to binding site location rather than binding affinities or rate of association (Jacquemin and Saint-Remy, 2004). Many of the various monoclonal antibodies were found to elicit antithrombotic effective in animal models, while some polyclonal antibodies were found to reduce thrombogenicity of ruptured atherosclerotic plaques in humans (Steffel *et al.*, 2006b). In an escalating dosage study a chimeric monoclonal anti-tissue factor antibody reduced thrombin production in patients diagnosed with stable coronary disease (Steffel *et al.*, 2006a).

 An alternative form of tissue factor, which exhibits reduced catalytic activity, has also been developed (Steffel *et al.*, 2006b). Recombinant tissue factor is a soluble protein similar to the extracellular domain of the wild type tissue factor, except for the alanine substitution of two lysine residues. Recombinant tissue factor displays clear antithrombotic properties without disturbing normal haemostasis and do not cause bleeding complications in animal models. Redesign of recombinant tissue factor has produced an increased affinity to Factor VIIa and a 20 fold higher antithrombotic activity than the original recombinant tissue factor (Frédérick *et al.*, 2005).
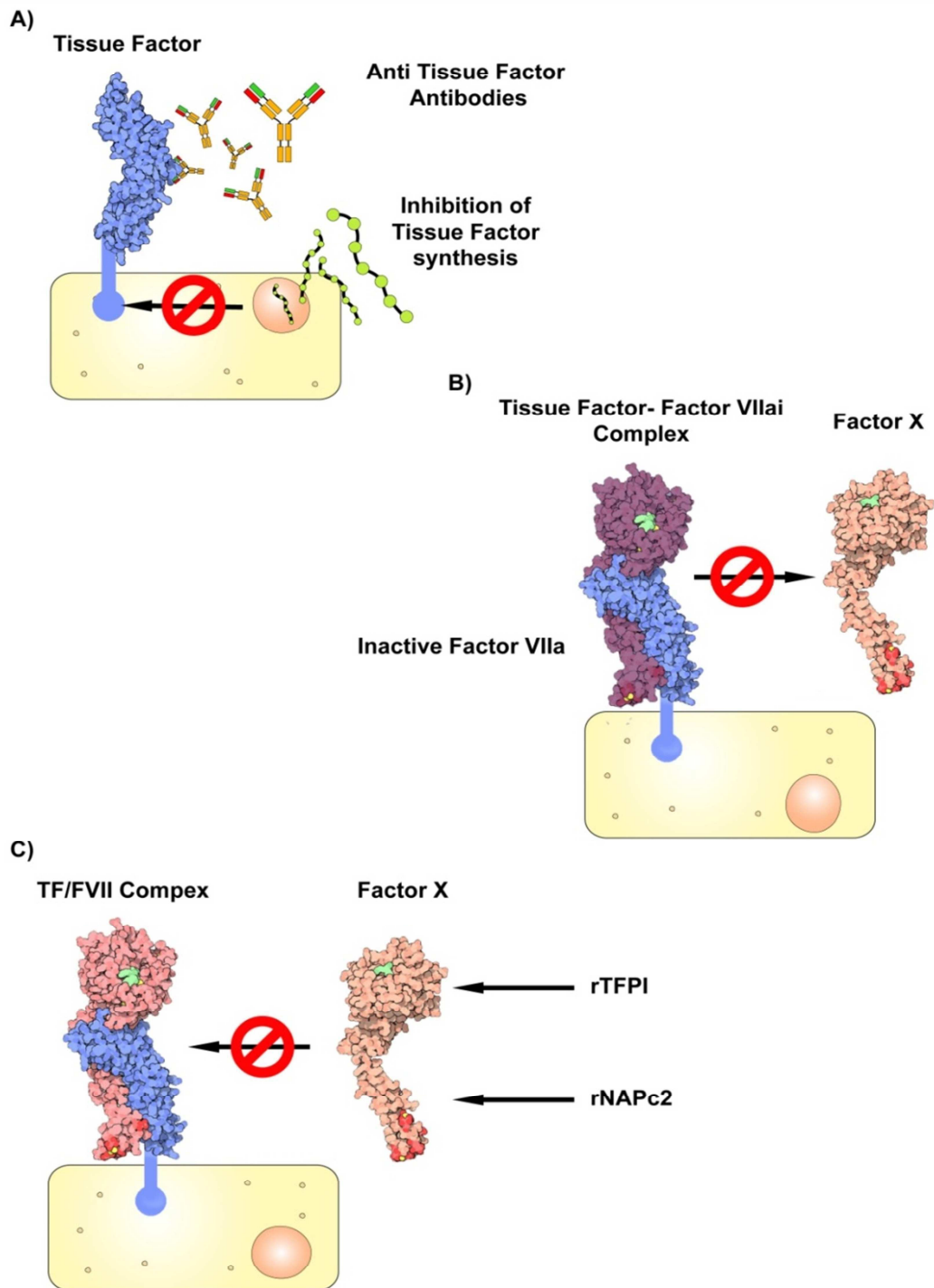
Figure 2.5: Therapeutic approaches to interfere with tissue factor

Several alternative therapeutic strategies have been developed to interfere with the actions of tissue factor. A) Different gene-silencing molecular strategies have been applied in order to prevent tissue factor synthesis as well as an array of mono- and polyclonal anti-tissue factor antibodies. B) Catalytic inactive Factor VII (Factor VIIai) has been developed that renders the tissue factor-factor VII complex is unable to activate Factor X and Factor IX. C) Recombinant tissue factor Pathway Inhibitor (rTFPI) and recombinant Nematode Anticoagulant Protein c2 (rNAPc2) bind with Factor X and prevent activation by the tissue factor-Factor VII(a) complex.

Derived from Steffel *et al.* (2006a).

2.7.3 Active site inactivated factor VIIa

Active Site Inactivated Factor VIIa (Factor VIIai) is capable of binding to tissue factor but it lacks catalytic activity (Steffel *et al.*, 2006b). Factor VIIai elicits an anticoagulant effect by competing with catalytically functional FVII for tissue factor (Gresle and Agnelli, 2002; Hirsh, 2003). Interestingly, incorporation of a Factor VIIa active site inhibitor leads to a 5 fold increase in binding affinity (Frédérick *et al.*, 2005; Carlsson *et al.*, 2006). This significantly enhances the efficacy of FVIIai inhibitor. Factor VIIai also displays good anticoagulant capabilities in baboon sepsis models as well as other animal models (Hirsh, 2003; Jacquemin and Saint-Remy, 2004; Frédérick *et al.*, 2005; Steffel *et al.*, 2006b). In clinical trials FVIIai also displayed dose-dependent inhibition of *ex-vivo* thrombin formation and fibrin deposition on perfusion chambers. In addition it also displayed prolonged anticoagulant effect (> 24h) following single intravenous injection (Frédérick *et al.*, 2005).

2.7.4 Recombinant tissue factor pathway inhibitor

Recombinant tissue factor Pathway Inhibitor (rTFPI) functions in a similar mechanism to physiological TFPI by inactivating Factor X and forming a complex that continues to deactivate tissue factor/factor VIIa complex (Gresle and Agnelli, 2002). Recombinant TFPI was found to elicit an antithrombotic effect during balloon injury in animal models and reduced atherosclerotic plaque thrombogenicity in humans (Steffel *et al.*, 2006b). Similar results were also obtained with the use of adenovirus gene transfer of rTFPI. The over-expression of rTFPI provided a clear antithrombotic effect without affecting systemic coagulation (Steffel *et al.*, 2006b). Although initial results were positive in various settings, phase III studies in patients suffering from sepsis and mild coagulopathy failed to reveal the therapeutic benefit of rTFPI due to increased bleeding tendencies (Steffel *et al.*, 2006b). Further studies are required to evaluate the therapeutic safety and efficacy of rTFPI.

<u>2.7.5 Nematode anticoagulant protein c2</u>

Recombinant Nematode Anticoagulant c2 (rNAPc2) is a potent inhibitor of the tissue factor/Factor VIIa complex (de Pont *et al.*, 2004). It has a long half-life and only needs to be administered every other day after subcutaneous injection (Gresle and Agnelli, 2002). rNAPc2 was originally identified and isolated from the canine hookworm *Ascylostoma caninum* (Frédérick *et al.*, 2005). The recombinant form (rNAPc2) is currently the most extensively evaluated novel inhibitor of the tissue factor/Factor VIIa complex (Gresle and Agnelli, 2002). rNAPc2 interferes with tissue factor function by binding with both Factor X or Xa before forming an inhibitory complex with tissue factor/Factor VIIa (Steffel *et al.*, 2006b). Phase I studies in the prevention of venous thromboembolism in patients who received unilateral total knee replacement surgery produced positive results with a lowered rate of deep vein thrombosis without bleeding complications (Gresle and Agnelli, 2002; Frédérick *et al.*, 2005). Other phase I studies revealed that administration of rNAPc2 completely blocked endotoxin induced thrombin generation without effecting key fibrinolytic modulators such as plasmin-antiplasmin complexes and plasminogen activator inhibitor type 1 (de Pont *et al.*, 2004). rNAPc2 was also found to be safe and effective in phase II studies investigating escalating doses in low-risk cardiac patients undergoing angioplasty (Frédérick *et al.*, 2005; Steffel *et al.*, 2006b). rNAPc2 harbours a lot of potential as a novel anticoagulant therapy and may prove to resolve many of the shortcomings of current anticoagulants.

2.8. Introduction to Current Study

Currently available monoclonal antibodies are primarily of rodent origin (Steffel *et al.*, 2006b). The use of rodent antibodies in human therapy presents numerous problems, the most significant of which is immunogenicity. The use of human monoclonal antibodies would address this limitation, but it has proven difficult to generate large amounts of human monoclonal antibodies by conventional technologies (Nicolaides *et al.*, 2006). By using the technology of antibody phage display, large amounts of antibody fragments can be amplified in *Escherichia coli* strains and purified using antibody purification chromatography.

Bacteriophages or "phages" are viruses that specifically infect and replicate within bacteria (Smith and Petrenko, 1997; Cox, 2001). Phages are used as molecular expression vectors as they are able to accommodate foreign DNA in the phage genome (Smith and Petrenko, 1997; Willats, 2002). As the phage infects and replicates within the bacterial host, the inserted foreign DNA is expressed as well (Kennel *et al.*, 2004).

The technology of phage-display was first introduced by G. Smith (Azzazy and Highsmith, 2002). Phage-display differs from conventional expression vectors, as the foreign DNA is spliced into a specific gene coding for the expression of surface or "coat" protein of the phage particle (Nissim *et al.*, 1994; Smith and Petrenko, 1997). Generally, foreign DNA is spliced into the coat protein III (pIII), although foreign DNA can also inserted into coat protein VI or VIII (pVI or pVIII) of the M13 bacteriophage (Hoogenboom *et al.*, 1998; Kretzschmar and von Rüden, 2002). This incorporation of foreign DNA into the gene of a coat protein leads to the expression or "display" of a hybrid or "fusion" protein on the outer surface of the phage particle (Smith and Petrenko, 1997; Ravn *et al.*, 2004). Due to the position of the displayed protein and it subsequent accessibility to the solvent, it behaves essentially as it would if it were totally detached from the phage surface (Smith and Petrenko, 1997). In order to maximize phage display ability to identify a specific protein with a desired phenotype, it utilizes large phage repositories or "phage libraries". These libraries consist of a large number of phages which express a vast variety of fusion proteins or peptides on their surfaces (up to 10^{10} fragments) (Azzazy and Highsmith, 2002). The ability to screen the large array of fragments coupled with the ability to unite the phenotypic functionality with the genetic information of the desired fusion protein within a single phage particle, makes phage display ideally suited for the identification and subsequent modification of a desired protein, peptide or antibody (Smith and Petrenko, 1997; Hoogenboom *et al.*, 1998; Azzazy and Highsmith, 2002; Kretzschmar and von Rüden, 2002; Willats, 2002).

Furthermore, antibodies and antibody fragments represent the fastest growing segment of the biopharma market (Nicolaides *et al.*, 2006). This is mainly due to their drug safety profiles and because it could elicit clinical benefits by antagonizing a specific antigen without the common side effects that are prevalent with small

chemical entities due to their non-specific effects on homeostatic biochemical pathways.

Keeping this in mind, our research group utilized phage display technology to explore the potential of the inhibition of tissue factor as novel antithrombotic therapy in a previous study. We identified a 26 kDa scFv from the Tomlinson I + J phage libraries (obtained from MRC HGMP Resource Centre, UK) that inhibits the coagulatory function of tissue factor. Preliminary *in vitro* testing showed that the scFv, from here on referred to as Tissue Factor Inhibitor scFv (TFI-scFv), extended prothrombin times in a dose dependent manner and reduced the maximum amount of thrombin generated in human plasma (Meiring *et al.*, 2009).

Although we were able to produce sufficient quantities of TFI-scFv in order to perform preliminary analysis, yields were low. The low yields were due to the fact that phage display technique is primarily concerned with the identification of a specific scFv and although it was possible to express soluble scFv, the system was never intended for the large scale production. Apart from the low yields it was also expensive and time consuming to produce soluble scFv. This was mainly due to the reliance on protein A affinity chromatography for purification. Protein A isolated from *Staphylococcus aureus* has a very high binding specificity for human immunoglobulin (Björck and Kronval, 1984). Although this property of protein A affinity chromatography is well suited for the purification of scFv's there are a number of disadvantages. Protein A affinity chromatography is expensive, it suffers from difficulties associated with scale-up and is highly labour intensive and is time consuming. In addition, due to the biological nature of protein A, it rapidly loses ability to bind the target protein. As a result protein A columns have a limited lifetime, which add to the overall cost of the purification process.

Consequently, protein A affinity chromatography is not ideally suited for our applications as we would need to produce and purify larger volumes of TFI-scFv than previously produced. The ability to obtain higher concentrations of TFI-scFv is of great importance as it would enable us to perform additional *in vitro* analysis, investigate the anticoagulant effect and safety under higher dosages and enable us as well to examining the anticoagulant efficacy of TFI-scFv in animal models.

The aim of the current study is the optimization of an expression mechanism with the goals of firstly producing higher concentrations functional soluble TFI-scFv and secondly to reduce the cost involved in the purification process.

2.9 References

ARDIT, E., GOLDBERG, M., GOMEZ-ORELLANA, I. & MAJURU, S. 2006. Oral heparin: status review. *Thrombosis Journal*, 4, 6.

AZZAZY, H.M.E. & HIGHSMITH, W.E. 2002. Phage display technology: clinical applications and recent innovations. *Clinical Biochemistry*, 35, 425-445.

BATES, S.M. & WEITZ, J.I. 1998. Direct Thrombin Inhibitors for Treatment of Arterial Thrombosis: Potential Differences Between Bivalirudin and Hirudin. *American Journal of Cardiology*, 82, 12P-18P.

BECK, L. & D'AMORE, P.A. 1997. Vascular development: Cellular and Milecular Regulation. *FASEB*, 11, 365-373.

BELTING, M., DORRELL, M., SANDGREN, S., AGUILAR, E., AHAMED, J., DORFLEUTNER, A., CARMELIET, P., MEULLER, B., FRIEDLANDER, M. & RUF, W. 2004. Regulation of angiogenesis by tissue factor cytoplasmic domain signalling. *Nature Medicine*, 10, 502 – 509.

BIRÓ, E., STURK-MAQUENLIN, K.N., VOGEL, G.M.T., MEULEMAN, D.G., SMIT, M.J., HACK, C.E., STRUK, A. & NIEUWLAND, R. 2003. Human cell-derived microparticles promote thrombus formation *in vivo* in tissue factor-dependent manner. *Journal of Thrombosis and Haemostasis*, 1, 2561-2568.

BJÖRCK, L. & KRONVAL, G. 1984. Purification and some properties of streptococcal protein G, a novel IgG-binding reagent. *The Journal of Immunology*, 133, 969-975.

BOGNANOV, V.Y., BALASUBRAMANIAN, V., HATHCOCK, J., VELE, O., LIED, M. & NEMERSON, Y. 2003. Alternatively spliced human tissue factor: a circulating, soluble, thrombogenic protein. *Nature Medicine*, 9, 458-462.

BRÜGGEMAN, L.W., DRIJFHOUT, J.W., REITSMA, P.H. & SPEK, C.A. 2006. Alternatively spliced tissue factor in mice: induced by Streptococcus pneumonia. *Thrombosis and Haemostasis*, 4, 918-920.

BUGGE, T.H., XIAO, Q., KOMBRINCK, K.W., FLICK, M.J., HOLMBACK, K., DANON, M.S., COLBERT, M.C., WITTE, D.P., FUJIKAWA, K., DAVIE, E.W. & DEGEN, J.L. 1996. Fatal embryonic bleeding events in mice lacking tissue factor, the cell-associated initiator of blood coagulation. *Medical Sciences*, 93, 6258-6263.

BUTENAS, S., BOUCHARD, B.A., BRUMMEL-ZIEDINS, K.E., PARHIMI-SEREN, B. & MANN, K.G. 2005. Tissue Factor activity in whole blood. *Blood*, 105, 2764-2770.

BUTENAS, S., DEE, J.D. & MANN, K.G. 2003. The Function of factor Xi in tissue factor-initiated thrombin generation. *Journal of Thrombosis and Haemostasis*, 1, 2103-2111.

BUTENAS, S., ORFEO, T., BRUMMEL-ZIEDINS, K.E. & MANN, K.G. 2007. Tissue factor in thrombosis and hemorrhage. *Surgery*, 142, S2-S14.

BUTENAS, S., ORFEO, T. & MANN, K.G. 2008. Tissue factor activity and function in blood coagulation. *Thrombosis Research*, 122, S42-S46.

CARLSSON, K., PERSSON, E., CARLSSON, U. & SVENSSON, M. 2006. Inhibitors of factor VIIa affect the interface between the protease domain and tissue factor. *Biochemical and Biophysical Research Communication*, 349, 1111-1116.

CARMELIET, P. & COLLEN, D. 1998. Molecule in focus: Tissue factor. *The International Journal of Biochemistry & Cell Biology*, 30, 661-667.

CARMELIET, P., MACKMAN, N., MOONS, L., LUTHER, T., VAN VLAENDEREN, P., VAN VLAENDEREN, I., DEMUNCK, H., KASPER, M., BREIER, G., EVRARD, P., MULLER, M., RISAU, W., EDGINGTON, T. & COLLEN, D. 1996. The role of tissue factor in embryonic blood vessel development. *Nature*, 383, 73-75.

CENSAREK, P., BOBBE, A., GRANDOCH, M., SCHRÖR, K. & WEBER, A.A. 2007. Alternatively spliced human tissue factor (asHTF) is not pro-coagulant. *Thrombosis and Haemostasis*, 97, 11-14.

CINES, D., POLLAK, E., BUCK, C., LOSCALZO, J., GA, Z., MCEVER, R., POBER, J., WICK, J., KONKLE, B., SCHWARTZ, B., BARNATHAN, E., MCCREA, K., HUG, B., SCHMIDT, A. & STERN, D. 1998. Endothelial Cells in Physiology and the Pathophysiology of Vascular Disorders. *Blood*, 91, 3527-3561.

COPELAND, C. & CHERYL, K. 2009. A tale of two Anticoagulants: Warfarin and Heparin. *Journal of Surgical Education*, 66, 176-181.

COX, M.M. 2001. Historical overview: searching for replication help in all of the rec places. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 8173-8180.

CUI, M., PARRY, G., OETH, P., LARSON, H., SMITH, M., HUANG, R., ADAMSON, E. & MACKMAN, N. 1995. Transcriptional Regulation of the Tissue Factor Gene in Human Epithelial Cells Is Mediated by Sp1 and EGR-1. *The Journal of Biological Chemistry*, 271, 2731-2739.

DAHLBÄCK, B. 2000. Blood Coagulation. *Lancet*, 355, 1627-1632.

DE PONT, A.C.J., MOONS, A.H.M., DE JONGE, E., MEIJERS, J.C.M., VLASUK, G.P., ROTE, W.E., BÜLLER, H.R., VAN DER POL, M. & LEVI, M.L. 2004. Recombinant nematode anticoagulant protein c2, an inhibitor of tissue factor/factor VIIa, attenuates coagulation and the interleukin-10 response in human endotoxemia. *Thrombosis and Haemostasis*, 2, 65-70.

DIAMANT, M., NIEUWLAND, R., PABLO, R.F., STURK, A., SMIT, J.W.A. & RADDER, J.K. 2002. Elevated Numbers of Tissue Factor Exposing Microparticles Correlates With Components of the Metabolic Syndromes in Uncomplicated Type 2 Diabetes Mellitus. *Circulation*, 106, 2442-2447.

ELKELBOOM, J., WHITE, H. & YUSUF, S. 2003. The Evolving Role of Direct Thrombin Inhibitors In Acute Coronary Syndromes. *Journal of the American College of Cardiology*, 41, 70S-28S.

ENGELMANN, B. 2006. Initiation of coagulation by tissue factor carriers in blood. *Blood Cells, Molecules, and Diseases*, 36, 188-190.

FÖRSTER, Y., MEYE, A., ALBRECHT, S. & SCHWENZER, B. 2006. Tissue factor and tumor: Clinical and laboratory aspects. *Clinical Chimiea Acta*, 364, 12-21.

FRANCHINI, M. 2005. Heparin-induced thrombocytopenia- an update. *Thrombosis Journal*, 3, 14.

FRANCIS, C.W. 2004. Ximelagatran: a new oral anticoagulant. *Best Practice & Research Clinical Haematology*, 17, 139–152.

GEORGE, F.D. 2008. Microparticles in vascular diseases. *Thrombosis Research*, 122, S55-S59.

FREDENBURG, J.C., STRAFFORD, A.R., POSISIL, C.H. & WEITZ, J.I. 2004. Modes and consequences of thrombin's interaction with fibrin. *Biophysical Chemistry*, 112, 277-284.

FRÉDÉRICK, R., POCHET, C., CHARLIER, C. & MASEREEL, B. 2005. Modulators of the Coagulation Cascade: Focus and Recent Advances in Inhibitors of Tissue Factor, Factor VIIa and their complex. *Current Medicinal Chemistry*, 12, 397-417.

GIESEN, P.L.A., RAUCH, U., BOHRMANN, B., KLING, D., ROQUÉ, M., FALLON, J.T., DADIMAN, J.J., HIMBER, J., RIEDERER, M.A. & NEMERSON, Y. 1999. Blood-borne tissue factor: Another view of thrombosis. *Medical Sciences*, 96, 2311-2315.

GOLINO, P. 2002. The inhibitors of the tissue factor:factor VII pathway. *Thrombosis Research*, 106, V257-V265.

GRESLE, P. & AGNLLI, G. 2002. Novel approaches to the treatment of Thrombosis. *Trends in Pharmacological Sciences*, 23, 25-32.

HELLSTERN, P., BEECK, H., FELLHAUER, A., FISCHER, A. & FALLER-STÖCKL, B. 1997. Measurement of factor VII and of activated factor VII in Healthy individuals and prothrombin complex concentrates. *Thrombosis Research*, 86, 493-504.

HEMBROUG, T. A., SWARTZ, G.M., PAPATHANSSIU, A., VLASUK, G.P., ROTE, W.E., GREEN, S.J. & PRIBLUDA, V.S. 2003. Tissue Factor/ Factor VIIa Inhibitors Block Angiogenesis and Tumor Growth Through a Non-hemostatic Mechanism. *Cancer Research*, 63, 2997-3000.

HIRSH, J. 2003. Current anticoagulant therapy - unmet clinical needs. *Thrombosis Research*, 109, S1-S8.

HOBBS, J., ZAKARIJA, A., CUNDIFF, D., DOLL, J., HYMEN, E., CORNWELL, M., CRAWFORD, S., LIU, N., SIGNAEVSKY, M. & SOFF, G. 2007. Alternative spliced human tissue factor promotes growth and angiogenesis in pancreatic cancer tumor model. *Thrombosis Research*, 120, S13-S21.

HOFFMAN, M. 2008. Some things I thought I knew about tissue factor that turn out to be wrong. *Thrombosis Research*, 122, S73-S77.

HOOGENBOOM, H.R., DE BRUINE, A.,P , HUFTON, S.E., HOET, R.M., ARENDS, J. & ROOVERS, R.C. 1998. Antibody phage display technology and its applications. *Immunotechnology*, 4, 1–20.

HOTAMISLIGIL, G.S., ARNER, P., CARO, J.F., ATKINSON, R.L. & SPIEGELMAN, B.M. 1995. Increased Adipose Tissue Expression of Tumao Necrosis Factor-α in Human Obesity and Insulin Resistance. *Journal of Clinical Investigation*, 95, 2409-2415.

JACQUEMIN, M. & SAINT-REMY, J.M. 2004. The Use of Antibodies to Coagulation Factors for Anticoagulant Therapy. *Current Medicinal Chemistry*, 11, 2291-2296.

KAKKAR, A.K., LEVINE, M., PINEDO, H.M., WOLFF, R. & WONG, J. 2003. Venous Thrombosis in Cancer Patients: Insights from the FRONTLINE Survey. *The Oncologist*, 8, 381-388.

KAMIHURA, Y., WADA, H., NOBORI, T., KOBAYASHI, T., SASE, T., NISHIKAWA, M., ISHIKURA, K., YAMADA, N., ABE, Y., NISHIOKA, J., NAKANO, T. & SHIKU, H. 2005. Elevated levels of leukocyte tissue factor mRNA in patients with venous thromboembolism. *Thrombosis Research*, 116, 307-312.

KEELING, D. 2006. Duration of anticoagulation: decision making based on absolute risk. *Blood Reviews*, 20, 173-178.

KENNEL, S.J., LANKFORD, T., FOOTE, L., WALL, M. & DAVERN, S. 2004. Phage display selection of scFv to murine endothelial cell membranes. *Hybridoma and Hybridomics*, 23, 205-11.

KERN, P.A., RANGANATHAN, S., LI, C., WOOD, L. & RANGANATHAN, G. 2001. Adipose tissue tumor necrosis factor and interleukin-6 expression in human obesity and insulin resistance. *American Journal of Physiological Endocrinolical Metabolism*, 280, E745-E751.

KIRCHHOFER, D. & NEMERSON, Y. 1996. Initiation of blood coagulation: the tissue factor/factor VIIa complex. *Current Opinion in Biotechnology*, 7, 386-391.

KRETZSCHMAR, T. & VON RÜDEN, T. 2002. Antibody discovery: phage display. *Current Opinion in Biotechnology*, 13, 598–602.

KIRCHHOFER, D., RIEDERER, M. & BAUMGARTNER, H.R. 1997. Specific Accumulation of Circulating Monocytes and Polymorphonclear Leukocytes on Platelet Thrombi in a Vascular Injury Model. *Blood*, 89, 1270-1278.

KRIKUN, G., LOCKWOOD, C. & PAIDAS, M. 2009. Tissue factor and the endometrium: Form physiology to pathology. *Thrombosis Research*, 124, 393-396.

LECHNER, D. & WELTERMANN, A. 2008. Circulating tissue factor-exposing microparticles. *Thrombosis Research*, 122, S47-S54.

LEE, K.W., BLANN, A.D. & LIP, G.Y.H. 2006. Inter-relationship of indices of endothelial damage/dysfunction [circulating endothelial cells, von Willebrand factor and flow-mediated dilatation] to tissue factor and interleukin-6 in acute coronary syndromes. *International Journal of Cardiology*, 111, 302-308.

LEVI, M. & VAN DER POLL, T. 2005. Two-Way Interaction Between Inflammation and Coagulation. *Trends in Cardiovascular Medicine*, 15, 254-259.

LEVINE, M.N., LEE, A. Y. & KAKKAR, A.K. 2003. From Trousseau to targeted therapy: New insights and innovations in thrombosis and cancer. *Journal of Thrombosis and Haemostasis*, 1, 1456-1463.

LIM, H.S., BANN, A.D. & LIP, G.Y.H. 2004. Soluble CD40 Ligand, Soluble P-Selectin, Interleukin-6, and Tissue Factor in Diabetes Mellitus: Relationship to Cardiovascular Disease and Risk factor Intervention. *Circulation*, 109, 2524-2528.

LIN, M., FANNY, A., CHEN, H., PARRY, G.C.N., MACKMAN, N., SHY, J.Y. & CHIEN, S. 1996. Shear Stress Induction of the Tissue Factor Gene. *The American Society for Clinical Investigation*, 99, 737-744.

LLORENTE-CORTÉS, V., OTERO-VINAS, M., CAMINO-LOPEZ, S., LLAMPAYAS, O. & BADIMON, L. 2004. Aggregated Low-Density Lipoprotein Uptake Induces Membrane Tissue Factor Procoagulant Activity and Microparticle Release In Human Vascular Smooth Muscle Cells. *Circulation*, 110, 452-459.

LÖSCHE, W. 2005. Platelets and tissue factor. *Platelets*, 16, 313-319.

MACKMAN, N. 1995. Regulation of the tissue factor gene. *FASEB*, 9, 883-889.

MACKMAN, N. 2004. Role of Tissue Factor in Hemostasis, Thrombosis, and Vascular Development. *Arteriosclerosis Thrombosis and Vascular Biology*, 24, 1015-1022.

MACKMAN, N. 2006. The role of tissue factor in hemostasis and thrombosis. *Blood Cells, Molecules, and Diseases*, 36, 104-107.

MACKMAN, N. 2007. Alternatively spliced tissue factor - One cut to many? *Journal of Thrombosis and Haemostasis*, 97: 5-8.

MACKMAN, N. 2008. Tissue-specific hemostasis: role of tissue factor. *Journal of Thrombosis and Haemostasis*, 6, 303-305.

MACKMAN, N. & MORRISSEY, J. 2008. Tissue Factor and factor VIIa: Understanding the molecular mechanism. *Thrombosis Research*, 122, S1-S2.

MARSIK, C., QUEHENBERGER, P., MACKMAN, N., OSTERUD, B., LUTHER, T. & JILMA, B. 2003. Validation of novel tissue factor assay in experimental human endotoxemia. *Thrombosis Research*, 111, 311-315.

MCVEY, L, 1999. Tissue Factor pathway. *Baillière's Clinical Haematology*, 12, 361-372.

MEIRING, S.M., LITTHAUER, D., HÁRSFALVI, J., WYK, V. V., BADENHORST, P. N. & KOTZÉ, H.F. 2002. In Vitro effect of a thrombin inhibition peptide selected by phage display technology. *Thrombosis Research*, 107, 365-371.

MEIRING, S.M., VERMEULEN, J. & BADENHORST, P.N. 2009. Development of an Inhibitory Antibody Fragment to Human Tissue Factor using Phage Display Technology. *Drug Development Research*,70, 199-205.

MEZZANO, D., MATUS, V., SÁEZ, C., PEREIRA, J. & PANES, O. 2008. Tissue factor storage, synthesis and function in normal and activated human platelets. *Thrombosis Research*, 122, S31-S36.

MIGDALSKI, A., KOTSCHY, M. & JAWIEN, A. 2005. Tissue factor, Tissue Factor Pathway Inhibitor and Vascular Endothelial Growth Factor-A in Carotid Atherosclerotic Plaques. *European Journal of Vascular and Endovascular Surgery*, 30, 41-47.

MORRISSEY, J. 2003. Tissue factor: in at the start....and the finish? *Thrombosis and Haemostasis*, 1, 878-880.

MUELLER, R.L. 2004. First-generation agents: aspirin, heparin and coumarins. *Best Practice & Research Clinical Haematology*, 17, 23-53.

NADIR, Y., BRENNER, B., ZETSER, A., ILAN, N., SHAFAT, I., ZCHZRIA, E., GOLDSMITDT, O. & VLODAVSKY, I. 2006. Heparanase induces tissue factor expression in vascular endothelial and cancer cells. *Thrombosis and Haemostasis*, 4, 2443-2451.

NEMERSON, Y. 1988. Tissue Factor and Hemostasis. *Blood*, 71, 1-8.

NISSIM, A., HOOGENBOOM, H.R., TOMLINSON, I.M., FLYNN, G., MIDGLEY, C., LANE, D. & WINTER, G. 1994. Antibody fragments from a 'single pot' phage display library as immunochemical reagents. *The EMBO Journal*, 13, 692-8.

NOFER, J., KEHREL, B., FOBKER, M., LEVKAU, B., ASSMANN, G. & ECKARDSTAIN, A.V. 2002. HDL and arteriosclerosis: beyond reverse cholesterol transport. *Atherosclerosis*, 161, 1-16.

NOMURA, S., OZAKI, Y. & IKEDA, Y. 2008. Function and role of microparticles in various clinical settings. *Trombosis Research*, 32, 130-146.

NORRIS, D.P., BROCKDORFF, N. & RASTAN, S. 1991. Methylation status of CpG-rich islands on active and inactive mouse X chromosomes. *Mammalian Genome*, 1, 78-93.

NORRIS, L.A. 2003. Blood Coagulation. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 17, 369-383.

O'BRIEN, D.P., ANDERSON, J.S., MARTIN, D.M.A., BYFIELD, P.G.H. & TUDDENHAM, E.G.D. 1993. Structural requirements for the interaction between Tissue Factor and Factor VII: characterization of chymotrypsin-derived Tissue Factor polypeptides. *Journal of Biochemistry*, 292, 7-12.

ØSTERUD, B., BREIMO, E.S. & OLSEN, J.O. 2007. Blood borne tissue factor revisited. *Thrombosis Research*, 122, 432–434.

PENG, Z., CAI, X., ZHANG, Y., KONG, D., GUO, H., LIANG, W., TANG, Q., SANG, H. & MA, D. 2007. A novel anti-tissue factor monoclonal antibody with anticoagulant potency derived from synthesized multiple antigenic peptide through blocking FX combination with TF. *Thrombosis Research*, 121, 85-93.

PRICE, G.C., THOMPSON, S.A. & KAN, P.C.A. 2004. Tissue factor and tissue factor pathway inhibitor. *Anaesthesia*, 59, 483-492.

QIANNING, L., YIMIN, Y., DAJUN, Y., RONGHUAN, C., ZILI, G., YONG, L., ZUJAN, Z. & JIAN, Z. 2008. Phosphorothioate oligonucleotide inhibit tissue factor expression in endothelial cells induced by blood flow shear stress in rats. *Journal of Medical Colleges of PLA*, 23, 154-161.

RAUCH, U., BONDERMAN, D., BONDERMAN, B., BADIMON, J.J., HIMBER, J., RIEDERER, M. & NEMERSON, Y. 2000. Transfer of tissue factor from leukocytes to platelets is mediated by CD15 and tissue factor. *Blood*, 96, 170-175.

RAUCH, U. & NEMERSON, Y. 2000. Circulating tissue factor and thrombosis. *Current Opinion in Hematology*, 7, 273-277.

RAUCH, U., SZOTOWSKI, B. & ANTONIAK, S. 2006. Alternatively Spliced Tissue Factor: A Previously Unknown Piece in the Puzzle of Hemostasis. *Trends in Cardiovascular Medicine*, 2006, 177-182.

RAVN, P., DANIELCZYK, A., JENSEN, K. B., KRISTENSEN, P., CHRISTENSEN, P. A., LARSEN, M., KARSTEN, U. & GOLETZ, S. 2004. Multivalent scFv display of phagemid repertoires for the selection of carbohydrate-specific antibodies and its application to the Thomsen-Friedenreich antigen. *Journal of Molecular Biology*, 343, 985-96.

RØNNING, H., RISØEN, U., ÖRNING, L., SLETTEN, K. & SAKARIASSEN, K. 1996. Synthetic peptide analogues of tissue factor and factor VII which inhibit factor Xa formation by the tissue factor/factor VIIa complex. *Thrombosis Research*, 84, 73-81.

SAMAD, F., PANDEY, M. & LOSKUTOFF, D. 2001. Regulation of tissue factor gene expression in obesity. *Blood*, 98, 3353-3358.

SELJEFLOT, I., HURLEN, M. & ARNESEN, H. 2004. Increased levels of soluble tissue factor during long-term treatment with warfarin in patients after an acute myocardial infarction. *Journal of Thrombosis and Haemostasis*, 2, 726-730.

SMITH, G.P. & PETRENKO, V.A. 1997. Phage Display. *Chemical reviews*, 97, 391–410.

SOMMEIJER, D.W., HANSEN, H.R., VAN OERLE, R., HAMULYAK, K., VAN ZANTEIN, A.P., MEESTERS, E., SPONK, H.M.H. & TENCATE, H. 2006. Soluble Tissue factor is a candidate marker for the progression of microvascular disease in patients with Type 2 diabetes. *Journal of Thrombosis and Haemostasis*, 4, 574-580.

STEFFEL, J., AKHMEDOV, A., FANHDRICH, C., RUSCHITZKA, F., LÜSCHER, T. T. & TANNER, F. 2006a. Differential effect of celecoxib on tissue factor expression in human endothelial and vascular smooth muscle cells. *Biochemical and Biophysical Research Communications*, 349: 597-603.

STEFFEL, J., LÜSCHER, T.F. & TANNER, F.C. 2006b. Tissue Factor in Cardiovascular Diseases: Molecular Mechanism and Clinical Implications. *Circulation*, 113, 722-731.

STEPHENS, A.C. & RIVERS, R.P.A. 1997. Suppression of human monocyte tissue factor synthesis by antisense oligodeoxynucleotide. *Thrombosis Research*, 85, 387-389.

STURK-MAQUELIN, K.N., NIEUWLAND, R., ROMIJN, F., EIJSMAN, L., HACK, C. & STURK, A. 2003. Pro- and non coagulant forms of non-cell-bound tissue factor in vivo. *Journal of Thrombosis and Haemostasis*, 1, 1920-1926.

TILLY, R.E., HOLSCHER, T., BELANI, R., NIEVA, J. & MACKMAN, N. 2008. Tissue factor activity is increased in a combined platelet and microparticle sample from cancer patients. *Thrombosis Research*. 122, 604-609.

VAN VEEN, J., GATT, A. & MAKRIS, M. 2008. Thrombin generation testing in routine clinical practice: are we there yet? *British Journal of Haematology*, 142, 889-903.

VIEIRA, L.M., DUSSE, L.M.S., FERNANDES, A.P., MARTINS-FILHO, O.A., DE BASTOS, M., FERREIRA, M.F.R., COOPER, A.J., LWALEED, B.A. & CARVALHO, M.G. 2007. Monocytes and plasma tissue factor levels in normal individuals and patients with deep venous thrombosis of the lower limbs: Potential diagnostic tools? *Thrombosis Research*, 199, 157-165.

VILES-GONZALEZ, J.F. & BADIMON, J.J. 2004. Atherothrombosis: the role of tissue factor. *The International Journal of Biochemistry & Cell Biology*, 36, 25-30.

VOS, H. 2006. Inherited defects of coagulation Factor V: the thrombotic side. *Thrombosis and Haemostasis*, 4, 35-40.

WATANABE, T., YASUDA, M. & YAMAMOTO, T. 1999. Angiogenesis Induced by Tissue Factor in Vitro and in Vivo. *Thrombosis Research*, 96, 183-189.

WEITZ, J.I. 2009. Unanswered questions in venous thrombosis. *Thrombosis Research*, 123, S2-S10.

WEITZ, J.I. & CROWTHER, M. 2002. Direct thrombin inhibitors. *Thrombosis Research*, 106, V275-V284.

WILLATS, W.G.T. 2002. Phage display: practicalities and prospects. *Plant Molecular Biology*, 50, 837–854.

ZWICKER, J.I., FURIE, B.C. & FURIE, B. 2007. Cancer-associated thrombosis. *Oncology/Hematology*, 62, 126-136.

CHAPTER 3

Large scale expression of tissue factor inhibitor

3.1 Introduction

Phage display provides a tool for antibody engineering that is unobtainable by traditional methods (Watkins and Ouwehand, 2000). Antibodies are among the most powerful therapeutic and diagnostic tools available today (Ghose *et al.*, 2006). Unsurprisingly, they are also the fastest growing class of therapeutic molecules in medicine. Recombinant antibody fragments such as scFv are becoming more popular as an alternative to full length antibodies due to their inherent smaller size and extreme versatility in their application (Weisser and Hall, 2009). Phage display technology is a powerful and effective molecular technique for the identification and generation of antibody fragments against a wide variety of proteins as well as unconventional antigens such as peptides, DNA products, single stranded RNAs as well as lipids (Willats, 2002; Wu *et al.*, 2007; Marcus *et al.*, 2008).

The Tomlinson I + J human single fold scFv libraries are among the most widely utilized libraries for the display of the antibody variable regions on the surface of filamentous phage (Barderas *et al.*, 2006). These scFv comprises of a single polypeptide with the Heavy- and Light chain variable domains of the fragment antigen-binding (Fab) region of human immunoglobulin (IgG) linked to each other by a flexible Glycine-Serine linker sequence. Single chain variable fragment DNA constructs are cloned into the minor protein III (pIII) coat protein of the phagemid due to its tolerance to large inserts (Smith and Petrenko, 1997). The construct also contain a C-terminal His- and myc-tag that is followed by an amber STOP codon. The libraries consist of 1.47×10^8 and 1.37×10^8 unique scFv constructs respectively. Constructs are cloned into the pIT2 phagemid vector and transformed into *E. coli* TG1. By co-infecting *E. coli* TG1 (suppressive strain) with the KM13 Helper phage, large quantities of soluble phage displaying scFv-pIII fusion proteins are expressed into the culture media (Yan *et al.*, 2004). This large pool of phagemids is screened against target molecules in order to isolate desired binders (also called panning).
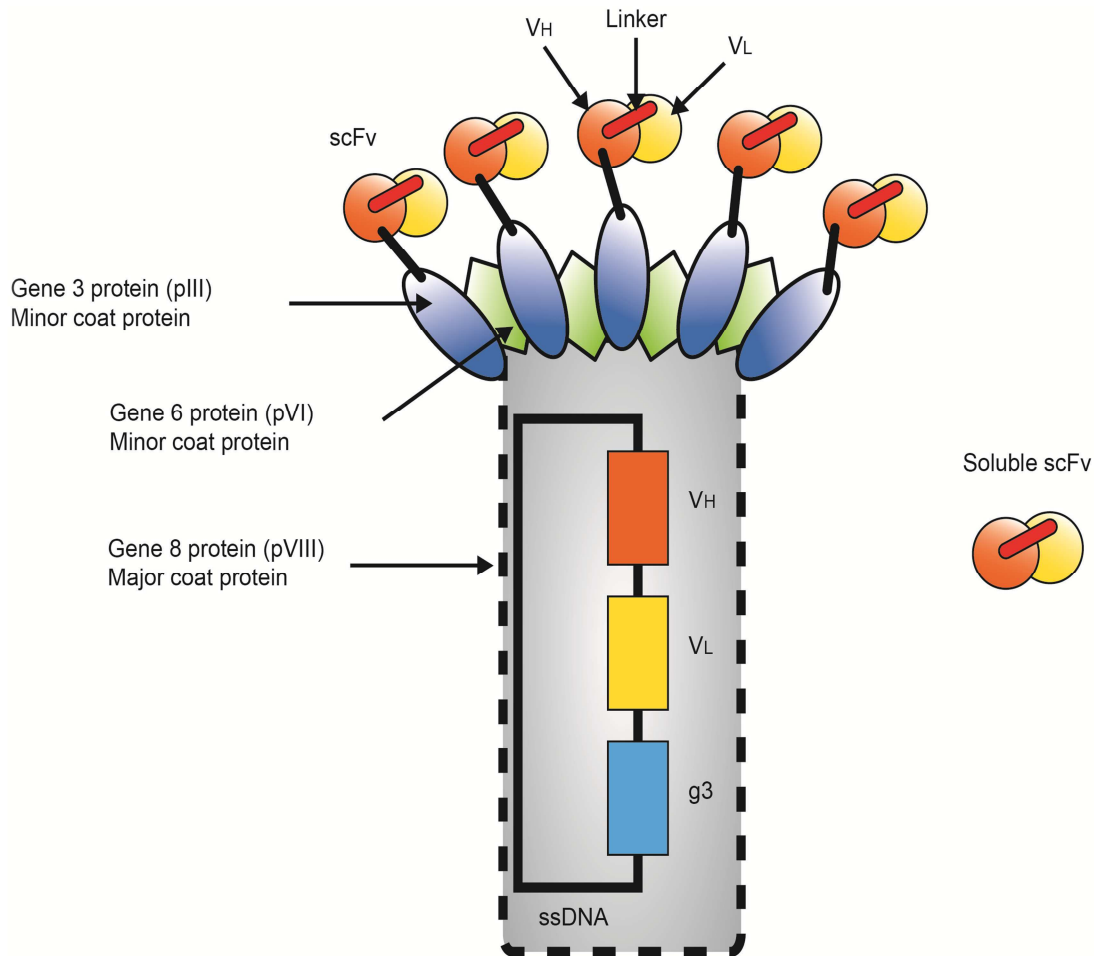
Figure 3.1: The M13 filamentous phage

The heavy chain- (V_H) and Light chain variable regions (V_L) connected with short linker sequence and spliced into the coat protein III (g3) gene. Translation of the single stranded DNA (ssDNA) results in the expression or displays the scFv–coat protein III (scFv-pIII) fusion products on the surface of the M13 bacteriophage. Due to the position of the displayed protein and it subsequent accessibility to the solvent, it behaves essentially as it would if when expressed as a soluble fragment (Soluble scFv).

Compiled from Nissim *et al.* (1994); Smith and Petrenko (1997); Hoogenboom *et al.* (1998); Kretzschmar and von Ruden (2002).

Finally soluble scFv fragments, totally detached from the phagemid (See: Figure 3.1), are produced by infecting *E. coli* HB2151 (non-suppressive strain) with the desired isolated phagemids (Barderas *et al.*, 2006; Eteshola, 2010). The recognition of the c-terminal amber codon by the non-suppressive strain prevents the co-expression of the downstream coat protein (pIII) and the subsequent production of the scFv-pIII fusion protein. After translation the soluble scFv are transported to periplasm and passively released into the culture media. The scFv are purified from

the culture media by means of protein A affinity chromatography as recommended by Tomlinson I + J procedure (See: Appendix A).

3.2 Aim

The aim of this study was to increase the yield of functional TFI-scFv expressed by *E. coli* HB2151 through the up-scaling the prescribed Tomlinson I + J method and to determine the purity of the Tissue Factor Inhibitor TFI after protein A purification.

3.3 Materials and Methods

3.3.1 Production of soluble TFI-scFv

Large quantities of KM13 helper phage were produced in *E. coli* TG1 and utilized in the production of phagemids expressing the tissue factor inhibitor fusion protein (TFI-phagemid). Isolated TFI-phagemid was used to transfect *E. coli* HB2151 that then inturn expressed soluble tissue factor inhibitor single chain variable fragment (TFI-scFv) into the culture media.

3.3.1.1 Production of KM13 helper phage

Helper phage was produced according to the protocol G as specified in the product documentation (Tomlinson I + J: Human Single Fold Libraries I + J) (LifeSciences, UK).

Briefly, 10 μL of *E. coli* TG1 was transferred to 10 mL of 2TY media (16 g.L^{-1} tryptone, 10 g.L^{-1} yeast extract and 5 g.L^{-1} NaCl) and incubated until an optical density reading of 0.4 at 600 nm was obtained. All cultures were aerobically incubated in an orbital shake-incubator (160 rpm) at 37 °C, unless indicated differently. Optical density was measured with the GeneQuant pro RNA/DNA calculator at 600 nm (AEC-Amersham, UK). When the desired optical density (0.4 at OD$_{600}$) was reached, five 200 μL aliquots of the culture media were collected. The aliquots were infected with 10 μL serial dilutions (10^{-1} to 10^{-6}) of KM13 helper phage

stock and incubated for 30 min at 37 ºC in a waterbath. A 3 mL volume of molten (42 ºC) H-Top bacteriological-agar (10 $g.L^{-1}$ tryptone, 8 $g.L^{-1}$ NaCl and 7 $g.L^{-1}$ bacteriological-agar) was added to each infected aliquot, mixed gently and poured onto warm TYE agar plates (15 $g.L^{-1}$ bacteriological-agar, 10 $g.L^{-1}$ tryptone, 8 $g.L^{-1}$ NaCl and 5 $g.L^{-1}$ yeast extract). Plates were allowed to set and then incubated overnight at 37 ºC. One small plaque was selected and transferred to 5 ml of fresh *E. coli* TG1 culture at OD_{600} reading of 0.4. The culture was incubated for 2 hours at 37 ºC. A 5 mL aliquot of infected culture was transferred to 500 mL of 2TY media in a 2 L Erlenmeyer flask and incubated for 1 hour at 37 ºC. At this point the culture was supplemented with kanamycin (50 $mg.L^{-1}$) and incubated overnight at 30 ºC. The overnight culture was centrifuged at 10800 x *g* for 15 minutes at room temperature and the supernatant retained. The phages were precipitated on ice by adding a volume of 100 mL polyethylene glycol (PEG) – NaCl solution (20 $g.L^{-1}$ PEG 6000 and 35 $g.L^{-1}$ NaCl) to 400 mL of culture and placed on ice for 1 hour. The culture was centrifuged at 10800 x *g* for 30 minutes at room temperature and the supernatant discarded. The remaining pellet containing the phages was resuspended in 8 mL 20 mM phosphate buffered saline (PBS) (8 $g.L^{-1}$ NaCl, 1.44 $g.L^{-1}$ Na_2HPO_4, 0.24 $g.L^{-1}$ KH_2PO_4 and 0.2 $g.L^{-1}$ KCl at pH 7.2). An additional 2 mL PEG - NaCl solution was added to the resuspended pellet and kept on ice for 20 minutes. The PEG - NaCl was removed by centrifugation at 3300 x *g* for 30 minutes at room temperature, followed by a brief additional centrifugation and aspiration of the remaining residual PEG-NaCl. Bacterial debris was removed by resuspending the remaining pellet in 5 mL PBS followed by centrifugation at 11600 x *g* for 10 minutes at room temperature. The supernatant containing the helper phage was collected, supplemented with 15 % glycerol and stored as phage stock at -80 ºC until later use or stored at 4 ºC when used within one month.

3.3.1.2 Production of TFI-phagemid

Tissue factor inhibiting phage was produced according to the protocol A as stipulated in the product documentation (TomlinsonI + J: Human Single Fold Libraries I + J).

Frozen *E. coli* TG1 cell stock stored at -80 ℃ was streaked out on a M9 minimal plate (See: Appendix B) and incubated for 16 hours. A single colony was selected

and inoculated into 10 ml 2TY media. The preculture was incubated until an optical density (OD_{600}) of 0.4 - 0.6 was reached. After the culture reached the desired optical density (0.4 - 0.6 at OD_{600}), 1 ml of the cultured media was collected and infected with 10 µL of the TFI-phagemid (2.9 x 10^{11}). The infected culture was incubated for 30 minutes at 37 ºC. The *E. coli* TG1 cells were harvested by means of centrifugation at 5000 x *g* at room temperature for 5 minutes. The cells were resuspended in 200 µL 2TY media and plated onto TYE plates (15 $g.L^{-1}$ bacteriological-agar, 10 $g.L^{-1}$ tryptone, 8 $g.L^{-1}$ NaCl and 5 $g.L^{-1}$ yeast extract) containing 10 $g.L^{-1}$ D-glucose and 100 $mg.L^{-1}$ ampicillin. The agar plates were incubated for 16 hours at 37 ºC. A single colony was selected and used to inoculate 50 ml 2TY media containing 10 $g.L^{-1}$ D-glucose and 100 $mg.L^{-1}$ ampicillin. The culture was incubated until an OD_{600} of 0.4 - 0.6 was reached, at which point 10 ml of the cultured media was harvested and infected with approximately 5 x 10^{10} helper phage for 30 minutes at 37 ºC. Infected *E. coli* TG1 cells were harvested by means of centrifugation at 5000 x *g* for 5 minutes at room temperature. The cell pellets were subsequently resuspended in 2TY media and used to inoculate 50 ml 2TY media containing 1 $g.L^{-1}$ D-glucose, 100 $mg.L^{-1}$ ampicillin and 50 $mg.L^{-1}$ kanamycin. The culture was incubated for 16 hours at 30 ºC. The following day the supernatant containing the phage was retained by means of centrifugation at 5000 x *g* for 15 minutes at room temperature. The phages were precipitated by the addition of 20 % (v/v) PEG – NaCl solution (15 $g.L^{-1}$ polyethylene glycol 6000 and 35 $g.L^{-1}$ NaCl). This mixture was then incubated on ice for 1 hour and the supernatant discarded after centrifugation at 5000 x *g* for 30 minutes at room temperature. The remaining phage pellet was resuspended in 2 mL 20 mM PBS buffer and centrifuged for 10000 x *g* for 10 minutes at room temperature and the supernatant containing TFI-phage was collected and frozen at -80 ºC as a 15 % glycerol (v/v) TFI-phage stock until later use.

3.3.1.3 Production of TFI-scFv

Soluble TFI-scFv was produced according to the protocol F as stipulated in the product documentation (Human Single Fold Libraries I + J - TomlinsonI & J).

Briefly, 15 % glycerol stock of *E. coli* HB2151 stored at -80 ºC was streaked out on a M9 minimal plate (See: Appendix B) and incubated for 16 hours. A single colony was selected and transferred to 10 ml 2TY media. The culture was incubated until an OD_{600} reading of 0.4 - 0.6 was reached, at which point 200 µl of the culture was harvested. Harvested *E. coli* HB2151 cells were infected with 1 % (v/v) of a 15 % glycerol TFI-phage stock and incubated for 30 minutes at 37 ºC. Infected cells were harvested by means of centrifugation at 5000 x *g* for 5 minutes at 4 ºC. The cells were then resuspended in 2TY media and a dilution series (10^{-1} to 10^{-6}) was prepared and plated onto TYE plates containing 10 g.L^{-1} D-glucose and 100 mg.L^{-1} ampicillin. The agar plates were incubated overnight for 16 hours at 37 ºC.

A single colony was selected and used to inoculate 10 ml 2TY media containing 1 g.L^{-1} D-glucose and 100 mg.L^{-1} ampicillin. The preculture was incubated for 16 hours (overnight) at 37 ℃. A volume of 1 mL of precultu re was used to inoculate 250 mL of 2TY media in a 2 L Erlenmeyer flask containing 1 g.L^{-1} D-glucose and 100 mg.L^{-1} ampicillin. The culture was incubated until an OD_{600} reading of 0.9 was reached, at which point Isopropyl-β-D-thiogalactoside (IPTG) was added to a final concentration of 0.1 mM in order to induce TFI-scFv expression. The infected culture was then incubated in an orbital shake-incubator (160 rpm) for an additional 16 hours at 30 ºC. The supernatant containing TFI-scFv was retained by removing the *E. coli* cells by centrifugation at 5000 x *g* for 10 minutes at 4 ºC. The retained supernatant was stored at 4 ºC until later use.

3.3.2 Protein A affinity chromatography purification of TFI-scFv

TFI-scFv was purified by means of protein A affinity chromatography. A 5 mL nProtein A Sepharose 4 Fast Flow column was obtained from GE Healthcare and the TFI-scFv was purified according to the manufacture's (GE Healthcare, RSA) guidelines. All purifications took place at room temperature and were performed within a week after expression of TFI-scFv (See: Section 3.3.2.2).

3.3.2.1 Packing of nProtein A Sepharose 4 Fast Flow resin

The column resin material was equilibrated to room temperature for 3 hours. During this time a 20 mm x 200 mm glass column (SIGMA, USA) was prepared by flushing it with 20 mL 70% ethanol followed by 20 mL distilled water and finally 20 mL binding buffer (6.057 g.L^{-1} tris-HCl buffer at pH 7.0). The resin material was equilibrated for 3 hours, degassed and poured into the glass column. Care was taken to prevent the formation of air pockets. The column was immediately filled with 50 mL binding buffer and the column closed at the top end. The resin material within the column was gently mixed by inverting the glass column twice and clamped vertically in place. The column was left in this position overnight at 4 ºC in order for the resin to settle.

3.3.2.2 nProtein A Sepharose 4 Fast Flow purification of TFI-scFv

After the resin material settled, the column was equilibrated to room temperature for 3 hours. The excess binding buffer was drained from the column. Care was taken not to let the column run dry in order to prevent the formation of air pockets. The column was then flushed with an additional 15 mL binding buffer.

The culture media supernatant containing soluble TFI-scFv (See: Section 3.3.1.3) was diluted in a 1:1 ratio with binding buffer. A total volume of 50 mL of this mixture was loaded onto the column and allowed to flow completely into the resin. The supernatant – binding buffer mixture was passed through the column by means of gravity. The column was then flushed with binding buffer, in order to remove all excess unbound particles, until absorbance (measured at 280 nm) of the effluent was lower than 0.1. Absorbance was measured with GeneQuant pro RNA/DNA calculator (AEC-Amersham, RSA). Bound TFI-scFv was eluted from the protein A column through the addition of 15 mL elution buffer (7.507 g.L^{-1} glycine at pH 3.0) that was passed through the column by means of gravity and collected in 15 x 1 mL fractions. Eluted fractions were immediately neutralized with the addition of 100 μL neutralization buffer (121.14 g.L^{-1} tris at pH 9.0) per 1 mL of eluate and the absorbance determined at 280 nm. Those fractions with the highest A_{280} measurements where pooled together and stored at 4 ºC until later use.

3.3.2.3 nProtein A Sepharose 4 Fast Flow maintenance and cleaning

The protein A column was cleaned with 15 mL elution buffer and re-equilibrated with 15 mL binding buffer before every purification round. Two rounds of purification were performed each day, after which any strongly bound hydrophobic proteins, lipoproteins and lipids were removed by exposing the protein A column to 70 % ethanol overnight at 4 ºC. The following day the protein A column was flushed in order to remove the ethanol containing unwanted substances and re-equilibrated with 25 mL binding buffer before use.

In the case of severe reduction in column performance possibly caused by the presence of precipitated or denatured substances, the column was washed with 10 mL 6 M guanidine hydrochloride (GuHCl) as recommended in the product's cleaning-in-place protocol. The protein A column was immediately flushed with 25 mL binding buffer before use.

3.3.2.4 Desalting and concentration of purified TFI-scFv

The collected fractions from the various protein A purification rounds were pooled together, desalted and concentrated using Amicon Ultra-15 centrifugal filtration unit with a 10000 molecular weight cut off (10K MWCO) (Millipore, USA). Desalting and concentration procedures were performed according to the product protocol (Amicon Ultra-15 centrifugal filtration devices).

The pooled fractions (15 mL) containing TFI-scFv were transferred to the Amicon Ultra-15 centrifugal filtration unit, capped and centrifuged at 5000 x *g* for 15 minutes at 4 ºC in a fix angle centrifuge (Beckman-Coulter, USA). A 5 mL volume tris buffered saline (TBS) (6.057 g.L^{-1} tris and 8.744 g.L^{-1} NaCl at pH 7.4) was added to the concentrated TFI-scFv fraction in the Ultra-15 centrifugal filtration unit, mixed gently with a 200 µL pipette and centrifuged using the same conditions as described above. This was repeated until all the pooled sample was desalted and concentrated in TBS. The final TFI-scFv sample was collected and stored at 4 ºC until later use.

3.3.3 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed in order to assess purity of the final TFI-scFv fractions. Sample separation was performed using the "Mighty small" miniature slab gel electrophoresis unit SE 200 according the manufacturers' (Hoefer Scientific Instruments) specifications. The sample was separated using 12.5 % resolving gel (See: Table 3.1) and a 4% stacking gel (See: Table 3.2) (Laemmli, 1970). The relative molecular mass of the protein was estimated by comparing the electrophoretic mobility with those of known proteins.

3.3.3.1 SPS-PAGE Gel Preparation

The resolving gel was prepared as shown in Table 3.1 in the Hoefer casting chamber and layered with 1 mL of H_2O saturated tetra-amyl alcohol before polymerisation. The gel was solidified for approximately 20 minutes at room temperature and the H_2O saturated tetra-amyl alcohol discarded. Any residual H_2O saturated tetra-amyl alcohol was removed by flushing the top end of the resolving gel with 10 mL distilled water.

Table 3.1: 12.5 % SDS-PAGE resolving gel

Reagent	Volume per Gel
30 % Acrylamide/0.8% Bisacrylamide solution	3.333 mL
3 M Tris-HCl pH 8.8 solution	1.065 mL
10 Sodium dodecyl sulphate (SDS)	200 µL
10 % Ammonium Persulphate	27 µL
N, N, N', N'-tetramethylethylenediamine (TEMED)	5 µL
Distilled Water	3.602 mL
Total	8 mL

Once the H_2O saturated tetra-amyl alcohol was completely removed the stacking gel was prepared as shown in Table 3.2 was applied to the top the resolving gel and a 10 well comb was inserted before polymerisation for approximately 15 minutes at room temperature.

Table 3.2: 4 % SDS-PAGE stacking gel

Reagent	Volume per Gel
30 % Acrylamide/0.8% Bisacrylamide solution	400 µL
0.5 M Tris-HCl pH 6.8 solution	210 µL
10 Sodium dodecyl sulphate (SDS)	30 µL
10 % Ammonium Persulphate	30 µL
N, N, N', N'-tetramethylethylenediamine (TEMED)	3 µL
Distilled H_2O	2.327 mL
Total	3.000 mL

Once the stacking gel was completely solidified the 10 well comb was removed and the SDS-Page gel was clamped into the Hoefer electrophoresis chamber taking care to prevent any leakage. The electrophoresis chamber was placed in to the Hoefer miniVE Vertical Electrophoresis reservoir and submerged in 1 L SDS-PAGE electrophoresis buffer (30.3 $g.L^{-1}$ tris-HCl, 144 $g.L^{-1}$ glycine and 10 $g.L^{-1}$ SDS at pH 8.3).

3.3.3.2 Protein Sample Preparation

The protein sample was prepared as follows: 20 µL protein sample, 16 µL loading buffer (24 $g.L^{-1}$ tris-HCl, 1.8 $g.L^{-1}$ ethylenediaminetetraacetic (EDTA), 0.34 sucrose and 0.1 g Bromophenol blue at pH 8.8) and 4 µL denaturing buffer (100 $g.L^{-1}$ SDS and 7.71 $g.L^{-1}$ dithiothreitol) was transferred to a sterile 1.5 mL micro-centrifuge tube and mixed. The sample was heated for 5 minutes at 96 ºC and 10 µL aliquot of the sample was loaded into the third well. A volume of 10 µL of the PageRuler[TM] Plus Prestained protein ladder (Fermentas, Canada) comprising of protein standards: 250 000 Da, to 10 000 Da, was loaded in the first well. Hoefer miniVE Vertical Electrophoresis (Harvard Bioscience, USA) chamber was connected to an external power supply (Wealtec, UK) and the gel was run at 100 Volts for approximately 1.5 hours.

3.3.3.3 SDS-PGAE gel staining and de-staining

The gel was stained with Coomassie Brilliant Blue R250 (BioRad, USA) using the Fairbanks method (Fairbanks *et al.*, 1971). Once the electrophoresis was complete the SDS-PAGE gel was stained overnight with 100 mL Fairbanks buffer A (25 %

(v/v) isopropanol, 10 % (v/v) acetic acid and 0.5 g.$^{L-1}$ Coomassie® Brilliant Blue R-250 with gentle stirring. The following day the gel was briefly boiled and left to cool for 10 minutes at which point the remaining Fairbanks buffer A was discarded. The gel was then rinsed in 50 mL distilled water transferred to 100 mL of Fairbanks buffer B (10 % (v/v) isopropanol, 10 % (v/v) acetic acid and 0.05 g.L^{-1} Coomassie Brilliant Blue R-250). The gel was then briefly boiled for a second time, left to cool for 10 minutes, the remaining Fairbanks buffer B was discarded and the gel rinsed in 50 mL distilled water. Once the gel was rinsed it was transferred to 100 mL of Fairbanks buffer C (10 % (v/v) acetic Acid and 0.02 g.$^{L-1}$ Coomassie® Brilliant Blue R-250), boiled briefly, left to cool for 10 minutes, the remaining Fairbanks buffer C was discarded and the gel rinsed with 50 mL distilled water. The gel was then transferred to 100 ml of Fairbanks buffer D (10 % (v/v) acetic Acid), boiled briefly, left to cool for 10 minutes, the remaining Fairbanks buffer D was discarded and the gel placed in 100 ml distilled water.

3.3.4 Prothrombin Time coagulation analysis

Prothrombin Time (PT) coagulation test was performed with modification to the standard protocol and guideline stipulated in the Thromborel® S (Human thromboplastin containing calcium) (Siemens, Germany), on the Stago STA 4 Coagulation analyzer (Diagnostia, France).

Briefly, 2 blood samples of 4.5 mL from 5 normal healthy volunteers were collected in 0.5 mL of a 32 g.L^{-1} sodiumcitrate solution (Saarchem). Platelet poor plasma was obtained by means of centrifugation at 2000 x g, for 10 minutes at room temperature. After centrifugation, platelet poor plasma samples were pooled together and stored at -80 ºC until later use. Prothrombin Time coagulation test was preformed with an escalating TFI-scFv dosage range.

Table 3.3: Prothrombin Time coagulation test

Reagent	A	B	C	D	E
TFI-scFv	0 µL	12.5 µL	25 µL	50 µL	100µL
TBS	100 µL	87.5 µL	75 µL	50 µL	0 µL
Thromborel	100 µL	100 µL	100 µL	100 µL	100 µL
PPP	50 µL	50 µL	50 µL	50 µL	50 µL

A volume of 100 µL Thromborel was incubated for 10 minutes at 37 ºC with TFI-scFv and TBS (See: Table 3.3). After the 10 minute incubation the PT coagulation test was initiated with the addition of 50 µL platelet poor plasma (PPP) (37 ºC) and the time to coagulation determined. All PT coagulation tests were performed in triplicate, data was analyzed and plotted with GraphPad Prism v 5.03.

3.4 Results

3.4.1 Protein A affinity chromatography purification of TFI-scFv

A total of 10 purification rounds were performed using a single 5 mL nProtein A Sepharose 4 Fast Flow column.

The absorbance of each of the eluted fractions was determined using GeneQuant pro RNA/DNA calculator at 280nm and is presented in Table 3.4. The purification profile for each round was plotted using GraphPad Prism v 5.0.3 and is summarized in Figure 3.4. After the completion of each round only those fractions corresponding to desired protein peaks (Indicated in **red**) were collected and pooled together. No fractions were collected during the 7th purification round, due to the lack of a clearly defined protein peak. The 2nd round of purification resulted in a protein fraction with the highest final absorbance of all the purification rounds. As a result the effectiveness of the purification round 1 to 10 was calculated as a percentage relative to round 2. The final absorbance values of the pooled fractions and relative effectiveness of each round are presented in Table 3.5.

Table 3.4: Absorbance (A$_{280}$) values of eluted fraction

Fractions	Purification Rounds									
-	1	2	3	4	5	6	7	8	9	10
1	0.092	0.098	0.093	0.087	0.088	0.083	0.102	0.105	0.101	0.099
2	0.153	0.205	0.142	0.138	0.137	0.213	0.238	0.187	0.186	**0.124**
3	0.191	0.344	0.236	**0.348**	**0.433**	**0.337**	0.307	**0.28**	**0.283**	**0.186**
4	**0.277**	**0.434**	**0.426**	**0.428**	**0.54**	**0.507**	0.338	**0.331**	**0.412**	**0.286**
5	**0.319**	**0.597**	**0.684**	**0.73**	**0.743**	**0.609**	0.353	**0.434**	**0.457**	**0.357**
6	**0.374**	**0.809**	**0.675**	**0.904**	**0.915**	**0.695**	0.328	**0.453**	**0.441**	**0.395**
7	**0.424**	**0.872**	**0.637**	**0.786**	**0.669**	**0.697**	0.325	**0.425**	**0.435**	**0.389**
8	**0.398**	**0.814**	**0.538**	**0.545**	**0.481**	**0.622**	0.328	**0.321**	**0.389**	**0.329**
9	**0.347**	**0.678**	**0.394**	**0.383**	**0.333**	**0.533**	0.331	0.263	**0.323**	**0.249**
10	**0.292**	**0.418**	**0.289**	**0.291**	**0.267**	**0.466**	0.283	0.183	**0.275**	0.212
11	0.224	0.375	0.243	0.22	0.244	**0.327**	0.222	0.144	0.155	0.175
12	0.218	0.328	0.179	0.204	0.195	0.255	0.198	0.107	0.118	0.148
13	0.185	0.249	0.159	0.172	0.156	0.21	0.15	0.093	0.091	0.071
14	0.134	0.203	0.135	0.144	0.132	0.176	0.129	0.089	0.069	0.089
15	0.113	0.133	0.119	0.124	0.114	0.126	0.107	0.083	0.079	0.079

The eluted fractions corresponding to desired protein peaks (Indicated in **red**) were collected and pooled together for further analysis.

Table 3.5: Final Absorbance and relative effectiveness

-	Purification Rounds									
-	1	2	3	4	5	6	7	8	9	10
Final A$_{280}$	0.35	0.66	0.48	0.55	0.55	0.56	na*	0.37	0.38	0.31
Relative Effectiveness (%)	53	100	73	83	84	80	na*	54	56	48

* - Not applicable

The first round of purification produced a weak protein peak (See: Figure 3.2) with a final absorbance of only 53 % of the pooled fraction obtained during the 2nd round of purification. The 2nd purification round produced the highest final absorbance and a clearly defined protein peak. Although the final absorbance of the pooled fractions for rounds 3 to 6 were each slightly less than the 2nd round the protein peaks observed were all similar. The 3rd, 4th, 5th and 6th rounds of purification all had high relative effectiveness values of 73 %, 83 % 84% and 80% respectively, in comparison to round 2. After the 7th round of purification it became clear that the

protein A column was no longer performing optimally. No clear protein peak was obtained and as a result no fractions were collected during this round. In an attempt to restore the functionality observed in rounds 2 to 6 the column was washed with 6 M guanidine hydrochloride as recommended in order to remove precipitated or denatured protein present within the protein A column that may interfere with the column's ability to bind TFI-scFv. Although, the 8th 9th and 10th rounds of purification produced a slightly improved protein peak than round 7 but it is clear that the attempt to remove the unwanted precipitated and denatured proteins did little to restore the protein A column to its former effectiveness. The 8th, 9th and 10th rounds of purification only had a relative effectiveness of 54%, 56 % and 48 % respectively and protein peaks similar to that of the 1st round of purification. The consecutive purification profiles (Absorbance at 280 nm) of TFI-scFv isolated from culture supernatant are summarised in Figure 3.2. After the completion of the 10th round of purification, the collected fractions from each round were all combined, desalted and concentrated (See: Section 3.3.5). The final protein concentration was calculated as 0.234 mg.mL^{-1} using the GeneQuant pro RNA/DNA calculator at 280nm.
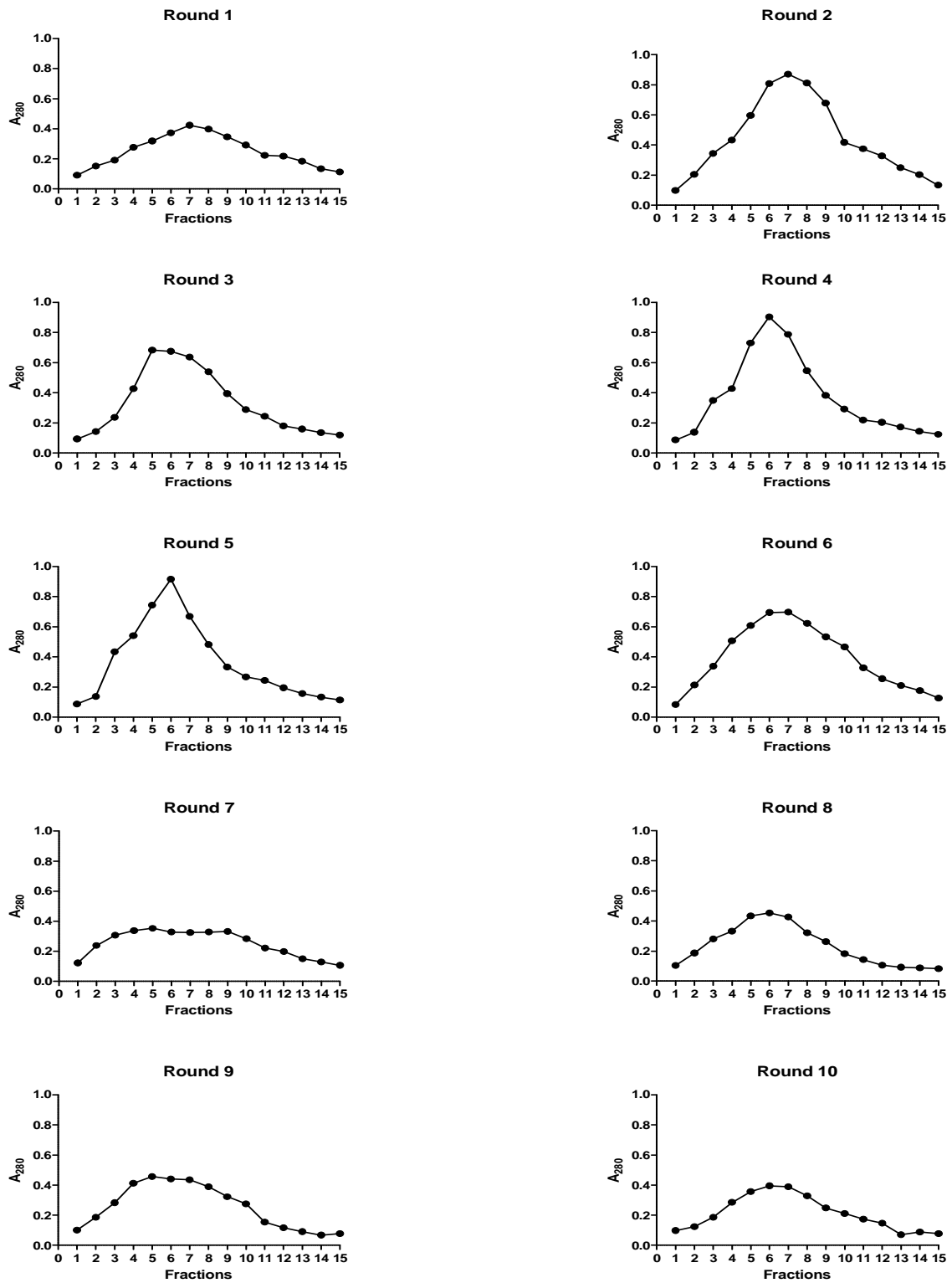
Figure 3.2: Protein A purification profiles

Protein A purification profiles (Absorbance at 280 nm) of TFI-scFv from culture supernatant obtained during ten consecutive rounds of purification.

3.4.2 SDS-PAGE analysis

The final TFI-scFv sample was analysed by means of 12.5 % SDS-PAGE in order to confirm the presence of the TFI-scFv fragment and to determine the purity of the sample.
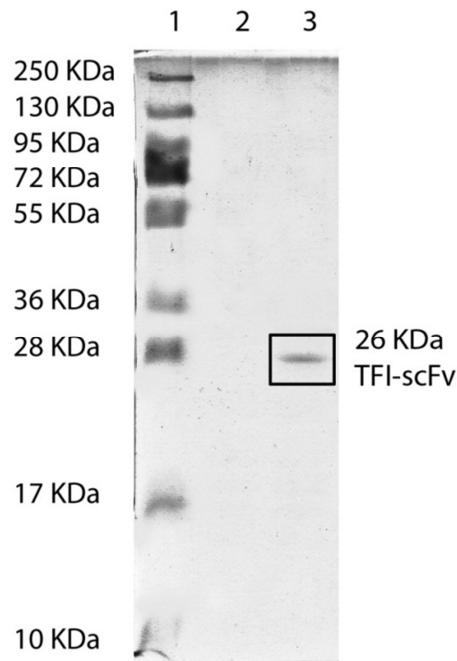
Figure 3.3: SDS-PAGE analysis of Protein A purified TFI-ScFv

The 12.5% SDS-PAGE analysis of final TFI-scFv sample. Lane 1 contains the pre-stained ladder (250 KDa – 10 KDa), Lane 2 is empty and lane 3 contains the 26 KDa TFI-scFv indicated with a black square.

A 10 µL aliquot of the protein ladder was loaded in lane 1 and no sample was loaded in lane 2. A 10 µL volume of the final TFI-scFv sample was loaded in lane 3 as indicated in Figure 3.3. The SDS-PAGE confirmed the presence of a single 26 KDa TFI-scFv fragment in lane 3. The lack of any additional protein bands in lane also confirms that a pure sample of TFI-scFv was isolated by means of protein A affinity chromatography.

3.4.3 Prothrombin Time coagulation Analysis

The functionality and inhibitory effect of escalating dosages of the purified TFI-scFv on human platelet poor plasma was assessed with the Prothrombin Time coagulation test. All tests were performed in triplicate and the time to coagulation

recorded. Final time to coagulation at the various concentrations of TFI-scFv (points A, B, C, D, and E) was calculated as the average of the three recorded values (See: Table 3.6). The extension in coagulation time and the normal patient's prothrombin time range are indicated in Figure 3.4.

Table 3.6: Prothrombin Time coagulation test

Reaction	TFI-scFv[#]	Round 1*	Round 2*	Round 3*	Average*	Time Extension*
A	0.0000	14.90	14.40	14.20	14.50	0.00
B	0.0117	15.00	14.50	15.60	15.03	0.53
C	0.0234	19.60	17.00	18.30	18.30	3.80
D	0.0468	28.70	29.60	31.00	29.77	15.27
E	0.0939	35.40	33.90	36.80	35.37	20.87

[#] - Concentration measured in $\mu g.mL^{-1}$
* - Time measured in seconds
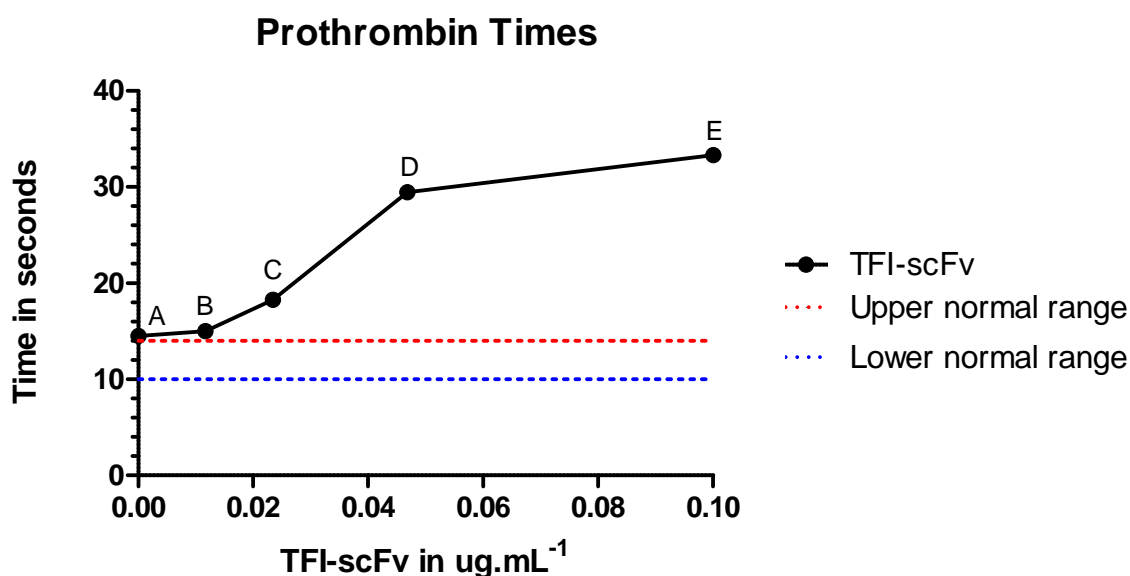
Figure 3.4: Inhibition effect of TFI-scFv

Averages for data points (A to E) were calculated and plotted as time against TFI concentration using GraphPad Prism v 5.0.3. The upper- (▪▪▪) and lower normal (▪▪▪) limits for normal values of PT coagulation test using Thromborel reagent are indicated.

Prolongation of the time to coagulation indicates that TFI-scFv was functional and capable of inhibiting coagulation in a dose depended manner.

3.5 Discussion

3.5.1 Protein A purification

The difference between the effectiveness of the first two consecutive rounds of purification is interesting. The final absorbance of the pooled fractions from the second round is practically a 2 fold improvement on the first round. The second round of purification also produced a much improved protein peak in comparison with the initial round. The improved protein peak obtained during the 2nd round is due to a clear disassociation point between the target protein and column. A clearly defined disassociation point indicates that the protein A column is reactive to the variation in the surrounding environment and is therefore performing optimally. This is important as the protein A column's ability to bind as well as release the target protein is controlled by manipulation of the pH value of the environment. Both purification rounds were performed on the same day with the same TFI-scFv sample and same buffers. This eliminates the possibility of any variation in TFI-scFv concentration of the sample as well as the variation in buffer composition. This discrepancy might have occurred due to the possibility that the protein A column had not been adequately equilibrated initially. Only after the post equilibration steps of the 1st round did the column became fully equilibrated and subsequently capable to bind more protein as observed in the 2nd round.

Despite the first round, the similarity of protein peaks and high level of effectiveness observed in rounds 2 to 6 indicates that a single nProtein A Sepharose 4 Fast Flow column was able to repeatedly purify TFI-scFv with similar results. However during the 7th round of purification, it became clear that the protein A column was no longer functioning optimally. This sudden reduction in effectiveness is interesting since a single TFI-scFv sample was used throughout all purification rounds and the attempts to remove residual proteins that may interfere with the binding of TFI-scFv restored little of the column's former effectiveness. Although GuHCl eluate was not analysed for the presence of unwanted proteins, one can speculate that the sudden lack of a protein peak observed in round 7 as well as the reduction in protein peaks observed in rounds 8 to 10 was not due to absence of target protein, nor due to the presence of interfering proteins that prevented Protein A from binding to TFI-scFv. The

reduction in capacity is most likely due to the biological nature of the sepharose linked protein A, as it rapidly loses ability to bind the target protein after consecutive use (Hober *et al.*, 2007). This phenomenon can be accelerated by the repeated cleaning required due to exposure to large volumes of bacterial culture supernatant (Follman and Fahrner, 2004).

Although the protein A affinity chromatography was able to produce consistent results from round two through to six, it is clear that the column was not functioning optimally after the sixth round of purification. Apart from the high cost, protein A resin are over 30 times more expensive than some ion exchange resins, this occurrence of gradual reduction in binding capacity after repeated usage is also major drawback with protein A purification (Follman and Fahrner, 2004; Ghose *et al.*, 2006; Hober *et al.*, 2007). An additional complication associated with protein A purification is the potential of ligand leakage, where the protein A itself is removed from the column resin. This is especially a concern when purifying therapeutic agents such as TFI-scFv, as protein A molecules may cause unwanted immunological and physiological responses in humans. An additional complication that may arise due to ligand contamination of sample is inactivation of target antibody due re-association between the isolated antibody and protein A after purification. If ligand leakage is detected, the sample would require additional chromatography steps to remove the protein A, which adds to overall cost involved (Follman and Fahrner, 2004). These are some of the concerns that render Protein A affinity chromatography unsuitable for our requirements and intended future applications.

3.5.2 SDS-PAGE analysis

The SDS-PAGE analysis of the purified TFI-scFv sample confirmed the presence of a single 26 KDa fragment. The slight discrepancy in size is due to the slight difference in motility of pre-stained protein ladders in comparison with unstained ladders. As a result pre-stained ladders migrate slightly faster than the unstained protein leading to the apparent molecular weight of a pre-stained protein marker being frequently less than that of an unstained protein marker. This accounts for the approximate 2 KDa to 5 Kda discrepancy observed in the expected size of the purified TFI-scFv.

The lack of any additional bands confirms that an uncontaminated sample of TFI-scFv was purified from the culture supernatant. The lack of a 42 KDa protein band is especially important as it confirms that protein A ligand leakage did not occur. The degradation of the column occurs after repeated use due to the low chemical and proteolytic stability of protein A. As a result, protein A ligand leakage occurs after repeated cleaning of the column under harsh conditions. This leads to the contamination of the purified sample by either the whole ligand or fragments of the ligand (Verdoliva *et al.*, 2002). The presence of protein A in the purified TFI-scFv sample is especially of concern due to the intended therapeutic nature of TFI. Protein A contamination, if not detected or even present at low levels would elicit an immune response in humans and mammals and so doing complicate any future characterization of TF-scFv (Atkins *et al.*, 2008). High purity of the TFI-scFv sample was expected due to the high specificity of protein A for the scfv (Ghose *et al.*, 2006; Hober *et al.*, 2007). The lack of protein A leakage observed during ten rounds of purification was ideal as no additional purification procedures were required to remove the contaminants. Although, the TFI-scFv purification was not complicated by protein A leakage, it still remains a constant concern especially with the repeated use of a single protein A column, in an up-scale setting.

In total, approximately 0.450 mg of TFI-scFv was purified from 250 mL of bacterial culture. In this case, the purification process was limited to only 250 mL due to the use of gravitational chromatography. Gravitational based chromatography is extremely time consuming and therefore it is the main limiting factor in an up-scale setting where large sample volumes are involved. If one extrapolates these results, the total protein yield from one litre of bacterial cell culture on would theoretically result in roughly 2 mg protein. However, if purified as described in this chapter, this would have taken approximately four weeks to complete as well as drastically increased the total cost involved. The TFI-scFv yield obtained in this experiment is similar to general yields obtained by the Tomlinson I + J expression procedure as reported in literature. Lower yields were reported by Kennel *et al.* (2004), who were able to produce 100 µg to 1 mg of soluble scFv from a litre (approximately 6 g wet weight) isolated from the periplasm of *E. coli* TG1. Interestingly, this research group reported that higher yields could be obtained by expressing the single antibody fragments in *E. coli* HB2151, but warned that complications with scFv folding may

arise (Kennel *et al.*, 2004). Eteshola (2010) was able to produce 44 µg to 1.8 mg of soluble scFv, pooled together from the supernatant-, periplasmic- as well as osmotic shock fractions of *E. coli* HB2151, per litre of culture (Eteshola, 2010). Marcus and co-workers (2006) utilized Tomlinson I +J Human scFv library to produce 300 µg soluble scFv per litre of culture, isolated from the periplasm of *E. coli* HB2151 (Marcus *et al.* 2006). Barderas *et al.* (2006) reported a relatively high yield ranging between 1 and 3 mg soluble scFv per litre of cell culture, also isolated from the periplasmic fraction of *E. coli* TG1 (Barderas *et al.*, 2006).

Although the Tomlinson I + J expression procedure produced, in our setting, a relatively high protein yield it is still far from the ideal expression system. It is clear from literature that despite the various modifications to the pIT 2 expression system, it is not optimized for truly large to industrial scale protein expression. General published protein yields range between 0.1 to 5.0 mg per litre of *E. coli* culture (Eteshola, 2010). The expression of the scFvs into the culture supernatant also complicates the purification process as it is difficult and more time consuming to process the large and bulky sample volume. Furthermore, due to the highly diluted nature of the target protein within such a large sample volume, additional concentration steps are required after purification. One of the popular strategies to circumvent this complication is to extract scFv from the periplasmic fraction of *E. coli* HB2151. Although this will greatly reduce the total sample volume, a large amount of scFv is lost due to movement of scFv from the periplasma across the cell wall to the culture supernatant. These are some of the contributing factors that render the Tomlinson I + J expression system unsuitable for obtaining the high TFI-scFv yields required for intended future applications. As a result other expression systems must be explored in order to increase protein yield.

3.5.3 Prothrombin time coagulation analysis

The PT coagulation test is a quick and effective test to detect abnormalities within the extrinsic pathway of coagulation. The PT coagulation test determines the time it takes for blood plasma to coagulate once it has been exposed to tissue factor. Thromborel contains tissue factor in association with a lipid layer, which is responsible for the initiation of the coagulation reaction. The PT coagulation test

provides information concerning the interactions of coagulation factors V, VII, X and prothrombin (McVey, 1999; Norris, 2003). Abnormalities of these coagulation factors result in a prolongation of the PT. The PT coagulation test is therefore ideally suited to assess the functionality of a tissue factor inhibitor such as TFI-scFv.

The standard PT coagulation test was modified for higher sensitivity in order to detect the inhibition effect of TFI-scFv at low levels. Point A, in Figure 3.4, served as the baseline and positive control as no TFI-scFv was present. Modification to the standard PT coagulation test resulted in a baseline coagulation time of 14.5 seconds. This is only 0.5 seconds above the upper limit for normal patients (10 to 14 seconds) at our routine haematology laboratory (National Health Laboratory services, Free State, Bloemfontein, RSA) using Thromborel reagent. At point B only a slight inhibition effect was observed. The baseline coagulation time was extended by 0.5 seconds when 0.0117 µg.mL^{-1} TFI-scFv was present. Although some inhibition is present it is clear that TFI-scFv has very little inhibition effect at these low concentrations. At point C, TFI-scFv inhibition effect was slightly improved, baseline coagulation time was extended by 3.8 seconds at a TFI-scFv concentration of 0.0234 µg.mL^{-1}. At point D, 0.0468 µg.mL^{-1} TFI-scFv extended the baseline coagulation time with 15.3 seconds, effectively doubling the time to coagulation. At point E the baseline coagulation time was extended with 20.87 seconds when the reaction was inhibited with 0.0939 µg.mL^{-1} TFI-scFv. It is clear that an inhibition plateau was reached at 0.05 µg.mL^{-1} TFI. Coagulation at point E was extended with 5.5 seconds longer than the inhibition observed in point D, where approximately half (0.0468 µg.mL^{-1}) TFI was present. A possible explanation for the observed plateau would be that the maximum effective TFI-scFv fragment concentration was reached. Once this concentration is reached, all higher dosages beyond this point will not increase inhibition. The plateau of inhibition indicates that the TFI-scFv does not exhibit excessive anticoagulant characteristics. This is in agreement with published findings that tissue factor inhibitors are generally less aggressive (Jacquemin and Saint-Remy, 2004). The less aggressive nature of TFI suggests that it is less likely to cause bleeding complications. Bleeding complications are currently the main drawback of commercial anticoagulants resulting in a reluctance of pharmaceutical companies to produce novel anticoagulants today. Although the inhibitory effect of

TFI is initially not as pronounced, the PT coagulation test confirmed that the TFI was functional and able to inhibit the action of tissue factor in a dose dependent manner.

3.6 Conclusion

Through the up-scaling of the Tomlinson I + J procedure a relatively high yield of functional TFI-scFv was expressed. An uncontaminated TFI-scFv sample was purified by means of protein A affinity chromatography. Despite the success of up-scaling the experiment there are definitive disadvantages to the TomlinsonI + J procedure. Firstly, the Tomlinson I + J human single fold scFv libraries were primarily designed for the effective identification and isolation of a desired phagemid displaying scFv-protein III fusion protein on the surface of the phagemid. It is not designed as an over-expression mechanism for the large scale production of the soluble scFv. It is therefore unsurprising that the expression of the pIT 2 vector in *E. coli* HB2151 was unable to produce sufficient TFI-scfv required for additional characterisation. Secondly, the financial- as well as practical complications associated with protein A affinity chromatography also complicate the optimal and cost effective isolation of large quantities of TFI-scFv. As a result, more cost effective alternative methods were investigated in this research in order to produce and isolate higher yields of functional TFI-scFv for our intended future applications.

3.7 References

ATKINS, K.L., BURMAN, J.D., CHAMBERLAIN, E.S., COOPER, J.E., POUTREL, B., BAGBY, S., JENKINS, T.A, FEIL, E.J. & VAN DEN ELSEN, J.M.H. 2008. S. aureus IgG-binding proteins SpA and Sbi: host specificity and mechanisms of immune complex formation. *Molecular Immunology*, 45, 1600-1611.

BARDERAS, R., SHOCHAT, S., MARTÍNEZ-TORRECUADRADA, J., ALTSCHUH, D., MELOEN, R. & IGNACIO CASAL, J. 2006. A fast mutagenesis procedure to recover soluble and functional scFvs containing amber stop codons from synthetic and semisynthetic antibody libraries. *Journal of Immunological Methods,* 312, 182-189.

ETESHOLA, E. 2010. Isolation of scFv fragments specific for monokine induced by interferon-gamma (MIG) using phage display. *Journal of Immunological Methods, 358*, 104-110.

FAIRBANKS, G., STECK, T.L. & WALLACH, D.F. 1971. Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. *Biochemistry*, 10, 2606-17.

FOLLMAN, D. & FAHRNER, R. 2004. Factorial screening of antibody purification processes using three chromatography steps without protein A. *Journal of Chromatography*, A 1024, 79-85.

GHOSE, S., HUBBARD, B. & CRAMER, S.M. 2006. Evaluation and comparison of alternatives to Protein A chromatography Mimetic and hydrophobic charge induction chromatographic stationary phases. *Journal of Chromatography*, A 1122, 144-152.

HOBER, S., NORD, K. & LINHULT, M. 2007. Protein A chromatography for antibody purification. *Journal of chromatography, Analytical Technologies in the Biomedical and Life Sciences,* 848, 40-47.

HOOGENBOOM, H.R., DE BRUÏNE, A P., HUFTON, S.E., HOET, R.M., ARENDS, J.W. & ROOVERS, R.C. 1998. Antibody phage display technology and its applications. *Immunotechnology : An International Journal of Immunological Engineering*, 4, 1-20.

JACQUEMIN, M. & SAINT-REMY, J.M. 2004. The use of antibodies to coagulation factors for anticoagulant therapy. *Current Medicinal* Chemistry, 11, 2291-2296.

KENNEL, S.J., LANKFORD, T., FOOTE, L., WALL, M. & DAVERN, S. 2004. Phage display selection of scFv to murine endothelial cell membranes. *Hybridoma and Hybridomics*, 23, 205-211.

KRETZSCHMAR, T. & VON RUDEN, T. 2002. Antibody discovery: phage display. *Current Opinion in Biotechnology*, 13, 598–602.

MARCUS, W., WANG, H., LOHR, D., SIERKS, M. & LINDSAY, S. 2006. Isolation of a scFv targeting BRG1 using phage display with characterization by AFM. *Biochemical and Biophysical Research Communications*, 342, 1123–1129.

MARCUS, W.D., WANG, H., LINDSAY, S.M. & SIERKS, M.R. 2008. Characterization of an antibody scFv that recognizes fibrillar insulin and [beta]-amyloid using atomic force microscopy. *Nanomedicine: Nanotechnology, Biology and Medicine*, 4, 1–7.

MCVEY, J.H. 1999. Tissue factor pathway. *Best Practice & Research Clinical Haematology*, 12, 361–372.

NISSIM, A, HOOGENBOOM, H.R., TOMLINSON, I.M., FLYNN, G., MIDGLEY, C., LANE, D. & WINTER, G. 1994. Antibody fragments from a "single pot" phage display library as immunochemical reagents. *The EMBO Journal*, 13, 692-698.

NORRIS, L.A. 2003. Blood Coagulation. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 17, 369-383.

SMITH, G.P. & PETRENKO, V. 1997. Phage Display. *Chemical Reviews*, 97, 391-410.

VERDOLIVA, A., PANNONE, F., ROSSI, M., CATELLO, S. & MANFREDI, V. 2002. Affinity purification of polyclonal antibodies using a new all-D synthetic peptide ligand: comparison with protein A and protein G. *Journal of Immunological Methods*, 271, 77-88.

WATKINS, N.A. & OUWEHAND, W.H. 2000. Introduction to antibody engineering and phage display. *Vox Sanguinis,* 78, 72-79.

WEISSER, N.E. & HALL, J.C. 2009. Applications of single-chain variable fragment antibodies in therapeutics and diagnostics. *Biotechnology Advances*, 27, 502-520.

WILLATS, W.G.T. 2002. Phage display: practicalities and prospects. *Plant Molecular Biology*, 50, 837-854.

WU, S., KE, A. & DOUDNA, J. 2007. A fast and efficient procedure to produce scFvs specific for large macromolecular complexes. *Journal of Immunological Methods*, 318, 95-101.

YAN, J.P., KO, J.H. & QI, Y.P. 2004. Generation and characterization of a novel single-chain antibody fragment specific against human fibrin clots from phage display antibody library. *Thrombosis Research*, 114, 205–211.

CHAPTER 4

Purification of TFI-scFv by means of nickel affinity chromatography

4.1 Introduction

As stated previously, the TFI-scFv was isolated from the TomlinsonI + J phage library. It is therefore important to discuss the structure and composition of the phagemid vector utilised during phage display. Phage display was first introduced by G. Smith in 1985 (Azzazy and Highsmith, 2002). In the years since its introduction, antibody phage display has evolved into a well-accepted technology. It is a robust technology that provides a convenient system for the selection and expression of antibodies against a large range of epitopes. Phage display derived antibodies have also proven their safety and efficacy in clinical trials (Kretzschmar and von Ruden, 2002). The Tomlinson I + J Library scFv phagemid library contains a synthetic V-gene (VH–VL) cloned into the pIT2 phagemid vector (See: Figure 4.1) (Wu *et al.*, 2007; Eteshola, 2010).

Synthetic diversity in these libraries is generated through the rearrangement of the heavy- and light chain (VH and VL) segments in vitro and by generating random heavy chain complimentary determining regions by using PCR and randomised oligonucleotides (Wind *et al.*, 1997; Barderas *et al.*, 2006). The phagemid vector (pIT2) contains a lac promoter and a pelB leader sequence up stream of the heavy- and light chain inserts, followed by Histidine- and Myc tags, an amber (TAG) stop codon and the gene encoding phage coat protein (pIII) (Aburatani *et al.*, 2002). Expression in *E. coli* HB2151 will make use of amber stop codons and lead to the expression of only the scFv consisting of pelB/Vh/linker/VL/His-tag/Myc-tag. The utilisation of the amber codon prevents the co-expression of the downstream coat protein (pIII) and the subsequent production of the scFv-pIII fusion protein (See: Figure 4.1).
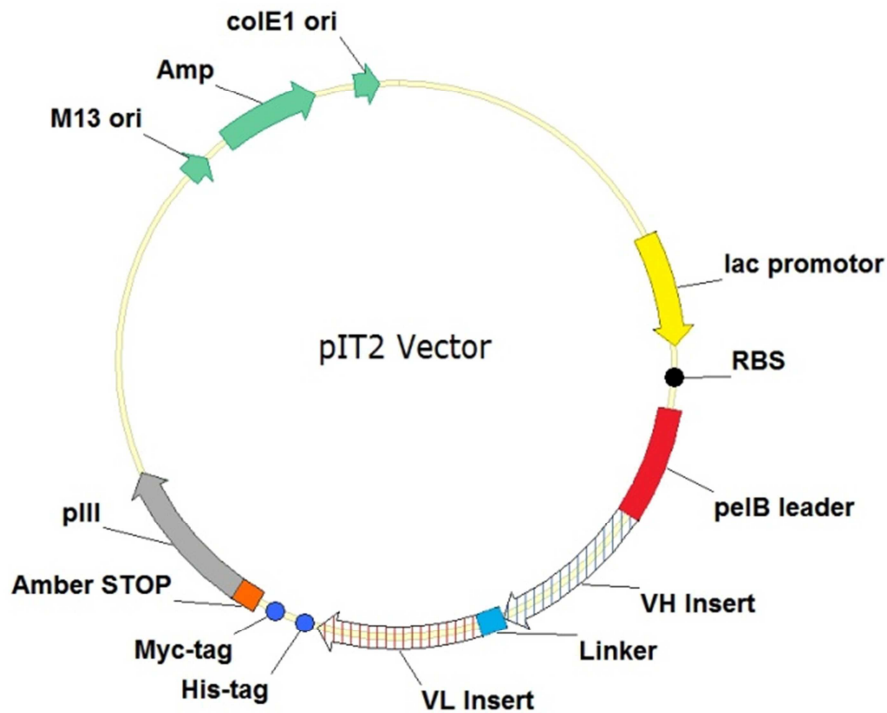
Figure 4.1: pIT2 phagemid vector map

Expression is regulated via the upstream lac promoter (⟶). Ribosomal binding site (RBS) is indicated with a (●). The expressed scFV consist of the pelB Leader sequence (pelB Leader - ▬), heavy chain insert (VH Insert - ⟩)**,** glycine-serine linker sequence (Linker - ▬), and light chain insert (VL Insert - ⟩), polyhistidine sequence (His-tag - ●) and c-myc sequence (Myc-tag - ●). A stop codon (Amber STOP - ▬) is incorporated upstream of the phagemid coat protein III (pIII- ⟶). M13 phagemid origin of replication (M13 ori), ampicillin resistance gene (Amp) and ecoli origin of replication (colE1 ori) are indicated by (⟶).

Recombinant proteins are generally modified to contain either an N- or C-terminus poly-histidine tag to facilitate simple protein purification by means of immobilized metal-affinity chromatography (IMAC) (Nissim *et al*., 1994; Azzazy and Highsmith, 2002; Guo *et al*., 2006). The pIT2 phagemid vector incorporates a His-tag downstream of the expressed scFv located at the C-terminus. IMAC was originally developed for protein purification by Porath and co-workers in the 1970's (Porath *et al*., 1975). The IMAC technique utilizes chelating compounds that are covalently bonded on solid chromatographic resin to capture metal ions such as Nickel (Ni^{2+}), Zink (Zn^{2+}), Cobalt (Co^{2+}) and Copper (Cu^{2+}).
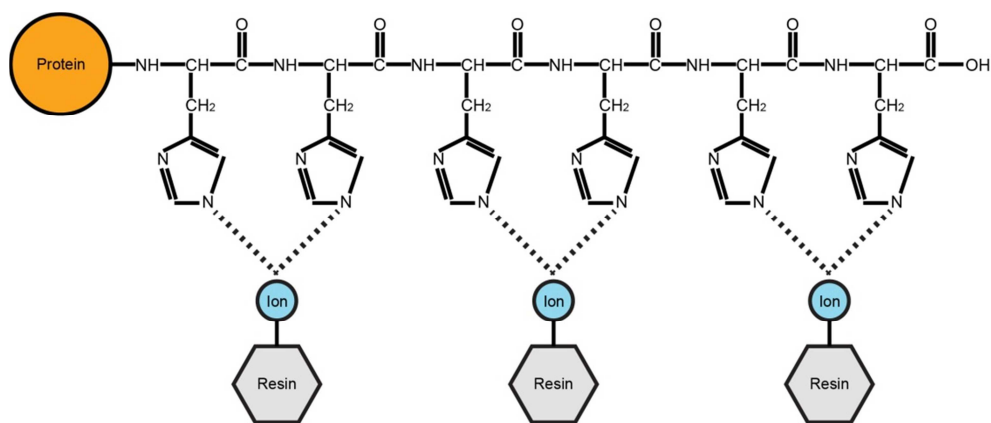
Figure 4.2: Principle of IMAC purifications

The association of a poly-histidine tagged protein to immobilized ions. Derived from Astatke (2011).

The immobilized metal ions serve as affinity ligands for proteins containing exposed poly-histidine residues (See: Figure 4.2) (Smith *et al.*, 1988; Zhao *et al.*, 2009; Gaberc-Porekar & Menart, 2001). The target protein containing the His-tag is chemically adsorbed, while unwanted proteins elute from the column during the washing phase. Target protein is eluted by protonation, ligand exchange or extraction of the immobilised metal ion by a strong metal chelator (Smith *et al.*, 1988). Despite being less specific than Protein A affinity chromatography, IMAC does have a number of advantages over bio-specific affinity chromatography. The advantages of IMAC include; improved ligand stability, high protein loading, milder elution conditions, simple regeneration and low cost (Gaberc-Porekar and Menart, 2001). As a result IMAC is ideally suited for the rapid purification of a target protein where substantial purity is required at low cost (Gaberc-Porekar and Menart, 2001).

4.2 Aim

The aim of this part of the study was to purify TFI-scFv from *E. coli* HB2151 culture media by means of a nickel based immobilized metal affinity chromatography as a more economical alternative to protein A affinity chromatography.

4.3 Materials and Methods

KM13 helper phage was produced in *E. coli* TG1 and utilized in the production of phagemids expression tissue factor inhibitor fusion protein (TFI-phagemid). Isolated TFI-phagemid was used to transfect *E. coli* HB2151 that was induced to express soluble tissue factor inhibitor single chain variable fragment (TFI-scFv) into the culture supernatant.

4.3.1.1 Production of KM13 helper phage

The helper phage was produced according to the protocol G as specified in the product documentation (Tomlinson I + J: Human Single Fold Libraries I + J) as previously described in Section 3.3.1.1.

4.3.1.2 Production of TFI-phagemid

The tissue factor inhibiting phage was produced according to the protocol A as stipulated in the product documentation (TomlinsonI + J: Human Single Fold Libraries I + J) as previously described in Section 3.3.1.2.

4.3.1.3 Production of TFI-scFv

The soluble TFI-scFv was produced according to the protocol F as stipulated in the product documentation (Human Single Fold Libraries I + J - TomlinsonI & J) as previously described in Section 3.3.1.3.

4.3.2 Immobilized metal ion affinity chromatography (IMAC) purification of HB2151 expressed TFI-scFv

The nickel charged sepharose HisTrap™ FF 5 mL column was prepared according to the manufacturer's specifications (GE Healthcare, USA). All buffers were prepared with high quality distilled water and filtered with a 0.22 µm syringe-driven filter unit (Millipore, USA) before use.

The ÄKTA Prime Plus (Amersham Biosciences, UK) system was flushed and loaded with distilled water. Afterwards, the pre-packed HisTrap™ FF column was connected to the ÄKTA Prime Plus system. Care was taken to prevent the formation of air pockets. The HisTrap™ FF column was flushed with 15 mL distilled water and equilibrated with 25 mL binding buffer (4.186 g.L^{-1} MOPS, 29.22 g.L^{-1} NaCl, 1.36 g.L^{-1} imidazole at pH 7.4) at the flow rate of 5 mL per minute. The harvested supernatant containing the TFI-scFv was prepared by diluting it with binding buffer in a 1:1 ratio and applied to the HisTrap™ FF column at a flow rate of 10 mL per minute. After the sample was loaded, the HisTrap™ FF column was flushed with 50 mL binding buffer in order to remove all unbound particles. The poly-histidine tagged TFI-scFv was then eluted with an imidazole gradient (100 mL), ranging from 20 mM to 500 mM, provided by the ÄKTA Prime Plus system through the combination of the binding buffer and the elution buffer (4.186 g.L^{-1} MOPS, 29.22 g.L^{-1} NaCl, 34.04 g.L^{-1} imidazole at pH 7.4). The eluted phase was collected in 5 mL fractions and stored at 4 °C until later use. Once the elution phase was completed, the His-Trap™ FF column was flushed with 50 mL binding buffer and visually inspected. It was then compared with an unused recharged column, a column subjected to a blank run performed with only the buffers and a column subjected to 2 L of sterile LB media that served as the negative control. The blank run and negative control were performed at the same flow rates and volumes as described above.

4.3.3 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SDS-PAGE was performed on culture supernatant and the negative control in order to confirm the expression of a 26 KDa fragment. A volume of 500 µL culture supernatant before and after purification as well as negative control supernatant was concentrated in the Vacufuge Concentrator 5301 (Eppendorf, Germany) to a final volume of 50 µL. SDS-PAGE was performed as previously described in Section 3.3.3.

4.3.4 TFI-scFv sequencing

The synthetic V-gene coding for TFI-scFv was isolated from TFI-phagemid by means of PCR. The DNA amplicon was purified from agarose gel and utilized for sequencing.

4.3.4.1 TFI-phagemid DNA isolation

A volume of 200 µL polyethylene glycol 20000 – sodium chloride (PEG/NaCl) solution (300 g.L^{-1} PEG and 58.44 g.L^{-1} NaCl) was added to 50 µL TFI-phage (See: Section 4.3.1.2). The content was mixed gently by inversion of the micro-centrifuge tubes and placed for 10 minutes on ice. The mixture was then centrifuged at 13000 x *g* for 10 minutes at room temperature and the supernatant discarded. The residual supernatant was removed by briefly centrifuging the pellet for an additional 30 seconds and aspirated with a 200 µL pipette. The remaining pellet was then dissolved in 100 µL 0.4 M sodium iodine buffer (60 g.L^{-1} NaI, 130 mg.L^{-1} tris and 30 mg.L^{-1} EDTA at pH 8.0) and 250 µL cold absolute ethanol (-20 ºC). The dissolved pellet was placed at room temperature for 10 minutes and then centrifuged at 13000 x *g* for 10 minutes at room temperature. The pellet was retained and the supernatant discarded. A volume of 200 µL cold 70 % [v/v] ethanol (-20 ºC) was added to the pellet and the contents was gently mixed by inversion of the micro-centrifugation tubes. The mixture was then centrifuged briefly at 13000 x *g* at room temperature and the supernatant removed. The sample was dried in an evaporation chamber for 2-5 minutes. After all the ethanol was evaporated the remaining DNA pellet was dissolved in 30 µL MilliQ sterile water and stored at -20 ºC until later use. The DNA concentration was determined at 260nm with GeneQuant pro RNA/DNA calculator.

4.3.4.2 TFI-phagemid DNA Polymerase Chain Reaction amplification

Polymerase Chain Reaction (PCR) was performed using a specific DNA primer pair (Integrated DNA technologies, Belgium): LMB3 (forward primer): 5' CAG GAA ACA GCT ATG AC 3' & pHEN (reverse primer): 5' CTA TGC GGC CCC ATT CA 3' as specified in protocol I of the product protocol (Human Single Fold Libraries I + J - TomlinsonI & J). All PCR amplifications were performed using the G-Storm thermal

cycler (Vacutec, RSA) in 20 µL total volume reactions prepared as shown in Table 4.1.

Table 4.1: PCR Reaction

Reagent	Concentration	Volume per reaction
Target DNA	Approximately 50 ng.µL^{-1}	1 µL
Phusion High-Fidelity Reaction Buffer	(7.5 mM MgCl)	2 µL
LMB3 Forward primer	100 pmol.µL^{-1}	0.2 µL
pHEN Reverse primer	100 pmol.µL^{-1}	0.2 µL
Phusion High-Fidelity DNA polymerase	2 U.µL^{-1}	0.2 µL
MilliQ sterile water	-	14.4 µL
dNTPs	2.5 mM of each	2 µL
Total		20 µL

The negative control was prepared as described above by replacing 1 µL TFI-phagemid DNA with 1 µL MilliQ sterile water. Thermo-cycling parameters were configured to the following conditions: A hot start preheated to 110 ⁰C followed by a denaturing cycle of 10 minutes at 94 ⁰C. This was followed by 32 cycles of: 15 seconds at 95 ⁰C, 15 seconds at 60 ⁰C and 1 minute at 72 ⁰C. The PCR was completed with a cycle of 10 minutes at 72 ⁰C and finally maintained at 4 ⁰C until removed.

4.3.4.3 Agarose gel electrophoresis confirmation of PCR product

The presence of 935 base pair (bp) PCR product was confirmed for documentation purpose by means of an 3 % agarose gel (30 g.L^{-1} agarose, 0.6 µg.mL^{-1} ethidium bromide, 6.3 g.L^{-1} Tris-HCL, 372.22 mg.L^{-1} EDTA, 1.26 mg.L^{-1} glacial acetic acid at pH 8) electrophoresis with TAE electrophoresis buffer (6.3 g.L^{-1} Tris-HCL, 372.22 mg.L^{-1} EDTA, 1.26 mg.L^{-1} glacial acetic acid at pH 8) at 100 V for approximately 1 hour and 30 minutes. A 10 µL aliquot of PCR product was mixed with 2 µl loading dye (Fermentas, Canada,) and loaded onto an agarose gel. MassRuler^{TM} DNA ladder comprising of DNA standards ranging in size from 10000 bp to 80 bp (Fermentas, Canada) was used to confirm correct band size. PCR products were

visualized with UV-Light and photographed using the ChemiDoc XRS (BioRad, USA) gel documentation system.

4.3.4.4 Agarose gel electrophoresis isolation of PCR product

The 935 bp PCR product was confirmed by means of an 3 % agarose gel (30 g.L^{-1} Agarose, 50 µL.L^{-1} Goldview (Guangzhou GeneShun Biotech, China), 6.3 g.L^{-1} Tris-HCL, 372.22 mg.L^{-1} EDTA, 1.26 mg.L^{-1} Glacial acetic acid at pH 8) electrophoresis with TAE electrophoresis buffer at 100 V for approximately 30 minutes. A volume of 20 µL PCR product was mixed with 2 µl loading dye (Fermentas, Canada) and loaded onto an agarose gel. The MassRuler^{TM} DNA ladder was used to confirm correct band size. PCR product was visualized on the darkreader transilluminators (Clare Chemical Research, USA) with blue light (490 nm). The bands corresponding to the desired amplicon size of 935 bp were purified from the agarose gel using the Biospin Gel Extraction kit (Bioflux, Japan) according to the product protocol.

Briefly, the desired fragment was excised with a sterilised scalpel, transferred to 1.5 micro-centrifuge tubes and weighed. A volume of 3 µL extraction buffer was added for each 1 mg of excised agarose material. The excised agarose was incubated at 50 ⁰C and vortexed every two to three minutes until completely melted. The melted material was added to the Bioflux Spin Columns, centrifuged at 6000 x g for 60 seconds at room temperature where after the flow through was discarded. A volume of 500 µL extraction buffer was added to each of the Spin columns. The Spin columns were then centrifuged at 12000 x g for 60 seconds at room temperature after which the flow through was discarded. The Spin columns were washed with 750 µL wash buffer, centrifuged at 12000 x g for 60 seconds at room temperature where after the flow through was discarded. Columns were centrifuged for an additional 12000 x g for 60 seconds at room temperature in order to remove any remaining residual liquid. Purified DNA amplicon was eluted using 50 µL MilliQ sterile water and the DNA concentration was determined with the NanoDrop Spectrophotometer ND-1000 (Thermo Fisher Scientific, USA). Isolated DNA samples were stored at – 20 ⁰C until later use.

Bidirectional DNA sequencing of PCR products were performed with the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA) using the LMB3 (forward primer): 5' CAG GAA ACA GCT ATG AC 3' & pHEN (reverse primer): 5' CTA TGC GGC CCC ATT CA 3' according to the manufacturer's guidelines.

All sequencing reactions were performed at 1/16th reaction size using the G-Storm thermal cycler (Vacutec, RSA). Sequencing reactions were performed in a total volume of 10 µL prepared as shown in Table 4.2. A positive control was prepared as 1/16th in 10 µL total volume as shown in Table 4.3.

Table 4.2: TFI sequencing reactions

Reagent	Concentration	Volume per reaction
Target DNA	Approximately 10 - 40 ng.µL^{-1}	1 µL
BigDye Sequencing Buffer		2 µL
Ready Reaction Premix		1 µL
Sequencing primer (LMB or pHEN)	1 pmol.µL^{-1}	3.2 µL
MilliQ sterile water	-	2.8 µL
Total Volume		10 µL

Table 4.3: Sequencing control reaction

Reagent	Concentration	Volume per reaction
Control plasmid {pGEM-3Zf(+)}	0.2 µg.µL^{-1}	2 µL
BigDye Sequencing Buffer		2 µL
Ready Reaction Premix		0.5 µL
Sequencing Control Primer	0.8 pmol.µL^{-1}	4 µL
MilliQ sterile water	-	1.5 µL
Total Volume		10 µL

Thermocycling parameters for reactions were configured to the following conditions: a denaturing cycle of 1 minute at 94 ºC, followed by 25 cycles of 10 seconds at 96 ºC, 5 seconds at 50 ºC and 4 minute at 60 ºC. After completion the reaction was maintained at 4 ºC.

The post sequencing cleanup was performed according to the EDTA/Ethanol precipitation protocol as specified in the BigDye® Terminator v3.1 Cycle Sequencing Kit.

Briefly, the sequencing reactions volumes were adjusted to 20 µL with MilliQ sterile water and transferred to a 1.5 mL micro-centrifuge tube containing 5 µL EDTA (125 mM) and 60 µL absolute ethanol. Samples were vortexed for 5 seconds and precipitated at room temperature for 15 minutes. Samples were then centrifuged at 20000 x g for 15 minutes at 4 ºC and completely aspirated. A volume of 200 µL 70% ethanol was transferred to each sample and centrifuged at 20000 x g for 5 minutes at 4 ºC. Samples were then completely aspirated and dried in the Vacufuge Concentrator 5301 (Eppendorf). Purified samples were stored in the dark at 4 ºC until processed on the 3130xl Genetic Analyser HITACHI (Applied Biosystems). Sequencing data was imported to Geneious v 4.8.5 and contiguous sequence was compiled. The finalised consensus TFI-scFv DNA sequence was aligned with empty pIT2 plasmid in order to identify the synthetic V-gene insert using internet based Lalign software (http://www.ch.embnet.org/software/LALIGN_form.html). The DNA sequence was imported to Vector NTI v 9.0 and the relevant genetic motifs were identified.

4.4 Results

4.4.1 Immobilised Metal Affinity chromatography

The purification of the TFI from the *E. coli* HB2151 culture media was attempted using the nickel affinity chromatography due to the presence of C-terminus His-tag on TFI-scFv. The column was washed with binding buffer in order to remove unbound protein between points A and B as indicated in Figure 4.3. Target proteins were eluted using imidazole gradient (20mM to 500 mM). Due to imidazole having a higher affinity for the nickel than the His-tag present on the protein, it will displace target proteins bound to the column. The imidazole gradient was introduced between point B and C.
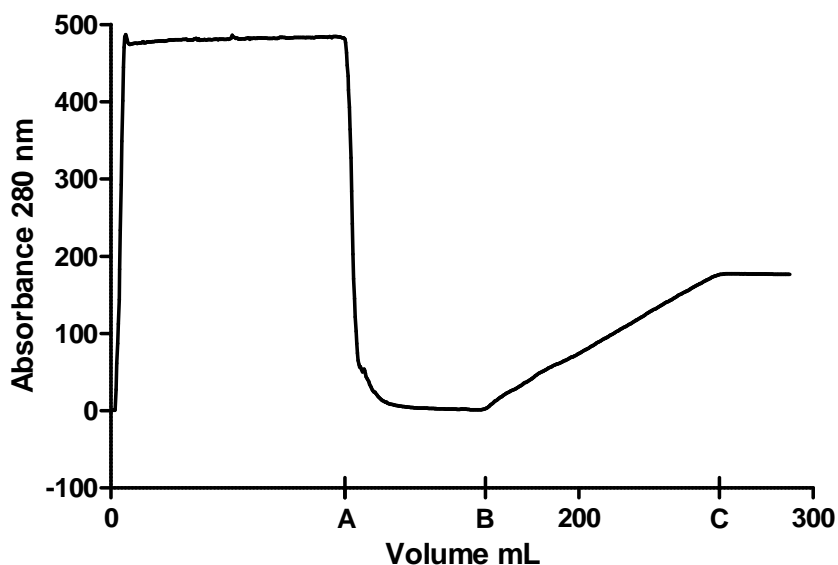
Figure 4.3: Purification profile for TFI-scFv expressed in *E. coli* HB2151

The purification profile was obtained by continuously monitoring (each 0.333 µL) the absorbance of the eluate at 280 nm with the Äkta Prime Plus System.

No clear TFI-scFv protein peak was obtained, indicating that nickel affinity chromatography was unable to purify TFI-scFv from the culture media.

4.4.2 HIS-Trap™ FF column after purification

The His-Trap™ FF column utilised during the purification attempt (D) was visually inspected after the purification and compared to the recharged column (A), a column exposed to a blank run consisting of only binding- and elution buffers (B) and a column exposed to 2 L sterile LB media (C) that served as a negative control as indicated in Figure 4.4. The loss of nickel was clear by reduction of the blue colour when compared with the recharged column (A) and the column used during the purification (D). The blue colour of the column that was only exposed to the buffers (B) indicated that the binding- and elution buffers will not remove the nickel from the column, as expected. There was a slight discolouring in the column subjected to 2 L sterile LB media (C). Although it seems like the uncultured LB media will remove the nickel from the column to a slight degree, it still does not account for the severe loss of nickel observed in the column used during the purification (D).
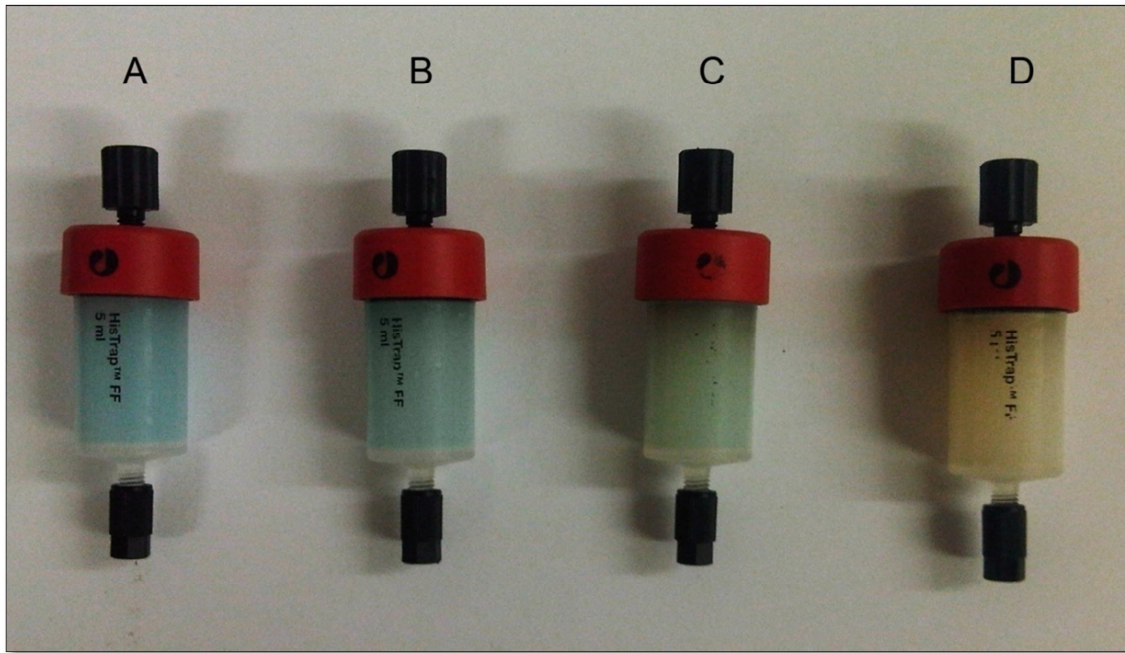
Figure 4.4: Comparison of nickel content of His-Trap^{TM} FF columns

Comparison of a recharged HisTrap^{TM} FF column (A) and columns subjected to a blank run (B), uncultured LB media (C) and the column utilized during TFI-scFv purification attempt (D).

4.4.3 SDS-PAGE analysis of unbound supernatant fraction

The nonbinding eluate was collected and subjected to 12.5 % SDS-PAGE in order to confirm the presence of a desired 26 KDa TFI-scFv fragment. A volume of 10 µL protein ladder was loaded in lane 1 and 20 µL culture supernatant was loaded in lane 2, 20 µL non-binding eluate was loaded in lane 3 and a negative control in lane 4 as indicated in Figure 4.5.
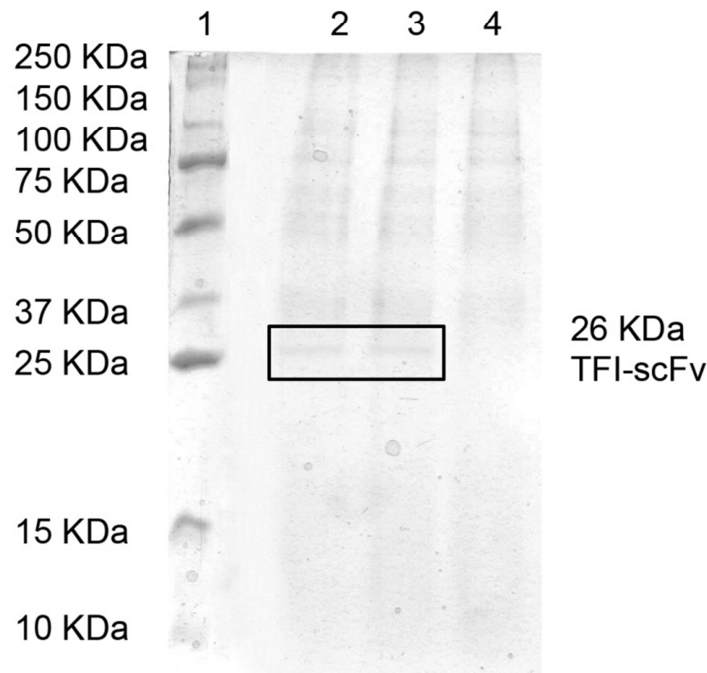
Figure 4.5: SDS-PAGE analysis of culture media

The SDS-PAGE analysis of culture media before and after nickel affinity chromatography purification. The presence of 26 KDa TFI-scFv is indicated by the frame in lane 2 and 3. The 26 KDa TFI-scFv is absent in the negative control in lane 4.

The presence of the 26 KDa fragment in the culture supernatant (Lane 2) combined with the absence of this fragment in the negative control (Lane 4) confirms the successful expression of TFI-scFv. The presence of the TFI-scFv in the eluate (Lane 3) indicates that the nickel affinity chromatography was unable to isolate the target protein from culture media.

4.4.4 TFI-scFv PCR

The PCR product was subjected to 3 % agarose gel electrophoresis and visualised with ethidium bromide in order to confirm the amplification of the TFI-scFv DNA fragment. The MassRuler™ DNA ladder was loaded in lane 1, the PCR product was loaded in lanes 2 and 3 and the negative control in lane 4 as illustrated in Figure 4.6.
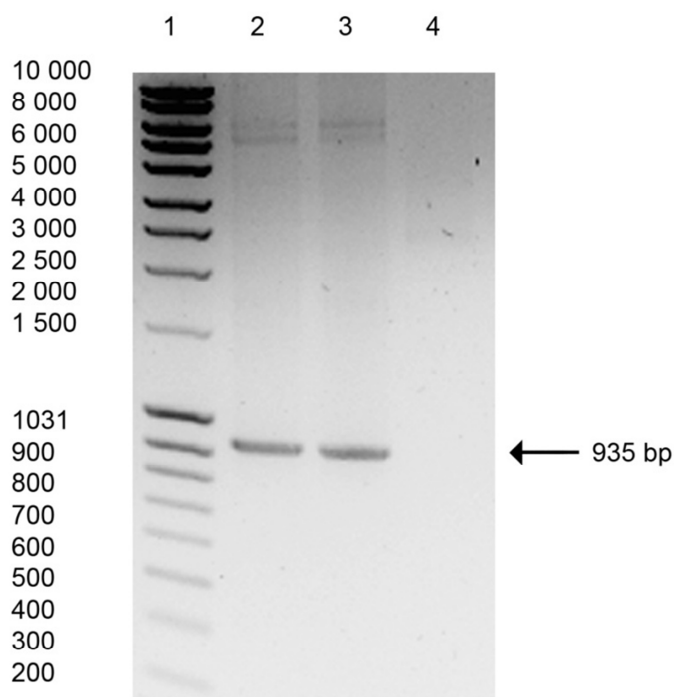
Figure 4.6: Agarose gel electrophoresis of PCR product

The 3 % agarose gel electrophoresis of PCR product. Lane 1 was loaded with DNA ladder and lane 2 and 3 contains the 935 bp TFI-scFv DNA fragment.

The agarose gel electrophoresis confirmed the successful amplification of TFI-scFv synthetic gene (935 bp) using TFI-phagemid template DNA.

4.4.5 TFI-scFv Sequencing

A global alignment (Lalign ExPASy) between TFI-scFv sequence (TFI) and pIT2 Vector reference sequence (pIT2), obtained from the product protocol, was performed and the IgG heavy- and light chain inserts were identified. Internet based Lalign alignment (http://www.ch.embnet.org/software/LALIGN_form.html) results are presented in Table 4.4. The restriction sites (▬) as well as the IgG heavy- (RED) and light chain (BLACK) inserts present in the TFI are indicated.

Table 4.4: TFI- and pIT2 reference sequence global alignment

```
Readseq version 2.1.30 (12-May-2010) Readseq version 2.1.30 (12-May-
        2010)lalign output for pIT2 Reference Sequence vs. pIT2-TFI
        [ISREC-Server] Date: Tue Jul 12 13:52:46 2011


ALIGN calculates a global alignment of two sequences version 2.2u

pIT2 Reference Sequence 332 bp  TFI 938 bp

35.4% identity in 938 nt overlap

Global score: -792


                 10        20        30        40        50        60
pIT2 R CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
TFI  T CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGA
                 10        20        30        40        50        60

                 70        80        90       100       110       120
pIT2 R CAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
TFI  T CAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
                 70        80        90       100       110       120
              NcoI
              130       140
pIT2 R AGCCGGCCATGGCCGAGGTG----------------------------------------
       :::::::::::::::::::::
TFI  T AGCCGGCCATGGCCGAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGG
                 130       140       150       160       170       180


pIT2 R ------------------------------------------------------------

TFI  T GGTCCCTGAGACTCTCCTGTGCAGCCTCTGGATTCACCTTTAGCAGCTATGCCATGAGCT
                 190       200       210       220       230       240

pIT2 R ------------------------------------------------------------

TFI  T GGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGGTCTCATCTATTAATCCGTTGGGTT
                 250       260       270       280       290       300


pIT2 R ------------------------------------------------------------

TFI  T GGAAGACACGTTACGCAGACTCCGTGAAGGGCCGGTTCACCATCTCCAGAGACAATTCCA
                 310       320       330       340       350       360

pIT2 R ------------------------------------------------------------

TFI  T AGAACACGCTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCCGTATATTACT
                 370       380       390       400       410       420
                                                                 XhoI
                                      150       160       170       180
pIT2 R -------------------TTTGACTACTGGGGCCAGGGAACCCTGGTCACCGTCTCGA
                          ::::::::::::::::::::::::::::::::::::::::::::
pIT2-T GTGCGAAAACTTCGTCTAGGTTTGACTACTGGGGCCAGGGAACCCTGGTCACCGTCTCGA
                 430       440       450       460       470       480
```

```
                                                       SalI
            190       200       210       220       230       240
pIT2 R GCGGTGGAGGCGGTTCAGGCGGAGGTGGCAGCGGCGGTGGCGGGTCGACGGACATCCAGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
TFI  T GCGGTGGAGGCGGTTCAGGCGGAGGTGGCAGCGGCGGTGGCGGGTCGACGGACATCCAGA
            490       500       510       520       530       540


pIT2 R TGACCCAG----------------------------------------------------
       ::::::::
TFI  T TGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAGTCACCATCACTTGCC
            550       560       570       580       590       600


pIT2 R ------------------------------------------------------------
TFI  T GGGCAAGTCAGAGCATTAGCAGCTATTTAAATTGGTATCAGCAGAAACCAGGGAAAGCCC
            610       620       630       640       650       660


pIT2 R ------------------------------------------------------------
TFI  T CTAAGCTCCTGATCTATGCTGCATCCAGTTTGCAAAGTGGGGTCCCATCAAGGTTCAGTG
            670       680       690       700       710       720


pIT2 R ------------------------------------------------------------
TFI  T GCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGTCTGCAACCTGAAGATTTTG
            730       740       750       760       770       780


pIT2 R ------------------------------------------------------------
TFI  T CAACTTACTACTGTCAACAGAGTTACAGTACCCCTAATACGTTCGGCCAAGGGACCAAGG
            790       800       810       820       830       840

                 NotI           His-tag
                 250       260       270       280       290
pIT2 R -------------GCGGCCGCACATCATCATCACCATCACGGGGCCGCAGAACAAAAAC
                    :::::::::::::::::::::::::::::::::::::::::::::::::::
TFI  T TGGAAATCAAACGGGCGGCCGCACATCATCATCACCATCACGGGGCCGCAGAACAAAAAC
            850       860       870       880       890       900

            300       310       320       330
pIT2 R TCATCTCAGAAGAGGATCTGAATGGGGCCGCATAGACT
       ::::::::::::::::::::::::::::::::::::::::
TFI  T TCATCTCAGAAGAGGATCTGAATGGGGCCGCATAGACT
            910       920       930
```

The relevant genetic motifs present in the TFI-scFv sequence were identified using Vector NTI v9.0 and indicated in Figure 4.7.
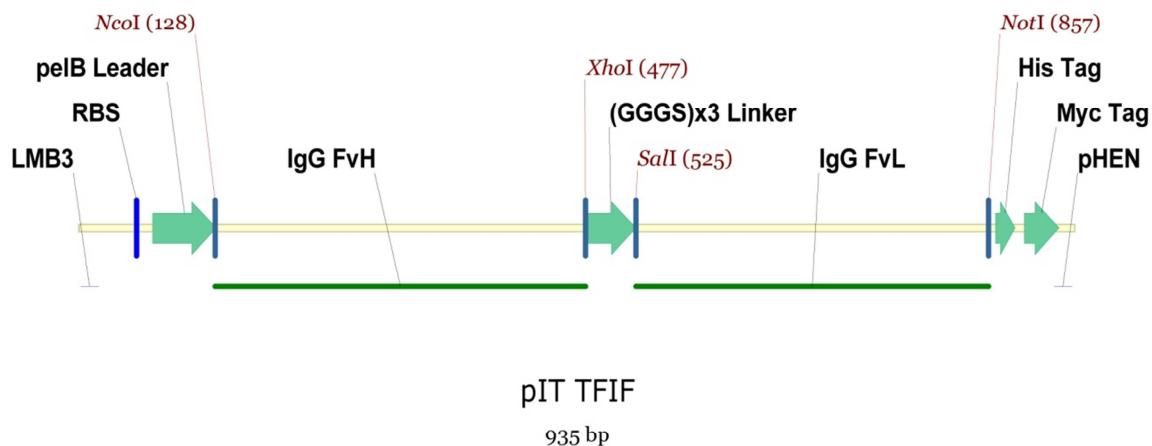
Figure 4.7: Structure of pIT2-TFI sequence

Finalised TFI sequence consists of 935 bp with the forward (LMB3) and reverse (pHEN) primer binding sites indicated as (――). Ribosomal binding site (RBS) is indicated with a (**|**). Human immunoglobulin G heavy- and light chain inserts are indicated as (▬). The pelB Leader sequence (pelB Leader), glycine-serine linker sequence repeat ((GGGSx3) Linker), polyhistidine sequence (HIS Tag) and c-myc gene (Myc Tag) are indicated by (➡). Restriction sites (*NcoI, XhoI SalI* and *NotI*) are indicated as (**|**).

The finalised TFI-scFv DNA sequence was subjected to a six frame translation using internet based ExPASy Translate Tool (http://web.expasy.org/translate). An open reading frame was identified, translating to a 288 amino acid sequence as presented in Table 4.5. The pelB leader sequence (MKYLLPTAAAGLLLLAAQPA), glycine-serine linker sequence (GGGGSGGGGSGGGGS), the six histidine repeat (HHHHHH) and myc-tag (EQKLISEEDLNGAA) at the C-terminus are indicated.

Table 4.5: TFI-scFv Protein Sequence

```
Translation of nucleotide sequence generated by ExPASy

   MKYLLPTAAA GLLLLAAQPA MAEVQLLESG GGLVQPGGSL RLSCAASGFT FSSYAMSWVR

   QAPGKGLEWV SSINPLGWKT RYADSVKGRF TISRDNSKNT LYLQMNSLRA EDTAVYYCAK

   TSSRFDYWGQ GTLVTVSSGG GGSGGGGSGG GGSTDIQMTQ SPSSLSASVG DRVTITCRAS

   QSISSYLNWY QQKPGKAPKL LIYAASSLQS GVPSRFSGSG SGTDFTLTIS SLQPEDFATY

   YCQQSYSTPN TFGQGTKVEI KRAAAHHHHH HGAAEQKLIS EEDLNGAA
```

4.5 Discussion

4.5.1 Nickel affinity chromatography

Even though the Tomlinson I + J protocol claims that protein purification by either Protein A or L yields better results than purification by agents that target the Myc- or His-tag, the total lack of target protein came as a surprise. No protein was eluted from the column during the imidazole elution gradient between point B at 20 mM and point C at 500 mM (See: Figure 4.3). The increase in absorbance is due to the increasing imidazole concentration and not the presence of target protein. Imidazole presents absorbance at 280 nm and as is evident from the linear increase of the absorbance observed between points B and C. A large culture volume was prepared in order to compensate for the relatively low expression levels of TFI-scFv in *E. coli* HB2151. Despite the relatively low expression rate of the pIT2 - *E. coli* HB2151 expression system it was expected, based on the results obtained in chapter 3, that a sufficient amount of protein would be expressed in 2 L of culture in order to purify enough protein for additional downstream testing.

Although the purification attempt was unsuccessful the utilisation of the ÄKTA Prime Plus system in combination with the HIS-Trap^{TM} FF column greatly improved the precision and speed in which the sample could be processed and analysed. A relatively large sample volume of 2 L was processed at a flow rate of 5 mL per minute, effectively reducing weeks required for Protein A purification (As carried out in Chapter 3) to a few hours on the HIS-Trap^{TM} FF column. The ease of use in an up-scale setting was part of the allure that initially prompted the utilisation of nickel affinity chromatography.

There are several factors that may have contributed to the inability of nickel affinity chromatography to isolate TFI-scFv from the culture supernatant. The absence or severely low level of TFI-scFv present in the supernatant caused by unsuccessful IPTG induction of protein expression would result in unsuccessful purification. As a result, the expression of TFI-scFv was investigated by means of SDS-PAGE analysis of the culture supernatant before and after nickel affinity chromatography (See: Section 4.4.3). Another possible factor that may prevent successful purification is the

positioning of the His-tag relative to the TFI-scFv protein. If the C-terminus His-tag is folded in such a manner that it becomes buried within the protein itself, it would be inaccessible and therefore unable to interact with the nickel column. The complete lack of the C-terminus His-tag would also result in the unsuccessful nickel affinity chromatography purification of TFI-scFv. The pIT2 construct was sequenced in order to confirm that the inability of the TFI-scFv to interact with the column was not due to mutation of TFI-scFv phagemid DNA (See: Section 4.4.5). DNA sequencing also confirmed that TFI-scFv contained an in-frame C-terminus His-tag (See: Table 4.5). Even though the above mentioned factors could all prevent the successful purification of TFI-scFv, the most likely explanation is that the TFI-scFv was unable to bind to the His-Trap^{TM} FF column due to the severe loss of nickel from the column.

4.5.2 HIS-Trap^{TM} FF column

On closer visual inspection of the His-Trap^{TM} FF column after purification, it became clear that the nickel was stripped from the column. The loss of nickel is evident from reduction of the blue colour between the recharged column (A) and the column used for TFI-scFv purification (D) (See: Figure 4.4). The blue colour of column B indicates that the binding and elution buffers used during purification do not remove the nickel, as one would expect. This confirms that there are no chelating agents present in the buffers used during the purification. Although nickel ligand leakage can occur during nickel affinity chromatography, it is stated to be very low (in range of up to 1 ppm) when performing nickel affinity chromatography at neutral to slightly alkaline pH (pH 7 to 8) in the presence of 0.5–1.0 M NaCl (Gaberc-Porekar and Menart, 2001). The addition of NaCl to the buffers eliminates the marginal ion-exchange effects that may lead to the unwanted retention of proteins. The use of an imidazole gradient ranging from 20 mM to 500 mM has little effect on nickel leakage (Goubran-Botros and Vijayalakshmi, 1991). In situations where imidazole is not used to elute target protein bound, the extraction of the metal ion bound protein from the column is usually performed by lowering of pH within the range of 2.5 to 7.5. An alternative method to strip the metal ion is through the addition of stronger chelators like ethylene glycol tetraacetic acid (EGTA), EDTA or citrate (Gaberc-Porekar and Menart, 2001).

The slight discolouring of the column C indicates that some of the nickel was removed by LB media. Due to the complex composition of yeast extract used in LB media it is very difficult to pinpoint what exactly may cause the removal of the nickel. Although the media itself seems to be capable of removing some of the nickel from the column, the media alone does not account for the severe loss of nickel from the column used during purification. The severe loss of nickel from the column indicates the presence of a nickel chelating agent, possibly produced by *E. coli* HB2151, in the culture supernatant.

Although nickel is toxic at high levels it remains an essential micronutrient and a catalytic co-factor of enzymes found in *E. coli* (Abelson, 1950; Chivers and Sauer, 2000). Nickel is ubiquitous in microbial environments but is usually present at low levels (Li and Zamble, 2009). As a result, *E. coli* express a variety of proteins that specifically contribute to nickel homeostasis by assembling metallocenters, importing or exporting the metal, transporting nickel ions within the cell, or regulating the expression of the factors involved (Mulrooney and Hausinger, 2003; Li and Zamble, 2009)

It is believed that *E. coli* produce nickel-metalophores that facilitate the uptake of nickel from the environment similar to the Fe^{3+}-siderophores responsible for iron uptake (Dosanjh and Michel, 2006; Li and Zamble, 2009). The presence of a nickel-metalophore in combination with the nickel stripping effect of the LB media could explain the severe loss of nickel from the column as well as the subsequent failure of nickel affinity based purification of TFI-scFv from the bacterial culture supernatant.

4.5.3 SDS-PAGE analysis

The SDS-PAGE analysis confirmed the successful expression of the 26 KDa TFI-scFv into the culture media as well as the presence of TFI-scFv in the unbound eluate after the purification attempt. This confirms that the nickel affinity purification was unable to isolate TFI-scFv from the *E. coli* HB2151 culture media as it is still present in the unbound eluate. This is in agreement with the purification profile obtained on the Äkta Prime Plus System (See: Section 4.4.3) and supports the

hypothesis that the nickel affinity chromatography failed due to the inability of TFI-scFv to interact with the immobilised nickel and not due to the lack of target protein.

The results obtained suggest that the nickel affinity chromatography purification of TFI-scFv from raw culture supernatant would require additional steps in order to be feasible. The simplest solution would be to substitute the culture media with a buffer of choice by means of dialysis. Although this would remove any possible chelating agents from the sample, it would be highly unpractical due to the large sample volume utilised. This is especially true if the production process were to be sufficiently up-scaled in order to produce 5 to 10 mg of TFI-scFv per purification round. The complete dialysis of 2 L or more of culture supernatant would require large amounts of buffer and would therefore considerably increase the time as well as the cost involved in the purification process. Another popular approach to the purification of scFv by means of nickel affinity chromatography is by isolating the periplasmic fraction (Marcus *et al.*, 2006; Hu *et al.*, 2007; Zhao *et al.*, 2009). The major advantage of this is that the supernatant is discarded which eliminates complications associated with nickel stripping agents. An additional advantage to localisation of the expressed protein to the periplasm is that the sample volume is substantially reduced. The difficulty in isolating the target protein from the periplasmic fraction is that the target protein's movement is not restricted to the periplasm. Once the expressed protein is transported to the periplasm it will be passively moved across the cell wall into the culture media (Sørensen and Mortensen, 2005). It is this movement of the target protein from the periplasm to supernatant that enables the purification of the culture media as performed in Chapter 3. By targeting the periplasm, a fraction of expressed protein will always inheritably be lost to the culture media. This coupled with the relative low expression rate of the pIT2 vector in *E. coli* HB2151 would further complicate periplasmic purification of TFI-scFv.

4.5.4 TFI-scFv DNA analysis

The TFI-scFv DNA was isolated from TFI-phagemid and the synthetic TFI-scFv gene sequenced in order to confirm the presence of an in-frame His-tag. Sequencing of the isolated DNA (utilising the LMB3 (forward primer) and the pHEN (reverse

primer)) produced a 935 bp DNA fragment (See: Figure 4.6). Amplification of the empty pIT2 Vector utilising the same primer set produced a 332 bp fragment. This sequence was provided in the product protocol and served as reference. The IgG heavy and light chain inserts are clearly visible from the global alignment between the TFI-scFv sequence and the pIT2 Vector reference sequence. The exact position of the heavy and light chain IgG inserts were identified by locating restriction sites used to construct the synthetic V-gene. The heavy chain was inserted between *Nco*I- and *Xho*I restriction sites and the light chain between *Sal*I-and *Not*I restriction sites (See: Table 4.4). Analysis of genetic motifs in the sequence shows that the synthetic V-gene is expressed in the *E. coli* HB2151 cytoplasm containing *Erwinia carotovora* pectate lyase B (pelB) leader peptide incorporating the ATG (START) codon. The pelB leader peptide facilitates the transport of the expressed protein from *E. coli* HB2151 cytoplasm to the periplasm (Kim *et al.*, 2000; Thie *et al.*, 2008). This transport system is incorporated into the synthetic gene as most antibody fragments are unable to fold correctly in the reducing environment of *E. coli* cytoplasm.

In order to facilitate correct folding of the antibody fragment the pelB leader sequence is incorporated at the N-terminal of the synthetic gene to transport the expressed protein to the oxidising environment of the periplasm (Pannekoek *et al.*, 1995; Thie *et al.*, 2008). Upon arrival in the periplasm, the *pelB* leader peptide is proteolytically cleaved from the expressed protein, yielding the mature protein with the original amino-terminus (Pannekoek *et al.*, 1995). The heavy- and light chain inserts are connected to one another by means of a linker sequence coding for a 15 amino acid Glysine-Serine (GGGGS x 3) repeat. A glycine-rich linker is more flexible than a linker of comparable length composed of non-glycine residues (Robinson and Sauer, 1998). This property of the linker sequence allows it to connect the V_H and V_L domains while providing enough flexibility to keep the adjacent domains in close proximity to one another in order to form an antigen binding site (Megeed *et al.*, 2006). The DNA section coding for the 6 histidine repeat, required for nickel affinity chromatography, is located downstream from the IgG light chain insert. The position of the His-tag in terms of its placement at the N- or C-terminus of the protein, as well as its relative position up- or downstream of the functional domains can influence the expression levels, functionality, as well as the ability to purify the expressed protein by means of nickel affinity chromatography (Goubran-Botros and Vijayalakshmi,

1991; Gaberc-Porekar and Menart, 2001; Klose *et al.*, 2004; Kabir *et al.*, 2010). The myc-tag sequence located downstream from the polyhistidine repeats can be utilized for the detection of scFv by anti-myc antibodies. A TAG (STOP) codon is incorporated shortly after the myc-tag.

Translation of the TFI-scFv DNA open frame sequence produced a 288 amino acid sequence including the pelB leader peptide, linker sequence and six histidine repeat at the C-terminus (See: Table 4.5). The presence of the His-tag once again supports the hypothesis that the nickel affinity chromatography failed due to the inability of TFI-scFv to bind to the immobilised nickel and not due to the lack of the His-tag itself. Although the pIT2 – *E. coli* HB2151 expression system is capable of successfully producing TFI-scFv, it is clear that in order to benefit from the advantages of nickel affinity purification the expression system must be modified in such a way that it is no longer secreted into the culture supernatant. The modification provides the opportunity not only to reduce total sample volume, but also to improve the expression of TFI-scFv by utilizing expression vectors and *E. coli* strains optimised for the over-expression of recombinant proteins.

4.6 Conclusion

TFI-scFv was successfully expressed with a His-tag at the C-terminus as required for nickel affinity chromatography. The nickel affinity chromatography was unable to isolate TFI-scFv possibly due the presence of a nickel chelating factors in the culture supernatant that stripped the nickel from the column. The loss of nickel from the column rendered the purification of TFI-scFv from culture supernatant impossible. Thus, in order to benefit from the many advantages that nickel affinity chromatography present, the expression of TFI-scFv must be modified in such a way that TFI-scFv is not expelled into the culture supernatant. Apart from solving the complications associated with the purification of TFI-scFv from the culture media, it also provides the ideal opportunity to reduce total sample volume and improve TFI-scFv expression through the utilisation of vectors and *E. coli* strains optimised for cytoplasmic over-expression of recombinant genes.

4.7 References

ABELSON, P. 1950. Ion antagonisms in microorganisms: interference of normal magnesium metabolism by nickel, cobalt, cadmium, zinc, and manganese. *Journal of Bacteriology*, 60, 401-413.

ABURATANI, T., UEDA, H. & NAGAMUNE, T. 2002. Importance of a CDR H3 basal residue in V(H)/V(L) interaction of human antibodies. *Journal of Biochemistry*, 132, 775-782.

ASTATKE, M. 2011. Comparison of His Detector Nickel-NTA Conjugates with a Single-Step Antibody Method for the Detection of His-Tagged Proteins. *Nature Methods Application Notes*

Available at: *http://www.nature.com/app_notes/nmeth/2008/081305/full/an5492.html*

[Accessed September 1, 2010].

AZZAZY, H.M.E. & HIGHSMITH, W.E. 2002. Phage display technology: clinical applications and recent innovations. *Clinical Biochemistry*, 35, 425-445.

BARDERAS, R., SHOCHAT, S., MARTÍNEZ-TORRECUADRADA, J., ALTSCHUH, D., MELOEN, R. & IGNACIO-CASAL, J. 2006. A fast mutagenesis procedure to recover soluble and functional scFvs containing amber stop codons from synthetic and semi-synthetic antibody libraries. *Journal of Immunological Methods*, 312, 182-189.

CHIVERS, P.T. & SAUER, R.T. 2000. Regulation of high affinity nickel uptake in bacteria. Ni2+-Dependent interaction of NikR with wild-type and mutant operator sites. *The Journal of Biological Chemistry*, 275, 19735-19741.

DOSANJH, N.S. & MICHEL, S.L.J. 2006. Microbial nickel metalloregulation: NikRs for nickel ions. *Current Opinion in Chemical Biology*, 10, 123-130.

ETESHOLA, E. 2010. Isolation of scFv fragments specific for monokine induced by interferon-gamma (MIG) using phage display. *Journal of Immunological Methods*, 358, 104-110.

GABERC-POREKAR, V. & MENART, V. 2001. Perspectives of immobilized-metal affinity chromatography. *Journal of Biochemical and Biophysical Methods*, 49, 335-360.

GOUBRAN-BOTROS, H. & VIJAYALAKSHMI, M. 1991. Immobilized metal ion affinity electrophoresis: a preliminary report. *Electrophoresis*, 12, 1028-1032.

GUO, Z., BI, F., TANG, Y., ZHANG, J., YUAN, D., XIA, Z. & LIU, J.N. 2006. Preparation and characterization of scFv for affinity purification of reteplase. *Journal of Biochemical and Biophysical Methods*, 67, 27-36.

HU, X., O'HARA, L., WHITE, S., MAGNER, E., KANE, M. & WALL, J.G. 2007. Optimisation of production of a domoic acid-binding scFv antibody fragment in Escherichia coli using molecular chaperones and functional immobilisation on a mesoporous silicate support. *Protein Expression and Purification*, 52, 194-201.

KABIR, M.E., KRISHNASWAMY, S., MIYAMOTO, M., FURUICHI, Y. & KOMIYAMA, T. 2010. Purification and functional characterization of a Camelid-like single-domain antimycotic antibody by engineering in affinity tag. *Protein Expression and Purification*, 72, 59-65.

KIM, S.H., TITLOW, C.C. & MARGOLIES, M.N. 2000. An approach for preventing recombination-deletion of the 40-50 antidioxin antibody V(H) gene from the phage display vector pComb3. *Gene*, 241, 19-25.

KLOSE, J., WENDT, N., KUBALD, S., KRAUSE, E., FECHNER, K., BEYERMANN, M., BIENERT, M., RUDOLPH, R. & ROTHEMUND, S. 2004. Hexa-histidine tag position influences disulfide structure but not binding behaviour of in vitro folded N-terminal domain of rat corticotrophin-releasing factor receptor type 2a. *Protein Science*, 13, 2470-2475.

KRETZSCHMAR, T. & VON RUDEN, T. 2002. Antibody discovery: phage display. *Current Opinion in Biotechnology*, 13, 598–602.

LI, Y. & ZAMBLE, D.B. 2009. Nickel homeostasis and nickel regulation: an overview. *Chemical Reviews*, 109, 4617-4643.

MARCUS, W., WANG, H., LOHR, D., SIERKS, M. & LINDSAY, S. 2006. Isolation of an scFv targeting BRG1 using phage display with characterization by AFM. *Biochemical and Biophysical Research Communications*, 342, 1123–1129.

MEGEED, Z., WINTERS, R.M. & YARMUSH, M.L. 2006. Modulation of single-chain antibody affinity with temperature-responsive elastin-like polypeptide linkers. *Biomacromolecules*, 7, 999-1004.

MULROONEY, S.B. & HAUSINGER, R.P. 2003. Nickel uptake and utilization by microorganisms. *FEMS Microbiology Reviews*, 27, 239-261.

NISSIM, A, HOOGENBOOM, H.R., TOMLINSON, I.M., FLYNN, G., MIDGLEY, C., LANE, D. & WINTER, G. 1994. Antibody fragments from a "single pot" phage display library as immunochemical reagents. *The EMBO Journal*, 13, 692-8.

PANNEKOEK, H., MEIJER, M., GAARDSVOLL, H. & ZONNEVELD, A.J. 1995. Functional display of proteins, mutant proteins, fragments of proteins and peptides on the surface of filamentous (bacterio) phages: A review. *Cytotechnology*, 18, 107-112.

PORATH, J., CARLSSON, J., OLSSON, I. & BELFRAGE, G. 1975. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature*, 598–599.

ROBINSON, C.R. & SAUER, R.T. 1998. Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 5929-34.

SMITH, M.C., FURMAN, T.C., INGOLIA, T.D. & PIDGEON, C. 1988. Chelating peptide-immobilized metal ion affinity chromatography. A new concept in affinity chromatography for recombinant proteins. *The Journal of Biological Chemistry*, 263, 7211-7215.

SØRENSEN, H.P. & MORTENSEN, K.K. 2005. Advanced genetic strategies for recombinant protein expression in Escherichia coli. *Journal of Biotechnology*, 115, 113-128.

THIE, H., SCHIRRMANN, T., PASCHKE, M., DÜBEL, S. & HUST, M. 2008. SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in E. coli. *New Biotechnology*, 25, 49-54.

WIND, T., STAUSBØL-GRØN, B., KJÆR, S., KAHNS, L., JENSEN, K.H. & CLARK, B.F.C. 1997. Retrieval of phage displayed scFv fragments using direct bacterial elution. *Journal of Immunological Methods,* 209, 75-83.

WU, S., KE, A. & DOUDNA, J.A, 2007. A fast and efficient procedure to produce scFvs specific for large macromolecular complexes. *Journal of Immunological Methods*, 318, 95-101.

ZHAO, Q., CHAN, Y.-W., LEE, S.S.T. & CHEUNG, W.T. 2009. One-step expression and purification of single-chain variable antibody fragment using an improved hexahistidine tag phagemid vector. *Protein Expression and Purification*, 68, 190-195.

CHAPTER 5

Cytoplasmic expression of TFI-scFv

5.1 Introduction

As concluded in the previous chapter, it is necessary to redirect the accumulation of expressed TFI-scFv from the culture media to the bacterial cytoplasm in order to facilitate purification by means of nickel affinity chromatography. It is therefore important to discuss the mechanisms utilised to achieve periplasmic expression.

The ribosomal translation of mRNA occurs within the cytoplasm of the *E. coli* cell. As a result, those proteins required in the extracellular environment or in other cell compartments need to be transported across the cytoplasmic membrane. Secretory proteins in general contain an N-terminal signalling peptide that is recognised by a transmembrane translocase (Manting & Driessen, 2000). Consequently, the transport of recombinant proteins expressed in the cytoplasm through the cytoplasmic membrane to the periplasm of *E. coli* is achieved through the incorporation of these signalling peptides upstream of the recombinant gene (Gu *et al.,* 2002; Sørensen & Mortensen, 2005; Tapryal *et al.,* 2010). Many signal peptides have been successfully utilised for the translocation of the recombinant protein to the periplasm in *E. coli.* These predominantly consist of prokaryotic signal sequences, such as the PhoA-, OmpA- OmpT-, LamB- and Omp signals from *E. coli* itself, protein A from *Staphylococcus aureus* and pelB from *Erwinia carotovora*, although eukaryotic murine RNase was well as human growth hormone signal have also been successfully utilised (Makrides, 1996; Sørensen and Mortensen, 2005). The pIT2- and pET expression vectors both incorporate the 22 amino acid pelB leader peptide (Lei *et al.*, 1987). The pelB leader peptide directs the expressed recombinant protein to a specific translocase complex imbedded in the inner membrane of *E. coli* (See: Figure 5.1) (Manting and Driessen, 2000, Sletta *et al.*, 2007). Once the expressed protein reaches the periplasm, the pelB leader peptide is cleaved by signal peptidases yielding the mature protein with the original amino-terminus (Pannekoek *et al.*, 1995).

Periplasmic expression is often preferred to cytoplasmic expression as the periplasm contains only 4 %, approximately 100 proteins, of the total protein contents. The low level of background cell protein greatly simplifies the purification process (Makrides, 1996; Sørensen and Mortensen, 2005). It is also an oxidizing environment and contains foldase enzymes such as thiol-disulfide isomerase that facilitate formation of disulfide bonds and the folding of the expressed proteins. Although it has been shown that disulfide bridges can form in the reducing environment of the cytoplasm, the periplasmic expression is generally utilised for recombinant proteins rich in cysteine residues (Rudolph & Lilie, 1996; Lilie *et al.,* 1998; Baneyx, 1999).
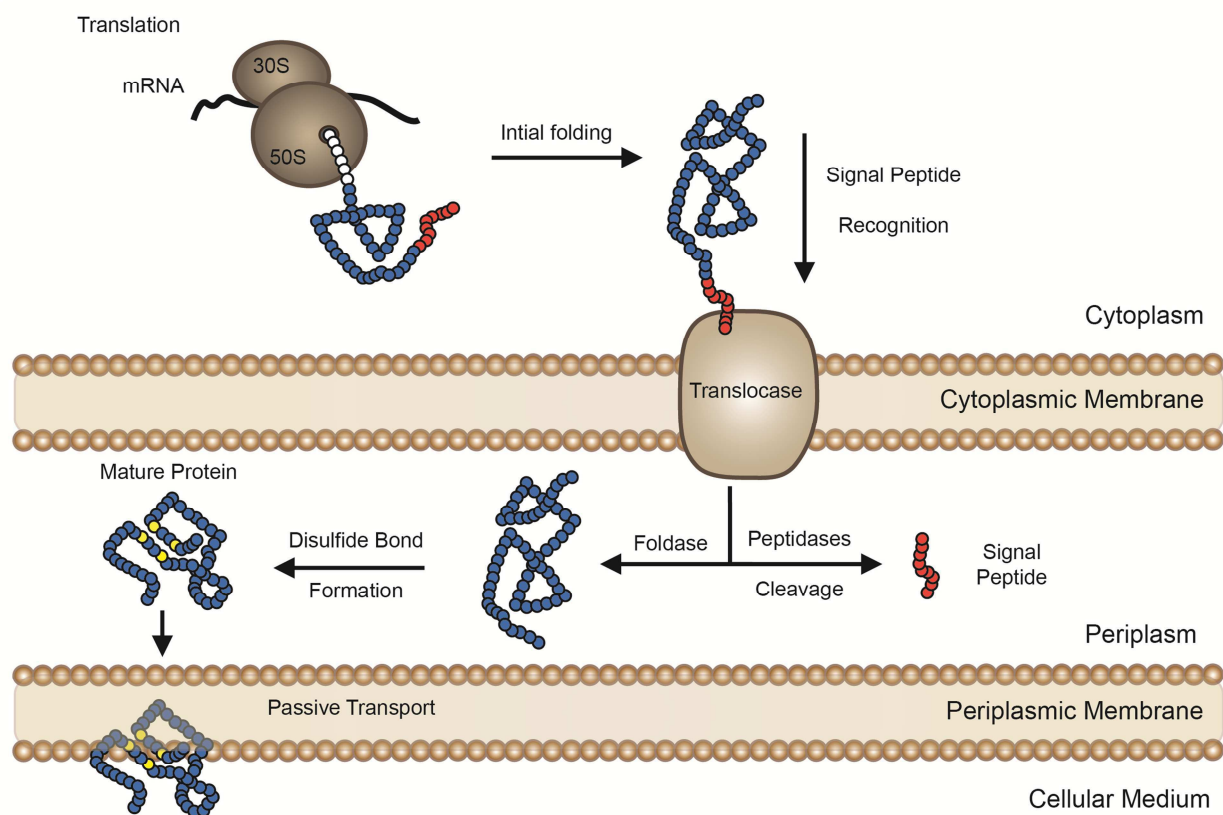
Figure 5.1: Protein folding and secretion in *E. coli*

Translation of mRNA by ribosomes and initial folding of the recombinant protein occurs in the cytoplasm. The N-terminal signal peptide is recognised by the translocase and the protein is transported across the cytoplasmic membrane. The signal peptide is removed by peptidases once the recombinant protein reaches the periplasm. Protein folding facilitated by foldase enzymes and disulfide bonds (◯) formation occurs within the oxidizing environment of the periplasm. The mature protein is passively transported the cellular medium.

Adapted from Pannekoek *et al.* (1995); Sørensen and Mortensen (2005) and Thie *et al.* (2008).

A major disadvantage of periplasmic expression is the loss of recombinant protein due to diffusion or leakage of the target protein into the extra cellular medium (Sørensen and Mortensen, 2005). Conversely, the purification of the target protein from the cellular medium is complicated by the lack of an active transport system for the transport of recombinant proteins from the periplasm to the cellular medium. As a result *E. coli* cells undergoing periplasmic expression require an additional osmotic shock step in order to stimulate the passive transport of the target protein to the supernatant (Velappan *et al.*, 2007; Li and Lu, 2009). Apart from being time consuming, an inherent amount of target protein will always be lost to the periplasm. As a result, the cytoplasmic expression of recombinant proteins is preferred as it provides higher protein yields and requires less complicated plasmid design (Hannig and Makrides, 1998; Sørensen and Mortensen, 2005).

5.2 Aims

The main aim of this chapter was to remove the pelB leader peptide from the tissue factor inhibitor - single chain variable fragment (TFI-scFv) in order to redirect the expression to the cytoplasm while atempting increase protein yield through the utilisation of the over-expression vectors and cell lines. In addition effect of alternative positions of the c-terminal His-tag was on nickel affinity chromatography was investigated.

5.3 Materials and Methods

5.3.1 Isolation of TFI-scFv synthetic V-gene

Helper phage was produced to facilitate the production of TFI-phage. The synthetic V-gene coding for TFI-scFv was isolated from TFI-phage and amplified with custom primer sets in order to incorporate *Nde*I and *EcoR*I restriction sites at the 5' and 3' ends respectively.

5.3.1.1 Production of KM13 helper phage

The helper phage was produced according to the protocol G as specified in the product documentation (Tomlinson I + J: Human Single Fold Libraries I + J) as previously described in Section 3.3.1.1.

5.3.1.2 Production of TFI-phagemid

The tissue factor inhibiting phage was produced according to the protocol A as stipulated in the product documentation (TomlinsonI + J: Human Single Fold Libraries I + J) as previously described in Section 3.3.1.2.

5.3.1.3 TFI-phagemid DNA isolation

DNA was isolated from the TFI-phagemid as previously described in Section 4.3.4.1

5.3.1.4 TFI Phagemid DNA polymerase chain reaction amplification

A two primer set were designed based on the sequencing results of pIT2 TFI-scFv construct obtained in Chapter 4 and PCR simulated using Vector NTI v10.0. Polymerase Chain Reactions (PCR) were performed on isolated TFI-phagemid DNA. Two primers set (Integrated DNA technologies, Belguim) were utilised in order to amplify the synthetic V-gene construct from the pIT2 vector. The mutual forward primer, reverse primer and targeted DNA fragment size are indicated in Table 5.1.

Table 5.1: Custom primer sets

Primer Set	Forward Primer	Reverse Primer	Simulated Amplicon Size
1	scFv F_NdeI 5' CAT ATG GCC GAG GTG CAG CTG 3'	scFv R1_EcoRI 5' GAA TTC TAG TGA TGG TGA TGA 3'	764 bp
2	scFv F_NdeI 5' CAT ATG GCC GAG GTG CAG CTG 3'	scFv R2_EcoRI 5' GAA TTC CGT TTG ATT TCC ACC T3'	734 bp

All PCR amplifications were performed in duplicate as previously described in Section 4.3.4.2

5.3.1.5 Agarose gel electrophoresis conformation of PCR product

The presence of the 764 and 734 base pair (bp) PCR products were confirmed for documentation purposes by means of a 0.8% agarose gel as previously described in Section 4.3.4.3.

5.3.1.6 Agarose gel electrophoresis isolation of PCR product

PCR products were confirmed by means of a 0.8 % agarose gel and isolated as previously described in Section 4.3.4.4.

5.3.2 Cloning PCR products into pSMART propagation vector

Isolated PCR products were cloned into the pSMART cloning vector (Lucigen, USA) by means of blunt end ligation. Both of the pSMART constructs were transformed into *E. coli* TOP10 (Invitrogen, USA) by means of heat shock for the *in vivo* amplification of constructs. After propagation of the pSMART plasmid, the plasmid was isolated and subjected to double digestion with *Nde*I and *EcoR*I restriction enzymes in order to verify the presence of the 764 and 734 bp inserts.

5.3.2.1 Phosphorylation of PCR products

The 5'-phosphorylation of PCR products were performed using the T4 Polynucleotide Kinase Kit (Fermentas, Canada) according to product protocol. Phosporylation reactions for both the 764 bp and 734 bp fragments were assembled as indicated in Table 5.2.

Table 5.2: 5'-Phosphorylation of PCR products

Reagent	Concentration	Volume per reaction
Target DNA	Approximately 1500 ng	Dry Sample
Polynucleotide Kinase	10 u/µl (Weis Units/µL)	1 µL
Reaction Buffer A	10 X	2 µL
ATP	10 mM	1 µL
MilliQ sterile water	-	15 µL
Total		20 µL

In short, both phosphorylation reactions were assembled, gently mixed and briefly centrifuged for 30 seconds at room temperature. Reactions were then incubated for 20 minutes at 37 ℃. Phosphorylation reactions were finalized by incubation at 75 ℃ for 10 minutes.

5.3.2.2 Ligation of PCR products into pSMART Vector

The blunt-end cloning of PCR products were performed using the Rapid DNA Ligation Kit (Fermentas, Canada) according to product protocol. Blunt-end ligation reactions for both the 764 bp and 734 bp fragments were assembled as indicated in Table 5.3.

Table 5.3: Blunt-end ligation of constructs into pSMART

Reagent	Concentration	Volume per reaction
pSMART Vector	50 ng.µL^{-1}	1 µL
PEG 4000	50 % (v/v)	1 µL
Rapid Ligation Buffer	10 X	1 µL
T4 DNA Ligase,	5u/µl (Weis Units/µL)	1 µL
MilliQ sterile water	-	3 µL
Phosphorylated PCR product	25 ng.µL^{-1}	3 µL
Total		10 µL

All ligation reactions were incubated at room temperature for one hour. Both pSMART plasmids containing the 763 bp- (Construct 1) or 734 bp (Construct 2) insert were collected and stored at – 20 ºC until later use.

5.3.2.3 Heat shock transformation and propagation of pSMART constructs into *E. coli* TOP10

In brief, 50 µL aliquots of *E. coli* TOP10 competent cell stock (-80 ℃) were thawed for 5 minutes on ice. The complete ligation reaction for pSMART construct 1 and 2 were separately transferred to the thawed *E. coli* TOP10 aliquots and incubated for 30 minutes on ice. The cultures were then heat shocked for 45 seconds at 42 ℃ in order to facilitate the uptake of the pSMART constructs by the cells. Cultures were immediately cooled on ice for 2 minutes and supplemented with 250 µL SOC media (See: Appendix B).   Transformation reactions were incubated at 37 ℃ for 1 hour, plated out onto LB media plates (15 $g.L^{-1}$ agar, 10 $g.L^{-1}$ tryptone, 10 $g.L^{-1}$ NaCl, 5 $g.L^{-1}$ yeast extract and 30 $mg.L^{-1}$ kanamycin) and incubated overnight at 37 ℃. The following day 6 *E. coli* TOP10 single colonies for each of the constructs were selected at random and individually transferred to sterile test tubes containing 5 ml LB media (10 $g.L^{-1}$ tryptone, 10 $g.L^{-1}$ NaCl, 5 $g.L^{-1}$ yeast extract and 30 $mg.L^{-1}$ kanamycin). All cultures were aerobically incubated in an orbital shake-incubator (160 rpm) overnight at 37 ℃ and stored at 4 ℃ unt il later use.

5.3.2.4 pSMART plasmid isolation

The pSMART plasmids were isolated from transformed *E. coli* TOP10 cells using the Biospin Plasmid DNA Extraction Kit (Bioflux, Japan) according to manufacturer's specifications. All centrifugations were performed at room temperature using an eppendorf centrifuge 5424 (Eppendorf, Germany).

Briefly, 2 mL *E. coli* TOP10 cells were harvested from the 5 mL overnight culture in 1.5 mL microcentrifuge tubes by means of repeated centrifugation at 10000 x *g* for 30 seconds. Harvested cells were resuspended in 250 µL resuspension buffer, containing RNase, until no cell aggregates were visible. A volume of 250 µL lysis buffer was added to the resuspended cells and the content gently mixed by inversion of the microcentrifuge tubes. A volume of 350 µL neutralization buffer was then added to the lysis solution and the content gently mixed by inversion of the microcentrifuge tubes. Cellular debris was removed by means of centrifugation at 14000 x *g* for 10 minutes and the supernatant applied to Bioflux spin columns. The

columns were centrifuged at 6000 x g for 60 seconds and the flow through discarded. The columns were then washed twice with 650 µL wash buffer by centrifugation at 12000 x *g* for 60 seconds and the flow through discarded. The remaining residual wash buffer was removed with an additional centrifugation at 12000 x *g* for 60 seconds. A volume of 50 µL sterile water was applied to the Bioflux spin columns that were then incubated for 60 seconds at 37 ℃. The isolated pSMART plasmids where collected in sterile microcentrifuge tubes by means of centrifugation at 12000 x *g* for 60 seconds and stored at -20 ℃ until later us e.

5.3.2.5 Double restriction digestion of pSMART plasmid

The isolated pSMART constructs were subjected to a double restriction digest using *Nde*I and *EcoR*I restriction enzymes (Fermentas, Canada) and restriction buffer O (Fermentas, Canada) in order to confirm the presence of the desired inserts as indicated in Table 5.4.

Table 5.4: Double restriction digestion reaction (Miniprep)

Reagent	Concentration	Volume per reaction
pSMART plasmid DNA (Construct 1 and 2)	Approximately 130 ng.mL^{-1}	2 µL
*EcoR*I	5 000 U.mL^{-1}	1 µL
*Nde*I	5 000 U.mL^{-1}	1 µL
Restriction Buffer O	10 x	1 µL
MilliQ sterile water	-	5 µL
Total		10 µL

Digestion reactions were incubated for 1 hour at 37 ℃ and stored at 4 ℃ until later use.

5.3.2.6 Agarose gel electrophoresis of digested pSMART plasmid

The presence of the pSMART constructs containing the 764 or 734 bp TFI-gene inserts were confirmed for documentation purposes by 0.8 % agarose gel of the double digested products as previously described in Section 4.3.4.3. The undigested pSMART plasmids containing the desired 764 or 734 bp inserts were identified and stored at -20 ℃ until later use.

5.3.4 Cloning of constructs into pET22b (+) expression vector

Both constructs 1 and 2 were isolated from pSMART cloning vector (Lucigen, USA) and cloned into pET22b (+) (Novagen) expression vector by means of sticky end ligation. These new constructs were transformed to *E. coli* TOP10 (Invitrogen, USA) and propagated *in vivo*. After propagation of the pET22 constructs the plasmids were isolated and subjected to double digestion with NdeI and EcoRI restriction enzymes in order to verify the presence of the 764 (Construct 1) and 734 bp (Construct 2) inserts.

5.3.4.1 Heat shock transformation of pET22 plasmid into *E. coli* TOP10

A volume of 2 µL pET22 plasmid (approximately 10 ng) was transferred to *E. coli* TOP10 as previously described in Section 5.3.2.3 with the replacement of 30 mg.L^{-1} kanamycin with 100 mg.L^{-1} ampicillin.

5.3.4.2 pET22 plasmid isolation

The pET22 plasmid was isolated from transformed cells using the Biospin Plasmid DNA Extraction Kit (Bioflux, Japan) according to manufacturer's specifications as previously described in Section 5.3.2.4. The isolated pET22 plasmids were collected in sterile 1.5 mL microcentrifuge tubes by means of centrifugation at 12000 x *g* for 60 seconds, dried in Vacufuge Concentrator 5301 (Eppendorf, Germany) and stored at – 22 ºC until later use.

5.3.4.3 Double restriction digestion of pSMART and pET22 plasmid

The isolated pSMART- (See: Section 5.3.3.2) and dried pET22 plasmid were subjected to a double restriction digest using *Nde*I and *EcoR*I restriction enzymes (Fermentas, Canada) and restriction buffer O (Fermentas, Canada) as indicated in Table 5.5 and 5.6 respectively.

Table 5.5: Double restriction digestion reaction of pSMART plasmid

Reagent	Concentration	Volume per reaction
pSMART plasmid DNA (Construct 1 and 2)	Approximately 130 ng.mL^{-1}	15 µL
*EcoR*I	5 000 U/ml)	5 µL
*Nde*I	5 000 U/ml)	5 µL
Restriction Buffer O	10 x	5 µL
MilliQ sterile water	-	20 µL
Total		50 µL

Table 5.6: Double restriction digestion reaction of pET22 plasmid

Reagent	Concentration	Volume per reaction
pET22 plasmid DNA	Approximately 1850 ng	Dry DNA sample
*EcoR*I	5 000 U/ml)	5 µL
*Nde*I	5 000 U/ml)	5 µL
Restriction Buffer O	10 x	4 µL
MilliQ sterile water	-	26 µL
Total		40 µL

Digestion reactions were incubated for 1 hour at 37 ℃ and stored at 4 ℃ until later use.

5.3.4.4 Agarose gel electrophoresis isolation constructs and pET22

The 764 and 734 bp inserts from pSMART and linearised pET22 plasmid were confirmed by means of a 0.8 % agarose gel and isolated as previously described in Section 4.3.4.4.

5.3.4.5 Ligation of constructs into pET22b (+) Vector

The sticky-end cloning of construct 1 and 2 into pET22 plasmid were performed using the Rapid DNA Ligation Kit (Fermentas, Canada) according to product protocol. Ligation reactions for both the 764 (Construct 1) and 734 bp (Construct 2) fragments were assembled as indicated in Table 5.7.

Table 5.7: Sticky-end ligation of constructs into pET22

Reagent	Concentration	Volume per reaction
MilliQ sterile water	-	-
Rapid Ligation Buffer	10 X	1 µL
T4 DNA Ligase,	5u/µl (Weis Units/µL)	1 µL
pET22bV (+) Vector	20 ng.µL^{-1}	3 µL
Construct 1 or 2	10 ng.µL^{-1}	5 µL
Total		10 µL

All ligation reactions were incubated at room temperature for 1 hour. The ligation reaction mixtures were then immediately collected for transformation into *E. coli* TOP10.

5.3.4.6 Heat shock transformation of pET22 constructs into *E. coli* TOP10

The complete ligation reaction mixtures were transferred to the thawed *E. coli* TOP10 as previously described in Section 5.3.4.1.

5.3.4.7 Isolation pET22 constructs

The pET22 constructs were isolated from transformed cells using the Biospin Plasmid DNA Extraction Kit (Bioflux, Japan) as previously described in Section 5.3.2.4. The isolated pET22 plasmids where collected in sterile 1.5 mL microcentrifuge tubes by means of centrifugation at 12000 x *g* for 60 seconds, dried in Vacufuge Concentrator 5301 (Eppendorf, Germany) and stored at – 22 ºC until later use.

5.3.4.8 Double restriction digestion of pET22 plasmid

The isolated pET22 plasmid was subjected to a double restriction digest using *Nde*I and *EcoR*I restriction enzymes (Fermentas, Canada) and restriction buffer O (Fermentas, Canada) as indicated in Table 5.8 in order to confirm the presence of the desired inserts.

Table 5.8: Double restriction digestion reaction (Miniprep)

Reagent	Concentration	Volume per reaction
pET22 plasmid DNA (Construct 1 and 2)	Approximately 130 ng.mL^{-1}	2 µL
*EcoR*I	5 000 U/ml	1 µL
*Nde*I	5 000 U/ml	1 µL
Restriction Buffer O	10 x	1 µL
MilliQ sterile water	-	5 µL
Total		10 µL

Digestion reactions were incubated for 1 hour at 37 ℃ and stored at 4 ℃ until later use.

5.3.4.9 Agarose gel electrophoresis of digested pET22 plasmid

The presence of the 764 and 734 bp inserts within the digested pET22 plasmid samples were confirmed for documentation purpose by means of a 0.8 % agarose gel as previously described in Section 4.3.4.3. The undigested pET22 constructs containing the desired 764 or 734 bp inserts were identified and stored at -20 ℃ until later use.

5.3.5 Cytoplasmic expression of pET22 constructs

Isolated pET22 plasmids containing either construct 1 or 2 were transformed into *E. coli* BL21 (DE3) (Lucigen, USA). Expression of the constructs was induced by means of auto-induction media.

5.3.5.1 Heat shock transformation of pET22 constructs into *E. coli* BL21

The pET22 construct isolates (See: Section 5.3.4.7) as well as empty pET22 plasmid (negative control) was transferred to *E. coli* BL21 as previously described in Section 5.3.4.1.

The successfully transformed *E. coli* BL21 colonies were collected by repeatedly flushing the agar plates with 2 to 3 mL LB-media containing 100 mg.L^{-1} ampicillin in order to dislodge colonies. The dislodged cells were used to inoculate a 50 mL LB media containing 100 mg.L^{-1} ampicillin. The precultures for the negative control as well as the pET22 constructs 1 and 2 were then aerobically incubated in an orbital shake-incubator (160 rpm) for 6 hours at 37 ℃. A volume of 1 mL preculture was used to inoculate 100 mL (10 x 100 mL = 1 L in total) auto-induction media (See: Appendix B) containing 100 mg.L^{-1} ampicillin. Cultures were then aerobically incubated in an orbital shake-incubator (160 rpm) for 24 hours at 37 ℃. The *E. coli* BL21 cells were harvested by means of centrifugation at 5000 x *g* for 10 minutes at 4 ℃ and the supernatant discarded. The harvested cells were weighed and stored at 4 ℃ until later use.

5.3.6 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SDS-PAGE was performed on *E. coli* Bl21 whole cell samples normalised to an OD$_{600}$ value of 1. SDS-PAGE was performed as previously described in Section 3.3.3.

5.3.7 Immobilized metal ion affinity chromatography (IMAC) purification of *E. coli* BL21 expressed TFI-scFv

5.3.7.1 Isolation of cytoplasmic fraction

The harvested *E. coli* BL21 cultures (See: Section 5.3.4) were resuspended in 200 mL cold (4 ℃) MOPS buffer (4.186 g.L^{-1} at pH 7.4), washed by means of centrifugation at 5000 x *g* for 10 minutes at 4 ℃ and the supernatant discarded. Cell pellets were then resuspended in 50 mL cold (4 ℃) disruption buffer (4.186 g.L^{-1} MOPS, 29.22 g.L^{-1} NaCl, 1.36 g.L^{-1} imidazole at pH 7.4) containing completely EDTA-free protease inhibitor (Roche, Switzerland) and placed on ice.

Collected cells were broken using a One Shot Constant Cell Disruption System (Constant Systems, UK) at 2000 bar. Disruption chamber was placed on ice for 1 minute between disruptions in order to prevent heat degradation of the expressed TFI-scFv. Cell disruption was performed twice per sample. The disrupted samples were collected and immediately placed on ice. The cellular debris was removed by centrifugation at 8000 x g for 10 minutes at 4 ℃. The remaining supernatant s were collected and subjected to ultra centrifugation at 100 000 x g for 1 hour and 30 minutes at 4 ℃ in order to remove insoluble protei n and membrane fraction. The remaining supernatants consisting of the soluble cytoplasmic fraction were collected and stored at 4 ℃ until late use.

5.3.7.2 Immobilized metal ion affinity chromatography (IMAC)

The nickel affinity chromatography purification of TFI-scFv from the isolated cytoplasmic fractions was performed as previously described in Section 4.3.2.

5.3.8 DNA Sequencing of pET22 constructs

Bidirectional DNA sequencing of isolated pET22 constructs were performed with the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA) with the T7 Promoter Primer (Integrated DNA technologies, Belgium) (forward primer): 5' TAA TAC GAC TCA CTA TAG GG 3' and T7 Terminator Primer (reverse primer): 5' GCT AGT TAT TGC TCA GCG G 3' (Integrated DNA technologies, Belgium) as previously described in Section 4.3.4.5 Sequencing data was imported to Geneious v 4.8.5 and a concensus sequence was compiled. The finalised consensus sequences were imported to Vector NTI v 10.0 and the genetic motives identified. The finalised concensus DNA sequences for construct 1 and 2 were translated using internet bases ExPASy Translate Tool (http://web.expasy.org/translate), aligned with the original TFI-scFv amino acid sequence using internet based ClustalW (http://www.genome.jp/tools/clustalw/) software and analysed for the presence of rare codons for expression in *E. coli* using internet based Rare Codon Calculator (http://people.mbi.ucla.edu/sumchan/ caltor.html).

5.4 Results

5.4.1 PCR amplification of synthetic TFI-gene

The PCR products were subjected to 0.8% agarose gel electrophoresis and visualised with ethidium bromide in order to confirm the amplification of the TFI-scFv gene fragments. PCR products for primer set 1 were loaded in lanes 1 and 2. PCR products primer set 2 was loaded in lanes 3 and 4 as indicated in Figure 5.2. The agarose gel electrophoresis confirmed the successful amplification of TFI-scFv synthetic gene using both custom primer sets. The primer set was highly specific as no additional fragments were amplified.
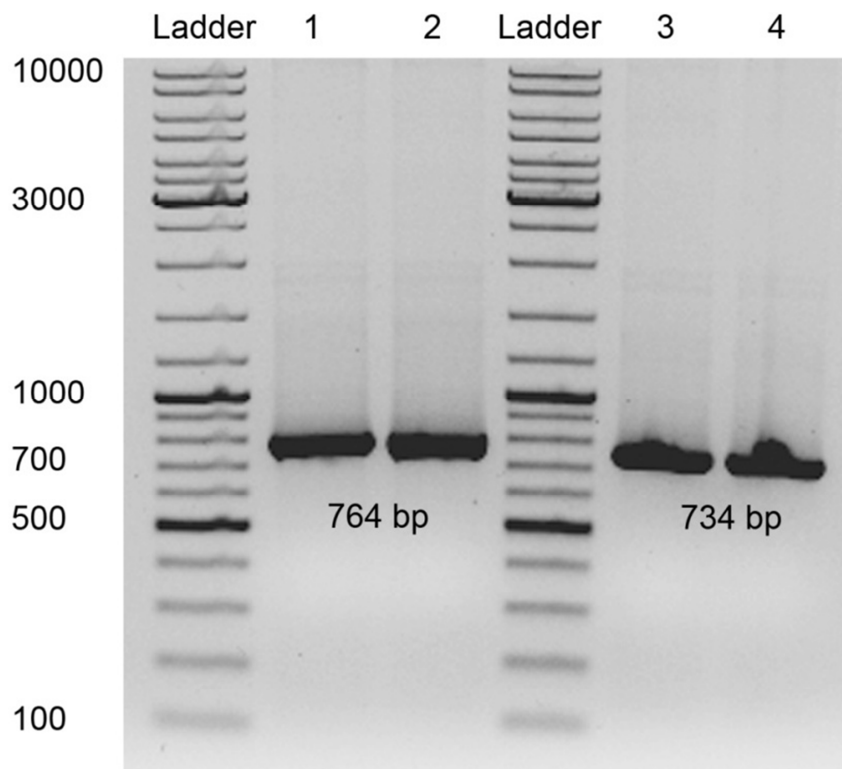
Figure 5.2: PCR amplification of synthetic V-gene from pIT 2 phagemid

The 0.8% agarose gel electrophoresis of both PCR products. Lane 1 and 2 contains the larger TFI-scFv DNA fragment amplified with primer set 1. Lanes 3 and 4 contains smaller TFI-scFv DNA fragment amplified with primer set 2.

5.4.2 Double digestion of pSMART constructs

The pSMART constructs were subjected to a double restriction digest using *Nde*I and *EcoR*I restriction enzymes in order to identify the positive transformants. Six pSMART digestion products were individually loaded in lanes 1 to 6 and analysed for the presence of 764 bp (Construct 1) and 734 bp (Construct 2) inserts respectively as indicated in Figure 5.3.
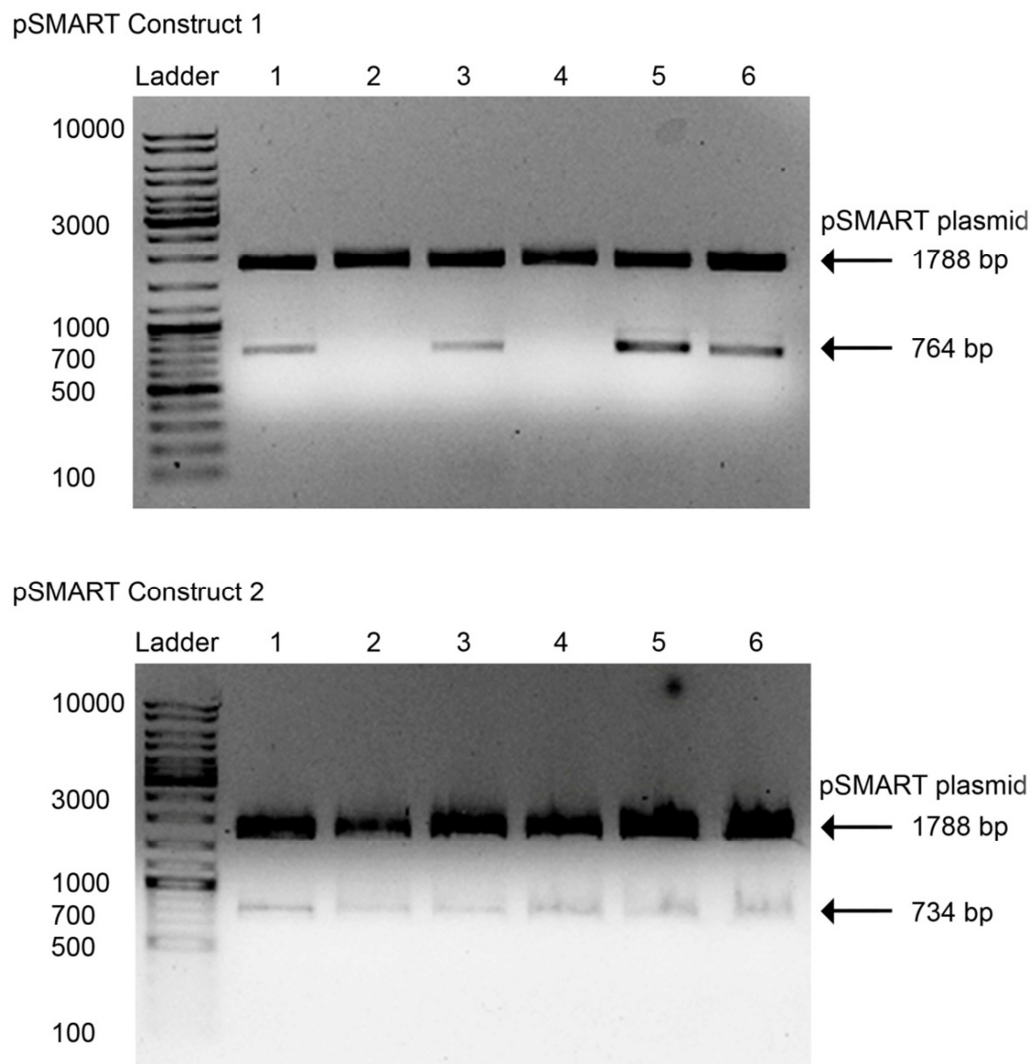
Figure 5.3: Double digestion of pSMART constructs

The 0.8% agarose gel electrophoresis of pSMART digestion products. Position of desired 764 bp and 734 bp TFI-scFv DNA fragments for constructs 1 and 2 respectively are indicated by arrows.

The agarose gel electrophoresis confirmed the presence pSMART plasmid in all of colonies selected for pSMART construct 1. The TFI-gene insert is absent in lanes 2

and 4. Lanes 1, 3, 5 and 6 contains the 763 bp TFI-gene insert. All of the colonies selected for pSMART construct 2 were positive for the presence of both the pSMART plasmid as well as the 734 bp TFI gene insert.

5.4.3 Double digestion of pET22 constructs

The pET22 constructs were subjected to a double restriction digest using *Nde*I and *EcoR*I restriction enzymes in order to confirm the presence TFI-scFv gene.
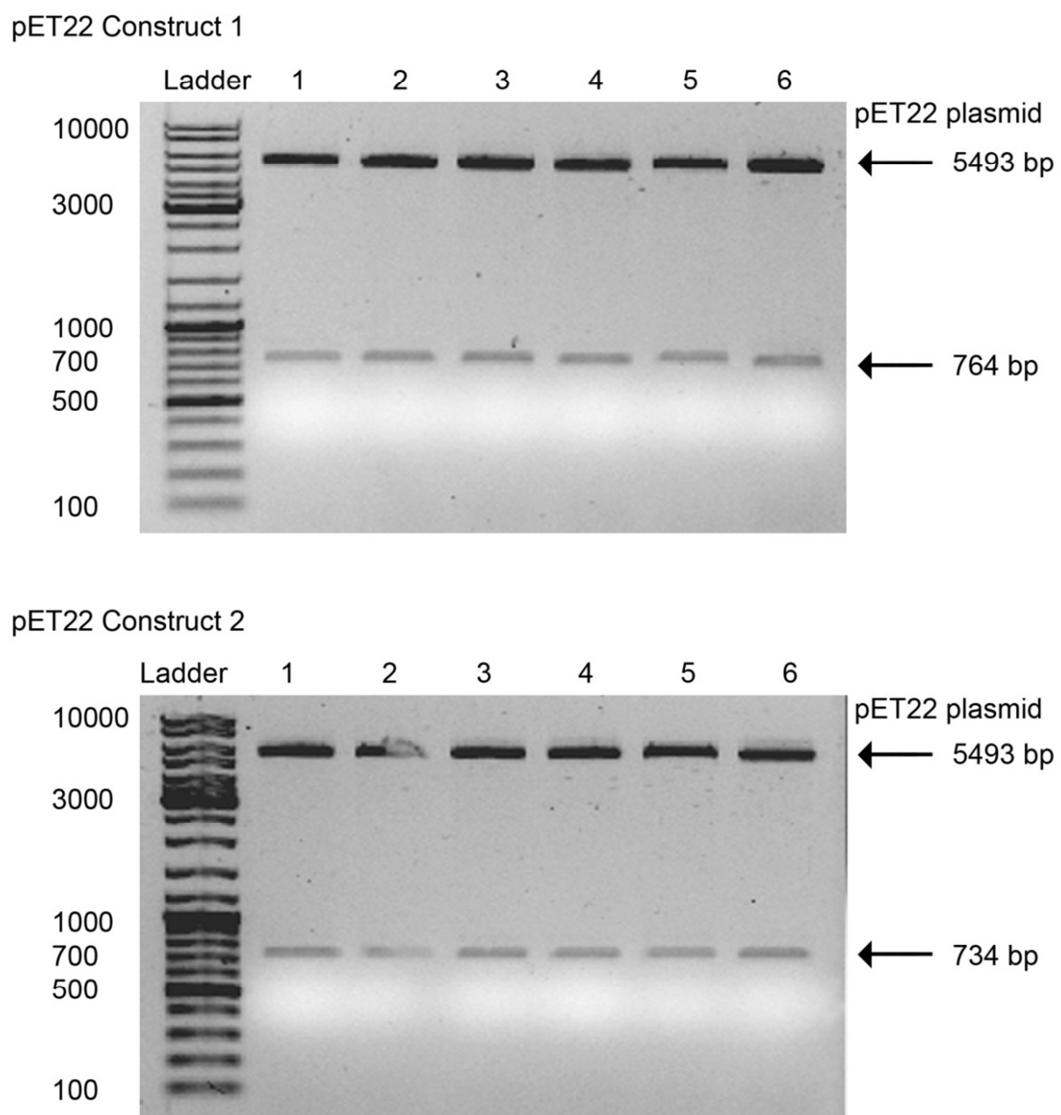
Figure 5.4: Double digestion of pET22 constructs

The 0.8% agarose gel electrophoresis of pET22 digestion products. Position of desired 764 bp and 734 bp TFI-scFv DNA fragments for constructs 1 and 2 respectively are indicated by arrows.

Six pET22 digestion products were individually loaded in lanes 1 to 6 and analysed for the presence of 764 bp (Construct 1) and 734 bp (Construct 2) inserts respectively as indicated in Figure 5.4.

The agarose gel electrophoresis confirmed the presence of pET22 plasmid in all of colonies selected for screening of pET22 construct 1. The 763 bp TFI-gene insert is present in lanes 1 to 6. All of the colonies selected for pET22 construct 2 were positive for the presence of both the pET22 plasmid as well as the 734 bp TFI gene insert.

5.4.4 SDS-PAGE analysis of TFI-scFv expression

Total *E. coli* Bl21 cells fractions were collected and subjected to 12.5 % SDS-PAGE in order to confirm the cytoplasmic expression of 26 KDa TFI-scFv fragments.
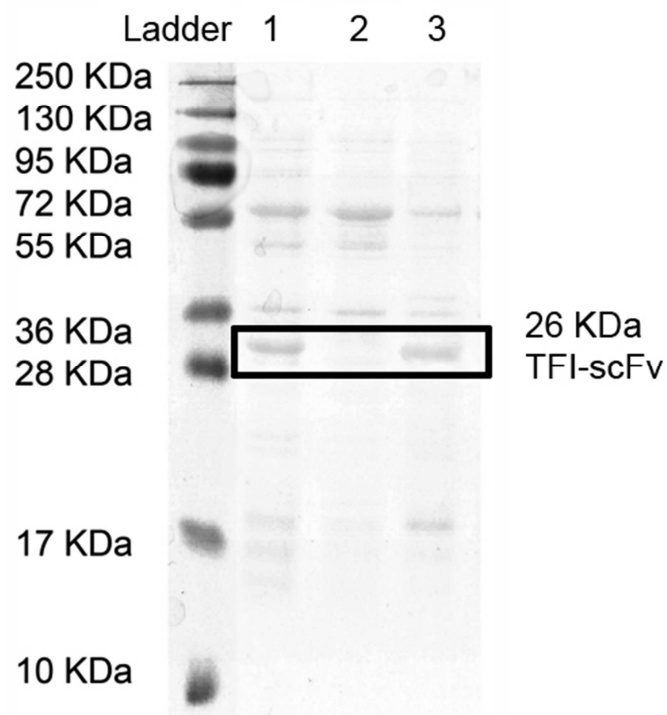
Figure 5.5: Whole cell SDS-PAGE analysis

SDS-PAGE analysis of *E. coli* BL21 culture harbouring pET22 constructs. The presence of 26 KDa TFI-scFv is indicated by frame in lane 1 and 3. The 26 KDa TFI-scFv is absent in the negative control in lane 2.

The *E. coli* Bl21 cells containing pET22 construct 1 were loaded into lane 1. Lane 2 was loaded with *E. coli* Bl21 cells harbouring empty pET22 (negative control). *E. coli* Bl21 cells containing pET22 construct 2 was loaded into lane 1 and 20 µL culture supernatant was loaded in lane 3, as indicated in Figure 5.5. The presence of 26 KDa fragment in *E. coli* BL21 harbouring pET22 construct 1 and 2 (Lane 1 and 3 respectively) combined with the absence of this fragment in the negative control (Lane 2) confirms the successful expression of both constructs.

5.4.5 Immobilised Metal Affinity chromatography

The nickel affinity chromatography purification profiles of TFI-scFv expressed in *E. coli* BL21 cytoplasm is indicated in Figure 5.6. The column was washed with binding buffer in order to remove unbound protein between points A and B. TFI-scFv constructs were eluted using imidazole gradient (20mM to 500 mM) introduced between point B and C.

Construct 1
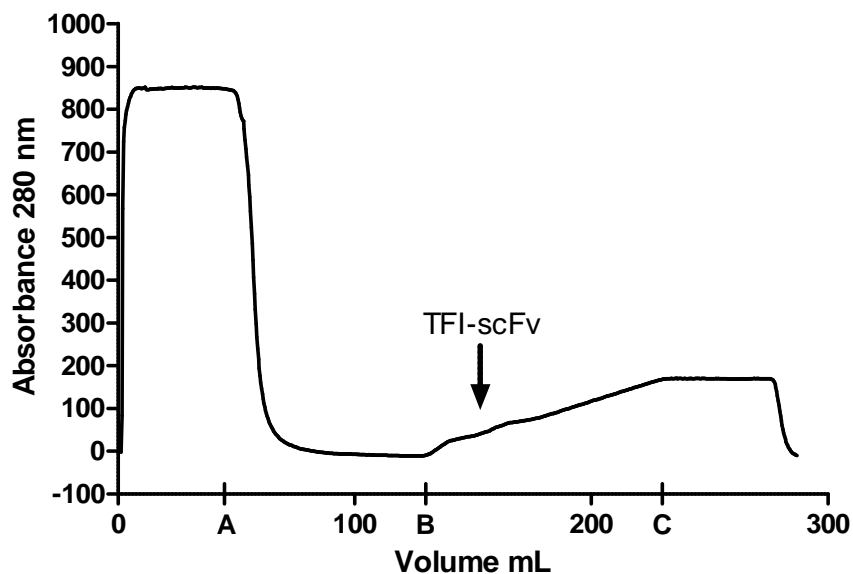
Figure 5.6 continues on next page.

Construct 2
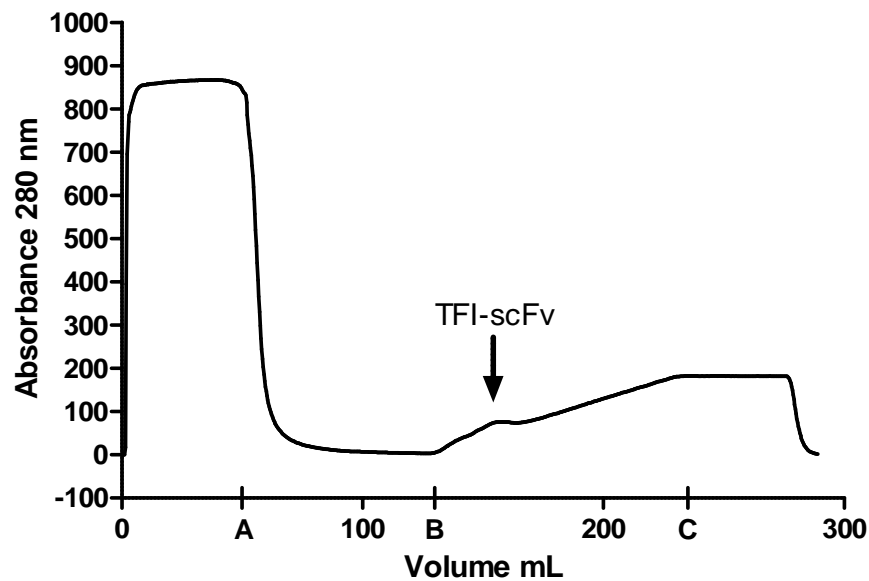
The purification profile for construct 1 indicates that extremely little TFI-scFv was purified by nickel affinity chromatography. Although the TFI-scFv protein peak is more pronounced for construct 2 than construct 1, the overall level of purified protein is still low. The slight disturbance in the early linearity of the immidazole gradients (TFI-scFv protein peaks) indicates that target protein is eluted at a relatively low concentration. Although the TFI-scFv peak is extremely small for construct 1 and only slightly improved for construct 2, it indicates the successful purification of both TFI-scFv constructs from *E. coli* BL21 cytoplasm by means of nickel affinity chromatography.

5.4.6 TFI-scFv sequencing analysis

The complete DNA sequences for both pET22 constructs were analysed in Vector NTI v10.0 and genetic motives identified as illustrated in Figure 5.7.
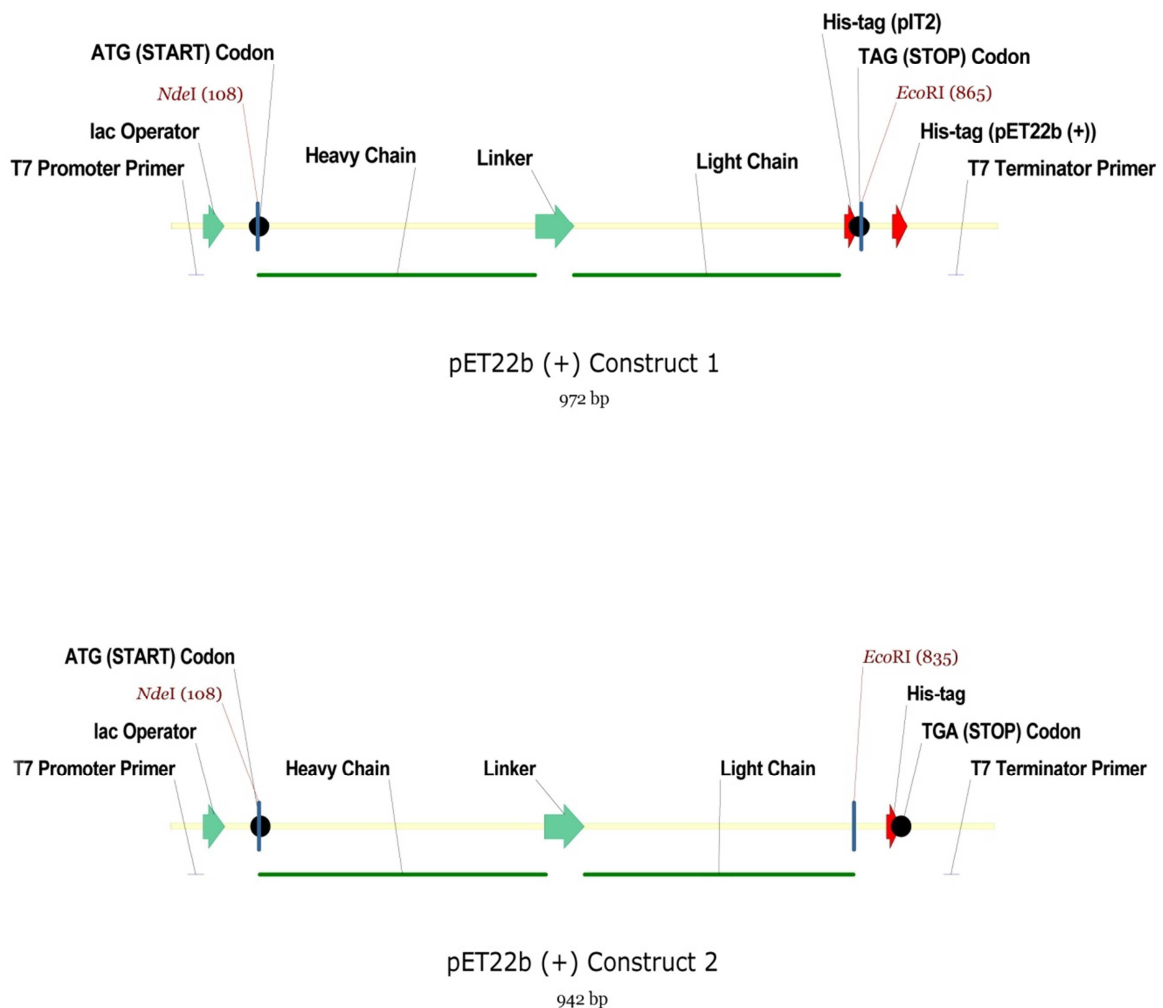
Figure 5.7: pET22 constructs DNA sequence

Finalised pET22 constructs sequence consist of 972 bp and 942 bp for construct 1 and 2 respectively. The T7 promoter and terminator primer binding sites are indicated as (━━). Human immunoglobulin G heavy- and light chain inserts are indicated as (━━). The lac operator and glycine-serine linker repeat sequence (Linker) are indicated by (➡). The polyhistidine sequences (His-Tag) are indicated as (➡). Position of start and stop codons with a (●). Restriction sites (*Nde*I and *EcoR*I) used during cloning are indicated as (|).

The DNA sequences were subjected to a six frame translation using internet based ExPASy Translate Tool (http://web.expasy.org/translate/) for comparison purposes. An open reading frame was identified, translating to a 251 and 261 amino acid sequence for construct 1 and 2 respectively. A multiple alignment between the pET22 constructs and original pIT2 construct was performed using internet bases ClustalW alignment software (http://www.genome.jp/tools/clustalw/) as presented in

Table 5.9. The pelB leader sequence (MKYLLPTAAAGLLLLAAQPA), glycine-serine linker sequence (GGGGSGGGGSGGGGS), hexa-histidine repeat (HHHHHH) and the myc-tag (EQKLISEEDLNGAA) at the C-terminus are indicated.

Table: 5.9: Amino Acid alignment of TFI-scFv constructs

```
CLUSTAL 2.1 Multiple Sequence Alignments

pIT2     : pIT2_Construct       288 Amino Acids
pET22_1 : pET22_Construct_1     251 Amino Acids
pET22_2 : pET22_Construct_2     261 Amino Acids

pIT2     MKYLLPTAAAGLLLLAAQPAMAEVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVR
pET22_1 -------------------MAEVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVR
pET22_2 -------------------MAEVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVR
                            *************************************

pIT2     QAPGKGLEWVSSINPLGWKTRYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK
pET22_1 QAPGKGLEWVSSINPLGWKTRYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK
pET22_2 QAPGKGLEWVSSINPLGWKTRYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK
          ************************************************************

pIT2_    TSSRFDYWGQGTLVTVSSGGGGSGGGGSGGGGSTDIQMTQSPSSLSASVGDRVTITCRAS
pET22_1 TSSRFDYWGQGTLVTVSSGGGGSGGGGSGGGGSTDIQMTQSPSSLSASVGDRVTITCRAS
pET22_2 TSSRFDYWGQGTLVTVSSGGGGSGGGGSGGGGSTDIQMTQSPSSLSASVGDRVTITCRAS
          ************************************************************

pIT2_    QSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATY
pET22_1 QSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATY
pET22_2 QSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATY
          ************************************************************

pIT2     YCQQSYSTPNTFGQGTKVEIKRAAAHHHHHHGAAEQKLISEEDLNGAA
pET22_1 YCQQSYSTPNTFGQGTKVEIKRAAAHHHHHH----------------
pET22_2 YCQQSYSTPNTFGQGTKVEIKRNSSSVDKLAAALEHHHHHH-------
          ********************* ::   .:
```

Alignment of the three constructs clearly shows the absence of the pelB leader peptide and myc-tag in the pET22 constructs. The His-tag of construct 2 is located in a 10 amino acid downstream position relative to construct 1. The heavy and light chain inserts remained unchanged as indicated by the (*).

The finalised sequences for both constructs where analysed for the presence of rare codons for expression in *E. coli* using internet based Rare Codon Calculator (http://people.mbi.ucla.edu/sumchan/caltor.html). The various rare codons within the finalised DNA sequence are highlighted and quantities summarised in Table 5.10.

Table 5.10: Rare Codon analysis of pET22 Construct 1

```
pET22 Construct 1          251 Amino acids


ATGGCCGAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGACTCTCCTGTGCA
GCCTCTGGATTCACCTTTAGCAGCTATGCCATGAGCTGGGTCCGCCAGGCTCCAGGGAAGGGCCTGGAGTGGGTC
TCATCTATTAATCCGTTGGGTTGGAAGACACGTTACGCAGACTCCGTGAAGGGCCGGTTCACCATCTCCAGAGAC
AATTCCAAGAACACGCTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCCGTATATTACTGTGCGAAA
ACTTCGTCTAGGTTTGACTACTGGGGCCAGGGAACCCTGGTCACCGTCTCGAGCGGTGGAGGCGGTTCAGGCGGA
GGTGGCAGCGGCGGTGGCGGGTCGACGGACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGA
GACAGAGTCACCATCACTTGCCGGGCAAGTCAGAGCATTAGCAGCTATTTAAATTGGTATCAGCAGAAACCAGGG
AAAGCCCCTAAGCTCCTGATCTATGCTGCATCCAGTTTGCAAAGTGGGGTCCCATCAAGGTTCAGTGGCAGTGGA
TCTGGGACAGATTTCACTCTCACCATCAGCAGTCTGCAACCTGAAGATTTTGCAACTTACTACTGTCAACAGAGT
TACAGTACCCCTAATACGTTCGGCCAAGGGACCAAGGTGGAAATCAAACGGCGGCCGCACATCATCATCACCAT
CACTAG


pET22 Construct 2          261 Amino Acids


ATGGCCGAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGACTCTCCTGTGCA
GCCTCTGGATTCACCTTTAGCAGCTATGCCATGAGCTGGGTCCGCCAGGCTCCAGGGAAGGGCCTGGAGTGGGTC
TCATCTATTAATCCGTTGGGTTGGAAGACACGTTACGCAGACTCCGTGAAGGGCCGGTTCACCATCTCCAGAGAC
AATTCCAAGAACACGCTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCCGTATATTACTGTGCGAAA
ACTTCGTCTAGGTTTGACTACTGGGGCCAGGGAACCCTGGTCACCGTCTCGAGCGGTGGAGGCGGTTCAGGCGGA
GGTGGCAGCGGCGGTGGCGGGTCGACGGACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGA
GACAGAGTCACCATCACTTGCCGGGCAAGTCAGAGCATTAGCAGCTATTTAAATTGGTATCAGCAGAAACCAGGG
AAAGCCCCTAAGCTCCTGATCTATGCTGCATCCAGTTTGCAAAGTGGGGTCCCATCAAGGTTCAGTGGCAGTGGA
TCTGGGACAGATTTCACTCTCACCATCAGCAGTCTGCAACCTGAAGATTTTGCAACTTACTACTGTCAACAGAGT
TACAGTACCCCTAATACGTTCGGCCAAGGGACCAAGGTGGAAATCAAACGGAATTCGAGCTCCGTCGACAAGCTT
GCGGCCGCACTCGAGCACCACCACCACCACCACTGA















Table continues on next page.
```

```
Summary of Rare Codon coposition


Percentage Rage codons for Construct 1 : 12%
Percentage Rare codons for Construct 2 : 11%
```

Amino Acid	Rare Codon	Occurrence
Arginine	AGA	4
	CGG	3
	AGG	2
	CGA	0
Glycine	GGG	10
	GGA	7
Isoleucine	AUA	0
Leucine	CUA	0
Proline	CCC	0
Threonine	ACG	4
Total		30

```
Consecutive Rare Codons

GGG GGA = 1
GGG GGG = 1
```

Due to the similarity of the two constructs, the position and quantity of rare codons are identical for both sequences. Analysis for the pET22 constructs indicated the presence of 30 rare codons including two instances of consecutive rare codon repeats near the N-terminal region of the DNA sequences.

5.5 Discussion

5.5.1 Cloning TFI-scFv gene

Primers were designed based on pIT2-TFI sequence obtained in Chapter 3. Using Vector NTI v10.0, two primer sets were designed consisting of a mutual forward primer and two dissimilar reverse primers.

The mutual forward primer (scFv F_NdeI) was designed to bind downstream of the pelB leader sequence in order to remove it from the new constructs as illustrated in Figure 5.8. The removal of the pelB leader peptide will prevent the transport of TFI-scFv to the culture media and enable the purification of TFI-scFv from the *E. coli* BL21 cytoplasm. The function of the pelB leader peptide is to facilitate the transport of the cytoplasmic expressed TFI-scFv to the periplasm region (Lei *et al.*, 1987). Once TFI-scFv reaches the periplasm, the pel B leader peptide is cleaved by signal peptidase (Kang *et al.*, 1991). As a result, the pelB leader peptide is absent from the mature TFI-scFv and the removal of the pelB leader peptide will not directly influence epitope recognition. The forward primer (--5' CAT ATG GCC GAG GTG CAG CTG 3'--) was design to contain a 5' *Nde*I (--5' CAT ATG 3'--) restriction site for the incorporation of the constructs into pET22 expression vector with the correct orientation. The *Nde*I restriction site also contains the ATG (START) codon for expression.

pIT2 Construct
938 bp
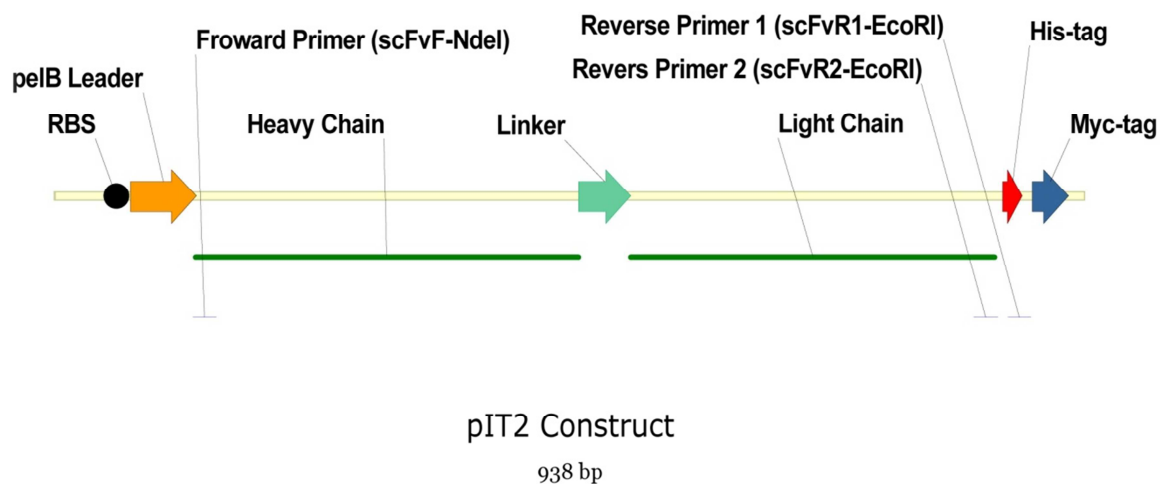
Figure 5.8: Binding sites of primer sets op TFI-phagemid

The primer binding sites are indicated as (▬). Human immunoglobulin G heavy- and light chain inserts are indicated as (▬). The pelB Leader sequence (pelB Leader) is indicated by (⇨), glycine-serine linker sequence repeat (Linker) by (⇨), polyhistidine sequence (His-Tag) by (⇨) and c-myc gene (Myc-tag) are indicated by (⇨).

Reverse primer 1 (scFv R1_EcoRI) was designed to include the C-terminus histidine tag of the original pIT2 plasmid that is required for nickel affinity chromatography. The reverse primer (--5' GAA TTC TAG TGA TGG TGA TGA 3'--) contains 3' *EcoR*I (--5' GAA TTC 3'--) restriction site for the incorporation of the constructs into pET22

expression vector in the correct orientation as well as a TAG (STOP) codon. Reverse primer 2 (scFv R2_EcoRI) was designed to bind 5 bases upstream of the pIT2 C-terminus histidine tag in order to remove the pIT2 His-tag from construct 2. The final construct would utilize the pET22 histidine tag for purification purposes. The reverse primer (--5' GAA TTC CGT TTG ATT TCC ACC T3'--) also contains a 3' *EcoR*I (--5' GAA TTC 3'--) restriction site for the incorporation of the constructs into pET22 expression vector in the correct orientation.

Simulated PCR amplification using primer set 1 and 2 predicted the amplification of a 764 bp and 734 bp fragment respectively. This was in agreement with the PCR product fragment sizes observed after 0.8% agarose gel electrophoresis (See: Figure 5.2). The lack of any additional DNA bands indicates the high specificity of the primer sets. These DNA fragments were isolated, incorporated into the pSMART cloning vector and transformed into *E. coli* TOP10 cloning host in order to propagate the constructs *in vivo*.

A total of 12 single *E. coli* TOP10 colonies, six for each of the pSMART constructs, were selected at random and analysed for the presence of the inserts. Double digestion of the constructs using *Nde*I and E*coR*I (See: Figure 5.3) enabled the identification of positive pSMART transformations that contain the 764 bp and 734 bp inserts for constructs 1 and 2 respectively. All of the six colonies analysed for the presence of pSMART construct 1 were positive for the presence of the pSMART. No false positive was selected, indicating that only *E. coli* TOP10 cells containing the pSMART plasmid (thereby gaining kanamycin resistance) were able to survive the exposure to 30 mg/L^{-1} kanamycin. Of the six colonies selected only those corresponding to lane 1, 3 5 and 6 were positive for the 764 bp insert. No additional DNA bands, excluding that of the pSMART plasmid, were present in any of the digestion products. All of the colonies analysed for the presence of construct 2 were positive for the presence of pSMART as well as the 734 bp insert. No additional DNA bands were present in any of the digestion products. The 764 bp DNA fragment in lane 5 of the construct 1 and 734 bp fragment in lane 1 of construct 2 was isolated, incorporated into the pET22b (+) expression vector and transformed to *E. coli* TOP10 cloning host in order to amplify the constructs *in vivo*.

A total of 12 single *E. coli* TOP10 colonies, six for each of the pET22 constructs, were selected at random and analysed for the presence of the 764 bp and 734 bp inserts for constructs 1 and 2 respectively. Double digestion of the constructs using *Nde*I and E*coR*I (See: Figure 5.4) restriction enzymes were performed on pET22 constructs isolated from *E. coli* Top10. All six of the colonies analysed for the presence of pET22 construct 1 were positive for the presence of the pET22 plasmid and 765 bp insert. The same was true for the colonies analysed for the presence of pET22 construct 2. All contained both the pET22 plasmid and the 734 bp insert. In the analysis of pET22 construct 1 and 2 no additional DNA fragments were present indicating that a pure sample was obtained. The pET22 constructs corresponding to lane 6 for construct 1 and lane 3 for construct 2 were selected for expression.

5.5.2 Cytoplasmic expression of pET22 constructs

The pET22b (+) plasmid utilises the lac operon for the expression of the inserted recombinant DNA. Induction of transcription can be accomplished by the addition of lactose or by the addition of a lactose analogue such as Isopropyl-β-D-thiogalactoside (IPTG) to the growth medium (Loomis and Magasanik, 1967; Pannekoek *et al.*, 1995; Borukhov and Lee, 2005). The lactose (1-4-*O*-β-d-galactopyranosyl-d-glucose) present in the culture media is converted by β-galactosidase to allolactose (1-6-*O*-β-d-galactopyranosyl-d-glucose) (Santillán and Mackey, 2004). The allolactose is the "true" inducer of the lac operon in the absence of glucose (Santillán *et al.*, 2007). Expression of the pET22 constructs by *E. coli* BL21 was induced through a combination of lactose and glucose media known as an auto-induction media. Auto-induced cultures typically have higher growth densities and as a result produce higher levels of target protein per millilitre of culture than cultures induced by IPTG (Studier, 2005). Li *et al.* (2011) reported that auto-induction produced a four time higher protein yield than an IPTG induction in a study comparing the expression of nine proteins by various plasmids in *E. coli* BL21 (Li *et al.,* 2011). Induction of the lac operon by auto-induction media is regulated by the availability of lactose and glucose in culture media. When glucose is present in the media it will prevent the uptake of lactose and the subsequent metabolic production of allolactose (Peti and Page, 2007). This is due to the fact that glucose is preferentially metabolized during early growth phase (Tyler *et al.*, 1967; Borukhov

117

and Lee, 2005; Li *et al.*, 2011). As a result the culture growth is enhanced due the presence of glucose. As the biomass increases however, the glucose will gradually become depleted (Studier, 2005). Once the glucose is present at sufficiently low levels the metabolic conversion of lactose to allolactose will occur and lead to the subsequent induction of the lac operon (Santillán *et al.*, 2007). The advantage of induction by auto-induction media is that recombinant protein expression is only initiated at a relatively high culture density (Peti and Page, 2007). An additional advantage to auto-induction media is that the cultures do not require regular spectrophotometric monitoring in order to determine the correct time to add IPTG in order to induce expression (Studier, 2005).

The *E. coli* BL21 (DE3) whole cell fractions containing pET22 construct 1, empty pET22 plasmid or construct 2 were normalised and subjected to 12.5 % SDS-PAGE in order to confirm the expression of TFI-scFv (See: Figure 5.5). The SDS-PAGE analysis confirmed the successful expression of both pET22 constructs by *E. coli* BL21. The expressed 26 KDa TFI-scFv falsely appeared to be of greater size (approximately 30 to 32 KDa). The apparent increase in fragment size is due to the increased motility of the pre-stained protein ladders and not due to genetic modification of TFI-scFv gene itself. This accounts for the approximate 5 KDa discrepancy in the expected TFI-scFv size. The 26 KDa TFI-scFv was not present in the negative control that contained the empty pET22 plasmid. The absence of a 26 KDa band in the negative control indicates that the additional 26 KDa band observed in the lane 1 and 3 is indeed that of TFI-scFv and not of another constitutively expressed protein of similar size. The presence of TFI-scFv in the whole cell fraction also confirms that the removal of the N-terminal pelB leader peptide resulted in the expression of TFI-scFv in the cytoplasm.

The similarities in band intensities between the two constructs indicate that the expression rates were similar. This was expected as the two constructs were of similar size and composition, differing by only 10 amino acids at the C-terminus, and were expressed in an identical manner. The similarity between the band intensity of the induced protein and that of the natively expressed *E. coli* proteins indicate that over-expression of TFI-scFv did not occur for either construct. Although protein expression rate is inherently depended on recombinant protein itself, the low levels

of expression was unexpected as the pET22 – *E. coli* BL21 expression system is widely utilised as it generally exhibits high level expression (Studier, 2005; Sørensen and Mortensen, 2005; Cabrita *et al.*, 2006). There are multiple parameters that are required for the successful expression of recombinant proteins in *E. coli*. These parameters include the efficiency of DNA transcription and translation, the stability of the expression vector and that of the transcribed RNA, localisation, proteolytic stability and folding of the expressed protein, as well as the efficiency of cell growth (Jonasson *et al.*, 2002; Sørensen and Mortensen, 2005).

Despite these many factors that may influence the successful expression of TFI-scFv, the low level of expression is most likely due to the heterologous expression of the heavy- and light chain inserts that are of human origin. The over-expression of a heterologous protein results in the production of large quantity of heterologous mRNA that will be encountered by the *E. coli* transcription machinery (Kane *et al.*, 1995). The efficiency of translation and eventual production of the target protein is dependent on the heterologous codon composition (Baca and Hol, 2000). The frequencies that certain codons appear in genes of human origin differ from that of genes of *E. coli* origin. This directly reflects a variation in the amount of specific tRNAs available in each organism during translation (Jonasson et al., 2002). The implication of this is that a specific codon that is frequently used in Humans and thus in high supply, will may be rarely used in *E. coli* and thus in low supply. The implication of this is that although both organisms have a ubiquitous set of 61 codons for the translation of mRNA, that the variation in tRNA availability between each organism would influence their ability to translate a specific mRNA sequence (Kane *et al.,* 1995; Kleber-Janke & Becker, 2000; Jonasson *et al.,* 2002). It is this phenomenon of variation in codon preference between organisms that complicates heterologous protein expression as the presence and distribution of rare codons within a DNA sequence is detrimental to its expression (Makrides, 1996; Hannig and Makrides, 1998).

5.5.3 Nickel affinity chromatography

The large culture volume (1 L) of auto-induction media utilized during expression yielded approximately 7 g (wet weight) *E. coli* BL21 for each construct. A large

volume was used in order to compensate for low levels of expression. The cytoplasmic fractions of both constructs were isolated and the TFI-scFv purified by means of nickel affinity chromatography using HIS-Trap™ FF column incorporation with the ÄKTA Prime Plus system. The ability to express TFI-scFv in the *E. coli* cytoplasm and thereby reduced sample volume size from 2L (as discussed in the Chapter 4) down to approximately 30 mL, greatly improved the processing of protein sample. Apart from solving the complications with nickel chelating agents in the culture supernatant, it also reduced the sample volume as well as time required to purify the TFI-scFv protein.

The linear increase in absorbance between points B and C in both purifications is due to the absorbance that is presented at 280 nm by imidazole. Although the protein peaks obtained during the purifications are minute, they are clearly visible as disturbances in the early stages of the imidazole gradient. The lack of a clearly defined protein peak is most likely linked to the low levels of expression that was previously observed (See: Section 5.4.4). It is also clear that the attempt to compensate for a low rate of expression by using large volumes of culture failed. With the current level of expression a substantially larger culture volume would be required in order to produce sufficient amounts of TFI-scFv for downstream testing.

Although very little of either construct was purified, it is interesting that the purification of construct 2 was more successful than that of construct 1. This was despite both constructs having similar DNA sequences, exhibiting similar expression rates and being purified in an identical fashion. It can be speculated that the improved purification of construct 2 might be due to the slightly further downstream position (10 amino acids) of the His-tag (See: Table 5.9). The position of the His-tag in terms of its placement near the functional domains can influence the ability to purify the expressed protein by means of nickel affinity chromatography (Goubran-Botros and Vijayalakshmi, 1991; Gaberc-Porekar and Menart, 2001; Klose *et al.*, 2004; Kabir *et al.*, 2010). During the folding of construct 1 the C-terminus His-tag might be folded in such a way that it becomes partially buried in the rest of the TFI-scFv protein itself and therefore unable to fully interact with the nickel column. By shifting the His-tag 10 amino acids downstream, it may be more exposed after the

folding of construct 2 and consequently have superior interaction with the nickel affinity column.

5.5.4 DNA analysis of the TFI-scFv constructs

The TFI-scFv constructs were sequenced in order to confirm the successful removal of the pel B leader peptide as well as the presence of an in-frame His-tag at the C-terminus. Additionally the sequencing results enabled the comparison of the two pET22 constructs with the original pIT2 construct in order to confirm that no mutation of the original synthetic gene occurred during the various cloning steps.

Both constructs were sequenced using the T7 promoter and T7 terminator primers in order to encompass the whole TFI-scFv synthetic gene (See: Figure 5.7). The sequencing of the constructs revealed 972 bp and 942 bp DNA fragments for construct 1 and 2 respectively. Sequencing results confirmed that the original pelB leader peptide (of pIT2 plasmid origin) was successfully removed from both constructs. The constructs were incorporated into the pET22 expression vector through the utilisation of the *Nde*I and *EcoR*I restriction sites, in such a manner that the pelB leader peptide in the cloning/expression region of pET22b (+) plasmid was also not retained. Both constructs contained a ATG (START) codon upstream of the heavy chain inserts. Construct 1 contained a TAG (STOP) codon just downstream of the His-tag retained from the original pIT2 construct, while Construct 2 utilised a TGA (STOP) codon downstream of the His-tag of pET22b (+) origin.

Translation of the TFI-scFv DNA open reading frames produced 251 amino acid and 261 amino acid sequence from construct 1 and 2 respectively with an in-frame C-terminus His-tag in both constructs. The removal and the pelB leader peptide and variation in C-terminal His-tag positions were clearly visible from the multiple alignment of the original pIT2 construct with the two pET22 constructs (See: Table 5.9). TFI-scFv construct 1 retained the original His-tag from the pIT2 plasmid but not the downstream myc-tag sequence due to the incorporation of a TAG (STOP) codon in the reverse primer scFv R1_EcoRI (See: Table 5.1) that binds just downstream of the original His-tag. TFI-scFv construct 2 did not retain the original His- or myc-tags. Due to the utilisation of the pET22 His-tag in construct 2, it was located 10 amino

acids downstream relative to the His-tag position of construct 1. Although speculative, the slightly more successful purification of construct 2 in comparison to that of construct 1 was most likely due to its His-tag positioned further downstream of the functional domains of TFI-scFv.

Insufficient amounts of TFI-scFv were purified in order to confirm the ability of the new constructs to inhibit tissue factor by prothrombin coagulation testing. It is important to keep in mind that despite the removal of the pelB leader peptide and additional modification to the original TFI-scFv construct, the heavy- and light chain inserts responsible for the recognition of epitopes on tissue factor have remained unchanged. Based purely on the DNA sequences, the antibodies should retain their ability to inhibit tissue factor, although the correct folding of the expressed scFv in the cytoplasm environment remained unconfirmed.

The rare codon analysis of the TFI-scFv constructs for expression in *E. coli* (See: Table 5.10) revealed that the rare codons were only present within the heavy- and light chain inserts. As a result, the rare codon compositions of the two constructs were identical as the heavy and light chain inserts remained unchanged in both constructs. Rare codon analysis revealed that a total of 30 rare codons were present within the two constructs. This constituted 12 % and 11.5% of the total codons required for the translation of construct 1 and 2 respectively. Genes with high rare codon contents have slower translation rates and ultimately low yields (Widmann *et al.*, 2008; Galperin and Cochrane, 2009). Although a high percentage of rare codons reduce the efficacy of translation, it is the presence of consecutive rare codons located near the N-terminal that is especially detrimental to translation (Kane *et al.*, 1995; Hannig and Makrides, 1998). The low level of expression observed (See: Section 5.4.4) is a result of the relatively high amount of rare codons present in the constructs.

5.6 Conclusion

The TFI-scFv synthetic gene was successfully isolated from the the pIT2 plasmid and cloned into the pET22b (+) expression vector. The expression of two TFI-scFv constructs was successfully directed to the cytoplasm of *E. coli* BL21 through the removal of the pelB leader peptide. The cytoplasmic expression resolved the complications encountered during nickel affinity chromatography due to the presence of nickel chelators in the culture supernatant. The cytoplasmic expression also resulted in reduction of the sample volume. Although both constructs were successfully purified by means of nickel affinity chromatography, the final protein yield was insufficient to perform functional assays. Despite the low yield, the purification of construct 2 was superior to that of construct 1 due to further downstream positioning of the His-tag. As a result it was decided to continue with only construct 2 in the following chapter. The inability of nickel affinity chromatography to purify sufficient amounts of protein was a result of low expression levels, despite the utilisation of a popular over-expression system. The low expression levels were due to the high frequency of rare codons present within the TFI-scFv gene. In order to fully benefit from the redirection of TFI-scFv to the cytoplasm, the expression of TFI-scFv must be modified to compensate for the low rate of expression caused by the presence of rare codons.

5.7 References

BACA, A.M., HOL, W.G. 2000. Overcoming codon bias: a method for high-level overexpression of Plasmodium and other AT-rich parasite genes in Escherichia coli. *International Journal for Parasitology,* 30, 113-8.

BANEYX, F. 1999. Recombinant protein expression in Escherichia coli. *Current Opinion in Biotechnology,* 10, 411–421.

BORUKHOV, S., LEE, J. 2005. RNA polymerase structure and function at lac operon. *Comptes Rendus Biologies,* 328, 576-87.

CABRITA, L.D., DAI, W., BOTTOMLEY, S.P. 2006. A family of E. coli expression vectors for laboratory scale and high throughput soluble protein production. *BMC Biotechnology,* 6, 12-20.

GABERC-POREKAR, V., MENART, V. 2001. Perspectives of immobilized-metal affinity chromatography. *Journal of Biochemical and Biophysical Methods,* 49, 335-60.

GALPERIN, M.Y., COCHRANE, G.R. 2009. Nucleic Acids Research annual Database Issue and the NAR online Molecular Biology Database Collection in 2009. *Nucleic Acids Research,* 37, D1-4.

GOUBRAN-BOTROS, H., VIJAYALAKSHMI, M. 1991. Immobilized metal ion affinity electrophoresis: a preliminary report. *Electrophoresis,* 12 1028-1032.

GU, Z., WEIDENHAUPT, M., IVANOVA, N., PAVLOV, M., XU, B., SU, Z.G., JANSON, J.C. 2002. Chromatographic methods for the isolation of, and refolding of proteins from, Escherichia coli inclusion bodies. *Protein Expression and Purification,* 25, 174-9.

HANNIG, G., MAKRIDES, S.C. 1998. Strategies for optimizing heterologous protein expression in Escherichia coli. *Trends in Biotechnology,* 16, 54-60.

JONASSON, P., LILJEQVIST, S., NYGREN, P.-A., STÅHL, S. 2002. Genetic design for facilitated production and recovery of recombinant proteins in Escherichia coli. *Biotechnology and Applied Biochemistry,* 35, 91-105.

KABIR, M.E., KRISHNASWAMY, S., MIYAMOTO, M., FURUICHI, Y., KOMIYAMA, T. 2010. Purification and functional characterization of a Camelid-like single-domain antimycotic antibody by engineering in affinity tag. *Protein Expression and Purification*, 72, 59-65.

KANE, J.F., 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. *Current Opnion Biology*, 6, 494-500.

KANG, A S., BARBAS, C.F., JANDA, K.D., BENKOVIC, S.J., LERNER, R.A. 1991. Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, 88, 4363-4366.

KLEBER-JANKE, T., BECKER, W.M. 2000. Use of modified BL21(DE3) Escherichia coli cells for high-level expression of recombinant peanut allergens affected by poor codon usage. *Protein Expression and Purification*, 19, 419-24.

KLOSE, J., WENDT, N., KUBALD, S., KRAUSE, E., FECHNER, K., BEYERMANN, M., BIENERT, M., RUDOLPH, R., ROTHEMUND, S., 2004. Hexa-histidin tag position influences disulfide structure but not binding behavior of in vitro folded N-terminal domain of rat corticotropin-releasing factor receptor type 2a. *Protein Science*, 13, 2470–2475.

LEI, S.P., LIN, H.C., WANG, S.S., CALLAWAY, J., WILCOX, G. 1987. Characterization of the Erwinia carotovora pelB gene and its product pectate lyase. *Journal of Bacteriology*, 169, 4379-83.

LI, C., LU, C.D. 2009. Unconventional integration of the bla gene from plasmid pIT2 during ISlacZ/hah transposon mutagenesis in Pseudomonas aeruginosa PAO1. *Current Microbiology*, 58, 472-477.

LI, Z., KESSLER, W., VAN DEN HEUVEL, J., RINAS, U. 2011. Simple defined autoinduction medium for high-level recombinant protein production using T7-based Escherichia coli expression systems. *Applied Microbiology and Biotechnology*, 1203-1213.

LILIE, H., SCHWARZ, E., RUDOLPH, R. 1998. Advances in refolding of proteins produced in E. coli. *Current Opinion in Biotechnology*, 9, 497-501.

LOOMIS, W.F., MAGASANIK, B. 1967. Glucose-lactose diauxie in Escherichia coli. *Journal of Bacteriology,* 93, 1397-1401.

MAKRIDES, S.C., 1996. Strategies for achieving high-level expression of genes in Escherichia coli. *Microbiological Reviews*, 60, 512-538.

MANTING, E.H., DRIESSEN, A.J.M. 2000. Escherichia coli translocase : the unravelling of a molecular machine. *Molecular Microbiology,* 37, 226-238.

PANNEKOEK, H., MEIJER, M., GAARDSVOLL, H., ZONNEVELD, A.J. 1995. Functional display of proteins, mutant proteins, fragments of proteins and peptides on the surface of filamentous (bacterio) phages: A review. *Cytotechnology*, 18, 107-112.

PETI, W., PAGE, R. 2007. Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost. *Protein Expression and Purification*, 51, 1-10.

RUDOLPH, R., LILIE, H. 1996. In vitro folding of inclusion body proteins. *The FASEB Journal*, 10, 49-56.

SANTILLÁN, M., MACKEY, M.C. 2004. Influence of catabolite repression and inducer exclusion on the bistable behaviour of the lac operon. *Biophysical Journal*, 86, 1282-1292.

SANTILLÁN, M., MACKEY, M.C. ZERON, E.S. 2007. Origin of bistability in the lac Operon. *Biophysical Journal*, 92, 3830-3842.

SLETTA, H., TONDERVIK, A., HAKVAG, S., AUNE, T., NEDAL, A., AUNE, R., EVENSEN, G., VALLA, S., ELLINGSEN, T., BRAUTASET, T. 2007. The presence of N-terminal secretion signal sequences leads to strong stimulation of the total expression levels of three tested medically important proteins during high-cell-density cultivations of Escherichia coli. *Applied and Environmental Microbiology*, 73, 906-912.

STUDIER, F.W. 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, 41, 207-234.

SØRENSEN, H.P., MORTENSEN, K.K. 2005. Advanced genetic strategies for recombinant protein expression in Escherichia coli. *Journal of Biotechnology*, 115, 113-128.

TAPRYAL, S., KRISHNAN, L., BATRA, J.K., KAUR, K.J., SALUNKE, D.M. 2010. Cloning, expression and efficient refolding of carbohydrate-peptide mimicry recognizing single chain antibody 2D10. *Protein Expression and Purification*, 72, 162-8.

THIE, H., SCHIRRMANN, T., PASCHKE, M., DÜBEL, S., HUST, M. 2008. SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in E. coli. *New Biotechnology*, 25, 49-54.

TYLER, B., LOOMIS, W.F., MAGASANIK, B. 1967. Transient repression of the lac operon. *Journal of Bacteriology*, 94, 2001-11.

VELAPPAN, N., MARTINEZ, J.S., VALERO, R., CHASTEEN, L., PONCE, L., BONDU-HAWKINS, V., KELLY, C., PAVLIK, P., HJELLE, B., BRADBURY, A.R.M. 2007. Selection and characterization of scFv antibodies against the Sin Nombre hantavirus nucleocapsid protein. *Journal of Immunological Methods*, 321, 60-69.

WIDMANN, M., CLAIRO, M., DIPPON, J., PLEISS, J. 2008. Analysis of the distribution of functionally relevant rare codons. *BMC Genomics* 9, 207-215.

CHAPTER 6

Rare-codon optimization of TFI-scFv

6.1 Introduction

The overexpression of heterologous proteins in *E. coli* is often unsuccessful due to the presence of rare codons in the target sequence. Apart from reducing final protein yield, rare codons have also been found to cause translational errors, especially in areas where rare codons are present and in clusters where doublets and triplets accumulate in large quantities (Sørensen and Mortensen, 2005; Redwan, 2006). Rare codon clusters lead to the premature termination of mRNA translation resulting in the production of truncated proteins (Roche & Sauer, 1999, Burgess-Brown *et al.,* 2008). Even the translation of mRNA containing a single rare codon can result in ribosomal stalling at these sites due to the requirement of amino acids supplied by minor tRNAs (Redwan, 2006). Ribosomal stalling can also occur at sites requiring amino acids supplied by major tRNAs due to the overutilization of that particular amino acid (Kane *et al.,* 1995; Wu *et al.,* 2004). Rare codon induced translation errors are generally observed as codon misreadings and processing errors such as amino acid substitution or frameshift events (Wu *et al.,* 2004).

In general, two strategies have been employed with various levels of success to address the complications associated with rare codons during heterologous protein expression in *E. coli* (Rosano and Ceccarelli, 2009). The first of these strategies addresses the availability of the minor tRNAs corresponding to rare codons by either the cloning of a desired wild type tRNA gene into the expression itself or using tRNA enhanced host strains (Burgess-Brown *et al.,* 2008). These host strains are usually modified to contain a compatible plasmid, commonly the pRARE- or pRARE2 plasmid commercialised by Novagen. This plasmid harbours a set of tRNA genes that are expressed in order to compensate for rare codons. By the co-expression of the pRARE plasmids, the availability of minor tRNAs that are required for rare codon translation during target protein expression, is improved (Makrides, 1996; Sørensen and Mortensen, 2005). The co-expression of plasmids harbouring rare tRNAs is an

inexpensive and generally effective strategy to compensate for rare codons (Baca and Hol, 2000; Redwan, 2006; Rosano and Ceccarelli, 2009).

The second strategy involved the modification of the target DNA sequence through site directed mutagenesis in order to replace rare codons with analogous codons that are frequently utilized in *E. coli* (Hannig & Makrides, 1998; Baca & Hol, 2000). The major disadvantage to this strategy is that the process becomes tedious and costly during extensive modification of the target gene (Baca and Hol, 2000; Redwan, 2006). This strategy would have been especially unpractical for the modification of the TFI-scFv DNA sequence as it contains 30 rare codons. A more attractive alternative is to simply synthesize a codon optimised target gene. Apart from being more cost effective, the synthesised codon optimized DNA sequence is also engineered to avoid regions of secondary mRNA structures and cis acting regions that could hamper gene expression (Redwan, 2006).

6.2 Aim

The aim of this section of the study was to improve the expression of TFI-scFv by compensating for the detrimental effect of rare codons within the TFI-scFv DNA sequence.

6.3 Materials and Methods

6.3.1 pRARE rare codon compensation analysis

It was decided to continue only with pET22 construct 2 (See: Section 5.3.4.7) due to its improved interaction with the nickel affinity column in comparison with pET22 construct 1. The distribution and extent of rare codon compensation for pET22 construct 2 was examined based on its rare codon analysis previously obtained in Section 5.4.6.

6.3.2 Co-expression of TFI-scFv and pRARE in *E. coli* BL21 (DE3)

The pET22 construct 2 was co-expressed with pRARE in *E. coli* BL21 (DE3) in order to improve the availability of tRNAs coding for rare codons.

6.3.2.1 Transformation of pET22 construct to *E. coli* Bl21 pRARE

pET22 construct 2, designed in Chapter 5 (See: Section 5.3.4.7), was transformed by means of heat-shock to chemical competent *E.col* Bl21 cells containing the pRARE plasmid that was kindly provided by Prof. J Albertyn, Department of Biotechnology, University of the Free State, RSA. This specific cell line was developed by the isolation of pRARE plasmids from *E. coli* Rosetta Gami (Novagen, USA) that were then transformed to *E. coli* BL21 (DE3) (Invitrogen, USA). Successful transformations were isolated, amplified and was made competent. The pET22 construct 2 was transformed to *E. coli* BL21-pRARE competent cells as previously described in Section 5.3.2.3 using 100 mg.L^{-1} ampicillin and 34 mg.L^{-1} chloramphenicol antibiotics where required.

6.3.2.2 Co-expression in *E. coli* BL21/pRARE/pET22 TFI-scFv

After completion of the transformation, all the surviving *E. coli* BL21 colonies containing both the pET22 construct and pRARE plasmid were collected as previously described in Section 5.3.5.2 using 100 mg.L^{-1} ampicillin and 34 mg.L^{-1} chloramphenicol antibiotics.

6.3.3 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SDS-PAGE was performed on whole cell samples of the *E. coli* BL21 co-expressing pRARE and pET22 construct 2, a negative control containing pRARE and an empty pET22 plasmid, *E. coli* BL21 expressing pET22 construct 2 and a negative control containing only empty pET22 plasmid. Whole cell fractions were normalised at OD$_{600}$. SDS-PAGE was performed as previously described in Section 3.3.3.

6.3.4 Purification of TFI-scFv

6.3.4.1 Isolation of soluble fraction

The soluble cytoplasmic fraction was isolated from harvested *E. coli* BL21 cultures (See: Section 6.3.2.2) as previously described in Section 5.3.7.1

6.3.4.2 Immobilized metal ion affinity chromatography (IMAC)

The nickel affinity chromatography purification of TFI-scFv from the isolated cytoplasmic fraction was performed as previously described in Section 4.3.2.

6.3.4.3 Size exclusion chromatography

The fractions corresponding to the TFI-scFv peak obtained during the nickel affinity chromatography were pooled together. The pooled fractions (approximately 35 mL) were concentrated to 3 mL using an Amicon stirrer cell (Millipore, USA) through a 10 kDa MWCO membrane (Osmonics, USA) using N_2 gas (Afrox, RSA) at 100 kPa at 4 ℃. The concentrated sample was loaded onto a (2.5 cm x 63 cm) Sephacryl S200HR column (Sigma, Switzerland) equilibrated with sodium phosphate (PBS) buffer ($8g.L^{-1}$ NaCl, $1.44g.L^{-1}$ Na_2HPO_4, 0.24 $g.L^{-1}$ KH_2PO_4 and $0.2g.L^{-1}$ KCl at pH7.4 using the ÄKTA Prime Plus (Amersham Biosciences, UK) system. Target protein was eluted using the same buffer at a flow speed of 1 $mL.min^{-1}$ and collected in 5 mL fractions and stored at 4 ℃ until later use. The fractions corresponding to the TFI-scFv peak were pooled together and concentrated to 10 mL using the Amicon stirrer cell as described above. At this point a buffer exchange was performed by adding 40 mL Tris Buffered Saline (TBS) buffer (8 $g.L^{-1}$ NaCl, 3 $g.L^{-1}$ tris base and 0.2 $g.L^{-1}$ KCl at pH 7.4). Buffer exchange was done in order to prepare the TFI-scFv sample for prothrombin analysis as the coagulation test is inhibited by phosphate buffers. The sample was again concentrated using the using the Amicon stirrer cell as described above.

6.3.5 Prothrombin time coagulation analysis

Prothrombin Time (PT) coagulation test was performed as previously described in Section 3.3.4.

6.3.6 GeneART® rare codon optimisation of TFI-scFv gene for expression in *E. coli*

Due to the high rare codon content in the TFI-scFv gene (See: Section 5.4.6) the sequence was optimised for expression in *E. coli* by GeneArt® (GeneArt, Germany). The TFI-scFv gene was synthesised to contain no rare codons or any *Nde*I-, *EcoR*I- and *Xho*I restriction sites within the gene sequence. A 5' *Nde*I- and 3' *Xho*I sticky end was incorporated for downstream cloning of the gene into pET22b (+) expression vector. The optimised TFI-scFv gene sequence was received cloned in a standard vector (Designated: pTFIopt) (See**:** Section 6.4.6) conveying kanamycin resistance. The pTFIopt vector and empty pET22b (+) was transformed into *E. coli* TOP10 (Invitrogen, USA) by means of heat shock for the *in vivo* amplification. After propagation of the plasmids, they were both isolated and subjected to double digestion with *Nde*I and *Xho*I restriction enzymes in order to isolate and prepare the TFIopt-scfv and pET22b (+) plasmid for cloning.

6.3.6.1 Heat shock transformation and propagation of pTFIopt-scFv and pET22b (+) in *E. coli* TOP10

A volume of 5 µL (0.5 ug) of the pTFIopt plasmid and 1 µL (50 ng.mL^{-1}) of the empty pET22b (+) were separately transferred to *E. coli* TOP10 cells as previously described in Section 5.3.2.3 using 30 mg.L^{-1} kanamycin and 100 mg.L^{-1} ampicillin for the selection of pTFIopt and pET22 plasmid respectively.

6.3.6.2 Isolation of plasmids

The pTFIopt and empty pET22b (+) plasmids were isolated from transformed *E. coli* TOP10 cells using the Biospin Plasmid DNA Extraction Kit (Bioflux, Japan) as previously described in Section 5.3.2.4.

6.3.6.3 Double Digestion of pTFIopt-scFv plasmid

The isolated plasmids (pTFIopt and pET22b (+)) were adjusted and dried in a Vacufuge Concentrator 5301 (Eppendorf, Germany) to contain approximately 2000 ng plasmid DNA. Dried plasmids were subjected to a double restriction digest using *Nde*I and *Xho* restriction enzymes (Fermentas, Canada) and restriction buffer O (Fermentas, Canada) in order to isolate the TFIopt-scFv gene from the pTFIopt plasmid and linearise pET22b (+) as indicated in Table 6.1.

Table 6.1: Double restriction digestion reaction

Reagent	Concentration	Volume per reaction
pTFIopt or pET22b (+)	Approximately 2000 ng	Dry sample
*EcoR*I	5 000 U.mL^{-1}	4 µL
*Xho*I	5 000 U.mL^{-1}	8 µL
Restriction Buffer O	10 x	4 µL
MilliQ sterile water	-	24 µL
Total		40 µL

Digestion reactions were incubated for 1 hour at 37 ℃ and stored at 4 ℃ until later use.

6.3.6.4 Isolation of TFIopt-scFv gene

The empty pET22b (+) plasmid and the 761 bp TFIopt-scFv optimised sequence was confirmed and isolated as previously described in Section 4.3.4.4.

6.3.7 Cloning of TFIopt-scFv gene into pET22b (+) expression vector

6.3.7.1 Ligation of constructs into pET22b (+) Vector

The sticky-end cloning of the TFIopt-scFv gene into the linear pET22b (+) plasmid was performed using the Rapid DNA Ligation Kit (Fermentas, Canada) according to product protocol. Ligation reactions for the 761 bp fragment was assembled as indicated in Table 6.2.

Table 6.2: Sticky-end ligation of constructs into pET22

Reagent	Concentration	Volume per reaction
MilliQ sterile water	-	3 µL
pET22b (+) Vector	25 ng.µL^{-1}	3 µL
TFIopt-scFv gene	25 ng.µL^{-1}	2 µL
T4 DNA Ligase,	5u/µl (Weis Units/µL)	1 µL
Rapid Ligation Buffer	10 X	1 µL
Total		10 µL

All ligation reactions were incubated overnight at 4 ℃. The following day the liagation reaction mixtures were collected for transformation into *E. coli* TOP10.

6.3.7.2 Heat shock transformation of pET22 construct into *E. coli* TOP10

The complete ligation reaction mixture was transferred to *E. coli* TOP10 as previously described in Section 5.3.2.3 using 100 mg.L^{-1} ampicillin for the selection of pET22 plasmid containing the TFIopt insert (Designated as pET22-TFIopt).

6.3.7.3 pET22-TFIopt plasmid isolation

The pET22-TFIopt plasmid was isolated from transformed cells using the Biospin Plasmid DNA Extraction Kit (Bioflux, Japan) according to manufacturer's specifications. All centrifugations were performed at room temperature using an Eppendorf centrifuge 5424 (Eppendorf, Germany).

Briefly, *E. coli* TOP10 cells were harvested from the 5 mL overnight culture in 1.5 mL microcentrifuge tubes by means of repeated centrifugation at 10000 x *g* for 30 seconds. Harvested cells were resuspended in 250 µL resuspension buffer, containing RNase, until no cell aggregates were visible. A volume of 250 µL lysis buffer was added to the resuspended cells and the content gently mixed by inversion of the microcentrifuge tubes. A volume of 350 µL neutralization buffer was then added to the lysis solution and the content gently mixed by inversion of the microcentrifuge tubes. The cellular debris was removed by means of centrifugation at 14000 x *g* for 10 minutes and the supernatant applied to Bioflux spin columns. Columns were centrifuged at 6000 x g for 60 seconds and the flow through

discarded. The columns were then washed twice with 650 µL wash buffer by means of centrifugation at 12000 x *g* for 60 seconds and the flow through discarded. The remaining residual wash buffer was aspirated after an additional centrifugation at 12000 x *g* for 60 seconds. A volume of 50 µL sterile water was applied to the Bioflux spin columns that were then incubated at room temperature for 80 seconds. The isolated pET22-TFIopt plasmids where collected in sterile 1.5 mL microcentrifuge tubes by means of centrifugation at 12000 x *g* for 60 seconds, dried in Vacufuge Concentrator 5301 (Eppendorf, Germany) and stored at – 22 ⁰C until later use.

6.3.7.4 Double restriction digestion of pET22 plasmid

The isolated pET22 plasmid was subjected to a double restriction digest using *Nde*I and *Xho*I restriction enzymes (Fermentas, Canada) and restriction buffer O (Fermentas, Canada) as indicated in Table 6.3 in order to confirm the presence of the desired inserts.

Table 6.3: Double restriction digestion reaction (Miniprep)

Reagent	Concentration	Volume per reaction
pET22 plasmid DNA	Approximately 55 ng.mL^{-1}	5 µL
*EcoR*I	5 000 U/ml)	1 µL
*Xho*I	5 000 U/ml)	2 µL
Restriction Buffer O	10 x	1 µL
MilliQ sterile water	-	1 µL
Total		10 µL

Digestion reactions were incubated for 1 hour at 37 ℃ and stored at 4 ℃ until later use.

6.3.7.5 Agarose gel electrophoresis of digested pET22-TFIopt plasmid

The presence 761 bp TFIopt-scFv gene and the digested plasmid were confirmed for documentation as previously described in Section 4.3.4.3.

6.3.8. DNA Sequencing of pET22-TFIopt construct

Bidirectional DNA sequencing of isolated pET22-TFIopt plasmid was performed with the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) with the T7 Promoter Primer (Integrated DNA technologies) (forward primer): 5' TAA TAC GAC TCA CTA TAG GG 3' and T7 Terminator Primer (reverse primer): 5' GCT AGT TAT TGC TCA GCG G 3' (Integrated DNA technologies) as previously described in Section 4.3.4.5. The finalised consensus DNA sequences were aligned with the original pET22b (+) construct 2 DNA sequence (See: Section 5.4.6) using internet based Lalign (http://www.ch.embnet.org/software/LALIGN_form.html) software, analysed for the presence of rare codons for expression in *E. coli* using internet based Rare Codon Calculator (http://people.mbi.ucla.edu/sumchan/caltor.html) and translated using internet based ExPASy Translate Tool (http://web.expasy.org/translate). The amino acid sequences aligned using ClustalW.

6.3.9 Cytoplasmic expression of pET22 constructs

Isolated pET22-TFIopt plasmid as well as empty pET22b (+) plasmid (negative control) were separately transformed into *E. coli* BL21 (DE3) (Lucigen). Expression of the constructs was induced by means of auto-induction media.

6.3.9.1 Heat shock transformation of pET22-TFIopt construct into *E. coli* BL21

The pET22-TFIopt plasmid isolates (See: Section 6.3.6.2) as well as the empty pET22 plasmid (negative control) were separately transferred to *E. coli* BL21 aliquots as previously described in Section 5.3.2.3 using 100 mg.L^{-1} ampicillin for the selection.

6.3.9.2 Expression of TFI-scFv

After the transformation, all the surviving *E. coli* BL21 colonies containing the pET22-TFIopt and empty pET22 plasmids were collected as previously described in Section 5.3.5.2 using 100 mg.L^{-1} ampicillin for selection.

6.3.9.3 Isolation of the *E. coli* cellular fractions

The harvested *E. coli* BL21 cultures containing the pET22-TFIopt plasmid and the negative control were both adjusted to 5 g (wet weight), resuspended in 50 mL cold (4 ℃) 20 mM MOPS buffer (4.186 g.L^{-1} at pH 7.4). The cells were washed and the supernatant removed by means of centrifugation at 5000 x *g* for 10 minutes at 4 ℃. The remaining cell pellets were resuspended in 20 mL cold (4 ℃) 20 mM MOPS buffer and placed on ice. A volume of 20 µL of both cultures were collected and stored at 4℃ for SDS-PAGE analysis. Collected cell s were broken using a One Shot Constant Cell Disruption System (Constant Systems) at 2000 bar. The disruption chamber was placed on ice for 1 minute between disruptions in order to prevent heat degradation of the expressed TFI-scFv. Cell disruption was performed twice per sample. The disrupted samples was collected and immediately placed on ice. The disrupted cells were centrifuged at 8000 x *g* for 10 minutes at 4 ℃ in order to separate the cellular debris from the supernatant. The pellets of cellular debris were dissolved in 3 mL cold (4℃) 20 mM MOPS buffer and stored at 4 ℃ SDS-PAGE analysis (Designated: Debris fraction). The isolated supernatants were collected and subjected to ultra centrifugation at 100000 x *g* for 1 hour and 30 minutes at 4 ℃ in order to separate the insoluble and soluble fractions. The insoluble fraction was resuspended in 5 mL cold (4℃) 20 mM MOPS buffer an d stored at 4 ℃ for SDS-PAGE analysis. The remaining soluble fractions were collected and stored at 4 ℃ until later use for SDS-PAGE analysis.

6.3.10 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

The cellular fractions isolated from *E. coli* BL21 expressing the pET22-TFIopt plasmid, as well as the cellular fractions isolated from negative control with empty pET22 plasmid, were prepared as indicated in Table 6.4. The SDS-PAGE analysis of was performed as previously described in Section 3.3.3.

Table 6.4: Whole cell SDS-PAGE preparation

Fraction	Volume	Loading Buffer	Denaturing Buffer	Total Volume
Whole cell	10 µL	8 µL	2 µL	20 µL
Cellular Debris	20 µL	16 µL	4 µL	40 µL
Insoluble Fraction	20 uL	16 µL	4 µL	40 µL
Soluble Fraction	20 µL	16 µL	4 µL	40 µL

6.4 Results

6.4.1 Rare codon compensation by pRARE

The pRARE plasmid only supplies tRNAs for AGG, AGA, AUA, CUA, CCC, and GGA rare codons. As a result not all rare codons present in the pET22 Construct 2 are compensated for by pRARE. The extent of rare codon compensation was examined based on internet based Rare Codon Calculator (http://people.mbi.ucla.edu/sumchan/caltor.html) analysis. The distribution of the compensated and well as uncompensated rare codons are presented in Table 6.5.

Table 6.5: Rare codon compensation analysis

```
pET22 Construct 2        261 Amino Acids
Rare codon composition


ATGGCCGAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGACTCTCCTGTGCA
GCCTCTGGATTCACCTTTAGCAGCTATGCCATGAGCTGGGTCCGCCAGGCTCCAGGGAAGGGCCTGGAGTGGGTC
TCATCTATTAATCCGTTGGGTTGGAAGACACGTTACGCAGACTCCGTGAAGGGCCGGTTCACCATCTCCAGAGAC
AATTCCAAGAACACGCTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCCGTATATTACTGTGCGAAA
ACTTCGTCTAGGTTTGACTACTGGGGCCAGGGAACCCTGGTCACCGTCTCGAGCGGTGGAGGCGGTTCAGGCGGA
GGTGGCAGCGGCGGTGGCGGGTCGACGGACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGA
GACAGAGTCACCATCACTTGCCGGGCAAGTCAGAGCATTAGCAGCTATTTAAATTGGTATCAGCAGAAACCAGGG
AAAGCCCCTAAGCTCCTGATCTATGCTGCATCCAGTTTGCAAAGTGGGGTCCCATCAAGGTTCAGTGGCAGTGGA
TCTGGGACAGATTTCACTCTCACCATCAGCAGTCTGCAACCTGAAGATTTTGCAACTTACTACTGTCAACAGAGT
TACAGTACCCCTAATACGTTCGGCCAAGGGACCAAGGTGGAAATCAAACGGAATTCGAGCTCCGTCGACAAGCTT
GCGGCCGCACTCGAGCACCACCACCACCACCACTGA



Table 6.5 continues on next page
```

Amino Acid	Rare Codons	Occurrence	Compensated rare codons	Remaining Uncompensated rare codons
Arginine	AGA	4	AGA	0
	CGG	3		3
	AGG	2	AGG	0
	CGA	0		0
Glycine	GGG	10		10
	GGA	7	GGA	0
Isoleucine	AUA	0	AUA	0
Leucine	CUA	0	CUA	0
Proline	CCC	0	CCC	0
Threonine	ACG	4		4
Total		30		17

Percentage rare codons for Construct 2: 11.5%

Percentage remaining uncompensated rare codons for construct 2: 6.5%

Consecutive Rare Codons

GGG GGG = 1

6.4.2 SDS-PAGE analysis of *E. coli* Bl21-pRARE TFI-scFv expression

The effect of rare codon compensation through the co-expression of pRARE on pET22 construct 2 by *E. coli* BL21 (DE3) was evaluated by means of 12.5 % SDS-PAGE. The whole cell fractions of *E. coli* BL21 (DE3) expressing only pET22 construct 2 (Lane 1), a negative control expressing only empty pET22 plasmid (Lane 2), *E. coli* BL21 (DE3) co-expressing pRARE and pET22 construct 2 (Lane 3) and a second negative control expressing pRARE and empty pET22 (Lane 4) were compared as indicated in Figure 6.1.
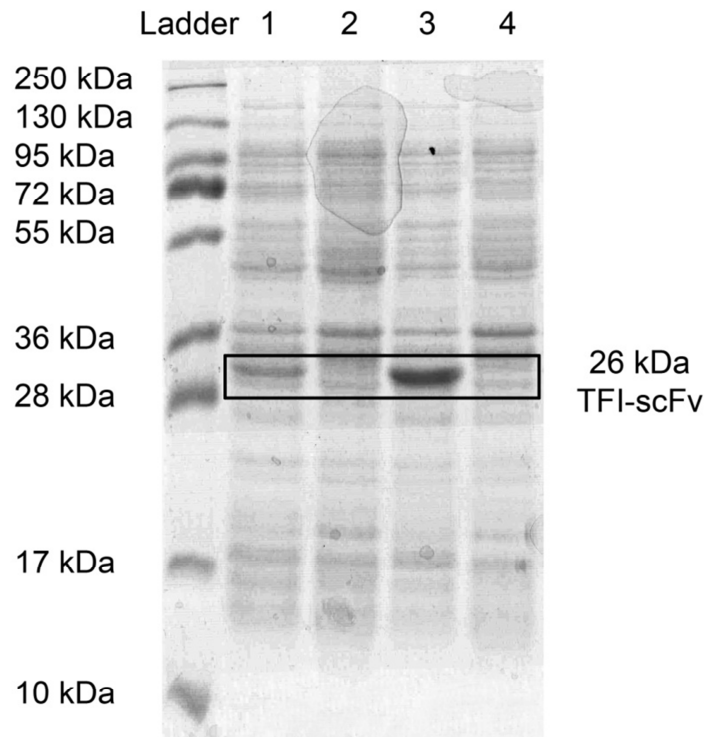
Figure 6.1: SDS-PAGE analysis of TFI-scFv and pRARE co-expression

The presence of 26 KDa TFI-scFv is indicated by frame in lane 1 and 3. The 26 KDa TFI-scFv is absent in both negative control in lane 2 and 4.

The co-expression of pET22 construct 2 and pRARE (Lane 3) yielded more TFI-scFv than the unaccompanied expression of pET22 construct 2 (Lane 1) indicating that the expression of the pRARE plasmid compensated for the rare codon requirements of TFI-scFv and thereby improved expression.

6.4.3 Immobilised Metal Affinity chromatography

The nickel affinity chromatography purification profile of TFI-scFv expressed by *E. coli* BL21 harbouring the pRARE plasmid is indicated in Figure 6.2. The column was washed with binding buffer in order to remove unbound protein between points A and B. TFI-scFv constructs were eluted using imidazole gradient (20 mM to 500 mM) introduced between point B and C.
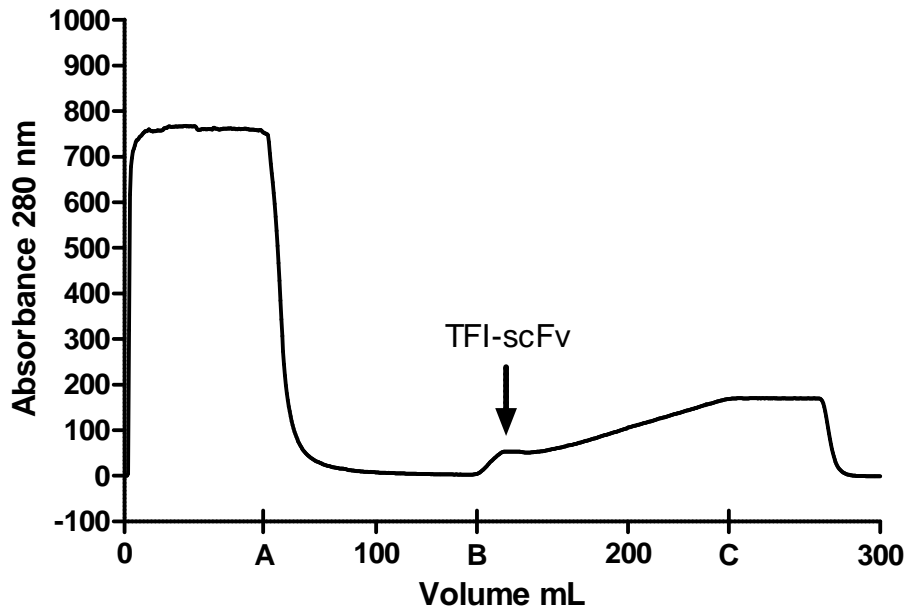
Figure 6.2: Purification profile for TFI-scFv expressed by _E. coli_ Bl21-pRARE

Nickel affinity purification profiles for TFI-scFv constructs obtained by continuously monitoring (0.333 µL increments) the absorbance at 280 nm of the eluate with the Äkta Prime Plus System. TFI-scFv peak is indicated by the arrow.

The protein peak, although minor, obtained during the initial stages of the imidazole gradient confirms that TFI-scFv was successfully purified by nickel affinity chromatography. The eluate fractions corresponding to the TFI-scFv protein peak were pooled together and concentrated.

6.4.4 Size exclusion chromatography

The purified TFI-scFv obtained after nickel affinity chromatography was subjected to size exclusion chromatography using a Sephacryl S200HR column in order desalt and remove imidazole from the isolated sample as indicated in Figure 6.3.

141
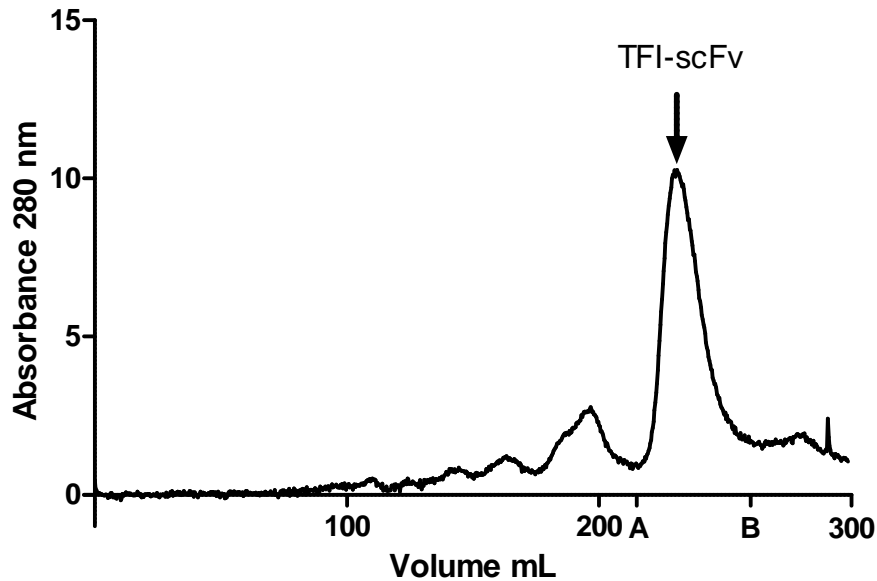
Figure 6.3: Size-exclusion chromatography purification profile

Size-exclusion chromatography purification of TFI-scFv construct 2 obtained by continuously monitoring (0.333 µL increments). The absorbance at 280 nm against volume of eluate is automatically measured and plotted by the Äkta Prime Plus System. TFI-scFv peaks are indicated by arrow.

The Fractions between point A and B corresponding to the TFI-scFv peak were collected and concentrated using the Amicon stirrer cell with TBS buffer.

6.4.5 Prothrombin Time coagulation analysis

The functionality and inhibitory effect of escalating dosages of the purified TFI-scFv on human platelet poor plasma was assessed with the Prothrombin Time coagulation test. All tests were performed in triplicate and the time to coagulation was recorded. Final time to coagulation at the various concentrations of TFI-scFv (points A, B, C, D, and E) was calculated as the average of the three recorded values (See: Table 6.6). The extension in coagulation time and normal range for the thromborel reagent is indicated in Figure 6.4.

Table 6.6: Prothrombin Time coagulation test

Reaction	TFI-scFv#	Round 1*	Round 2*	Round 3*	Average*	Time Extension*
A	0.0000	12.30	14.80	13.40	13.50	0.00
B	0.0141	14.80	13.90	13.10	13.90	0.40
C	0.0283	17.60	17.30	16.20	17.00	3.50
D	0.0565	26.70	28.40	29.30	28.13	14.63
E	0.1130	31.60	33.70	34.20	33.83	20.33

- Concentration measured in $\mu g.mL^{-1}$

* - Time measured in seconds
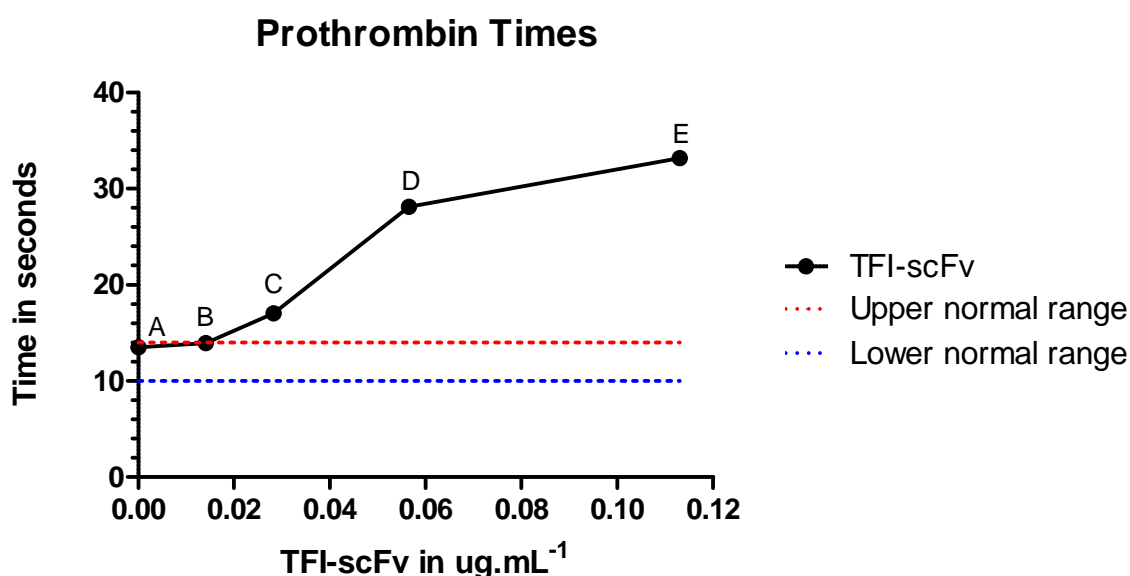
Figure 6.4: Inhibition effect of TFI-scFv

Averages for data points (A to E) were calculated and plotted as time against TFI concentration using GraphPad Prism v 5.0.3. The upper- (▪▪▪) and lower normal (▪▪▪) limits for normal values of PT coagulation test using Thromborel reagent are indicated.

The dose dependent extension of the time to coagulation indicates that the pET22 construct 2 co-expressed with the pRARE plasmid in the cytoplasm of *E. coli* BL21 was functional and capable of inhibiting coagulation.

6.4.6 GeneART® rare codon Analysis

The DNA sequence of pET22 construct 2 was sent to GeneART® (Germany) for rare codon analysis and optimization. Codon usage was adapted to the codon bias of *E. coli* genes. The TFIopt-scFv gene was synthesised to avoid regions of very high (>80 %) and very low (<30 %) GC content. Additionally, the cis-acting sequence motifs

that may have hampered expression such as internal TATA boxes, Chi-sites, RNA instability motifs and RNA secondary structures were avoided where possible. The pTFIopt construct as optimised by GeneART® is indicated in Figure 6.5.
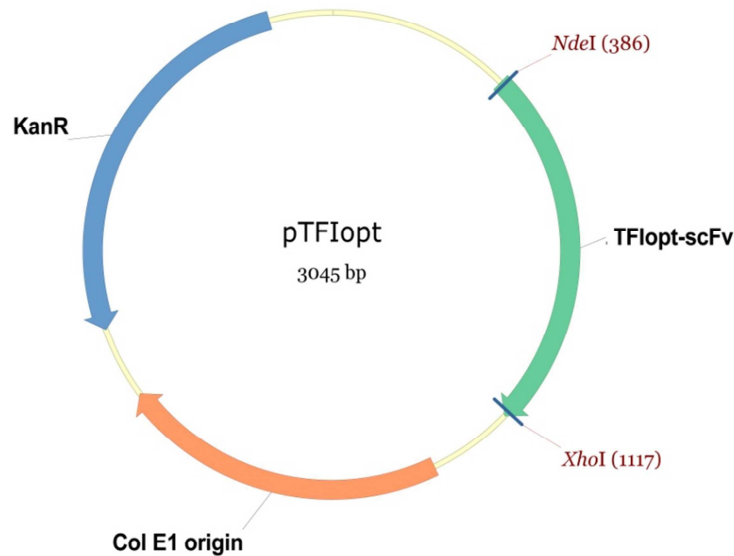
Figure 6.5: Structure of pTFIopt plasmid synthesised by GeneART®

The pTFIopt plasmid as synthesised by GeneART® containing the 741 bp rare codon optimised sequence (➡) for expression in *E. coli*. The TFIopt-scFv gene incorporates a 5' *Nde*I- and 3' *Xho*I sticky end was incorporated for downstream cloning indicated by (I). Kanamycin resistance marker is indicated by (➡).

6.4.7 Double digestion of pET22-TFIopt plasmid

The pET22-TFIopt plasmid was subjected to a double restriction digest using *Nde*I and *Xho*I restriction enzymes. Digestion products were separated using 0.8 % agarose gel electrophoresis and visualised with ethidium bromide in order to confirm the successful incorporation of the TFIopt-scFv gene into pET22b (+) expression vector. The pET22-TFIopt digestion products were individually loaded in lanes 1 to 5 and analysed for the presence of 761 bp as indicated in Figure 6.6.
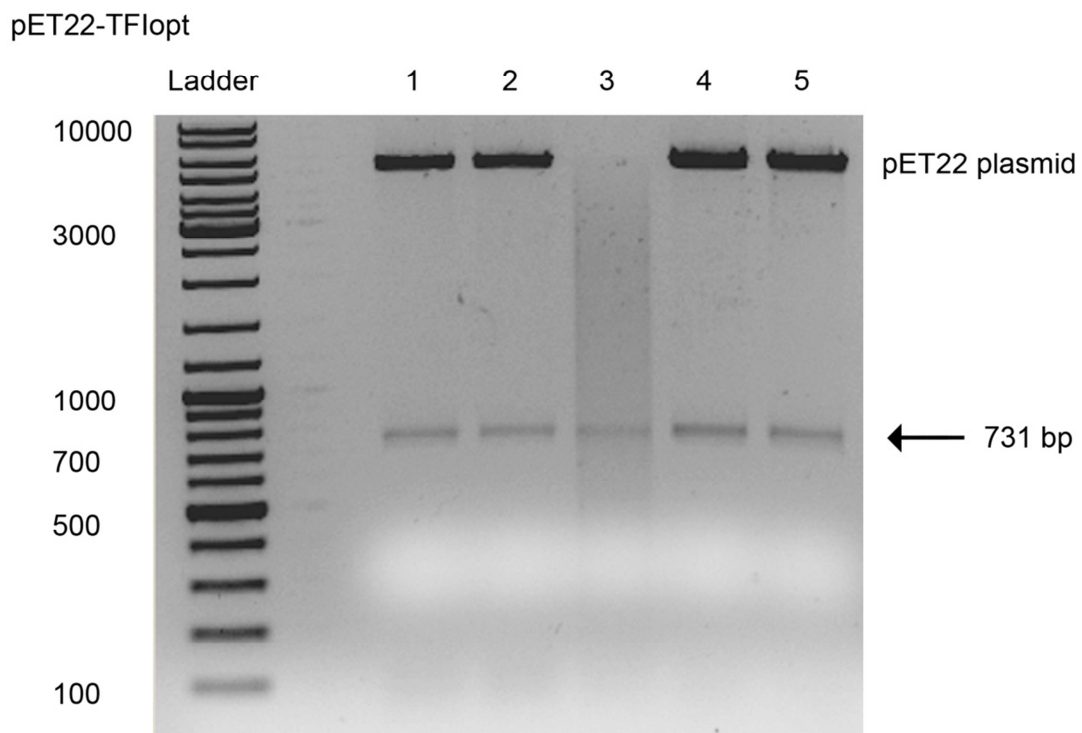
Figure 6.6: Double digestion of pET22-TFIopt plasmid

A 0.8% agarose gel electrophoresis of pET22-TFIopt digestion products The desired 761 bp optimised gene is indicated by the arrow.

The agarose gel electrophoresis confirmed the presence of the linear pET22 plasmid as well as the 761 bp TFIopt gene insert in lane 1, 2 4 and 5 of colonies selected for screening. Although the 761 bp TFIopt insert is clearly visible in lane 3, the lack of the pET22 band (approximately 5000 bp) with the presence of a smear indicates the degradation of the linear pET22 plasmid.

6.4.8 TFI-scFv Sequencing

A global alignment (Lalign ExPASy) between pET22 construct 2 and pTFIopt was performed and the IgG heavy- and light chain inserts identified. Internet based Lalign alignment (http://www.ch.embnet.org/software/LALIGN_form.html) results are presented in Table 6.7. The restriction sites (▬) as well as the IgG heavy- (RED) and light chain (BLACK) inserts present in both plasmids are indicated.

145

Table 6.7: pET22 construct 2 and pTFlopt global alignment

```
Readseq version 2.1.30 (12-May-2010) Readseq version 2.1.30 (12-May-2010)
        lalign output for pET22Construct2 vs. pTFIopt
        [ISREC-Server] Date: Wed Oct 12 17:22:18 2011
        ─────────────────────────────────────────────────────────
        ALIGN calculates a global alignment of two sequences version 2.2u
        Please cite: Myers and Miller, CABIOS (1989) 4:11-17
        pET22Construct2 942 bp vs. pTFIopt 942 bp
79.0% identity in 942 nt overlap
Global score: 2928


            10        20        30        40        50        60
pET22C TAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
pTFIop TAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTT
            10        20        30        40        50        60

                                      NdeI
            70        80        90       100       110       120
pET22C TGTTTAACTTTAAGAAGGAGATATACATATGGCCGAGGTGCAGCTGTTGGAGTCTGGGGG
       :::::::::::::::::::::::::::::::::::::::: ::  :: :::::: ::::    :: ::
pTFIop TGTTTAACTTTAAGAAGGAGATATACATATGGCAGAAGTTCAGCTGCTGGAAAGCGGTGG
            70        80        90       100       110       120

           130       140       150       160       170       180
pET22C AGGCTTGGTACAGCCTGGGGGGGTCCCTGAGACTCTCCTGTGCAGCCTCTGGATTCACCTT
        ::  :::: :::::::: ::     :::: : ::   ::::::::::      :: :: :::::
pTFIop TGGTCTGGTGCAGCCTGGTGGTAGCCTGCGTCTGAGCTGTGCAGCAAGCGGTTTTACCTT
           130       140       150       160       170       180

           190       200       210       220       230       240
pET22C TAGCAGCTATGCCATGAGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGGTCTC
       ::::::::::: :::::::::: :: ::::: :: :: :: :: ::::: :::::
pTFIop TAGCAGCTATGCAATGAGCTGGGTTCGTCAGGCACCGGGTAAAGGTCTGGAATGGGTTAG
           190       200       210       220       230       240

           250       260       270       280       290       300
pET22C ATCTATTAATCCGTTGGGTTGGAAGACACGTTACGCAGACTCCGTGAAGGGCCGGTTCAC
        ::::::::: :::::::::: :: ::::: :::::    ::: :: :: :: :: :: ::
pTFIop CAGCATTAATCCGCTGGGTTGGAAAACCCGTTATGCAGATAGCGTTAAAGGTCGTTTTAC
           250       260       270       280       290       300

           310       320       330       340       350       360
pET22C CATCTCCAGAGACAATTCCAAGAACACGCTGTATCTGCAAATGAACAGCCTGAGAGCCGA
       :::     : : ::::::   :::  ::::: :::::::::: ::::: ::::::  : :: ::
pTFIop CATTAGCCGTGACAATAGCAAAAACACCCTGTATCTGCAGATGAATAGCCTGCGTGCAGA
           310       320       330       340       350       360

           370       380       390       400       410       420
pET22C GGACACGGCCGTATATTACTGTGCGAAAACTTCGTCTAGGTTTGACTACTGGGGCCAGGG
       :: :: :: :: ::::: ::::: :::::          :  :::: :: ::::: :::::
pTFIop AGATACCGCAGTTTATTATTGTGCAAAAACCAGCAGCCGCTTTGATTATTGGGGTCAGGG
           370       380       390       400       410       420

                                            Linker Sequence
           430       440       450       460       470       480
pET22C AACCCTGGTCACCGTCTCGAGCGGTGGAGGCGGTTCAGGCGGAGGTGGCAGCGGCGGTGG
       :::::::: :::::          :::::  :: :::    :: :: :::::      :: :::::
pTFIop CACCCTGGTTACCGTTAGCTCAGGTGGTGGTGGTAGCGGTGGCGGTGGTTCTGGTGGTGG
           430       440       450       460       470       480
```

146

```
                490       500       510       520       530       540
pET22C  CGGGTCGACGGACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGA
        :::    ::  ::  :: ::::::::::::  ::    :  ::::    :::    :: :: ::
pTFIop  CGGTAGTACCGATATTCAGATGACCCAGAGCCCGAGCAGCCTGAGCGCAAGCGTTGGTGA
                490       500       510       520       530       540


                550       560       570       580       590       600
pET22C  CAGAGTCACCATCACTTGCCGGGCAAGTCAGAGCATTAGCAGCTATTTAAATTGGTATCA
         :  ::  ::::: ::  :: :: ::::: ::::::::::::: ::: :  :::::::::::
pTFIop  TCGTGTTACCATTACCTGTCGTGCAAGCCAGAGCATTAGCAGTTATCTGAATTGGTATCA
                550       560       570       580       590       600


                610       620       630       640       650       660
pET22C  GCAGAAACCAGGGAAAGCCCCTAAGCTCCTGATCTATGCTGCATCCAGTTTGCAAAGTGG
        :::::::::: ::  ::::: :: :: :: ::::: ::::: :::    :::   :::: :: ::
pTFIop  GCAGAAACCTGGTAAAGCACCGAAACTGCTGATTTATGCAGCAAGCAGCCTGCAGAGCGG
                610       620       630       640       650       660


                670       680       690       700       710       720
pET22C  GGTCCCATCAAGGTTCAGTGGCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAG
         :: ::        : :: :: :: :: ::     ::: :: ::::: :: :: ::::: :::::
pTFIop  TGTTCCGAGCCGTTTTAGCGGTAGCGGTAGTGGCACCGATTTTACCCTGACCATTAGCAG
                670       680       690       700       710       720


                730       740       750       760       770       780
pET22C  TCTGCAACCTGAAGATTTTGCAACTTACTACTGTCAACAGAGTTACAGTACCCCTAATAC
        :::::: ::  :::::::::::::::  :: :: ::::: ::::: :: :: ::::: :::::
pTFIop  TCTGCAGCCGGAAGATTTTGCAACCTATTATTGTCAGCAGAGCTATAGCACCCCGAATAC
                730       740       750       760       770       780


                790       800       810       820       830       840
pET22C  GTTCGGCCAAGGGACCAAGGTGGAAATCAAACGGAATTCGAGCTCCGTCGACAAGCTTGC
         :: ::::: ::  ::::: :: ::::: ::::: :::::::::::::::::::::::::::::
pTFIop  CTTTGGCCAGGGCACCAAAGTTGAAATTAAACGTAATTCGAGCTCCGTCGACAAGCTTGC
                790       800       810
                XhoI          His-tag
                850       860       870       880       890       900
pET22C  GGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAA
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
pTFIop  GGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAA
                850       860       870       880       890       900


                910       920       930       940
pET22C  GGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGC
        ::::::::::::::::::::::::::::::::::::::::::::
pTFIop  GGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGC
                910       920       930       940
```

147

The relevant genetic motifs present in the pET22-TFIopt sequence were identified in Vector NTI v10.0 and are indicated in Figure 6.7.
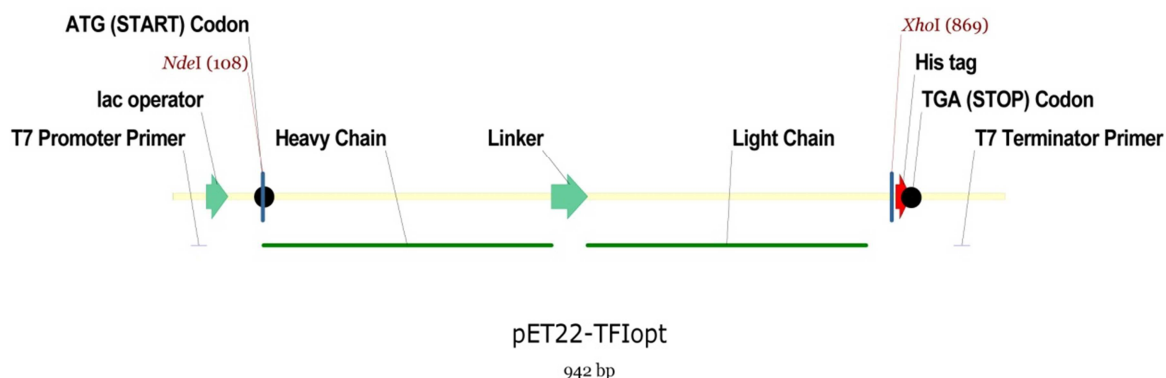
Figure 6.7: Structure of pET22-TFIopt

Finalised pET22-TFIopt DNA sequence consisting of 942 bp. The T7 promoter and terminator primer binding sites are indicated as (━). Human immunoglobulin G heavy- and light chain inserts are indicated as (━). The lac operator and glycine-serine linker repeat sequence (Linker) are indicated by (➡). The polyhistidine sequences (His-Tag) are indicated as (➡). Position of start and stop codons with a (●). Restriction sites (*Nde*I and *Xho*I) used during cloning are indicated as (|).

The compiled sequence was analysed for the presence of rare codons for expression in *E. coli* using internet based Rare Codon Calculator (http://people.mbi.ucla.edu/sumchan/caltor.html). The various rare codons within the finalised DNA sequence are highlighted and quantities summarised in Table 6.8.

Table 6.8: Rare Codon analysis of TFIopt-scFv

```
TFIopt-scFv        261 Amino Acids
ATGGCAGAAGTTCAGCTGCTGGAAAGCGGTGGTGGTCTGGTGCAGCCTGGTGGTAGCCTGCGTCTGAG
CTGTGCAGCAAGCGGTTTTACCTTTAGCAGCTATGCAATGAGCTGGGTTCGTCAGGCACCGGGTAAAG
GTCTGGAATGGGTTAGCAGCATTAATCCGCTGGGTTGGAAAACCCGTTATGCAGATAGCGTTAAAGGT
CGTTTTACCATTAGCCGTGACAATAGCAAAAACACCCTGTATCTGCAGATGAATAGCCTGCGTGCAGA
AGATACCGCAGTTTATTATTGTGCAAAAACCAGCAGCCGCTTTGATTATTGGGGTCAGGGCACCCTGG
TTACCGTTAGCTCAGGTGGTGGTGGTAGCGGTGGCGGTGGTTCTGGTGGTGGCGGTAGTACCGATATT
CAGATGACCCAGAGCCCGAGCAGCCTGAGCGCAAGCGTTGGTGATCGTGTTACCATTACCTGTCGTGC
AAGCCAGAGCATTAGCAGTTATCTGAATTGGTATCAGCAGAAACCTGGTAAAGCACCGAAACTGCTGA
TTTATGCAGCAAGCAGCCTGCAGAGCGGTGTTCCGAGCCGTTTTAGCGGTAGCGGTAGTGGCACCGAT
TTTACCCTGACCATTAGCAGTCTGCAGCCGGAAGATTTTGCAACCTATTATTGTCAGCAGAGCTATAG
CACCCCGAATACCTTTGGCCAGGGCACCAAAGTTGAAATTAAACGTAATTCGAGCTCCGTCGACAAGC
TTGCGGCCGCACTCGAGCACCACCACCACCACCACTGA
```

148

```
Summary of Rare Codon composition


Percentage Rare codons for TFIopt-scFv : 0%
```

Amino Acid	Rare Codon	Frequency of Occurrence
Arginine	AGA	0
	CGG	0
	AGG	0
	CGA	0
Glycine	GGG	0
	GGA	0
Isoleucine	AUA	0
Leucine	CUA	0
Proline	CCC	0
Threonine	ACG	0
Total		0

```
Consecutive Rare Codons
None
```

Analysis for the pET22-TFIopt plasmid indicated that all rare codons were removed by GeneART® sequence optimisation for expression in *E. coli*.

The pET22-TFIopt DNA sequence was subjected to a six frame translation using internet based ExPASy Translate Tool (http://web.expasy.org/translate/) and compared with pET22 construct 2 (See: Section 5.4.6). An open reading frame was identified, translating to a 261 amino acid sequence. An alignment between the pET22 construct 2 and pET22-TFIopt plasmid was performed using internet bases ClustalW alignment software (http://www.genome.jp/tools/clustalw/) as presented in Table 6.9.

Table 6.9: Amino Acid alignment of TFI-scFv constructs

```
CLUSTAL 2.1 Multiple Sequence Alignments

pET22_2: pET22 Construct 2      261 aa
pET22_T: pET22-TFIopt           261 aa


pET22_2     MAEVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSSINPLGWKT
pET22_T     MAEVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSSINPLGWKT
            ************************************************************


pET22_2     RYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAKTSSRFDYWGQGTLVTVSSGG
pET22_T     RYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAKTSSRFDYWGQGTLVTVSSGG
            ************************************************************


pET22_2     GGSGGGGSGGGGSTDIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKL
pET22_T     GGSGGGGSGGGGSTDIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKL
            ************************************************************


pET22_2     LIYAASSLQSGVPSRFSGSGSGSTDFTLTISSLQPEDFATYYCQQSYSTPNTFGQGTKVEI
pET22_T     LIYAASSLQSGVPSRFSGSGSGSTDFTLTISSLQPEDFATYYCQQSYSTPNTFGQGTKVEI
            ************************************************************


pET22_2     KRNSSSVDKLAAALEHHHHHH
pET22_T     KRNSSSVDKLAAALEHHHHHH
            ********************
```

Alignment of the two constructs clearly indicated that the expressed TFI-scFv is identical despite having dissimilarity in the DNA sequences the expressed protein remained unchanged.

6.4.9 SDS-PAGE analysis of pET22-TFIopt expression in *E. coli* BL21

The cellular fractions of *E. coli* BL21 expressing the pET22-TFIopt plasmid and the negative control expressing empty pET22 plasmid were collected and subjected to 12.5 % SDS-PAGE in order to confirm the cytoplasmic expression as well as analyse the level of expression of 26 KDa TFIopt-scFv. Whole cell fractions of the negative control and *E. coli* BL21 expression TFIopt-scFv were loaded in lane 1 and 2 respectively. Cellular debris of the negative control and *E. coli* BL21 expressing TFIopt-scFv obtained after cell disruption were loaded in to lane 3 and 4 respectively. Lane 6 and 7 were loaded with the negative control and *E. coli* BL21 expressing TFIopt-scFv insoluble fraction obtained after ultra-centrifugation respectively. The soluble fractions of the negative control and *E. coli* BL21

expression TFIopt-scFv were loaded in lane 7 and 8 respectively as indicated in Figure 6.8.
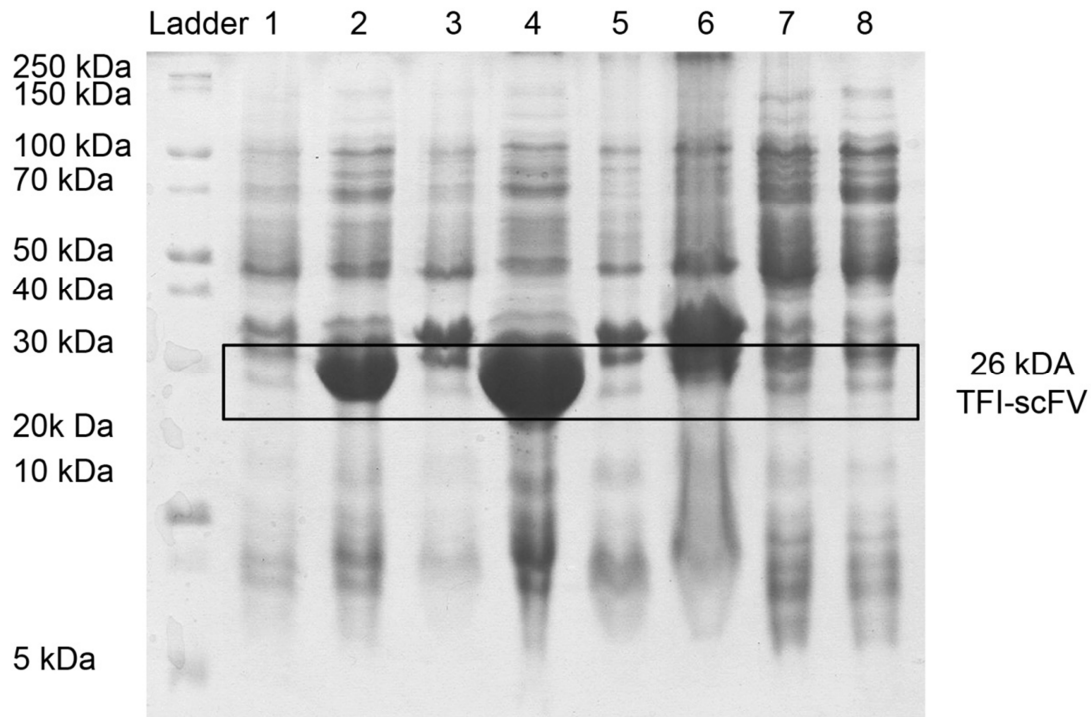
Figure 6.8: Whole cell SDS-PAGE analysis

The SDS-PAGE analysis of TFIopt-scFv expression by *E. coli* BL21. The desired fragment size of 26 kDa is indicated by the frame. TFI-scFv is present in lane 2 and 4 but is not present in lane 6 and 8 of the *E. coli* BL21 transformed with the pET22-TFIopt plasmid. The 26 TFI-scFv is absent in the control 1, 3, 5 and 7 of the negative control transformed with empty pET22 plasmid.

The presence of the 26 kDa fragment in *E. coli* BL21 whole cell fraction harbouring the pET22-TFIopt plasmid (lane 2) combined with the absence of this fragment in the negative control (Lane 1) confirmed the successful expression of the TFIopt-scFv gene. The extent and band intensity of TFIopt-scFv in lane 2 and 4 in comparison with the constitutively expressed bacterial proteins as well as the negative control indicated that the optimized TFI-scFv sequence was expressed at a high level. The insoluble fraction loaded (Lane 7 and 8) did not contain the 26 kDa TFi-scFv band. The soluble fraction of the negative control (Lane 7) and the *E. coli* BL21 expressing the pET22-TFIopt plasmid (Lane 8) were virtually identical. Both fractions lack the 26 kDa TFI-scFv band.

151

6.5 Discussion

6.5.1 Rare codon compensation by pRARE

Rare codon analysis of the pET22 construct 2 was performed in order to examine if the co-expression of the pRARE plasmid would improve the overall TFI-scFv yield. The pRARE plasmid only supplies tRNAs for AGG, AGA, AUA, CUA, CCC, and GGA rare codons that occur in *E. coli*. Of this specific set of tRNAs the pET22 construct 2 contain 13 rare codons in total that are compensated for. These codons include 6 instances of rare Arginine codons (2 instances of AGG and 4 instances of AGA) and 7 instances of the rare Glycine codon (GGA). The pET22 construct 2 do not contain any rare Isoleucine (AUA), Leucine (CUA) and Proline (CCC) codons. A total of 17 rare codons present in the pET22 construct 2 sequence are not compensated for by the pRARE plasmid. These include 3 instances of the rare Arginine codon (CGG), 10 instances of the rare Glycine codon (GGG) that include one consecutive repeat relatively near the N-terminal and 4 instances of the rare Threonine codon (AGG). The uncompensated rare codons constitute 6.5 % of the total 261 amino acids required for the translation of TFI-scfv. The effect of rare codon compensation on the expression rate and ultimately the total yield of the target protein is difficult to predict as there are many variables to consider. The rate of recombinant protein expression by the expression host is inherent to the nature of the host strain, the expression vector used, mRNA stability, and a variety of environmental factors such as temperature and culture media compensation (Studier, 2005; Sørensen and Mortensen, 2005)**.** Apart from these factors, the extent to which rare codon compensation can further improve target protein expression is dependent on several factors. This include the specific set of rare codons that are compensated for, the expression rate or availability of the rare tRNAs as well as the requirements for these rare tRNAs by the expressed target protein during translation. It is therefore important to remember that although rare codon compensation analysis presents a clear picture as to which rare codons are compensated for, their frequency of occurrence and position within the target sequence does not answer the above mentioned questions. As a result the rare codon compensation analysis should not be seen as an absolute compensation. It is

important to keep in mind that the rare codons that are compensated for are still present within the target sequence and may influence the expression rate.

6.5.2 The co-expression of pET22 construct 2 and pRARE

The whole cell fractions of *E. coli* BL21 (DE3) harbouring the pRARE plasmid co-expressed with pET22 construct 2, the negative control harbouring the pRARE and empty pET22 plasmid, *E. coli* BL21 (DE3) expression the pET22 construct 2 and the negative control harbouring empty pET22 were normalised and subjected to 12.5 % SDS-PAGE in order to examine the effect of rare codon compensation on the expression levels of TFI-scFv (See: Figure 6.1). The SDS-PAGE analysis confirmed the successful expression of the pET22 construct 2 by both *E. coli* BL21 cultures without (lane 1) and with (lane 2) the co-expression of the pRARE plasmid. The expressed 26 KDa TFI-scFv falsely appears to be of greater size of approximately 30 to 32 KDa due to the increased motility of the pre-stained protein ladders. This accounts for the approximate 5 KDa discrepancy in the expected TFI-scFv size. The 26 KDa TFI-scFv was not present in the negative control that contained empty pET22 plasmid. The absence of the 26 KDa band in the negative controls confirms that the additional 26 KDa band observed in the lane 1 and 3 is that of TFI-scFv and not of another constitutively expressed protein of similar size.

The expression of pET22 constructs without rare codon compensation (Lane 1) produced a 26 kDa band that was expressed slightly higher than the constitutive proteins. It is clear that the over expression of the pET22 construct 2 did not occur as was expected based on results previously obtained (See: Section 5.4.4). The co-expression of pET22 construct2 and pRARE (Lane 3) produced TFI-scFv which was clearly expressed at a higher rate than the constitutive proteins. By comparing the TFI-scFv band intensities of lane 1 and lane 3 it is clear that the co-expression of pRARE plasmid improved the expression of TFI-scFv. Although the expression was improved the over expression of TFI-scFv did not occur. This is most likely due to the presence of the uncompensated rare codons. Through the co-expression of pRARE plasmid only 13 of the 30 rare codons present in the TFI-scFv sequence were compensated for. The presence of 17 remaining rare codons is still detrimental to the high expression level of TFI-scFv.

6.5.3 Chromatography purification of TFI-scFv

Approximately 6 g (wet weight) *E. coli* BL21 (DE3) co-expressing the pET22 construct and pRARE was harvested from auto-induction media (1 L) utilized during the expression. The cytoplasmic fraction was isolated and the TFI-scFv purified by means of nickel affinity chromatography using HIS-Trap^{TM} FF column and the ÄKTA Prime Plus system. The linear increase in absorbance between points B and C in the purification profile was due to the absorbance that was presented by imidazole at 280 nm used to elute the target protein. A clearly visible protein peak in the early stages of the imidazole gradient was obtained. The fractions corresponding to the TFI-scFv peak (135 ml to 170 ml) were collected. The purification of TFI-scFv by means of nickel affinity chromatography is an improvement on previous purification attempts (See: Section 5.4.5). This improvement is most likely due to the higher level of TFI-scFv expression achieved through the co-expression of pRARE.

Although purely speculative, is seems that the level of expression observed (See: Section 6.5.2) does not correlate with the amount of TFI-scFv purified by means of nickel affinity chromatography. This may be due to an unsuccessful interaction of TFI-scFv with the immobilised nickel on the column which results in the TFI-scFv simply passing through the column. Another possibility might be that a portion of the expressed TFI-scFv within the bacterial cytoplasm is insoluble due to formation of protein aggregates known as inclusion bodies. The scFvs expressed in *E. coli* can be expressed in high levels but are notorious inclusion bodies producers (Tsumoto *et al.,* 1998; Gu *et al.,* 2002; Fursova *et al.,* 2009).

The fractions corresponding to the TFI-scFv peak were collected and subjected to size exclusion chromatography as an additional "polishing" step in order to desalt the isolated sample. The fractions corresponding to the only major protein peak (210 ml to 260 ml) were collected, a buffer exchange completed and finally concentrated.

6.5.4 Prothrombin time coagulation analysis

The final concentration of TFI-scFv was measured using the GeneQuant pro RNA/DNA calculator at 280 nm (AEC-Amersham, UK) and adjusted to match the

TFI-scFv concentration of the PT performed previously (See: Section 3.4.3) for ease of comparison. A modified PT coagulation test was performed in order to detect the inhibition effect of TFI-scFv at low levels.

Point A in Figure 6.4 served as the baseline and positive control for the reaction as no TFI-scFv was present. Modification to the standard PT coagulation test resulted in a baseline coagulation time of 13.5 seconds. This modification was done in order to improve the sensitivity of the test which is only 0.5 seconds below the upper limit for normal patients (10 to 14 seconds) at our routine haematology laboratory (National Health Laboratory services, Free State, Bloemfontein, RSA). At point B the baseline coagulation time was extended by an additional 0.4 seconds when 0.0141 $\mu g.mL^{-1}$ TFI-scFv was present. It is clear that TFI-scFv had very little inhibition effect at these low concentrations. At point C the inhibition effect of TFI-scFv was increasingly visible as the baseline coagulation time was extended by 3.5 seconds at a TFI-scFv concentration of 0.0283 $\mu g.mL^{-1}$. At point D, when a concentration of 0.0565 $\mu g.mL^{-1}$ TFI-scFv was present within the reaction the baseline coagulation time was extended by 14.6 seconds. At point E the baseline coagulation time was extended with 20.33 seconds when the reaction showed a plateau with 0.113 $\mu g.mL^{-1}$ TFI-scFv. Coagulation at point E was only extended with 5.7 seconds longer than the inhibition observed in point D, where only approximately half (0.0565 $\mu g.mL^{-1}$) TFI-scFv was present.

These results closely resemble that of the previous PT coagulation test performed with the original pIT2 construct (See: Section 3.4.3). As with the previous PT coagulation test, the inhibition effect of TFI-scFv was small at very low concentrations and only became clear at approximately 0.024 $\mu g.ml^{-1}$ TFI-scFv. Additionally, a clear plateau was again reached at approximately 0.05 $\mu g.mL^{-1}$ TFI-scFv. The PT coagulation test confirmed that the TFI-scFv produced through the co-expression of pET22 construct 2 and pRARE by *E. coli* Bl21 (DE3) was functional and able to hamper coagulation through the inhibition of tissue factor. The fact that pET22 construct 2 displayed similar characteristics as the original pIT2 construct confirms that the modification to the expression vectors, as performed in the previous chapter, did not alter TFI-scFv ability to inhibit tissue factor. Additionally it

also confirmed that the protein peaks isolated during the nickel affinity- and the size exclusion chromatography was indeed that of TFI-scFv.

6.5.5 Double digestion of pET22-TFIopt plasmid

Plasmids were isolated from 5 *E. coli* BL21 cultures and subjected to double digestion using *Nde*I and *Xho*I in order to confirm the presence of the TFIopt-scFv plasmid. All of the cultures were positive for linearised pET22b (+) plasmid (approximately 5000 bp) as well as the rare codon optimised TFIopt -scFv sequence (761 bp). The only exception was that of the culture analysed in lane 3. Although the TFIopt-scFv band was visible, the band corresponding to the larger pET22 plasmid was missing. The presence of a clearly visible smear in lane 3 suggests that DNA degradation occurred. The most likely cause for the degradation of the DNA sample may be due to DNase contamination. If the sample was contaminated with DNase, the pET22 plasmid as well as the TFIopt gene would have been randomly cleaved into smaller fragments of various sizes. It is this collection of fragments with various lengths that was responsible for the "smeared" profile observed in lane 3 when separated on a agarose gel. As a result the plasmid analysed in lane 3 was discarded in order to prevent any additional complications.

6.5.6 DNA analysis of the pET22-TFIopt construct

The pET22-TFIopt construct was sequenced and aligned against pET22 Construct 2 DNA sequence (See: Section 5.4.6) in order to identify the position of the heavy- and light chain IgG inserts. Additionally the obtained pET22-TFIopt sequence was analysed and the relevant genetic motives were identified.

The pET22-TFIopt construct was sequenced using the same T7 promoter and T7 terminator primers used for the sequencing of pET22 Construct 2 in the previous chapter. Bidirectional sequencing produced a 942 bp DNA fragment that shared only 79 % similarity with the original pET22 Construct 2 DNA sequence of the same size. The position of the IgG inserts, linker sequence as well as the His-tag were identified through the alignment of the two constructs (See: Table 6.7). By comparing the genetic motifs present in the pET22-TFIopt construct (See: Figure 6.7) with that of

pET22 Construct 2 (See: Figure 5.7) it is clear that the structure of the TFI-scFv gene remained unchanged even though the DNA sequence has been modified. The only apparent difference was the absence of an *EcoR*I restriction site at the 3' end of the light chain.

Rare Codon analysis of TFIopt-scFv gene was performed using the internet based Rare Codon Calculator (http://people.mbi.ucla.edu/sumchan/caltor.html). It is clear from the analysis that all rare codons have been replaced by major codons for expression in *E. coli.*

The similarity between pET22-TFIopt and pET22 construct 2 became even more apparent when the open reading frames were translated (See: Table 6.9). Translation of both constructs produced an identical 261 amino acid protein sequence with an in-frame hexa-histidine repeat at the C-terminus of the protein. It is important to remember that despite extensive modification of the TFI-scFv DNA sequence that the eventual amino acid sequence of the mature protein remained unchanged.

6.5.7 Cytoplasmic expression pET22-TFIopt

The SDS-PAGE analysis of the whole cell fractions (Lane 1 and 2) confirmed the successful over-expression of the TFI-scFv. It was clear that the optimised TFI-scFv sequence was expressed at a high rate due to the substantial amount of TFI-scFv produced by *E. coli* BL21 (DE3) as can been seen in lane 2. The high level of expression was even more apparent when examining the cellular debris (Lane 3 and 4). Apart from the presence of a large 26 kDa protein band, the band intensity was also considerably higher than that of the constitutively expressed bacterial proteins. Despite the high levels of expression it is clear that the TFI-scFv was not soluble as the 26 kDa band was absent in the soluble fraction (Lane 8). The insolubility was also confirmed by the presence of the TFI-scFv in the cellular debris fraction (Lane 4).

It is clear that although the rare codon optimisation of the TFI-scFv DNA sequence drastically improved the expression rate, it was also detrimental to the solubility of the antibody. The expression of Human recombinant proteins in *E. coli* often results in the production of insoluble aggregates (inclusion bodies) that accumulate within the host cell instead of a soluble native protein (Rudolph and Lilie, 1996; Peti and Page, 2007). The expression of insulin is a classic example of transgenic protein expression in *E. coli* resulting the production of inclusion bodies (Williams *et al.,* 1982). In reality most single-chain antibodies scFv form inclusion bodies when expressed in bacteria (Sinacola & Robinson, 2002). The formation of inclusion bodies has been observed during the overexpression of recombinant- as well as native proteins in various host systems. This suggest that inclusion body formation is a result of high expression rates, regardless of the expression system or protein used (Lilie *et al.,* 1998). In addition there is also accumulating evidence that the modulation of the translational speed influences folding events in the target protein (Lilie *et al.,* 1998; Cortazzo *et al.,* 2002; Rosano & Ceccarelli 2009). Native genes that contain a high percentage of rare codons are translated at a lower rate that genes with a low rare codon percentage (Widmann *et al.,* 2008). It has also been observed that rare codons have a tendency to occur in secondary protein structures such as turns, loops, and domain linkers (Cortazzo *et al.,* 2002). It is therefore believed that rare codons facilitate correct protein folding by causing ribosomal pausing at these sites. This allows the newly synthesised amino acid chain to adopt a required intermediate conformation necessary for eventual solubilisation (Rosano & Ceccarelli, 2009).

Although there are other factors which can also contribute to the formation of inclusion bodies it is clear that the formation of inclusion bodies in this case is most likely due to the rare codon optimisation of the TFI-scFv sequence. A similar conclusion was made by Rosano and Ceccarelli (2009) who found that a set of heterologous plant proteins with high rare codon contents became insoluble when overexpressed in a codon bias-adjusted *E. coli* strain. Similarly the synonymous codon substitution resulted in the removal of critical ribosomal stalling sites required for intermediate folding required for the solubilisation of the target protein. The key to obtaining the delicate balance between quantity and quality lay in the ability to identify non-critical rare codons that can be safely substituted with major codons

while retaining the critical rare codons required for ribosomal stalling (Widmann et al., 2008). Through the reintroduction of only the critical rare codons in to the pET22-TFIopt construct the expression rate will remain as high as possible but sufficiently reduced to facilitate correct folding of TFI-scFv in order to retain solubility.

Various other methods have been utilised in order to prevent the production of inclusion bodies. One of these methods utilised for the solubilisation of scFv is through the co-expression of molecular chaperones while using thioredoxin-deficient host strains in order to maintain a favourable redox potential (Hannig and Makrides, 1998; Jonasson et al., 2002; Hu et al., 2007). Molecular chaperones facilitate the correct folding by binding partially folded proteins and maintain them in a soluble conformation, thereby reducing the probability that two or more intermediates will interact with each other and lead to the production of inclusion bodies (Thomas et al., 1997). However, the general consensus regarding the co-expression of chaperones in *E. coli* is that the prevention of inclusion body formation appears to be protein specific (Makrides, 1996; Hannig and Makrides, 1998; Jonasson et al., 2002). An additional disadvantage, is that the chaperone system may become exhausted by the highly overexpressed target protein. This will ultimately result in the production of inclusion bodies regardless of the chaperones' ability to facilitate the correct folding of the target protein (Rudolph and Lilie, 1996).

The utilization of protein fusion systems has also been shown to facilitate the production of soluble recombinant proteins in *E. coli* (Hannig and Makrides, 1998). The target protein is fused to a highly soluble partner protein in order to improve the solubility of the target protein and thus in this way prevent the formation of inclusion bodies in the cytoplasm (Jonasson et al., 2002). The fusion partner can then be cleaved from the target protein with commercially available specific peptidases (Thomas et al., 1997). Proteins frequently used as solubilising fusion partners include: thioredoxin, ubiquitin, NusA, the IgG-binding domains ZZ from Staphylococcus *aureus*, albumin-binding protein, streptococcal protein, maltose-binding protein and glutathione S-transferase (Makrides, 1996; Thomas et al., 1997; Hannig and Makrides, 1998; Jonasson et al., 2002).

The *in vitro* folding of inclusion body proteins is an alternative method utilised to obtain soluble protein (Rudolph and Lilie, 1996; Lilie et al., 1998). The general strategy followed when attempting to refold inclusion body proteins includes three steps: isolation of inclusion bodies; solubilisation and the refolding of the solubilised protein (Lilie et al., 1998; Clark, 2001). The production of inclusion bodies can be advantageous as the target protein is highly enriched and protected against proteolysis by intracellular proteases due to their tightly aggregated nature (Clark, 2001). Additionally, inclusion bodies are easily isolated from soluble protein fraction through centrifugation, filtration or size-exclusion chromatography which simplifies the protein purification process (Li et al., 2004). Once the target protein is isolated the inclusion bodies are solubilised through the addition of denaturants such as guanidinium chloride or urea (Lilie et al., 1998). Extreme pH values in combination with high temperature have also been used to solubilise inclusion bodies (Clark, 2001). Once the inclusion bodies are solubilised the target protein is refolded through the removal denaturants by either dilution or by step wise exchange buffer to yield a soluble and functional target protein (Rudolph and Lilie, 1996; Lilie et al., 1998; Clark, 2001). Although the *in vitro* refolding of scFv targeting a variety of eptipopes has been reported, it still remains a complicated and generally expensive procedure requiring protein specific optimisation (Sanchez et al., 1999; Gu et al., 2002; Guo et al., 2006; Fursova et al., 2009; Tapryal et al., 2010; Kudou et al., 2011).

6.6 Conclusion

The expression level of TFI-scFv was improved by alleviating the detrimental effect of rare codons. This was achieved through the modification of the expression host as well as the optimisation of the TFI-scFv DNA sequence for expression in *E. coli*.

Even though the co-expression of the pRARE plasmid did not provide tRNAs for all the rare codons present within the TFI-scFv DNA sequence, the expression was considerably improved. More importantly, TFI-scFv was expressed as a soluble scFv and displayed a similar inhibition effect that the original pIT2 construct. Although functional TFI-scFv was produced through the co-expression of pET22 construct 2 and the pRARE plasmid, the level of expression is still insufficient for our eventual

goal to further characterisation of TFI-scFv in parallel plate flow chambers and animal models. Ultimately the co-expression of these plasmids would have to be up scaled to such an extent in order to provide a sufficient yield of TFI-scFv that purification would again become impractical and excessively labour intensive.

By synthesising a rare codon optimised sequence for expression in *E. coli,* the expression level TFI-scFv was even further improved. If one considers the substantial difference in expression level between the un- and optimised TFI-scFv sequences, the severity of the detrimental effect of rare codons on protein expression levels is quite remarkable. Unfortunately, despite the substantial improvement in final protein yield, the synonymous replacement of all rare codons in the TFI-scFv DNA sequence resulted in the production of insoluble aggregates known as inclusion bodies. This is most likely due to the removal of critical rare codons that provide ribosomal stalling sites that are required for the correct intermediate folding required for the solubilisation of TFI-scFv.

Although the loss of TFI-scFv solubility is not ideal, the high level of expression provides an ideal platform for the further development of TFI-scFv. The key to future success lies in the ability to solubilise the inclusion bodies either through the modification of the expression system or by the identification and reintroduction of only critical rare codons to the optimised TFI-scFv DNA sequence. Alternatively TFI-scFv inclusion bodies can be solubilised through the utilisation of a highly soluble fusion partner. An additional alternative that is often utilized for the solubilisation of inactive scFv is the *in vitro* refolding of inclusion bodies as they are easily isolated and protected against proteolysis. The large scale expression of the rare codon optimised TFI-scFv sequence in a soluble form would produce a sufficient quantity of TFI-scFv for characterization in animals.

6.7 References

BACA, A.M. & HOL, W.G. 2000. Overcoming codon bias: a method for high-level overexpression of Plasmodium and other AT-rich parasite genes in Escherichia coli. *International Journal for Parasitology*, 30, 113-8.

BURGESS-BROWN, N. A, SHARMA, S., SOBOTT, F., LOENARZ, C., OPPERMANN, U. & GILEADI, O. 2008. Codon optimization can improve expression of human genes in Escherichia coli: A multi-gene study. *Protein Expression and Purification*, 59, 94-102.

CLARK, E. 2001. Protein refolding for industrial processes. *Current Opinion in Biotechnology*, 12, 202–207.

CORTAZZO, P., CERVEÑANSKY, C., MARÍN, M., REISS, C., EHRLICH, R. & DEANA, A. 2002. Silent mutations affect in vivo protein folding in Escherichia coli. *Biochemical and Biophysical Research Communications*, 293, 537-41.

FURSOVA, K.K., LAMAN, A.G., MELNIK, B.S., SEMISOTNOV, G.V., KOPYLOV, P.K., KISELEVA, N.V., NESMEYANOV, V.A. & BROVKO, F.A. 2009. Refolding of scFv mini-antibodies using size-exclusion chromatography via arginine solution layer. *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences,* 877, 2045-51.

GU, Z., WEIDENHAUPT, M., IVANOVA, N., PAVLOV, M., XU, B., SU, Z.G. & JANSON, J.C. 2002. Chromatographic methods for the isolation of, and refolding of proteins from, Escherichia coli inclusion bodies. *Protein Expression and Purification*, 25, 174-9.

GUO, Z., BI, F., TANG, Y., ZHANG, J., YUAN, D., XIA, Z. & LIU, J.N. 2006. Preparation and characterization of scFv for affinity purification of reteplase. *Journal of Biochemical and Biophysical Methods*, 67, 27-36.

HANNIG, G. & MAKRIDES, S.C. 1998. Strategies for optimizing heterologous protein expression in Escherichia coli. *Trends in Biotechnology*, 16, 54-60.

HU, X., O'HARA, L., WHITE, S., MAGNER, E., KANE, M. & WALL, J.G. 2007. Optimisation of production of a domoic acid-binding scFv antibody fragment in Escherichia coli using molecular chaperones and functional immobilisation on a mesoporous silicate support. *Protein Expression and Purification*, 52, 194-201.

JONASSON, P., LILJEQVIST, S., NYGREN, P.-A. & STÅHL, S. 2002. Genetic design for facilitated production and recovery of recombinant proteins in Escherichia coli. *Biotechnology and Applied Biochemistry*, 35, 91-105.

KANE, J.F., PHARMACEUTICALS, S.B. & PRUSSIA, K. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. *Current Biology*, 8, 494-500.

KUDOU, M., EJIMA, D., SATO, H., YUMIOKA, R., ARAKAWA, T. & TSUMOTO, K. 2011. Refolding single-chain antibody (scFv) using lauroyl-L-glutamate as a solubilization detergent and arginine as a refolding additive. *Protein Expression and Purification,* 77, 68-74.

LI, M., SU, Z.G. & JANSON, J.C. 2004. In vitro protein refolding by chromatographic procedures. *Protein Expression and Purification,* 33, 1-10.

LILIE, H., SCHWARZ, E. & RUDOLPH, R., 1998. Advances in refolding of proteins produced in E. coli. *Current Opinion in Biotechnology*, 9, 497-501.

MAKRIDES, S.C. 1996. Strategies for achieving high-level expression of genes in Escherichia coli. *Microbiological Reviews*, 60, 512-38.

PETI, W. & PAGE, R. 2007. Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost. *Protein Expression and Purification*, 51, 1-10.

REDWAN, M. 2006. The optimal gene sequence for optimal protein expression in Escherichia coli: principle requirements. *Arab Journal of Biotechnology*, 9, 493-510.

ROCHE, E.D. & SAUER, R.T. 1999. SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. *The EMBO Journal*, 18, 4579-89.

ROSANO, G.L. & CECCARELLI, E. 2009. Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted Escherichia coli strain. *Microbial Cell Factories,* 8, 41.

RUDOLPH, R., LILIE, H. 1996. In vitro folding of inclusion body proteins. *The FASEB Journal*, 10, 49-56.

SANCHEZ, L., AYALA, M., FREYRE, F., PEDROSO, I., BELL, H., FALCÓN, V. & GAVILONDO, J.V. 1999. High cytoplasmic expression in E. coli, purification, and in vitro refolding of a single chain Fv antibody fragment against the hepatitis B surface antigen. *Journal of Biotechnology*, 72, 13–20.

SINACOLA, J.R. & ROBINSON, A.S. 2002. Rapid refolding and polishing of single-chain antibodies from Escherichia coli inclusion bodies. *Protein Expression and Purification*, 26, 301–308.

STUDIER, F.W. 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, 41, 207-234.

SØRENSEN, H.P. & MORTENSEN, K.K. 2005. Advanced genetic strategies for recombinant protein expression in Escherichia coli. *Journal of Biotechnology,* 115, 113-28.

TAPRYAL, S., KRISHNAN, L., BATRA, J.K., KAUR, K.J. & SALUNKE, D.M. 2010. Cloning, expression and efficient refolding of carbohydrate-peptide mimicry recognizing single chain antibody 2D10. *Protein Expression and Purification*, 72, 162-8.

THOMAS, J.G., AYLING, A. & BANEYX, F. 1997. Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from E. coli. To fold or to refold. *Applied Biochemistry and Biotechnology*, 66, 197-238.

TSUMOTO, K., SHINOKI, K., KONDO, H., UCHIKAWA, M., JUJI, T. & KUMAGAI, I. 1998. Highly efficient recovery of functional single-chain Fv fragments from inclusion bodies overexpressed in Escherichia coli by controlled introduction of oxidizing reagent--application to a human single-chain Fv fragment. *Journal of Immunological Methods*, 219, 119-29.

WIDMANN, M., CLAIRO, M., DIPPON, J. & PLEISS, J. 2008. Analysis of the distribution of functionally relevant rare codons. *BMC Genomics*, 9, 207-215.

WILLIAMS, D.C., VAN FRANK, R.M., MUTH, W.L. & BURNETT, J.P. 1982. Cytoplasmic inclusion bodies in Escherichia coli producing biosynthetic human insulin proteins. *Science*, 215, 687-9.

WU, X., JÖRNVALL, H., BERNDT, K.D. & OPPERMANN, U. 2004. Codon optimization reveals critical factors for high level expression of two rare codon

genes in Escherichia coli: RNA stability and secondary structure but not tRNA abundance. *Biochemical and Biophysical Research Communications*, 313, 89-96.

CHAPTER 7
General Conclusion

The original pIT2 construct, as isolated from the Human single fold scFv libraries I + J (TomlinsonI + J), was expressed in an up-scale version of the prescribed procedure in an attempt to increase TFI-scFv yield for additional characterisation. The up-scaling of the TomlinsonI + J procedure produced a comparatively high yield of functional TFI-scFv, which was purified by means of protein A affinity chromatography. Despite the relative success of the experiment there are definitive disadvantages to the up-scaling TomlinsonI + J procedure for the expression and purification of TFI-scFv. The primary concern is the inability of the expression system to produce a sufficient quantity of TFI-scFv. The expression system, specifically the expression of the pIT2 plasmid in *E. coli* HB2151, is not a recognised over-expression mechanism. Furthermore, the expression system is simply inadequate for the large scale production of TFI-scFv that is required for further characterisation. Additional concerns were the financial- as well as practical complications associated with protein A affinity chromatography. Although the high specificity of protein A affinity chromatography for human IgGs makes it the ideal purification system for TFI-scFv, the high cost involved would effectively prevent the continued research on TFI-scFv, even at a small scale. As a result IMAC, specifically immobilised nickel affinity chromatography, was investigated as a more cost effective alternative to Protein A purification for the isolation of TFI-scFv. The initial focus was on the processing of larger sample volumes rather than the improvement of the expression system in order to increase TFI-scFv yield.

The pIT2 construct is designed to incorporate a C-terminus His-tag for purification by means of IMAC. A large culture volume was utilised in order to compensate for the low level of expression of TFI-scFv. TFI-scFv was successfully expressed and accumulated in the culture media, which was collected and subjected to nickel affinity chromatography. Nickel affinity chromatography was unable to isolate TFI-scFv due to the presence of nickel chelating elements in the culture media. It was clear that in order to benefit from the advantages of nickel affinity chromatography the expression system had to be modified to prevent the accumulation of TFI-scFv in the culture media. This could be achieved by the removal of the pelB leader peptide.

Removal of the pelB leader peptide would result in the accumulation of TFI-scFv in the bacterial cytoplasm, consequently resolving the complications associated with the presence of nickel chelators in the culture media. The expression of TFI-scFv in the bacterial cytoplasm would also reduce difficulties associated with large sample volumes. Additionally, an ideal opportunity would be provided to improve TFI-scFv expression through the utilisation of vectors and *E. coli* strains optimised for over-expression of recombinant genes.

The TFI-scFv synthetic gene, excluding the pelB leader peptide, was successfully isolated from the original pIT2 construct and cloned into the pET22b (+) expression vector. Two constructs were created differing only in the position of their C-terminus His-tags. The expression of both constructs was successfully directed to the cytoplasm of *E. coli* BL21 through the removal of the pelB leader peptide. This facilitated the successful purification by nickel affinity chromatography. The purification of the construct with a further downstream positioning of the His-tag was superior, most likely due to a more exposed His-tag. Although nickel affinity chromatography was successful, the final protein yields were insufficient to perform functional assays. The utilization of a frequently used over-expression system, specifically the pET22b (+) expression vector in *E. coli* BL21, was unable to increase the yield of TFI-scFv. DNA analysis confirmed that the low expression levels were due to the high frequency of rare codons present within the TFI-scFv DNA sequence. As a result the two different methods of rare codon compensation were investigated in order to increase the expression level of TFI-scFv in *E. coli* BL21.

The expression level of TFI-scFv was improved by alleviating the detrimental effect of rare codons through the modification of the expression host. By adjusting the rare codon bias of the expression host through the co-expression of the pRARE plasmid, a clear improvement in TFI-scFv expression level was observed. This improvement was achieved despite the pRARE plasmid not compensating all the rare codons present within the TFI-scFv DNA sequence. Through the co-expression of the pRARE plasmid and the pET22 construct a soluble TFI-scFv was produced which displayed a similar inhibition effect that the original pIT2 construct. Although the expression was improved the final yield of TFI-scFv was still insufficient for our eventual goal of the further characterization of TFI-scFv in animal models. The rare

codon optimisation of the TFI-scFv DNA sequence for expression in *E. coli* also resulted in an increase in TFI-scFv yield. Despite the substantial improvement in final protein yield, the synonymous replacement of all rare codons in the TFI-scFv DNA sequence resulted in the production of inclusion bodies. The loss of solubility was most likely due to the removal of the rare codons that function as critical ribosomal stalling sites required for correct intermediate folding of TFI-scFv. Although the loss of TFI-scFv solubility is unwanted, the high level of expression provides an ideal platform for the further development of TFI-scFv.

The key to future success lies in the ability to solubilise the inclusion bodies. The solubilisation of TFI-scFv can be achieved by either modifying the expression system to produce a soluble TFI-scFv or by expressing a large quantity of insoluble protein which is subsequently refolded *in vitro* to produce functional TFI-scFv. An up-scale setting of either system would produce a sufficient quantity of functional TFI-scFv for additional characterization and testing in animal thrombosis models.

Summary

Tissue factor is a transmembrane glycoprotein that functions as the primary initiator of coagulation in response to mechanical or chemical damage. Due to its key position within the coagulation cascade it also plays an important role in the pathology of thrombosis and thrombotic complications associated with cardiovascular disease as well as in non-thrombotic disorders and diseases such as obesity, diabetes, cancer and HIV-AIDS. Recognising the potential in tissue factor inhibition as a novel approach to antithrombotic therapy, our laboratory utilized phage display technology in a previous study, in order to identify a 26 kDa single chain antibody fragment which functionally inhibits human tissue factor.

In the current study, the tissue factor inhibitory single chain variable fragment (TFI-scFv) was expressed by means of the pIT2 plasmid vector by *Escherichia coli* HB2151. This expression system was utilised in an up-scale setting in an attempt to improve the TFI-scFv yield. Although functional TFI-scFv was successfully purified from the culture media by means of Protein A affinity chromatography, the process was hampered by large sample volumes, low levels of expression as well as the high cost involved in Protein A purification.

Due to an initial focus on improving TFI-scFv yield through the processing of larger sample volumes rather than the improvement of the expression system, immobilised nickel affinity chromatography was investigated as a more cost effective alternative to Protein A affinity chromatography. It was found that the original expression system was incompatible with immobilised nickel affinity chromatography as the protein was expressed into the culture media. The culture media contained nickel chelating elements that stripped the nickel from the column and consequently prevented TFI-scFv purification.

Subsequently the TFI-scFv gene was isolated, cloned into an over-expression system and modified to redirect the expression to the bacterial cytoplasm. Although TFI-scFv was successfully redirected to the bacterial cytoplasm and purified by means of nickel affinity chromatography, it was found that expression was hampered due to the presence of rare codons. The detrimental effect of rare codons on TFI-

scFv yield was addressed through the modification of the expression host by the co-expression of the pRARE plasmid as well as by the rare codon optimization of the TFI-scFv gene sequence for expression in *E. coli*. Although the co-expression of the pRARE plasmid only slightly improved TFI-scFv yield, a sufficient amount of TFI-scFv was generated for functional testing. The modified TFI-scFv displayed a similar inhibition effect with reference to the original construct. The rare codon optimisation resulted in a substantial increase in TFI-scFv yield but consequently resulted in the loss of solubility and the production of inclusion bodies. Although the loss of TFI-scFv solubility is unwanted, the high level of expression achieved provides an ideal platform for the further development and characterization TFI-scFv in animal thrombosis models.

Keywords: Coagulation, Immobilised nickel affinity chromatography, Inclusion bodies, Over-expression systems, Protein A affinity chromatography, Rare codons, Single chain antibody fragment, Thrombosis, Tissue factor

Opsomming

Weefselfaktor is 'n transmembraan glikoproteïen wat dien as die primêre inisieërder van bloedstolling na aanleiding van meganiese- of chemiese skade. As gevolg van die belangrike posisie van weefselfaktor aan die begin van die stollingskaskade, speel dit ook 'n belangrike rol in die patologie van trombose en die trombotiese komplikasies wat geassosier word met kardiovaskulêre siektes, asook siektes soos vetsug, diabetes, kanker en MIV-vigs. As gevolg van die kardinale rol wat weefselfaktor in die bogenoemde siektes speel, besit die inhibisie van weefselfaktor as 'n benadering tot anti-trombotiese terapie baie potentiaal. In 'n vorige studie het ons navorsingslaboratorium van "phage display" tegnologie gebruik gemaak om 'n 26 kDa "single chain variable fragment" teen weefselfaktor te identifiseer wat die inisiëringsfunksie van menslike weefselfaktor inhibeer.

In die huidige studie was die teenligaam fragment "Tissue factor inhibitory single chain variable fragment - TFI-scFv" geproduseer deur die uitdrukking van die pIT2 vektor deur Escherichia coli HB2151. Hierdie uitdrukkingsisteem is op groter skaal aangewend in 'n poging om die TFI-scFv opbrengs te verbeter. Alhoewel funksionele TFI-scFv deur middel van Proteïen A affiniteitschromatografie suksesvol vanuit die kultuurmedia gesuiwer was, was die proses bemoeilik deur groot monster volumes, lae vlakke van uitdrukking, sowel as die hoë koste verbonde aan Proteïen A suiwering.

As gevolg van 'n aanvanklike fokus op die verbetering van TFI-scFv opbrengs deur middel van die verwerking van groter monster volumes, eerder as om die uitdrukking sisteem te verbeter, was geïmmobiliseerde nikkel affiniteitschromatografie ondersoek as 'n meer koste effektiewe alternatief tot proteïen A affiniteitschromatografie. Daar is gevind dat die oorspronklike uitdrukking sisteem nie versoenbaar met geïmmobiliseer nikkel affiniteitschromatografie was nie aangesien die uitgedrukte proteïen na die kultuurmedia vervoer was. Daar is gevind dat die kultuurmedia nikkel chelerende elemente bevat het wat die nikkel van die kolom afgestroop het. Die teenwoordigheid van nikkel chelerende elemente het

gevolglik die suiwering van TFI-scFv deur middel van geïmmobiliseer nikkel affiniteitschromatografie verhoed.

Die TFI-scFv geen was geïsoleer en gekloon in 'n hoë opbrengs uitdrukkingsisteem. Die geen was ook gemodifiseer om die lokalisering van die TFI-scFv te skuif van die bakteriële periplasma na die sitoplasma. Hoewel TFI-scFv suksesvol geherlokaliseer was na die bakteriële sitoplasma sowel as suksesvol gesuiwer was deur middel van nikkel affiniteitschromatografie, is daar gevind dat die uitdrukking belemmer was as gevolg van die teenwoordigheid van seldsame kodons in die DNS.

Die nadelige effek van die seldsame kodons op TFI-scFv proteïenopbrengs is aangespreek deur die modifisering van die uitdrukkingsgasheer deur middel van die mede-uitdrukking van die pRARE plasmied. Die nadelige effekte van seldsame kodons was ook aangespreek deur middel van die seldsame kodon optimisering van die TFI-scFv DNS vir uitdrukking in *E. coli.* Hoewel die mede-uitdrukking van die pRARE plasmied die TFI-scFv opbrengs slegs effens verbeter het, was 'n voldoende hoeveelheid TFI-scFv gegenereer vir funksionele analise. Die gemodifiseerde TFI-scFv het 'n soortgelyke inhibisie effek getoon as die oorspronklike konstruk. Die seldsame kodon optimisering het gelei tot 'n aansienlike toename in die TFI-scFv opbrengs, maar gevolglik gelei tot die verlies van oplosbaarheid weens die produksie van proteïen aggregate bekend as "inclusion bodies". Hoewel die verlies van TFI-scFv oplosbaarheid ongewens was, bied die hoë vlakke van uitdrukking wat bereik is 'n ideale platform vir die verdere ontwikkeling en karakterisering van TFI-scFv in trombotiese diermodelle.

<u>Appendix A</u>

Human Single Fold scFv Libraries I + J (Tomlinson I + J)

General Introduction to the Libraries

Over the past 10 years Greg Winter's lab at the MRC Laboratory of Molecular Biology and the MRC Centre for Protein Engineering (Cambridge, UK) has created a number of artificial libraries of antibodies that can be used to derive binders to almost any target molecule using phage display and selection. These binders can be used for all the same applications as conventional monoclonal antibodies (ELISA, Western blotting, FACS, immunohistochemistry etc) but can be isolated in a fraction of the time and without the need for animal immunisation. To date these so called "naïve" or "single pot" phage-antibody libraries have been used successfully in hundreds of molecular biology labs world-wide to derive highly specific antibody reagents to a wide range of different proteins, peptides or small molecule compounds.

The latest libraries (Tomlinson I and J) that are being distributed by the MRC HGMP Resource Centre each comprise over 100 million different scFv fragments cloned in an ampicillin resistant phagemid vector and transformed into TG1 *E. Coli* cells (scFv fragments comprise a single polypeptide with the VH and VL domains attached to one another by a flexible Glycine-Serine linker). By carefully following the protocol provided, large numbers of phagemids can be produced and used to select specific binders to target molecules that are attached to the surface of a tube or biotinylated and captured by streptavidin coated beads (so called "panning"). After each round of panning, the non-binders are washed away and the phagemids bound to the target molecule/s are eluted and amplified by infection into fresh TG1 cells. After producing new phagemids from the previous round of panning, the process can be repeated. Typically two or three rounds of panning are required to ensure that more than half the different scFvs in the selected population bind to the target molecule. The monoclonal scFvs can then be screened for binding (using a simple ELISA based protocol) and then used for further analysis of the target molecule. Since all the functional scFvs in the Tomlinson I and J libraries bind Proteins A and L, either of these secondary reagents can be used for detection, purification or immobilisation. Alternatively, secondary reagents that bind the attached myc or HIS6 tags can be used, although in our experience it is better to use the Protein A or L reagents.

Finally, we would like to emphasise that these libraries represent a valuable resource. Whether you are familiar with phage display or not we recommend that you perform test selections and subsequent ELISA screening using the anti-bovine serum albumin and anti-bovine ubiquitin controls provided. Only when these experiments have been successfully carried out should you defrost the libraries and start preparing library phage.

Derivation of libraries

Both libraries are based on a single human framework for V_H (V3-23/DP-47 and J_H4b) and V_κ (O12/O2/DPK9 and J_κ1) with side chain diversity incorporated at positions in the antigen binding site that make contacts to antigen in known structures and are highly diverse in the mature repertoire. The canonical structure (V_H: 1-3, V_κ: 2-1-1) encoded by this framework is by far the most common in the human antibody repertoire. The CDR3 of the heavy chain was designed to be as short as possible yet still able to form an antigen binding surface. Both libraries can be selected and affinity matured without knowing the sequences of selected clones. Libraries are in phagemid/scFv format and have been pre-screened for binding to Protein A and Protein L so that the majority of clones in the unselected libraries are functional.

Tomlinson I

Constructed in pIT2 (HIS myc tag). Diversified (DVT) side chains based mainly on those positions which are diverse in the primary repertoire (total of 18 residues - H50, H52, H52a, H53, H55, H56, H58, H95, H96, H97, H98, L50, L53, L91, L92, L93, L94 and L96). After selection, can be matured by incorporating additional diversity based on somatic mutation.

Library size (with insert):   **1.47×10^8**

Percentage insert: **96%**

Tomlinson J

As above but with NNK side chains.

Library size (with insert):   **1.37×10^8**

Percentage insert: **88%**

Read carefully before using the libraries

1. Check that you have received:

   a tube of Library I (~500 μl)

   a tube of Library J (~500 μl)

   a glycerol stock of a positive control anti-ubiquitin ScFv in bacterial strain TG1 (labeled TG1-anti ubi)

   a glycerol stock of a positive control anti-BSA ScFv in bacterial strain TG1 (labeled 13CG2)

   a glycerol stock of T-phage resistant *E. Coli.* TG1 for propagation of phage (labeled TG1Tr)

   (K12 Δ(lac-proAB) supE thi hsdD5/F' traD36 proA$^+$B lacI^q lacZΔM15)

   a glycerol stock of *E. Coli.* HB2151 for expression of antibody fragments

   (K12 ara Δ(lac-proAB) thi/F' proA$^+$B lacI^q lacZΔM15)

   Phage KM13^3 (~100 μl with 10^7 pfu/ml)

2. Check the library is still frozen and make sure you keep it frozen at -70°C until needed.

3. Make stock of KM13 according to Protocol G.

4. Run through the protocols using the control clones before you use the library: Streak the controls on TYE plates containing 100 μg/ml ampicillin and 1 % glucose. After overnight growth at 37°C in an incubator pick a single colony from each and grow these overnight (shaking at 37°C) in 5 ml 2xTY^1 containing 100 μg/ml ampicillin and 1 % glucose. Make phage for the positive and negative controls separately (use 500 μl of overnight in D1-D10). Use a 1:100 mixture of phage produced from the positive and the negative controls and perform one round of selection (C1-C11) using 100 μg/ml of ubiquitin^2 in PBS for coating. Check for enrichment of ubiquitin binders (should be over 50% after one round of selection) by monoclonal phage ELISA (E9-E14).

5. Wherever possible use devoted pipettes and disposable plastic ware. The use of polypropylene tubes is recommended as phage may adsorb non-specifically to other plastics.

6. For efficient infection of phage, *E. coli* must be grown at 37°C and be in log phase (OD at 600 nm of 0.4). To prepare this:

   i. Transfer a bacterial colony from a minimal media plate into 5 ml of 2xTY medium (no antibiotics or glucose). Grow shaking overnight at 37°C.

   ii. Next day dilute overnight 1:100 into fresh 2xTY medium. Grow shaking at 37°C until OD 600 is 0.4 (1.5-2 hrs)

7. All centrifugations, except those performed in a micro centrifuge, are performed at 4°C.

8. Libraries I and J must be used separately and preferably in parallel. This will ensure selecting the most antigen binding clones.

Day by day overview of library selection

In advance

Gather all equipment and reagents (product details are given in the notes at the end of all the protocols). Make sure you have all the necessary media and plates for bacterial growth (the large Bio-Assay plates need to be air-dried in a sterile environment for 2 hrs before use).

Plan your time - most of the daily procedures can be performed simultaneously, so read through each protocol carefully before starting.

	Steps	*Procedure*
Day 1 (5 hrs)	A1-A6	Grow libraries I and J and make phage
	B1-B3	Make secondary stock of libraries
Day 2 (6 hrs)	A7-A12	Grow libraries I and J and make phage (cont.)
	C1	Coat immunotubes for 1st round of selection
Day 3 (6.5 hrs)	C2-C11	1st round of selection
Day 4 (3 hrs)	D1-D6	Make phage from 1st round of selection
	C1	Coat immunotubes for 2nd round of selection
Day 5 (6.5 hrs)	D7-D11	Make phage from 1st round of selection (cont.)
	C2-C11	2nd round of selection
Day 6 (3 hrs)	D1-D6	Make phage from 2nd round of selection
	C1	Coat immunotubes for 3nd round of selection
Day 7 (6.5 hrs)	D7-D11	Make phage from 2nd round of selection (cont.)
	C2-C11	3rd round of selection
Day 8 (3 hrs)	D1-D6	Make phage from 3rd round of selection
	E1	Coat 96 well plate for polyclonal phage ELISA
Day 9 (6.5 hrs)	D7-D11	Make phage from 3rd round of selection (cont.)
	E2-E8	Polyclonal phage ELISA

Further characterisation of individual clones can be performed by monoclonal phage ELISA (protocol E), monoclonal ELISA using soluble scFv fragments (protocol F), PCR screening (to check for insert, protocol H) and sequencing (protocol I).

A. Growing the libraries

1. Add the library stock to 200 ml pre-warmed 2xTY containing 100 μg/ml ampicillin and 1 % glucose.

2. Grow shaking at 37°C until the OD 600 is 0.4 (1-2 hrs).

3. Take 50 ml of this and add 2×10^{11} KM13 helper phage[3]. (Use the remaining 150 ml to make a secondary bacterial stock of the library by following protocol B).

4. Incubate without shaking in a 37°C water bath for 30 min.

5. Spin at 3,000 g for 10 min (3,600 rpm in Centra 8 or equivalent). Resuspend in 100 ml of 2xTY containing 100 μg/ml ampicillin, 50 μg/ml kanamycin and 0.1% glucose.

6. Incubate shaking at 30°C overnight.

7. Spin the overnight culture at 3,300 g for 30 min (4,000 rpm in Centra 8 or equivalent).

8. Add 20 ml PEG/NaCl (20 % Polyethylene glycol 6000, 2.5 M NaCl) to 80 ml supernatant. Mix well and leave for 1 hr on ice.

9. Spin 3,300 g for 30 min (4,000 rpm in Centra 8 or equivalent). Pour away PEG/NaCl. Respin briefly and aspirate any remaining dregs of PEG/NaCl.

10 Resuspend the pellet in 4 ml PBS and spin at 11,600 g for 10 min in a micro centrifuge to remove any remaining bacterial debris.

11. Store the phage at 4°C for short term storage or in PBS with 15 % glycerol for longer term storage at -70°C.

12. To titre the phage stock dilute 1μl phage in 100μl PBS, 1μl of this in 100μl PBS and so on until there are 6 dilutions in total. Add 900μl of TG1 at an OD 600 of 0.4 to each tube and incubate at 37°C in a waterbath for 30 mins. Spot 10 μl of each dilution on a TYE[5] plate containing 100 μg/ml ampicillin and 1 % glucose and grow overnight at 37°C. Phage stock should be 10^{12}-10^{13}/ml, enough for at least 10 selections.

B. Growing secondary stocks of the libraries

1. Grow the remaining 150 ml from A3 for a further 2 hr shaking at 37°C.
2. Spin down the cells at 10,800 g for 10 min. Resuspend in 10 ml of 2xTY containing 15 % glycerol.
3. Store this secondary stock in 20x 500 μl aliquots at -70°C. Use one aliquot for each phage preparation according to protocol A - this will only be necessary if you wish to do more than 10 selections.

C. Selection on immunotubes

(Alternatively, phage can be selected using biotinylated antigen in solution or affinity chromatography. For details see Winter *et al.* (1994) *Annu. Rev. Immunol.* **12**, 433)

1.   Coat immunotube[6] overnight with 4 ml of the required antigen. The efficiency of coating can depend on the antigen concentration, the buffer and the temperature. Usually 10-100 μg/ml antigen in PBS is used.

2.   Next day wash tube 3 times with PBS (pour into the tube and then pour it immediately out again).

3.   Fill tube to brim with 2 % MPBS (2 % Marvel milk powder[7] in PBS). Incubate at rt. Standing on the bench for 2 hr to block.

4.   Wash tube 3 times with PBS.

5.   Add 10^{12} to 10^{13} phage from A11 in 4 ml of 2 % MPBS. Incubate for 60 min at rt. rotating using an under-and-over turntable and then stand for a further 60 min at rt. Throw away supernatant.

6.   Wash tubes 10 (round 1)-20 (rounds 2 and 3) times with PBS containing 0.1 % Tween 20.

7.   Shake out the excess PBS and elute phage by adding 500 μl of trypsin-PBS (50 μl of 10mg/ml trypsin stock solution[8] added to 450 μl PBS) and rotating for 10 min at rt using an under-and-over turntable.

8.   Take 1.75 ml of TG1 at an OD 600 of 0.4 and add 250 μl of the eluted phage (the remaining 250 μl should be stored at 4°C). Incubate for 30 min at 37°C in a water bath without shaking.

9.   Spot 10 μl, 10 μl of a $1:10^2$ dilution and 10 μl of a $1:10^4$ dilution on TYE plates containing 100 μg/ml ampicillin and 1 % glucose and grow overnight at 37°C to titre the phage.

10.   *In round 1 if using a complex antigen (eg cells, cell lysates etc):* take the remaining TG1 culture and spin at 11,600 g in a micro centrifuge for 5 min. Resuspend the pelleted bacteria in 1 ml of 2xTY and plate on a large square Bio-Assay dish[9] containing TYE, 100 μg/ml ampicillin and 1 % glucose.
       In round 1 if using a single hapten, carbohydrate or protein antigen and in subsequent rounds for all antigens: take the remaining TG1 culture and spin at 11,600 g in a micro centrifuge for 5 min. Resuspend the pelleted bacteria in 50 μl of 2xTY and plate on a regular TYE plate containing 100 μg/ml ampicillin and 1 % glucose.

11.   Grow plates at 37°C overnight.

The first round of selection is the most important. Any errors made at this point will be amplified in subsequent rounds of selection. After each round you should get back at least 100 infective phage. If you get less than this it is probable that a mistake has occurred. If so, try repeating the infection with a freshly grown TG1 culture (from a new overnight) at an OD 600 of 0.4 using the remaining 250 μl of eluted phage from C7. If this still yields less than 100 phage, repeat the selection starting at C1.

D. Further rounds of selection

1. After overnight growth add 7 ml of 2xTY 15 % glycerol to the large square Bio-Assay dish or 2mls to the regular plates and loosen the cells with a glass spreader, mixing them thoroughly. After inoculating 50 μl of the scraped bacteria to 50 ml of 2xTY containing 100 μg/ml ampicillin and 1 % glucose, store 1 ml of the remaining bacteria at -70°C in 15% glycerol.

2. Grow shaking at 37°C until the OD 600 is 0.4 (1-2 hrs).

3. Take 10 ml of this culture and add 5×10^{10} helper phage.

4 Incubate without shaking in a 37°C water bath for 30 min.

5. Spin at 3,000 g for 10 min. Resuspend in 50 ml of 2xTY containing 100 μg/ml ampicillin, 50 μg/ml kanamycin and 0.1% glucose.

6. Incubate shaking at 30°C overnight.

7. Spin the overnight culture at 3,300 g for 15 min.

8. Add 10 ml PEG/NaCl (20 % Polyethylene glycol 6000, 2.5 M NaCl) to 40 ml supernatant. Mix well and leave for 1 hr on ice.

9. Spin 3,300 g for 30 min. Pour away PEG/NaCl. Respin briefly and aspirate any remaining dregs of PEG/NaCl.

10. Resuspend the pellet in 2 ml PBS and spin at 11,600 g for 10 min in a micro centrifuge to remove the remaining bacterial debris.

11. Use 1 ml of this phage for the next round of selection (protocols C and D) and store 1 ml at 4°C.

12. Repeat selection for another 2 rounds.

E. Screening phage particles by ELISA

Populations of phage produced at each round of selection can be screened for binding by ELISA to identify "polyclonal" phage antibodies. Phage from single colonies can then be screened by ELISA to identify "monoclonal" phage antibodies. Alternatively, after a polyclonal phage ELISA you could proceed directly to making monoclonal soluble antibody fragments, see protocol F. In general, we have found that 2% Marvel in PBS is best for blocking during phage ELISA whereas 3% BSA in PBS is best for blocking during scFv ELISA.

(a) Polyclonal phage ELISA

1. Coat a 96 well flexible assay plate[10] overnight with 100 μl per well of antigen in the same buffer and at the same concentration as used for selection.
2. Wash wells 3 times with PBS. Plates can be immersed in a shallow bath containing PBS but you should check that all wells fill with wash solution (if they do not you may create false positives during later washes). Discard liquid by flipping plate over and then shaking it. Add 200 μl per well of 2 % MPBS (2 % Marvel in PBS) or 3% BSA-PBS (3% bovine serum albumin in PBS) to block and incubate for 2 hr at rt.
3. Wash wells 3 times with PBS. Add 10 μl PEG precipitated phage from the end of each round of selection in 100 μl 2 % MPBS (or 3 % BSA-PBS).
4. Incubate for 1 hr at rt. Discard phage solution and wash 3 times with PBS-0.1 % Tween 20.
5. Add 1 in 5000 dilution of HRP-anti-M13[11] in 2 % MPBS (or 3 % BSA-PBS). Incubate for 1 hr at rt., wash 3 times with PBS-0.1 % Tween 20.
6. Add 100 μl of substrate solution (100 μg/ml TMB[12] in 100 mM sodium acetate, pH 6.0. with 10 μl of 30 % hydrogen peroxide added to 50 ml of this solution directly before use) to each well and leave at rt. for 2-15 min. A blue colour should develop.
7. Stop the reaction by adding 50 μl 1 M sulphuric acid. The blue colour should turn yellow.
8. Read the OD at 650 nm and at 450 nm. Subtract OD 650 from OD 450.

(b) Monoclonal phage ELISA

9. Inoculate individual colonies from the titration plates from each round of selection (see C11) into 100 μl 2xTY containing 100 μg/ml ampicillin and 1 % glucose in 96 cell-well plates[13]. Grow shaking (250 rpm) overnight at 37°C.
10. Use a 96 well transfer device[14] to transfer a small inoculum (about 2 μl) from this plate to a second 96 cell-well plate containing 200 μl of 2xTY with 100 μg/ml ampicillin and 1 % glucose per well. Grow shaking (250 rpm) at 37°C for 2 hr. (Make glycerol stocks of the original 96-well plate, by adding glycerol to a final concentration of 15 %, and then storing the plates at -70°C).
11. After 2 hr incubation (of the second plate) add 25 μl 2xTY containing 100 μg/ml ampicillin, 1 % glucose and 10^9 helper phage.
12. Shake (250 rpm) at 37°C for 1 hr. Spin 1,800 g for 10 min, then aspirate off the supernatant.
13. Resuspend pellet in 200 μl 2xTY containing 100 μg/ml ampicillin and 50 μg/ml kanamycin. Grow shaking (250 rpm) overnight at 30°C.
14. Spin at 1,800 g for 10 min and use 50 μl of the supernatant in phage ELISA as detailed above.

F. Production of soluble antibody fragments

Individual colonies picked from the various rounds of selection (as plated on the dilution series) can be induced in TG-1 to produce soluble scFv (F2-F6). This will ensure the expression of all selected clones including those in which the scFvs contain TAG stop codons (TG-1 as able to suppress termination and introduce a glutamate residue at these positions). Unfortunately, since the TAG stop codon between the scFv and the gIII gene is also suppressed this leads to co-expression of the scFv-pIII fusion which tends to lower the overall levels of scFv expression, even in clones where there are no TAG stop codons in the scFv itself. To circumvent this problem, the selected phage can be used to infect HB2151 (a non-suppresor strain) which is then induced to give soluble expression of antibody fragments (scFv genes that do not contain TAG stop codons will now yield higher levels of soluble scFv than in TG-1, but those that contain TAG stop codons will not produce <u>any</u> soluble scFv) (F1-F6). The expressed scFvs can then be used in ELISA. Detection of bound scFv can be performed using either Protein A-HRP[15] or Protein L-HRP[16] conjugates.

1. From each selection take 10 μl of eluted phage and infect 200 μl exponentially growing HB2151 bacteria (OD 600 of 0.4) for 30 min at 37°C in a water bath. Plate 50 μl, 50 μl of a 1:10^2 dilution, 50 μl of a 1:10^4 dilution and 50 μl of a 1:10^6 dilution on TYE plates containing 100 μg/ml ampicillin and 1 % glucose and grow overnight at 37°C.

2. Pick individual colonies into 100 μl 2xTY 100 μg/ml ampicillin and 1 % glucose in 96 cell-well plates and grow shaking (250 rpm) overnight at 37°C.

3. Use a 96 well transfer device[14] to transfer a small inocula (about 2 μl) from this plate to a second 96 cell-well plate containing 200 μl 2xTY containing 100 μg/ml ampicillin and <u>0.1 % glucose</u> per well. Grow shaking (250 rpm) at 37°C until the OD 600 is approximately 0.9 (about 3 hr). (A stock can be made of the first plate, by adding glycerol to a final concentration of 15 % and storing at -70°C).

4. Once OD 0.9 is reached (wells look quite cloudy) add 25 μl 2xTY containing 100 μg/ml ampicillin and 9 mM IPTG (isopropyl β-D-thiogalactoside, final concentration 1 mM IPTG). Continue shaking (250 rpm) at 30°C overnight.

5. Coat a 96 well flexible assay plate overnight with 100 μl per well of antigen in the same buffer and at the same concentration as used for selection.

6. Spin the plate from step F4 at 1,800 g for 10 min and use 50 μl of the supernatant (take care not to transfer any bacteria) for ELISA in 3% BSA-PBS (final concentration) (protocol E) except using a 1:5000 dilution of Protein A-HRP[15] or Protein L-HRP[16] to detect binding in step E5.

G. Production of large quantities of helper phage

1. Infect 200 µl TG1 at an OD 600 of 0.4 with 10 µl of 100-fold serial dilutions of KM13 helper phage[3] (in order to get well separated plaques) in a 37°C water bath (without shaking) for 30 min. Add to 3 ml molten H-top agar (42°C) and pour onto warm TYE plates (no antibiotics). Allow to set and incubate overnight at 37°C.

2. Pick a <u>small</u> plaque into 5 ml of fresh TG1 at an OD 600 of 0.4. Grow for about 2 hr shaking at 37°C.

3. Add to 500 ml 2xTY in a 2 litre flask and grow shaking at 37°C for 1 hr. Add kanamycin to a final concentration of 50 µg/ml (no glucose). Grow overnight shaking at 30°C.

4. Spin overnight culture at 10,800 g for 15 min. Add 100 ml PEG/NaCl (20 % polyethylene glycol 6000, 2.5 M NaCl) to 400 ml supernatent and leave for 1 hr on ice.

5. Spin 10,800 g for 30 min. Pour away PEG/NaCl.

6. Resuspend the pellet in 8 ml PBS and add 2 ml PEG/NaCl. Mix well and leave for 20 minutes on ice.

7. Spin 3,300 g for 30 min. Respin briefly and aspirate any remaining dregs of PEG/NaCl.

8. Resuspend pellet in 5 ml PBS and spin at 11,600 g for 10 min in a micro centrifuge to remove any remaining bacterial debris.

9. Store the helper phage at 4°C for short term storage or in PBS with 15 % glycerol for longer term storage at -70°C.

10. To titre the helper phage take 45µl phage and add 5µl trypsin stock solution. Incubate for 30 mins at 37 °C. Dilute 1µl of trypsin treated phage in 1ml PBS and make five 100 fold serial dilutions of this in 1ml aliquots of PBS. Add 50µl of the six dilutions to six separate tube containing 1ml of TG1 at an OD 600 of 0.4. Mix, add 3 ml molten H-Top and pour evenly onto TYE plates. Perform the same dilution series using 1µl of non-trypsin treated phage. The titre of the trypsin treated phage should be 10^5-10^8 lower than for the non-trypsin treated phage. If not, pick another plaque and repeat helper phage preparation.

H. PCR screening selected clones

Once the libraries have been selected you may wish to check individual clones for the presence of full length VH and V_K insert. All PCRs are at annealing temperature of 55°C. 1 min extension for V_H or V_K on their own, 2 min extension for V_H and V_K together.

For V_H only use   LMB3:   CAG GAA ACA GCT ATG AC
                            link seq new: CGA CCC GCC ACC GCC GCT G
                            with insert = 527 bp
                            without insert = 227 bp

For V_K only use   DPK9 FR1 seq: CAT CTG TAG GAG ACA GAG TC
                            pHEN seq: CTA TGC GGC CCC ATT CA
                            with insert = 368 bp
                            without insert = no band

For V_H and V_K use LMB3:   CAG GAA ACA GCT ATG AC
                            pHEN seq: CTA TGC GGC CCC ATT CA
                            with insert = 935 bp
                            without insert = 329 bp

I. Sequencing selected clones

For sequencing of selected clones the following primers are recommended.

For V_H use link seq new   CGA CCC GCC ACC GCC GCT G

For V_K use pHEN seq   CTA TGC GGC CCC ATT CA

```
                                                     RBS
CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATA ATG AAA TAC CTA
---------------->                                                      M   K   Y   L
     LMB3
```

```
                                              SfiI        NcoI
TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCG GCC CAG CCG GCC ATG GCC GAG GTG TTT
 L   P   T   A   A   A   G   L   L   L   L   A   A   Q   P   A   M   A   E   V   F
```

```
                                     XhoI                              linker
GAC TAC TGG GGC CAG GGA ACC CTG GTC ACC GTC TCG AGC GGT GGA GGC GGT TCA GGC GGA GGT
 D   Y   W   G   Q   G   T   L   V   T   V   S   S   G   G   G   G   S   G   G   G
```

```
                 SalI                                         NotI           HIS-tag
GGC AGC GGC GGT GGC GGG TCG ACG GAC ATC CAG ATG ACC CAG GCG GCC GCA CAT CAT CAT CAC
 G   S   G   G   G   S   T   D   I   Q   M   T   Q   A   A   A   H   H   H   H
  <------------------------
       link seq new
```

```
                           myc-tag
CAT CAC GGG GCC GCA GAA CAA AAA CTC ATC TCA GAA GAG GAT CTG AAT GGG GCC GCA TAG ACT
 H   H   G   A   A   E   Q   K   L   I   S   E   E   D   L   N   G   A   A   *   T
                                                        <--------------------
                                                              pHEN seq
```

Notes

1.   2xTY is 16g Tryptone, 10g Yeast Extract and 5g NaCl in 1 litre.

2.   Bovine ubiquitin (5 mg) is available from Fluka, Chemika-BioChemika, Industriestrasse 25, CH-9470 Buchs, Switzerland, Tel +41 81 755 2511, Fax +41 81 756 5449.

3.   KM13 is the protease cleavable helper phage described in Kristensen and Winter, *Folding and Design* **3**, 321-328 (1998).

4.   PBS is 5.84 g NaCl, 4.72 g Na_2HPO_4 and 2.64 g $NaH_2PO_4.2H_2O$, pH 7.2, in 1 litre.

5.   TYE is 15g Bacto-Agar, 8g NaCl, 10g Tryptone, 5g Yeast Extract in 1 litre.

6.   Nunc Maxisorp immuno test tubes (Cat. No. 4-44202) are available from Gibco BRL, Life Technologies Ltd., P. O. Box 35, Trident House, Washington Road, Paisley, PA3 4EF, Scotland, U.K, Tel +44 141 814 6100, Fax +44 141 887 1167.

7.   'Marvel' is dried skimmed milk powder.

8.   Trypsin (T-1426 Type XIII from Bovine Pancreas - Sigma Chemical Company Ltd., Fancy Rd., Dorset, BH17 7NH, U.K, Tel +44 1202 733114; Fax +44 1202 715460) made up in 50mM Tris-HCl pH7.4, 1mM $CaCl_2$ and stored at -20°C

9.   Nunc Bio-Assay dish is available from Gibco BRL (see note 6).

10.   Falcon MicroTest III flexible 96 well flat bottomed assay plates are available from Becton Dickinson Labware, Becton Dickinson and Co., 2 Bridgewater Lane, Lincoln Park, NJ 07035, USA.

11.   HRP-anti-M13 is available from Amersham International plc, Amersham Place, Little Chalfont, Buckinghamshire, HP7 9NA, UK. Tel: +44 01494 544000; Fax: +44 01494 542929.

12.   TMB is 3,3',5,5'-tetramethylbenzidine and is available from Sigma (see Note 8). A 10 mg/ml stock solution can be made by dissolving the TMB in DMSO.

13.   Corning 'Cell Wells' flat-bottomed multiple well tissue culture treated plates are available from Corning Glass Works, Corning N.Y. 14831. USA.

14.   The 96 well transfer device is a piece of wood the size of a microtitre plate with a handle on one side and 96 metal pins (each 7 cm long with a concave end) on the other. This can be sterilised between bacterial transfer by immersion in a bath of ethanol and then by flaming (hold well away from your body and any flamable objects). If you haven't got one of these (or something similar) you will have to make one yourself or alternatively use a multichannel pipette for bacterial transfer.

15.   Horse Radish Peroxidase conjugated Protein A is available from Amersham International plc (see note 11)

14.   Horse Radish Peroxidase conjugated Protein L is available from Actigen Ltd, 5 Signet Court, Swanns Road, Cambridge, CB5 8LA, UK. Tel: +44 01223 319101; Fax: +44 01233 316443.

Appendix B

Media recipes

M9 minimal plates

1. To make M9 Salts aliquot 800ml H_2O and add
   - 64g Na_2HPO_4-$7H_2O$
   - 15g KH_2PO_4
   - 2.5g NaCl
   - 5.0g NH_4Cl
   - 15g Agarose
   - Stir until dissolved
   - Adjust to 1000ml with distilled H_2O
   - Sterilize by autoclaving
2. Measure ~700ml of distilled H_2O (sterile)
3. Add 200ml of M9 salts
4. Add 2ml of 1M $MgSO_4$ (sterile)
5. Add 20 ml of 20% glucose (or other carbon source)
6. Add 100ul of 1M $CaCl_2$ (sterile)
7. Adjust to 1000ml with distilled H_2O

SOC Media

1. Add the following to 900ml of distilled H_2O
   - 20g Bacto Tryptone
   - 5g Bacto Yeast Extract
   - 2ml of 5M NaCl.
   - 2.5ml of 1M KCl.
   - 10ml of 1M $MgCl_2$
   - 10ml of 1M $MgSO_4$
   - 20ml of 1M glucose
2. Adjust to 1L with distilled H_2O
3. Sterilize by autoclaving