PREPARATION AND IN VITRO CHARACTERISATION OF AN ANTI-TISSUE FACTOR SINGLE CHAIN VARIABLE FRAGMENT

by JAN-G VERMEULEN

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

Ph.D. (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

January 2017

Supervisor: Professor SM Meiring
Co-Supervisors: Professor FJ Burt & Professor E van Heerden

To the reader

It is with great pride that I present my thesis titled: The preparation and characterisation of a single chain variable fragment against human tissue factor. I feel exceptionally blessed and privileged to have been able to follow this dream of mine. I would therefore like to acknowledge and thank the following people for their continued support and guidance throughout this project.

To Professor Muriel Meiring for taking me under her wing all those years ago, in 2007 when I first enrolled at the Department of Haematology and Cell Biology. Thank you for all the constant support over all these years and for allowing me the freedom to grow as a scientist of my own making.

To Professor Esta van Heerden for all her support and for welcoming me to the Extreme Biochemistry group at the Department of Microbial- Biochemical and Food Biotechnology. Thank you for ease that you allowed me to integrate into you research group and making lab22 the “home” that is to me as well as so many other students.

To Professor Felicity Burt, thank you the time that I spent as you research assistant in the Department of Virology. I will always appreciate the knowledge and experience that I gained under your guidance. I would also especially like to thank you for all the reading and reviewing that you did for me.

I would also like to thank Leon du Preez, Armand Bester, Dr Dirk Opperman and Dr Jacqueline van Marwijk for the assistance and guidance that they provided through this project. In truth, very little of this project would have been possible were it not for your patience and teachings.

Lastly, I would like to thank my parents and friends for the support and inspiration.

Yours sincerely
Jan-G Vermeulen

Chapter Index

Chapter 1: Introduction to the study

List of abbreviations

Escherichia coli	E.coli
Human single chain antibody fragment	scFv
Human Genome Mapping Project Resource Centre	HGMP
Immobilised Metal Affinity Chromatography	IMAC
Medical Research Council	MRC

Chapter 1: Introduction to the study

1.1 Introduction

Tissue factor is a trans membrane glycoprotein that functions as the primary initiator of coagulation in response to mechanical or chemical damage. It plays a key role in the pathology of thrombosis and thrombotic complication associated with cardiovascular disease [1]. These disorders are annually responsible for 17 million deaths worldwide, 80% of which occur in middle- and low income countries such as South Africa [2]. Thrombotic complications associated with non-thrombotic disorders and diseases such as obesity, diabetes, cancer and HIV-AIDS drastically increases mortality [3–8]. Consequently, a significant challenge in the field of cardiovascular and thrombotic research is the development of safe and effective anticoagulant agents, especially at low cost, 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. By inhibiting the initiation step of coagulation, the formation of “new" thrombin is prevented and the positive feedback mechanism responsible for thrombin amplification is down regulated [9]. Tissue factor inhibition also limits the *vice versa* activation between coagulation and inflammation [9–11]. A further potential advantage is that tissue factor inhibitors only act at the site of vascular injury while leaving the physiological haemostasis intact [9]. Animal models have shown that the inhibition of tissue factor has little risk of haemorrhage [12–14]. 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 [15].

Tissue factor does not only initiate the coagulation system, but also plays an important role in the pathogenesis of arterial thrombosis, arthrosclerosis, obesity, type 2 diabetes and HIV-AIDS [16–20]. Apart from this role, it also contributes to the metastasis and angiogenesis observed in certain malignancies especially breast cancer [21–23]. 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 [24,25]. It is this *vice versa* activation between these two physiological processes that enhances the pathological severity of thrombosis and other diseases exhibiting thrombotic complications [10]. Thus, the inhibition of tissue factor is of great importance as it may provide a novel antithrombotic therapy that addresses the complications of these diseases.

Antibodies and antibody fragments represent the fastest growing segment of the bio-pharmacological market [26]. Currently available monoclonal antibodies against tissue factor are primarily of rodent origin [11]. 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 [26].

Phage-display was first introduced in 1985 by G. Smith [27]. Bacteriophages or "phages" are viruses that specifically infect and replicate within bacteria [27]. Phages are used as molecular expression vectors as they are able to accommodate foreign DNA in the phage genome [28]. As the phage infect and replicate within the bacterial host, the insert foreign DNA is also expressed as well [29]. By using phage display, large amounts of antibody fragments can be more effectively produced in *Escherichia coli* strains than mammalian cell lines. Recognising the potential in tissue factor inhibition as a novel approach to antithrombotic therapy, our laboratory utilised phage display technology to identify a series of human single chain antibody fragment (scFv) from the Tomlinson I + J Human Single Fold Phage Libraries (obtained from the MRC HGMP Resource Centre, London) [30]. The scFv that displayed the strongest inhibition (JTC5-scFv) of tissue factor coagulatory function was retained for the future research discussed in this thesis.

Preliminary *in vitro* studies have shown that this antibody fragment directed against tissue factor prolonged the plasma prothrombin clotting times in a dose dependent manner. 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 [30]. Although the initial findings were promising, the further characterisation of the tissue factor inhibiting scFv was hampered by low proteins yields as well as financial complication associated with the initial purification methods.

Thus, the focus of study shifted toward improvement of the expression of the scFv and the purification by a more financially viable method. Initially the scFv was expressed in an upscale setting of the prescribed Tomlinson I + J procedure in an attempt to increase antibody yield. Although scFv yield was comparable to those published in literature, it was still insufficient for our characterisation purposes. As a result, the scFv gene was isolated from the pIT2 vector, as utilised by the Tomlinson I + J system, and cloned into the pET22 over-expression vector in an attempt to further increase antibody yield. The new construct was expressed in *E. coli* BL21 cytoplasm appose to the periplasm. The antibody was purified using immobilised metal affinity chromatography (IMAC). Although antibody expression was somewhat improved the final yield of protein was much lower than expected. The utilisation of IMAC proved to be a clear improvement on the original protein A affinity chromatography, despite the lower than expected antibody yield.

1.2 Summary of chapter contents

The previous modifications to the antibody expression mechanisms as well as the utilisation of improved purification methods provided the ideal opportunity to explore alternative methods for the production of tissue factor inhibiting scFv. Broadly the aim of this study is therefore to improve the production of functional antibody as well as the characterisation of protein structure and inhibition mechanism though *in silico* modelling.

To start with, an in-depth literature review is presented in chapter 2. The literature review focus primarily on various roles of tissue factor during normal haemostasis as well as thrombosis. Special attention is given to the genetics and various forms of tissue factor, the structure of the tissue factor and the molecular- and physiological mechanisms involved in the initiation of the coagulation cascade. To conclude the literature review, the pathophysiological role of tissue factor and the various strategies that has been explored to inhibit the prothrombotic actions of tissue factor are discussed in brief.

The *in silico* characterisation of the tissue factor inhibiting single chain variable fragment (TFI-scFv) structure and inhibition mechanism are presented in chapter 3. This chapter focuses on the exploration of the three-dimensional structure of the antibody itself by means of homology modelling. The position and orientation of the complimentary determining regions of the scFv are identified and utilised to predict most likely interactions between the scFv antibody and tissue factor by means of protein-protein docking modelling. The predicted interactions were then refined my means of molecular dynamic simulation. This protein-protein docking model provides novel insights on the mechanism involved during the inhibition of the initiation coagulation.

In chapter 4, the optimisation of antibody gene through the use of synonymous codon substitution in order to improve the antibody yield is presented. The codon optimisation of the antibody gene resulted in an increase in protein expression but also in the production of protein aggregates known as inclusion bodies. The remainder of the chapter focuses on the solubilisation of the inclusion bodies in order to produce functional antibodies.

In chapter 5, the use of solubilisation expression systems as an alternative to *in vitro* refolding techniques for the production of soluble antibody anti body fragments. The focus of this chapter is on the utilization of cold shock expression systems, fusion partners and specialized *E.coli* strains for successful production of functional TFI-scFv in the cytoplasm. A comparison between the structure of the tissue factor inhibiting scFv and that of the fusion construct is made in an attempt to better understand the solubilisation process.

Lastly, a general discussion and a conclusion is summarised in chapter 6.

1.3 Regarding the thesis format

The chapters within this thesis are presented as sequential series of lone somewhat bulky standing articles. The purpose of this specific writing style is primarily to satisfy the new requirements of the University of the Free State and secondary to condense the work down to suitable length. As a result, some repetitiveness may be present as central recurring themes will have to be repeated and addressed in each chapter separately. Also, due to the nature this new required format, some additional attention is given to subjects that would normally would not be included or only be presented as supplementary information. I therefore hope that the work presented in this thesis

is sufficiently explanatory in nature to allow for a clear understanding of foreign concepts while still maintaining an ease of reading.

1.4 References

[1] N. Mackman, The role of tissue factor in hemostasis and thrombosis, Blood Cells, Mol. Dis. 36 (2006) 104–107.

[2] A. Callow, Cardiovascular disease 2005 - the global picture, Vascul. Pharmacol. 45 (2006) 302–307.

[3] G. Granata, T. Izzo, P. Di Micco, B. Bonamassa, G. Castaldo, V.G. Viggiano, U. Picillo, G. Castaldo, A. Niglio, Thromboembolic events and haematological diseases: a case of stroke as clinical onset of a paroxysmal nocturnal haemoglobinuria., Thromb. J. 2 (2004) 10. doi:10.1186/1477-9560-2-10.

[4] C. Milsom, J. Yu, L. May, B. Meehan, N. Magnus, The role of tumor-and host-related tissue factor pools in oncogene-driven tumor progression, Thrombosis. 120 (2007) S82–S91.

[5] L. Pantanowitz, B.J. Dezube, Monocytes tied to HIV-associated thrombosis, Blood. 115 (2010) 156–157. doi:10.1182/blood-2009-11-250597.

[6] J. Nofer, B. Kehrel, M. Fobker, B. Levkau, G. Assmann, A. von Eckardstein, HDL and arteriosclerosis : beyond reverse cholesterol transport, Atherosclerosis. 161 (2002) 1–16.

[7] A.K. Kakkar, M. Levine, H.M. Pinedo, R. Wolff, J. Wong, Venous Thrombosis in Cancer Patients: Insights from the FRONTLINE Survey, Oncologist. 8 (2003) 381–388.

[8] M.N. Levine, A.Y. Lee, A.K. Kakkar, From Trousseau to targeted therapy: New insights and innovations in thrombosis and cancer, J. Thromb. Haemost. 1 (2003) 1456–1463. doi:10.1046/j.1538-7836.2003.00275.x.

[9] P. Golino, The inhibitors of the tissue factor:factor VII pathway, Thromb. Res. 106 (2002) V257–V265.

[10] A.J. Chu, Tissue factor mediates inflammation., Arch. Biochem. Biophys. 440 (2005) 123–132. doi:10.1016/j.abb.2005.06.005.

[11] J. Steffel, A. Akhmedov, C. Fähndrich, F. Ruschitzka, T.F. Lüscher, F.C. Tanner, Differential effect of celecoxib on tissue factor expression in human endothelial and vascular smooth muscle cells, Biochem. Biophys. Res. Commun. 349 (2006) 597–603. doi:10.1016/j.bbrc.2006.08.075.

[12] S.R. Hanson, K.S. Sakariassen, Blood flow and antithrombotic drug effects, Am. Heart J. 135 (1998) S132-145.

[13] M. Roque´, E.D. Reis, V. Fuster, A. Padurean, J.T. Fallon, M.B. Taubman, J.H. Chesebro, J.J. Badimon, Inhibition of Tissue Factor Reduces Thrombus Formation and Intimal Hyperplasia After Porcine Coronary Angioplasty, J. Am. Coll. Cardiol. 36 (2000) 2303–2310.

[14] P. Carmeliet, D. Collen, Molecules in focus Tissue factor, Int. J. Biochem. Cell Biol. 30 (1998) 661–667. doi:10.1016/S1357-2725(97)00121-0.

[15] M. Miura, N. Seki, T. Koike, T. Ishihara, T. Niimi, F. Hirayama, T. Shigenaga, Y. Sakai-Moritani, T. Kawasaki, S. Sakamoto, M. Okada, M. Ohta, S. Tsukamoto, Potent and selective TF/FVIIa inhibitors containing a neutral P1 ligand., Bioorg. Med. Chem. 14 (2006) 7688–7705. doi:10.1016/j.bmc.2006.08.010.

[16] F. Samad, Regulation of tissue factor gene expression in obesity, Blood. 98 (2001) 3353–3358. doi:10.1182/blood.V98.12.3353.

[17] M. Diamant, Elevated Numbers of Tissue-Factor Exposing Microparticles Correlate With Components of the Metabolic Syndrome in Uncomplicated Type 2 Diabetes Mellitus, Circulation. 106 (2002) 2442–2447. doi:10.1161/01.CIR.0000036596.59665.C6.

[18] Y. Kamikura, H. Wada, T. Nobori, T. Kobayashi, T. Sase, M. Nishikawa, K. Ishikura, N. Yamada, Y. Abe, J. Nishioka, T. Nakano, H. Shiku, Elevated levels of leukocyte tissue factor mRNA in patients with venous thromboembolism., Thromb. Res. 116 (2005) 307–312. doi:10.1016/j.thromres.2004.12.013.

[19] J. Steffel, T.F. Lüscher, F.C. Tanner, Tissue factor in cardiovascular diseases: Molecular mechanisms and clinical implications, Circulation. 113 (2006) 722–731. doi:10.1161/CIRCULATIONAHA.105.567297.

[20] J.I. Zwicker, B.C. Furie, B. Furie, Cancer-associated thrombosis, Crit. Rev. Oncol. Hematol. 62 (2007) 126–136. doi:10.1016/j.critrevonc.2007.01.001.

[21] J. Rak, C. Milsom, N. Magnus, J. Yu, Tissue factor in tumour progression., Clin. Haematol. 22 (2009) 71–83. doi:10.1016/j.beha.2008.12.008.

[22] S. Butenas, Tissue Factor Structure and Function, Scientifica (Cairo). 2012 (2012) 1–15. doi:10.6064/2012/964862.

[23] J.E. Hobbs, A. Zakarija, D.L. Cundiff, J. a Doll, E. Hymen, M. Cornwell, S.E. Crawford, N. Liu, M. Signaevsky, G. a Soff, Alternatively spliced human tissue factor promotes tumor growth and angiogenesis in a pancreatic cancer tumor model., Thromb. Res. 120 Suppl (2007) S13-21. doi:10.1016/S0049-3848(07)70126-3.

[24] R. von Känel, R.A. Nelesen, P.J. Mills, M.G. Ziegler, J.E. Dimsdale, Relationship between heart rate variability, interleukin-6, and soluble tissue factor in healthy subjects., Brain. Behav. Immun. 22 (2008) 461–468. doi:10.1016/j.bbi.2007.09.009.

[25] M. Levi, T. van der Poll, Two-Way Interactions Between Inflammation and Coagulation, Trends Cardiovasc. Med. 15 (2005) 254–259. doi:10.1016/j.tcm.2005.07.004.

[26] N.C. Nicolaides, P.M. Sass, L. Grasso, Monoclonal antibodies: a morphing landscape for therapeutics, Drug Dev. Res. 67 (2006) 781–789. doi:10.1002/ddr.20149.

[27] H.M.. Azzazy, W.E. Highsmith, Phage display technology: clinical applications and recent innovations, Clin. Biochem. 35 (2002) 425–445. doi:10.1016/S0009-9120(02)00343-0.

[28] W.G.T. Willats, No Title, Plant Mol. Biol. 50 (2002) 837–854. doi:10.1023/A:1021215516430.

[29] S.J. Kennel, T. Lankford, L. Foote, M. Wall, S. Davern, Phage Display Selection of scFv to Murine Endothelial Cell Membranes, Hybrid. Hybridomics. 23 (2004) 205–211. doi:10.1089/1536859041651295.

[30] S.M. Meiring, J. Vermeulen, P.N. Badenhorst, Development of an inhibitory antibody fragment to human tissue factor using phage display technology, Drug Dev. Res. 70 (2009) 199–205. doi:10.1002/ddr.20295.

Chapter 2: Tissue factor: a review.

List of abbreviations

Active site inactivated factor VIIa	Factor VIIai
Activated Protein C	APC
Alternatively spliced tissue factor	TFas
Antithrombin	AT
Disseminated intravascular coagulation	DIC
Full length tissue factor	TFfl
Heparin induced thrombocytopenia	HIT
International normalised ratio	INR
microRNAs	miRNAs
Protease-activated receptor -2	PAR-2
Recombinant nematode anticoagulant c2	rNAPc2
Recombinant tissue factor pathway inhibitor	rTFPI
RNA interference	RNAi
Tissue factor-factor VIIa complex	TF:FVIIa
Tissue factor	TF
Tissue Factor Pathway Inhibitor	TFPI
Vascular endothelial growth factor	VEGF
Venous thromboembolism	VE
Von Willebrand factor	vWF

Chapter 2: Tissue factor: a review.

2.1 The role of tissue factor

Tissue factor (TF) (formerly known as thromboplastin) is a member of the cytokine receptor superfamily and binds both activated and un-activated forms of factor VII [1,2]. The tissue factor-factor VIIa (TF:FVIIa) complex has both procoagulant- as well as signalling functions [3,4]. Under normal haemostasis TF participates in the low level basal activation of the coagulation cascade, also referred to as “idling”, which occurs in healthy individuals [2]. Tissue factor is the primary physiological initiator of coagulation in response to mechanical- or chemical damage [5,6]. It is constitutively expressed by a variety of extra-vascular tissues especially in organ capsules, cells of the epidermis, mucosal epithelium and the vascular adventitial layer [7–9]. This specific distribution of tissue factor provides a primary haemostatic barrier surrounding blood vessels and organs for the immediate activation of blood coagulation in response to injury [3,4]. Intravascular cells, in contrast, do not express tissue factor constitutively but expression 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 and α-thrombin), hormones, endotoxins, virus infections, immune complexes, modified lipoproteins as well as biogenic amines [8,10–13].

Aberrant expression of tissue factor has been associated with arterial thrombosis, venous thrombosis as well as disseminated intravascular coagulation [2,14]. Atherosclerotic plaques in humans contain large amounts of tissue factor that is responsible for thrombosis after plaque rupture [2,15,16]. Elevated levels of tissue factor is also responsible for the high rate of venous thrombosis observed in patients with breast and pancreatic cancer [2,14].

Apart from its fundamental role in haemostasis and thrombosis, tissue factor also has several non-haemostatic roles [3]. Tissue factor mediates the activation of protease-activated receptor -2 (PAR-2) signalling by forming a complex with factor VIIa and factor Xa. PAR signalling functions in a variety of biological processes including, the induction of intracellular Ca^{2+} signals, inflammation, the regulation of metastatic behaviour of melanoma cells, cell migration as well as angiogenesis [3,17,18]. Tissue factor’s role in angiogenesis and vasculogenesis is especially important as it is an "immediate early gene" necessary for embryonic survival [19–21]. Tissue factor deficient blood vessels lack outer muscle wall and are unable to handle the increasing blood pressures as the embryo develop, resulting in blood vessel rupture and ultimately embryo death due to massive bleeding [10,19].

2.2 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 blood haemorrhage [22,23].The tissue factor (previously known as the extrinsic) pathway of blood coagulations can be summarized as an extracellular cascade reaction that is initiated by tissue

factor, that involves the activation of a series zymogen proteases: factor VIIa, factor Xa, and thrombin (factor IIa), that lead to the generation of fibrin in the presence of phospholipids and Ca^{2+} (Figure 2.1) [24].

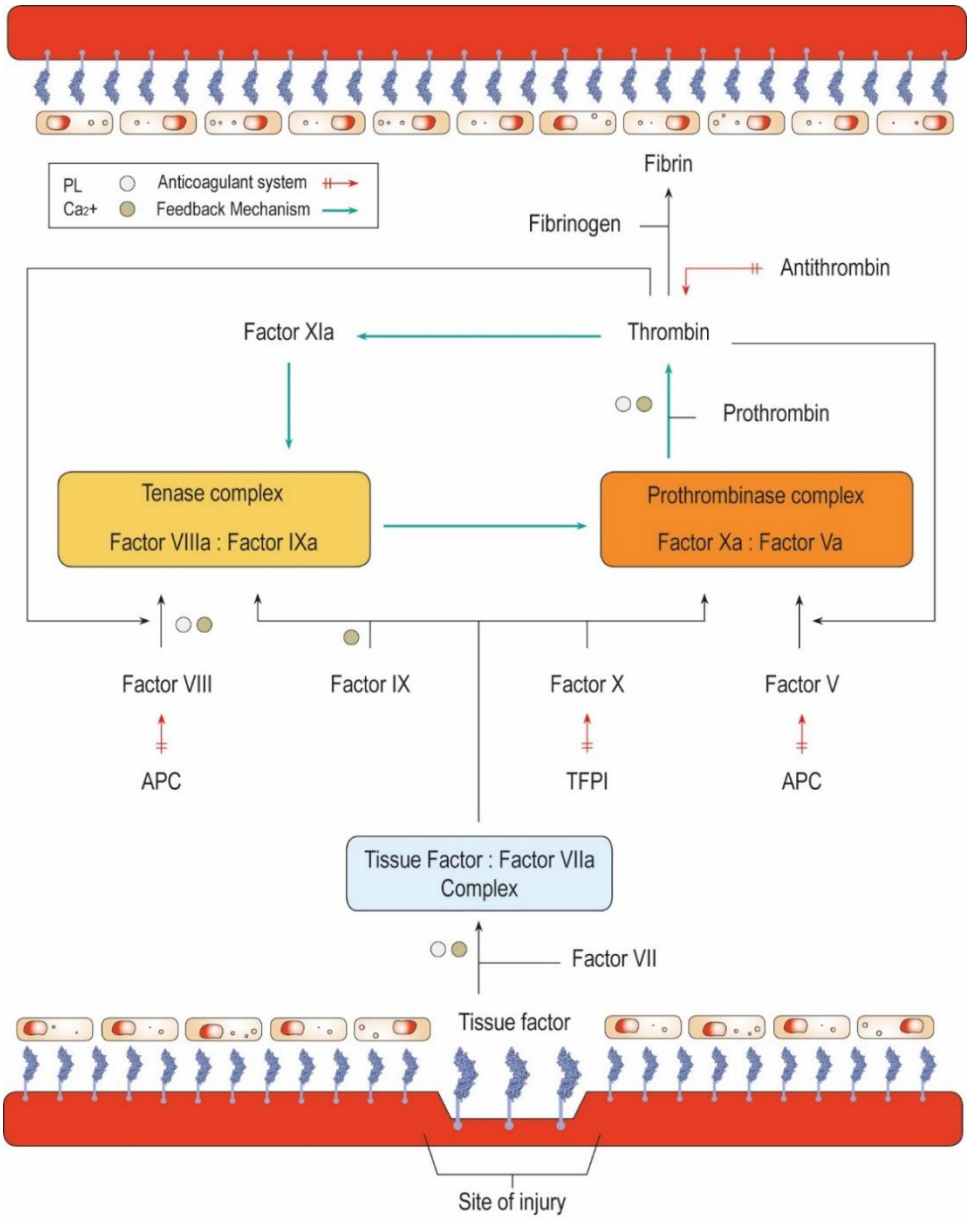


Figure 2.1: Summary of the tissue factor coagulation pathway.

Upon damage, constitutively expressed tissue factor is exposed to circulation zymogen factor VII. The association between tissue factor and factor VII results in the allosterical activation of factor VIIa. The tissue factor: factor VIIa complex (TF:FVIIa Complex) is responsible for the activation of both factor IX and X. Factor IXa associates with factor VIIIa (co-factor) to form the Tenase Complex which is responsible for the activation of additional factor X. The generated factor Xa associates with its factor Va (co-factor) to form the Prothrombinase complex that is responsible for the conversion of inactive prothrombin to the initial thrombin. Thrombin activates inactive fibrinogen to fibrin as well as factor XI. The positive feedback loop for the generation of the majority of thrombin is indicated by the arrows in teal. Thrombin also activates additional factor V and VII via a feedback mechanism. The anticoagulant systems limit the coagulation cascade by targeting activated factors V, VIII, IX, X and thrombin. The interaction of phospholipids PL and Ca^{2+} are indicated by coloured circles. Summarised from: [22,25,38,40]

The modern view of the extrinsic coagulation cascade includes the fundamental role of cellular surfaces [25,26]. The cellular model of coagulation occurs in three phases; 1) initiation 2) amplification and 3) propagation and involves a series of cascade reactions which requires tissue factor bearing cells and platelets [25,27,28]. The initiation phase (Figure 2.2 A) occurs on the surface of tissue factor bearing cells [25]. Tissue factor on injured sub endothelium or decrypted tissue factor on circulating vascular cells initiate coagulation through its interaction with factor VII [28–30]. Factor VII is a vitamin K dependent plasma protein circulating freely in both active (1%, Factor VIIa) and the inactive (99%, Factor VII) forms [23,31]. Tissue factor binds to both forms of factor VII and forms the active tissue factor - factor VIIa complex (TF:FVIIa). This complex is the most potent activator of the coagulation cascade known [23]. The TF:FVIIa complex rapidly activate more inactive factor VII to Factor VIIa to further support the initiation phase [25]. The active tissue factor – factor VIIa complex activates small amounts of both factor IX and factor X via proteolytic cleavage. The generated factor IXa can dissociate and move to the surfaces of nearby platelets or other cells. In contrast to factor IXa, any factor Xa that disassociates from the membrane surface is immediately inhibited by tissue factor pathway inhibitor and antithrombin [25,28]. This restricts factor Xa to the surfaces of tissue factor-bearing cells and prevent unwanted coagulation. The factor Xa located on the cell surface associate with factor V to form a prothrombinase complex which catalyses the conversion of prothrombin (factor II) to thrombin (factor IIa) [25,32,33]. The initiation phase of coagulation is responsible for the generation of a small amount of thrombin (< 4% of total thrombin) [16,29,34]. Under normal physiological condition the initiation stage of coagulation counteracted by tissue factor pathway inhibitor (TFPI) and antithrombin [35,36]. Two isoforms of TFPI have been identified that downregulate coagulation by directly associating with factor X and as such prevents the association of factor X with the TF:FVII complex [37]. Antithrombin (AT) is one major natural anticoagulant and function as a serine proteases inhibitor that targets factors IXa, Xa as well as thrombin [38,39].

During the amplification phase (Figure 2.2 B) the small amount of thrombin that was generated on the surfaces of tissue factor bearing cells diffuse away to perform several functions. One of these is to bind to the surfaces of platelets and activates them [25]. The activation of platelets results in morphological changes, shuffling of membrane phospholipids to create a procoagulant membrane surface, and release of granule contents which releases agonists to induce further platelet activation [25,42]. Thus thrombin also activates factor V and XI on the platelet surface [25,42]. In addition, thrombin activates factor VIII proteolytically which result in the release of Willebrand factor (vWF) to mediate additional platelet adhesion and aggregation [25]. The released FVIII is subsequently activated by thrombin to factor VIIIa which enhances the function of factor IXa serine proteinase that is involved in the more sustained generation of thrombin [25,42]. The protein C, S and thrombomodulin anticoagulant systems that limits the procoagulant activity to the site of injury [26]. The Protein C and S are vitamin K-dependant plasma protein that plays a major role in controlling coagulation [43]. Protein C is activated (APC) when thrombin complexes with thrombomodulin, an endothelial surface transmembrane glycoprotein [39]. Activated protein C combined with its cofactor protein S inactivates factors Va and VIIIa which in turn effectively limits further thrombin generation [44].

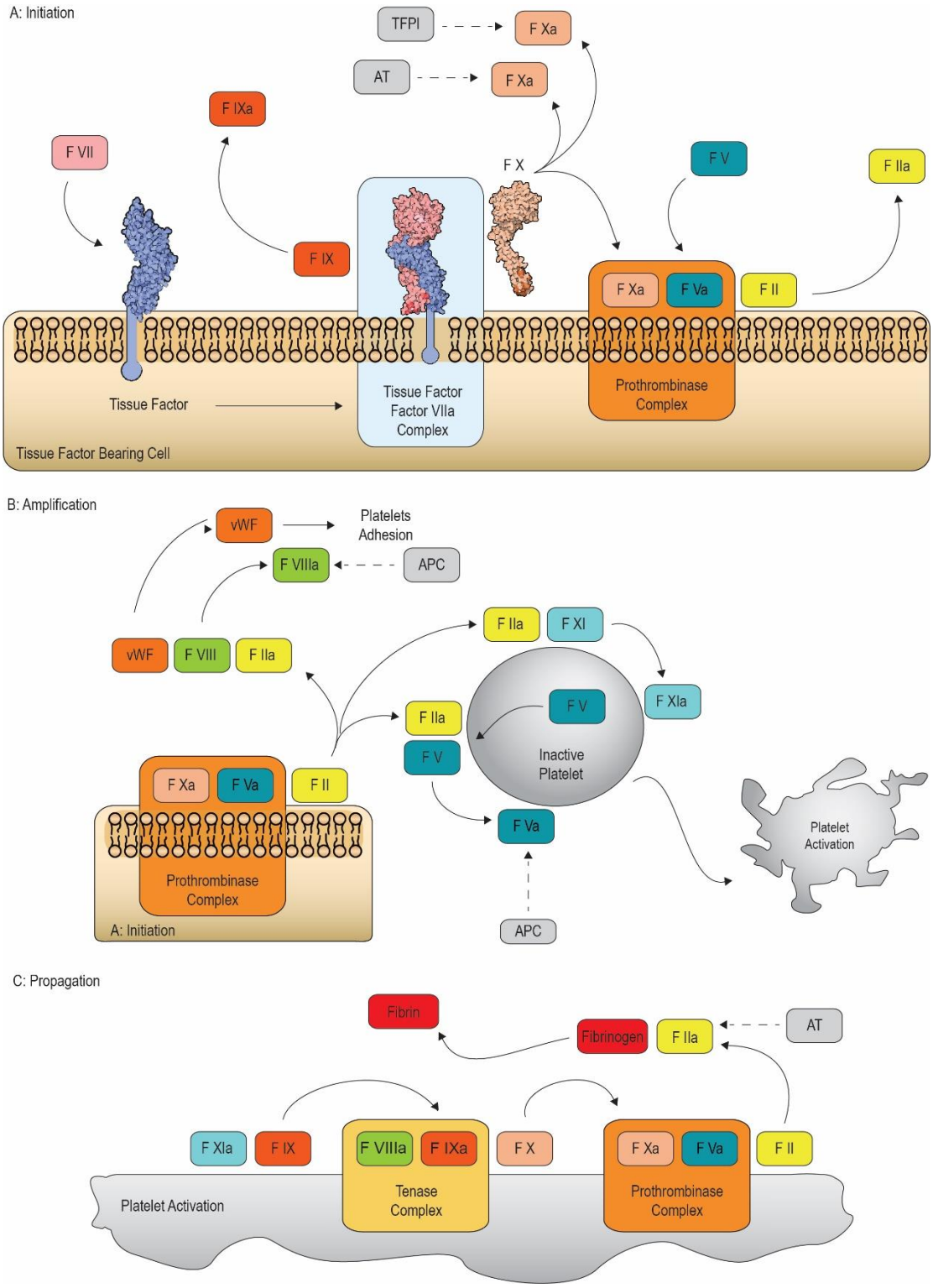


Figure 2.2: The cell-based model of haemostasis.

The coagulation cascade occurs in three overlapping stages: A) Initiation, which occurs on a tissue factor bearing cell; B) Amplification, in which platelets and cofactors are activated in preparation to large scale thrombin generation C) Propagation which occurs on the surface of platelet. The anticoagulant systems counteract the coagulation cascade in order to maintain haemostasis. Summarised from: [27,28,43]

The propagation phase (Figure 2.2 C) occurs on the surfaces of activated platelets [28]. Factor IXa that is generated during the initiation phase by the TF:FVIIa complex, now combines with the factor VIIIa (generated during the cleavage of vWF from factor VIII) on the platelet surface [28]. The formation of the intrinsic tenase complex (FIXa– FVIIIa) on the surface of activated platelet results in the rapid generation of factor Xa [27]. The factor XIa (generated during the amplification phase by thrombin) will facilitate the generation of additional factor IXa. The factor Xa generated on the surface of the activated platelet rapidly combines with factor Va to from the prothrombinase complex which rapidly coverts inactive prothrombin to active thrombin [25,41]. It is this prothrombinase complex that is responsible for the rapid generation of large quantities of thrombin [16,45,46]. This large quantities of thrombin is required for the rapid formation of the fibrin clot through the activation of fibrinogen to fibrin and factor XIII [32]. Fibrin assists in the activation of additional factor XIII which stabilizes the thrombus through cross-linking of the fibrin fibres [29]. 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 [47–49].

The underlying physiological function of the tissue factor pathway is therefore to localize the rapid generation of large quantities of thrombin in response to vascular trauma, in order to prevent bleeding [50]. However, although thrombin plays an essential role in coagulation, it is important to bear in mind that it is tissue factor which is responsible for the initiation of the coagulation cascade [34]. As a result, elevated levels of tissue factor expression have been implicated in the pathogenesis of various thrombotic disorders [1].

2.3 The genetic variants of tissue factor

Tissue factor exhibits a distinct non-uniform tissue distribution, with high levels of expression in vascularized organs such as lung, brain, and placenta; intermediate expression levels in the heart, kidney, intestine, testes, and uterus; and low levels in the spleen, thymus, and the liver [51]. The expression of the human tissue factor gene (F3 gene) is regulated by a number of distinct transcription factors that interact with distinct regions of the tissue factor promoter region. As a result, gene expression and regulation varies with relation to cell types as well as exposure to external and certain stimuli [8,52].The human tissue factor gene is located on chromosome one region 1p21-22 [53,54]. Only two transcript variants of tissue factor have been identified, the primary transcript (also referred to as full length tissue factor (TFfl)), and the alternative transcript know as alternatively spliced tissue factor (TFas) (Figure 2.3) [55]. The tissue factor gene is organized into six exons and five introns and spans a total of 12,4 kBp [56]. Exon 1 correlates to the translation initiation sites and encodes for a 32 amino acid propeptide which is proteolytically cleaved during transport to the plasma membrane [54]. Exons 2-5 encode the 219 amino acid extracellular domain which facilitates interactions with coagulation factors VII and X which is responsible for the mediation of the coagulant activity [11–13,57,58]. Exon 6 encodes the 23 amino acid transmembrane domain was well as 21 amino acid cytoplasmic tail [8,59]. The transmembrane domain anchors tissue factor to the cell surface

which is required to achieve full pro-coagulant activity, while cytoplasmic domain regulates PAR 2 signal transduction during angiogenesis [60–62].

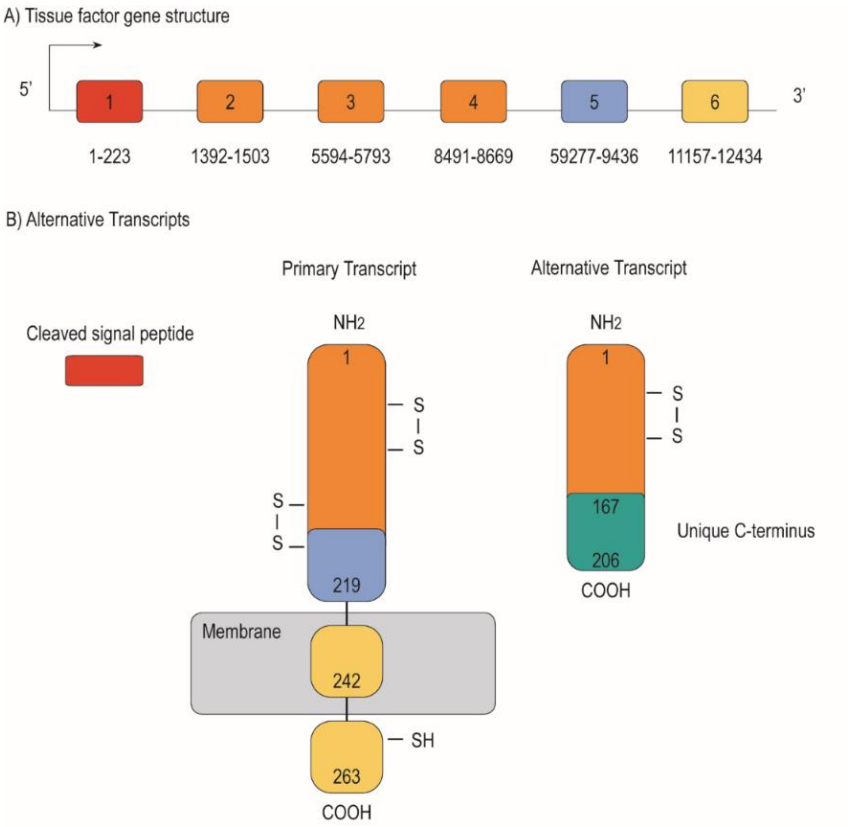


Figure 2.3: Tissue factor gene structure.

A) 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. B) 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 the 32 amino acid N-terminus which serves as a signal peptide in intercellular transport and is cleaved from the mature protein. The primary transcript consists of 3 domains; a 219 amino acid extracellular domain, a 23 amino acid membrane domain and a 21 amino acid cytoplasmic domain. 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 ( ). Summary of: [8,18,54,56,59,62,68,69]

As a result, the mature TFfl transmembrane protein consists of 263 amino acid residues [63]. The TFas is expressed as a 206 residue soluble protein [18,59,64]. During transcription exon 4 is spliced directly into exon 6 which results in an alteration of the reading frame [65]. Amino acid residues 1-166 of TFas are identical to the main transcript, while residues 167-206 represent a unique C-terminus [21,59]. This alteration to the protein structure result in the loss of the transmembrane domain [60]. As a consequence, TFas is believed to be unable to effectively initiate coagulation, since an association with a membrane is required in order to achieve full pro-coagulant activity

[66,67]. Some groups proposed that TFas is inactive while circulating within blood and only becomes pro-coagulant when they co-localize with platelets and are exposed to phospholipids [18].

2.4 Molecular structure of tissue factor

The full-length tissue factor is a 263 residue transmembrane glycoprotein has been classified as a member of the class two cytokine receptor super-family and type III fibronectin family. Proteins of the cytokine family are cell-surface molecules with a single transmembrane domain and a cytoplasmic domain with structural diversity [70]. A unique characteristic of this family is a distinct disulfide bond in the extracellular domain further dividing the group into class 1 and class 2 receptors of which the tissue factor forms part of the latter [56]. The extracellular domain of tissue factor, which is responsible for interactions with factor VII and factor X, is composed of two type III fibronectin like domains arranged at approximately 125 degrees and shows an extensive interface between domains [71]. The fibronectin type III motive is one of the most common found in higher organisms. The fibronectin type III structures generally mediate protein-protein interactions by serving as a 'spacer' in order to achieve biological function [72]. Fibronectin domains consist of approximately 100 residues that are arranged into two overlapping beta sheets with the top sheet containing three antiparallel beta strands and the bottom sheet containing four beta strands forming the characteristic antiparallel beta-sandwich. Interactions between hydrophobic amino acids on the inner side of the sandwich stabilises the structure [71,73]

The extracellular domain of tissue factor contains an extended finger like region that consists of a short alpha helix and beta sheet which protrudes from the side of the main molecule body at the intermolecule interface with factor VIIa. This loop regions is highly conserved for tissue factor amongst mammalian species [70]. Several alanine-screening mutagenesis studies have been performed to identify the tissue factor residues responsible for tissue factor and factor VIIa interaction as well as the interactions between the TF:FVIIa complex and factor X (Figure 2.4) [71]. The crystal structure of factor VII in complex with the extracellular part of tissue factor reveals an extensive interface between the two proteins encompassing all four domains of factor VIIa [74]. The protein-protein interactions between factor VII and TF involves residues from domain 1 and an extended loop (finger like region) of domain 2 of the extracellular region TF. When factor VII associates with tissue factor it adopts an extended conformation with the light chain wrapping around the central structure (scaffolding) of tissue factor [75]. This allosteric modification serves to stabilize the active site of factor VII which results in $2x10^7$ increase in activity when in association with appropriate membrane. In contrast, the proteolytic activity of factor VII towards factor X is negligible in the absence of membrane bound tissue factor [17,56,76,77]. The interaction between tissue factor – factor VII complex and factor X occurs via domain 2 on tissue factor (Figure 2.5). The zymogen forms of factor VII, IX and X are transformed into their active forms via proteolytic cleavage of the activation peptide (N-terminal part of the protease domain) [78]. It is the formation of the so called ternary complex that is responsible of the initiation phase of coagulation [36].

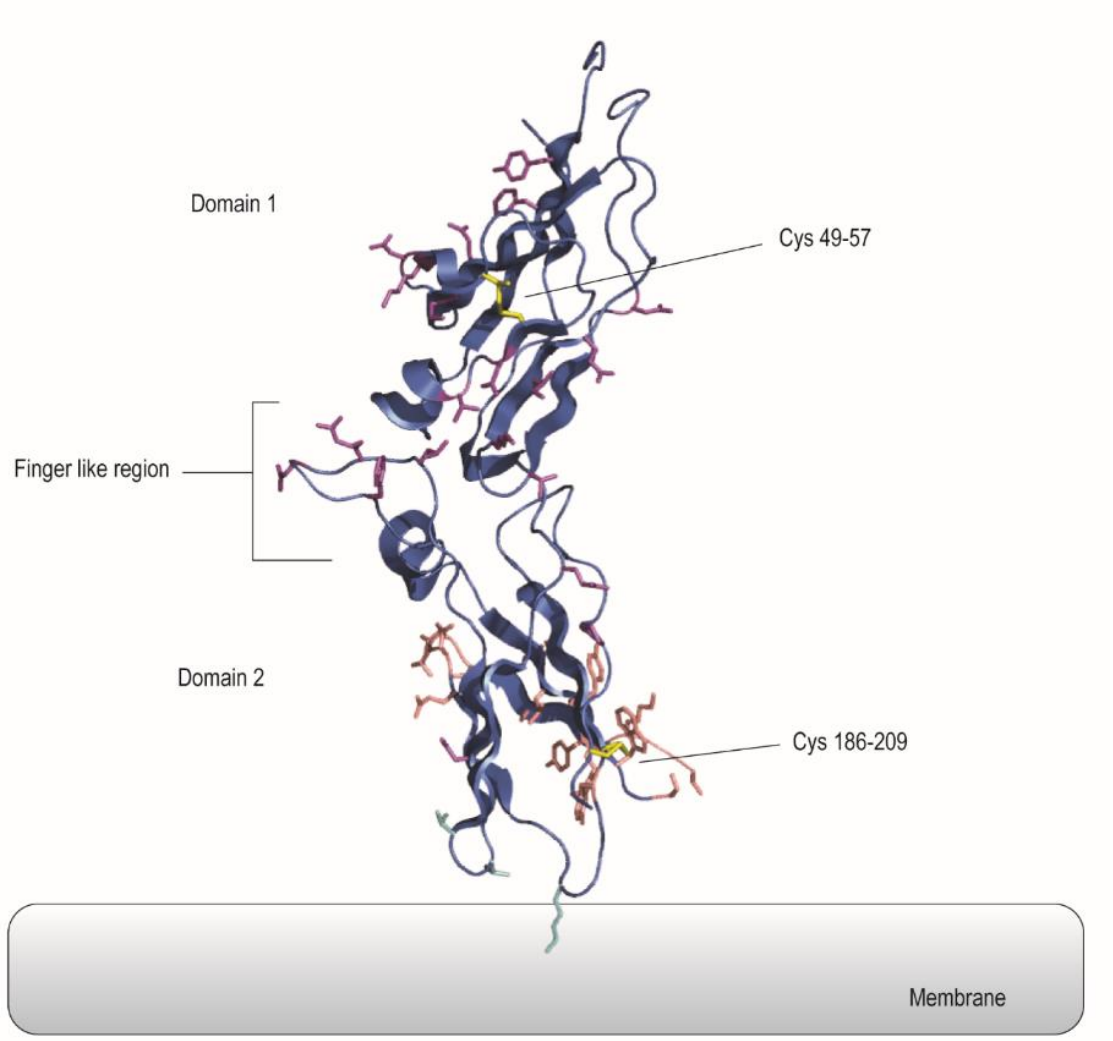


Figure 2:4: The extracellular domain of tissue factor.

The ribbon diagram of the structure of the extracellular domain of TF on an illustrated lipid membrane surface at 2.20 Å resolution (PDB no:1Boy). Highlighted in magenta are residues important for TF interaction with factor VIIa (Thr[17], Lys[20], Ile[22], Glu[24], Gln[37], Asp[44], Lys[46], Lys[48], Asp[58], Thr[60], Phe[76], Tyr[78], Gln[110], Leu[133], Arg[135], Phe[140], and Val[207]). Highlighted in cream are residues important for the interaction with factor X (Thr[154]-Glu[174] and Tyr[185]). Highlighted in cyan are the residues important for TF interaction with the membrane (Gln[118], Val[119], Thr[121], Lys[159], Asp[180], Lys[181], and Glu[183]). Indicated in yellow are the two disulfide bridges of tissue factor at positions Cys[49]-Cys[57] and Cys[186]- Cys[209].Summarised From: [56,79]

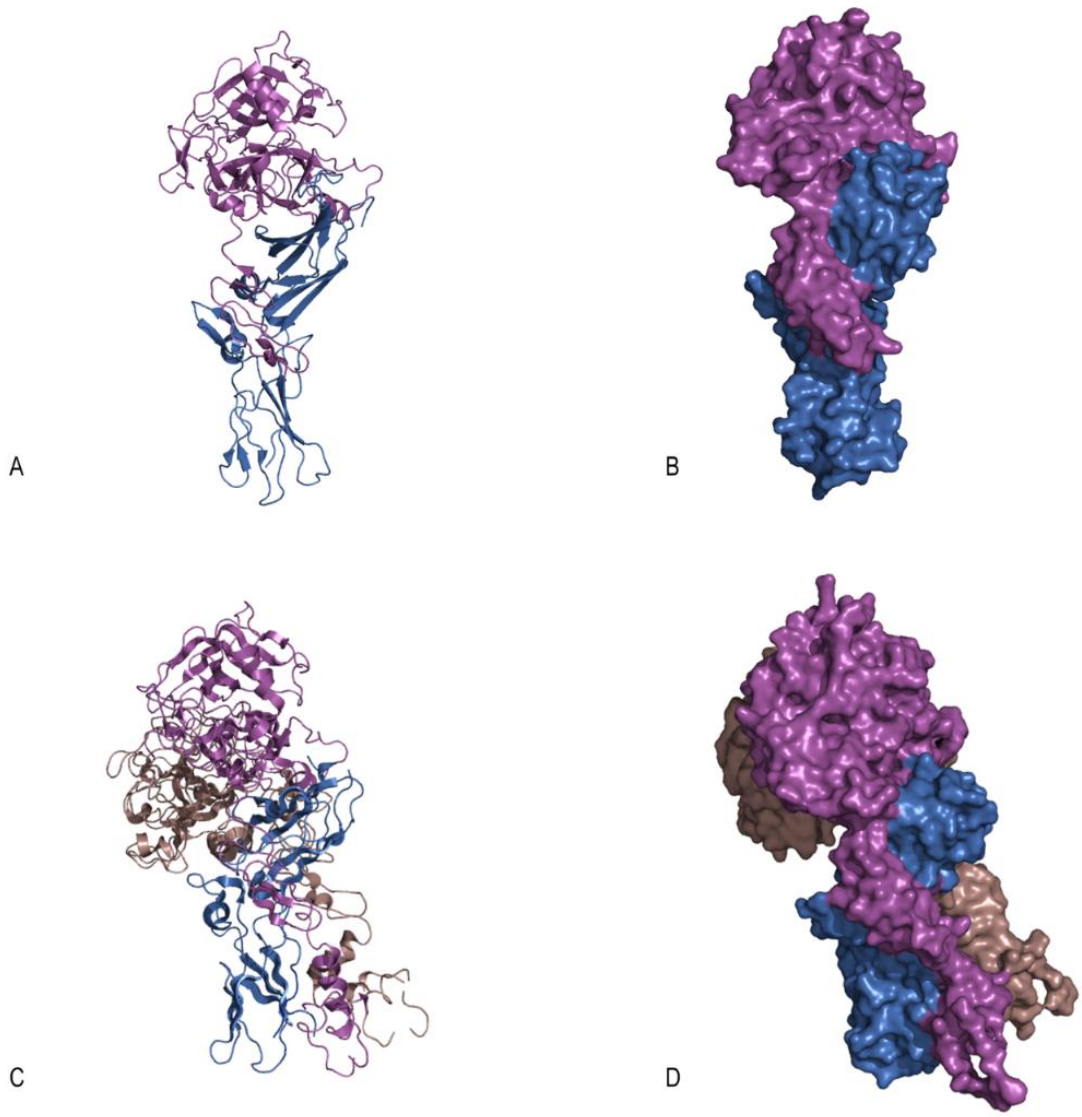


Figure 2.5: The ternary complex of coagulation.

*The cartoon structure of the ternary complex responsible for the initiation stage of coagulation. The membrane is presumed to be at the bottom of the protein structures A) The extracellular domain of tissue factor (**blue**) in complex with factor VIIa (**magenta**) at 2.90 Å resolution (PDB no: 1J9C). B) The surface structure of tissue factor – factor VIIa complex. C) A theoretical model of the ternary complex between tissue factor (**blue**), factor VII with N-terminal uncleaved (**magenta**) and factor X (**cream**) (See PDB no:1NL8 and supplementary video). D) The solvent accessible surface area of the ternary structure. Tissue factor serves as scaffold sandwiched between factor VII and X in order to facilitate the association between these proteins during the initiation face of coagulation. Summarised from* [74,78,80]

2.5 Tissue factor bearing microparticles

Although it was originally believed that tissue factor was predominantly present on the surface of extra vascular cells, it has subsequently also been detected circulating freely in association with microparticles in healthy subjects [69]. TF bearing microparticles are sub-micrometric (0.1 - 1µm) membrane fragments that are constitutively released from the surface of cells [81]. These membrane fragments lack a nucleolus and are incapable of protein synthesis [82].

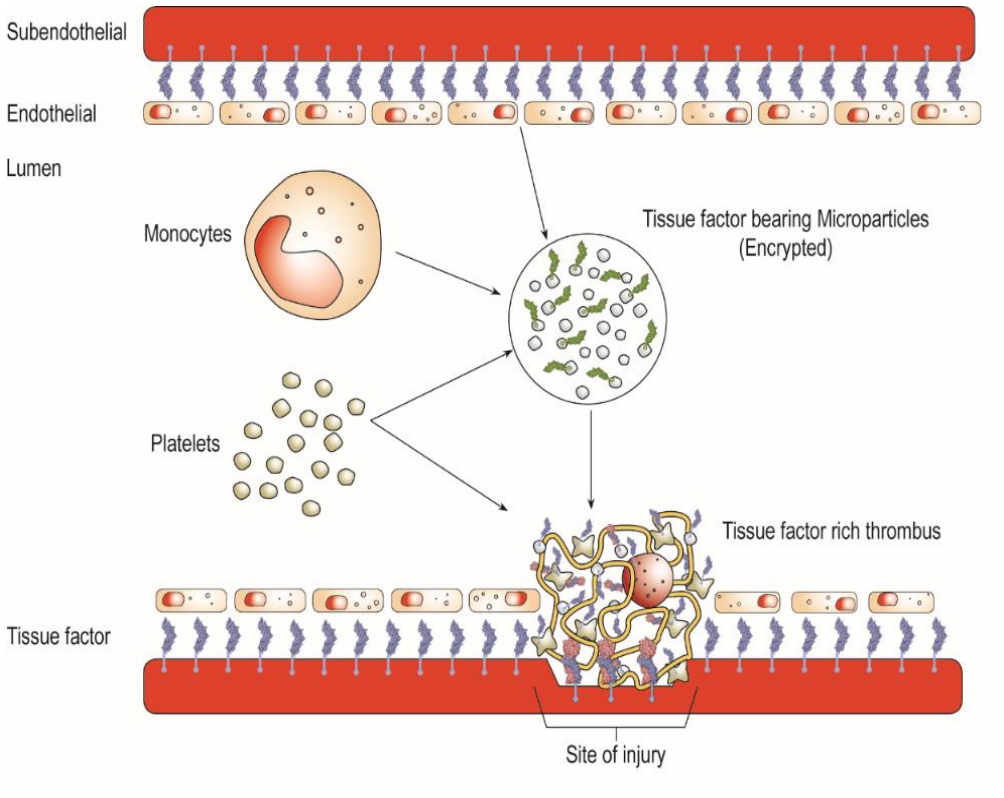


Figure 2.6: Sources of tissue factor at the site of injury.

Upon vessel damage, coagulation is initiated by constitutively expression of extra-vascular tissue factor. As the thrombus forms a tissue factor rich micro-environment is maintained by various sources of blood borne tissue factor. Platelets express tissue factor upon activation, inducing tissue factor expression in surrounding cells when incorporated into the thrombus and contribute to the pool of encrypted tissue factor bearing microparticles. Monocytes and other leukocytes also contribute to the pool of encrypted tissue factor bearing microparticles. 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. Summarised from: [5,6,68,82–85]

Cell activation, apoptosis and shear stress up-regulates the formation of microparticles [85]. Microparticle surface-exposed negatively charged phospholipids are highly prothrombotic and essential for thrombin generation [67]. They also contain many membrane-derived proteins that play a role in the regulation of host immunity and inflammatory responses [51]. Although much disputed, it has been suggested that up to 50% of total tissue factor is present in an inactive state in nonvascular cells [86]. They are believed to circulate in an inactive “encrypted” form [84,87]. This “encrypted form” prevents the inappropriate activation of the coagulation cascade [5,88].

Although microparticles originate from various cell types the majority are derived from platelets (80%), leucocytes, erythrocytes and endothelial cells [67,81]. This pool of circulating tissue factor bearing microparticles are responsible for the supplying of tissue factor to the site of injury during the initial phase of thrombus development [35,86]. Importantly, the levels of circulating tissue factor bearing microparticles are increased in a variety of pathologic states [89]. The sources of tissue factor bearing microparticles and their role in the clot formation is summarised in Figure 2.6.

2.6 Pathophysiological role of tissue factor

2.6.1 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 [1]. Hyperglycaemic conditions have also been associated with an increased tissue factor expression in monocytes, elevated levels of tissue factor-bearing microparticles and an increased activity of circulating tissue factor [90–92]. This aberrant expression, accompanied with increased tissue factor activity is believed to be responsible for the high cardiovascular mortality and thrombotic complications observed in type 2 diabetes mellitus [93].

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 and is broadly 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 [90,94,95]. 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 TF - FVIIa complex [96,97]. This leads to an increase in tissue factor expression and enhances thrombotic tendency which lead to cardiovascular complications [9,95].

2.6.2 Tissue factor in atherosclerosis

Tissue factor contributes to the pathophysiology of atherosclerosis and acute coronary syndromes (unstable angina and myocardial infarction) that are associated with plaque rupture [45,98–100]. 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 [1,9,57,101]. Blood borne tissue factor levels are also increased due to the release of inflammatory cytokines [1,102]. Inflammation further enhances the thrombogenicity of atherosclerotic plaques through the down regulation of the physiological anticoagulant systems [9]. Very high levels

of tissue factor are present within an atherosclerotic plaque. This is a result of 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 [57,99–101].

Once plaque rupture occurs the highly thrombotic core is exposed to the circulating clotting factors and platelets in the blood stream which lead to the rapid and excessive thrombin generation via the tissue factor coagulation pathway [57,68,100–102]. 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 [57,68,102]. Activated platelets also contribute to the pool of blood borne tissue factor by shedding parts of their plasma membrane as tissue factor-bearing microparticles [57]. 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 [98,101]. Tissue factor is also considered as the principle molecule linking fibrous plaque disruption to thrombo- and pulmonary embolism [12,98].

2.6.3 Malignancy and thrombosis

Although the molecular mechanism still remains unclear, there is a well-established association between certain malignancies and thrombosis [86,103]. 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 some malignancies [4,104–106]. Malignant disease increases the risk of disseminated intravascular coagulation (DIC) and venous thromboembolism (VE) [104,105,107]. The risk of thrombotic complications is further increased with the administration of hormone replacement and chemotherapies [106,108]. It was reported by Zwicker *et al.* (2007) 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 [104]. In breast cancer, for example, the aberrant expression of tissue factor stimulates tumour angiogenesis in part through the up-regulation of vascular endothelial growth factor (VEGF) [109]. Indeed, a variety of malignant cells exhibit enhanced tissue factor activity through excessive expression of tissue factor on lipid rich surfaces [103,104]. This partially explains why the risks of thrombotic complications are not uniform across all forms of tissue factor expressing malignant cells [106]. Pulmonary embolism is common in ovarian, pancreatic, brain and bone cancers but not in head- and neck cancers [106,108]. Tissue factor is also constitutively expressed by human leukaemia- and lymphoma cells [103,105]. The expression of tissue factor in tumour cells induces the expression of VEGF, which in turn further enhances tumour angiogenesis and tissue factor expression [4,106,109]. Many malignant cells elicit an inflammatory host response that stimulates tissue factor expression in monocytes via pro-inflammatory cytokines [106,110]. Elevated levels of

platelet- and tumour derived microparticles also contribute to thrombogenicity during cancer [106,107]. This aberrant expression coupled with an increase in blood borne tissue factor expression and enhanced thrombotic tendency contribute to pathological complications associated with malignancies.

2.7 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 [111,112]. Despite the popularity and effectiveness of these antithrombotic drugs, treatment is still far from ideal due to clinical as well as practical shortcomings [113].

2.7.1 Warfarin

The antithrombotic effect of warfarin is mediated via vitamin K-dependant coagulation factors (Prothrombin, factors VII, -IX and -X) and is therefore very sensitive to dietary fluctuations of vitamin K [114,115]. 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 [111,112,114,116]. As a result, patients undergoing warfarin treatment require close dosage monitoring in order to maintain the desired international normalised ratio (INR) [117]. 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 foetal abnormalities [114].

2.7.2 Heparin

Unfractionated heparin acts as an antithrombotic by enhancing the ability of anti-thrombin to neutralize the serine active site of thrombin as well as factors IXa, -Xa, -XIa, -XIIa, -XIIIa and plasmin [114,118]. However, heparin binds non-specifically to platelets, endothelial cells, macrophages as well as plasma proteins (platelet factor 4 and high molecular weight von Willebrand multimers). This leads to a reduction in the amount of viable heparin available to interact with functional anti-thrombin [32,114]. In addition, heparin as well as the low molecular weight heparin-anti-thrombin complex is unable to target thrombin bound to fibrin, factor Xa bound to platelets or the prothrombinase complex [119]. Some of the adverse effects of heparin treatment include; bleeding, heparin induced thrombocytopenia (HIT), osteoporosis, renal dysfunction and hematoma or skin necrosis [114,118]. As a result heparin treatment also requires careful monitoring to ensure that the correct dosage is administered [120]. This does not only increase the overall cost of treatment but also contributes to poor patient compliance [49,111]. Despite the relative effectiveness of these drugs, there is still a great need for the development of safe and effective therapies for long term treatment of thrombosis.

2.7.3 Direct thrombin and factor Xa inhibitors

Due to the critical role that thrombin plays within the coagulation cascade several direct thrombin (dabigatran, hirudin, melagatran) and factor X inhibitors (rivaroxaban, apixaban, edoxaban) have been developed in the last 15 years to specifically overcome the limitations of warfarin and heparin anti-coagulants [119,121,122]. These anticoagulants produce a more predictable anti-coagulation response which make them very attractive for use as treatment [47,119]. Direct thrombin inhibitors have potential efficacy advantages as they can inhibit both free thrombin as well as inactivated fibrin-bound thrombin and do not interact with platelet factor 4 or high molecular weight von Willebrand multimers [48,121]. Similarly, direct factor X inhibitors have capacity to inhibit prothrombinase-bound factor Xa and also do not interact with other coagulation factors [121].

Despite these advantages, many of these direct thrombin and factor X inhibitors show little to no improvement on the first generation anti-coagulants in terms of effectiveness [122]. They are still associated with bleeding complications, inability of drug reversal and still require routine laboratory monitoring. [119,122]. It has also been reported that low dosage treatment with anti-thrombin drug, melagatran, significantly worsens the coagulation status [123]. This only emphasises the importance of routine monitoring of the coagulation status of patients irrespective of the predictability of drug action. An additional complication with the use of direct thrombin and factor X inhibitors for routine treatment is the cost, as these drugs are still significantly more expensive to produce than the first generation anti-coagulations [121]. As a result, the search for new anticoagulants is still ongoing. 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 first generation as well as modern anti-coagulants.

2.8 Strategies to tissue factor inhibition

Tissue factor provides an ideal target for the development of novel anticoagulants due to its fundamental role in thrombotic disorders [32,36]. Tissue factor targeting agents interfere with the very first step of coagulation, thereby suppressing the downstream reactions within of the coagulation cascade [1,11,124]. 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 loop responsible for thrombin amplification [124]. Secondly, inhibitors of tissue factor may not affect the integrity of physiological haemostasis as it will only function at the site of vascular injury [11]. Thirdly, the inhibition of tissue factor has reduced bleeding tendency in comparison to first-generation anticoagulants [32,124]. Furthermore, it is believed that relatively low concentrations of tissue factor inhibitor would be required to sufficiently suppress an overactive coagulation. 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 in order to inhibit the initiation phase of coagulation.

2.8.1 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 oligodexinucleotides [125]. These short DNA sequences bind to the major grove 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 grove in the promoter region of the tissue factor gene [126]. A similar approach is the use of microRNAs (miRNAs) to regulate gene expression[127]. These hybridising antisense oligonucleotides bind to tissue factor mRNA oligonucleotides in a process termed RNA interference (RNAi) which prevents gene translation [128–130]. Studies have shown that miRNAs modulated the expression of tissue function under varied physiologic- and pathophysiologic-relevant settings [131]. A hairpin ribozyme that destroys the tissue factor mRNA in vascular smooth muscle has also been developed [1]. An indirect method of inhibiting tissue factor synthesis is by the use of curcumin. It is a naturally 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 [1]. Although these inhibitors are relatively effective *in vitro*, further research is required to determine their effectiveness *in vivo*.

2.8.2 Direct inhibition of tissue factor action

Apart from a multitude of synthetic peptide tissue factor inhibitors, a large array of tissue factor targeting poly-, monoclonal and antibody fragments have also been developed. Most of these antibodies elicit an anti-coagulatory function by binding to tissue factor and in so doing block the association of tissue factor with factor VII [132,133]. Other antibodies block the association of the tissue factor/factor VII with factor Xa [36]. These antibody inhibitors vary in their size, binding site and ability to inhibit the function of tissue factor [45,134]. Interestingly, the relative efficacy of a series of murine derived antibodies is was found to relate to binding site location rather than binding affinities or rate of association. The influence of the epitope location on antibody efficacy was attributed to correlation with key TF residues that are involved in cofactor interaction [134]. Many of the various monoclonal antibodies were found to elicit an antithrombotic effect in animal models, while some polyclonal antibodies were found to reduce thrombogenicity of ruptured atherosclerotic plaques in humans [1]. In an escalating dosage study, a chimeric monoclonal anti-tissue factor antibody reduced thrombin production in patients diagnosed with stable coronary disease with no serious adverse events observed [1,36].

An alternative form of tissue factor, which exhibits reduced catalytic activity has also been developed [1]. 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 key lysine residues (Lys 165 and Lys166) that are required for substrate interaction [36]. Recombinant tissue factor displayed clear antithrombotic properties by forming inactive TF:FVIIa complexes, without disturbing normal haemostasis and did 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 [32,36].

2.8.3 Active site inactivated factor VIIa

Active site inactivated factor VIIa (factor VIIai) is capable of binding to tissue factor but it lacks catalytic activity [1]. Factor VIIai elicits an anticoagulant effect by competing with catalytically functional FVII for tissue factor [45,49]. Interestingly, incorporation of a factor VIIa active site inhibitor leads to a 5 fold increase in binding affinity [32,135]. 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 [1,32,49,134]. In a series thrombosis studies in baboons, FVIIai prevented thrombosis at carotid endarterectomy before restoring flow in the operated vessel [136]. 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 infusion [32]. Similar results have been obtained with other animal models [24]. These studies indicate that inhibitors of TF-dependent coagulation produce antithrombotic effects. Moreover, because FVIIai eliminates thrombin production via TF activation at sites of vascular injury, no significant systemic homeostatic im-pairment occurs compared with other anticoagulant therapies (e.g., aspirin, heparin, or coumarin). The use of FVIIai, therefore, could be an effective and safe strategy for inhibiting the formation of microthrombi during surgery without the major complications of anticoagulation therapy.

2.8.4 Recombinant tissue factor pathway inhibitor

Recombinant tissue factor pathway inhibitor (rTFPI), produced in *Saccharomyces cerevisea*, 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 [45,137]. Recombinant TFPI was found to elicit an antithrombotic effect during balloon injury in animal models and reduced atherosclerotic plaque thrombogenicity in humans. 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. 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 [1] . Further studies are required to evaluate the therapeutic safety and effectiveness of rTFPI.

2.8.5 Nematode anticoagulant protein c2

Recombinant nematode anticoagulant c2 (rNAPc2) is a 85 amino-acid serine protease and a potent inhibitor of the TF:FVIIa complex [36,138]. It has a long half-life and only needs to be administered every other day after subcutaneous injection [45]. NAPc2 was originally identified and isolated from the canine hookworm *Ancylostoma caninum* [32]. The recombinant form (rNAPc2) is one of the most extensively evaluated novel inhibitors of the

TF:FVIIa complex [45]. rNAPc2 interferes with tissue factor function by binding with both factor X or -Xa before forming an inhibitory complex with TF:FVIIa [1,36]. 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 [32,45]. 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 [138]. rNAPc2 was also found to be safe and effective in phase II studies investigating escalating doses in low-risk cardiac patients undergoing angioplasty [1,32]. Similarly, alternative phase II study found that thrombin generation was in a dose-dependent manner with no difference in the rate of bleeding observed between patients receiving either rNAPc2 or placebo [133]. NAPc2 harbours a lot of potential as a novel anticoagulant therapy and may prove to resolve many of the shortcomings of current anticoagulants.

2.9 Conclusion

"It is natural that a striking feature such as coagulation of blood always inspired researches and that it is not surprising that since the beginnings of scientific research there were numerous hypotheses and theories that tried to explain this phenomena" wrote Marawitz in his overview in 1905 [139,140]. Tissue factor has been a key area of research for many years and it highly likely that this will continue to be the case as our body of knowledge and its role with in hemostasis grows. The pathophysiological role of tissue factor in thrombosis, inflammation as well as several other thrombosis related diseases such as diabetes, obesity, some cancers, and even irregular hemostasis observed in HIV patients is well understood. Tissue factor functions as the principle initiator of the coagulation cascade in response to physical or chemical damage. As a result, it provides an ideal target for the development of novel anticoagulants due to its fundamental role in these thrombotic disorders.

Several strategies have been followed to develop anti thrombotic agents that specifically targets tissue factor directly, the TF:VIIa complex, or the interaction of factor X with the TF:VIIa in an attempt to modulate the initiation stage coagulation. However, the successful development of a tissue factor inhibitor to supplement and eventually replace the anti-coagulants that are currently in use has not yet been achieved. Recognising the potential in tissue factor inhibition as a novel approach to antithrombotic therapy, our research group identify a human single chain antibody fragment (scFv) that inhibits the hemostatic functions of tissue factor. The use scFv as an anti-thrombotic agent has some advantages appose to monoclonal antibodies. As it is only a 1/6 the size of a complete IgG, the smaller size allows for the improved penetration of blood vessels and tissue barriers [141]. In addition, scFvs have more stable structures, shorter half-life, and because they are human origin they have reduced immunogenicity that their larger counterparts [142]. It is therefore important that the focus should remain on the development and characterization of novel anti tissue factor inhibitors to treat and prevent the complications associated with aberrant tissue factor expression.

2.10 References

[1] J. Steffel, T.F. Lüscher, F.C. Tanner, Tissue factor in cardiovascular diseases: Molecular mechanisms and clinical implications, Circulation. 113 (2006) 722–731. doi:10.1161/CIRCULATIONAHA.105.567297.

[2] N. MACKMAN, The many faces of tissue factor, J. Thromb. Haemost. 7 (2009) 136–139. doi:10.1111/j.1538-7836.2009.03368.x.

[3] L. Petersen, Tissue Factor-dependent Factor VIIa Signaling, Trends Cardiovasc. Med. 10 (2000) 47–52. doi:10.1016/S1050-1738(00)00041-4.

[4] M. Hoffman, Some things I thought I knew about tissue factor that turn out to be wrong, Thromb. Res. 122 (2008) 73–77. doi:10.1016/S0049-3848(08)70024-0.

[5] N. Mackman, Role of tissue factor in hemostasis and thrombosis, Blood Cells, Mol. Dis. 36 (2006) 104–107. doi:10.1016/j.bcmd.2005.12.008.

[6] S. Butenas, T. Orfeo, K.G. Mann, Tissue factor activity and function in blood coagulation, Thromb. Res. 122 (2008) 42–46. doi:10.1016/S0049-3848(08)70018-5.

[7] N. Mackman, Regulation of the tissue factor gene., FASEB J. 9 (1995) 883–9. http://www.ncbi.nlm.nih.gov/pubmed/7615158.

[8] P. Carmeliet, D. Collen, Molecules in focus Tissue factor, Int. J. Biochem. Cell Biol. 30 (1998) 661–667. doi:10.1016/S1357-2725(97)00121-0.

[9] M. Levi, T. van der Poll, Two-Way Interactions Between Inflammation and Coagulation, Trends Cardiovasc. Med. 15 (2005) 254–259. doi:10.1016/j.tcm.2005.07.004.

[10] T.H. Bugge, Q. Xiao, K.W. Kombrinck, M.J. Flick, K. Holmbäck, M.J. Danton, M.C. Colbert, D.P. Witte, K. Fujikawa, E.W. Davie, J.L. Degen, Fatal embryonic bleeding events in mice lacking tissue factor, the cell-associated initiator of blood coagulation., Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 6258–63. doi:10.1073/pnas.93.13.6258.

[11] P. Golino, The inhibitors of the tissue factor:factor VII pathway, Thromb. Res. 106 (2002) V257–V265. doi:10.1016/S0049-3848(02)00079-8.

[12] Y. Kamikura, H. Wada, T. Nobori, T. Kobayashi, T. Sase, M. Nishikawa, K. Ishikura, N. Yamada, Y. Abe, J. Nishioka, T. Nakano, H. Shiku, Elevated levels of leukocyte tissue factor mRNA in patients with venous thromboembolism, Thromb. Res. 116 (2005) 307–312. doi:10.1016/j.thromres.2004.12.013.

[13] Z.-C. Peng, X. Cai, Y.-G. Zhang, D.-S. Kong, H.-S. Guo, W. Liang, Q.-Q. Tang, H.-Y. Song, D. Ma, A novel anti-tissue factor monoclonal antibody with anticoagulant potency derived from synthesized multiple antigenic peptide through blocking FX combination with TF, Thromb. Res. 121 (2007) 85–93. doi:10.1016/j.thromres.2007.03.001.

[14] A.S. Wolberg, A.E. Mast, Tissue factor and factor VIIa - Hemostasis and beyond, Thromb. Res. 129 (2012) S1–S4. doi:10.1016/j.thromres.2012.02.017.

[15] T. Luther, Tissue Factor in the Heart Multiple Roles in Hemostasis, Thrombosis, and Inflammation, Trends Cardiovasc. Med. 11 (2001) 307–312. doi:10.1016/S1050-1738(01)00129-3.

[16] S. Butenas, T. Orfeo, K.E. Brummel-Ziedins, K.G. Mann, Tissue factor in thrombosis and hemorrhage., Surgery. 142 (2007) 2–14. doi:10.1016/j.surg.2007.06.032.

[17] D. Kirchhofer, Y. Nemerson, Initiation of blood coagulation: The tissue factor/factor VIIa complex, Curr. Opin. Biotechnol. 7 (1996) 386–391. doi:10.1016/S0958-1669(96)80112-1.

[18] V.Y. Bogdanov, V. Balasubramanian, J. Hathcock, O. Vele, M. Lieb, Y. Nemerson, Alternatively spliced human tissue factor: a circulating, soluble, thrombogenic protein., Nat. Med. 9 (2003) 458–462. doi:10.1038/nm841.

[19] L. Beck, P. a D'Amore, Vascular development: cellular and molecular regulation., FASEB J. (1997) 365–373. doi:10.1002/term.394.

[20] T. Watanabe, M. Yasuda, T. Yamamoto, Angiogenesis Induced by Tissue Factor in Vitro and in Vivo, Thromb. Res. 96 (1999) 183–189. doi:10.1016/S0049-3848(99)00101-2.

[21] N. Mackman, Role of Tissue Factor in Hemostasis, Thrombosis, and Vascular Development, Arterioscler. Thromb. Vasc. Biol. 24 (2004) 1015–1022. doi:10.1161/01.ATV.0000130465.23430.74.

[22] J.H. McVey, Tissue factor pathway., Baillie re's Best Pract. Res. Clin. Haematol. 12 (1999) 361–72. http://www.ncbi.nlm.nih.gov/pubmed/10856975.

[23] D.P. Norris, N. Brockdorff, S. Rastan, Methylation status of CpG-rich islands on active and inactive mouse X chromosomes, Mamm. Genome. 1 (1991) 78–83. doi:10.1007/BF02443782.

[24] A.J. Chu, Tissue factor mediates inflammation, Arch. Biochem. Biophys. 440 (2005) 123–132. doi:10.1016/j.abb.2005.06.005.

[25] S.A. Smith, The cell-based model of coagulation, J. Vet. Emerg. Crit. Care. 19 (2009) 3–10. doi:10.1111/j.1476-4431.2009.00389.x.

[26] M. Hoffman, Remodeling the Blood Coagulation Cascade, J. Thromb. Thrombolysis. 16 (2003) 17–20. doi:10.1023/B:THRO.0000014588.95061.28.

[27] M. Hoffman, D.M. Monroe, A cell-based model of hemostasis., Thromb. Haemost. 85 (2001) 958–65. doi:01060958 [pii].

[28] A.L. Ruseva, A.. Dimitrova, A NEW UNDERSTANDING OF THE COAGULATION PROCESS – THE CELL-BASED MODEL, J. Biomed. Clin. Res. 4 (2011) 17–22.

[29] B. Dahlbäck, Blood coagulation, Lancet. 355 (2000) 1627–1632. doi:10.1016/S0140-6736(00)02225-X.

[30] G.C. Price, S.A. Thompson, P.C.A. Kam, Tissue factor and tissue factor pathway inhibitor, Anaesthesia. 59 (2004) 483–492. doi:10.1111/j.1365-2044.2004.03679.x.

[31] P. Hellstern, H. Beeck, A. Fellhauer, A. Fischer, B. Faller-Stöckl, Measurement of factor VII and of activated factor VII in healthy individuals and in prothrombin complex concentrates, Thromb. Res. 86 (1997) 493–504. doi:10.1016/S0049-3848(97)00098-4.

[32] R. Frédérick, C. Charlier, B. Marsereel, L. Pochet, C. Charlier, B. Masereel, Modulators of the coagulation cascade: focus and recent advances in inhibitors of tissue factor, factor VIIa and their complex., Curr. Med. Chem. 12 (2005) 397–417. http://www.ncbi.nlm.nih.gov/pubmed/15720249.

[33] H.L. VOS, Inherited defects of coagulation Factor V: the thrombotic side, J. Thromb. Haemost. 4 (2006) 35–40. doi:10.1111/j.1538-7836.2005.01572.x.

[34] Y. Nemerson, Tissue factor and hemostasis., Blood. 71 (1988) 1–8. http://www.bloodjournal.org/content/71/1/1.abstract.

[35] B. Østerud, Tissue factor/TFPI and blood cells, Thromb. Res. 129 (2012) 274–278. doi:10.1016/j.thromres.2011.11.049.

[36] I. Ott, Inhibitors of the initiation of coagulation, Br. J. Clin. Pharmacol. 72 (2011) 547–552. doi:10.1111/j.1365-2125.2011.03960.x.

[37] J. Wood, P. Ellery, Biology of tissue factor pathway inhibitor, Blood. 123 (2014) 2934–2943. doi:10.1182/blood-2013-11-512764.

[38] S.K. Austin, Haemostasis, Medicine (Baltimore). 45 (2017) 204–208. doi:10.1016/j.mpmed.2017.01.013.

[39] T. Iba, I. Nagaoka, M. Boulat, The anticoagulant therapy for sepsis-associated disseminated intravascular coagulation, Thromb. Res. 131 (2013) 383–389. doi:10.1016/j.thromres.2013.03.012.

[40] S.A. Smith, R.J. Travers, J.H. Morrissey, How it all starts: Initiation of the clotting cascade, Crit. Rev. Biochem. Mol. Biol. 50 (2015) 326–336. doi:10.3109/10409238.2015.1050550.

[41] M. Hoffman, A cell-based model of coagulation and the role of factor VIIa, Blood Rev. 17 (2003) S1–S5. doi:10.1016/S0268-960X(03)90000-2.

[42] J.T.B. Crawley, S. Zanardelli, C.K.N.K. Chion, D.A. Lane, The central role of thrombin in hemostasis, J. Thromb. Haemost. 5 (2007) 95–101. doi:10.1111/j.1538-7836.2007.02500.x.

[43] F. Bock, K. Shahzad, N. Vergnolle, B. Isermann, Activated protein C based therapeutic strategies in chronic diseases, Thromb. Haemost. 111 (2014) 610–617. doi:10.1160/TH13-11-0967.

[44] J.H. Griffin, B. V Zlokovic, L.O. Mosnier, Review Article Activated protein C : biased for translation, Blood. 125 (2015) 2898–2908. doi:10.1182/blood-2015-02-355974.

[45] P. Gresele, G. Agnelli, Novel approaches to the treatment of thrombosis, Trends Pharmacol. Sci. 23 (2002) 25–32. doi:10.1016/S0165-6147(00)01885-X.

[46] S. Butenas, J.D. Dee, K.G. Mann, The function of factor XI in tissue factor-initiated thrombin generation, J. Thromb. Haemost. 1 (2003) 2103–2111. doi:10.1046/j.1538-7836.2003.00431.x.

[47] S.M. Bates, J.I. Weitz, Direct thrombin inhibitors for treatment of arterial thrombosis: potential differences between bivalirudin and hirudin., Am. J. Cardiol. 82 (1998) 12P–18P.

[48] J. Eikelboom, H. White, S. Yusuf, The evolving role of direct thrombin inhibitors in acute coronary syndromes., J. Am. Coll. Cardiol. 41 (2003) 70S–78S. doi:10.1016/S0735-1097(02)02687-6.

[49] J. Hirsh, Current anticoagulant therapy—unmet clinical needs, Thromb. Res. 109 (2003) S1–S8. doi:10.1016/S0049-3848(03)00250-0.

[50] J.C. Fredenburgh, A.R. Stafford, C.H. Pospisil, J.I. Weitz, Modes and consequences of thrombin's interaction with fibrin, Biophys. Chem. 112 (2004) 277–284. doi:10.1016/j.bpc.2004.07.031.

[51] B. Østerud, E. Bjørklid, Sources of Tissue Factor, Semin. Thromb. Hemost. 32 (2006) 011–023. doi:10.1055/s-2006-933336.

[52] E. Camerer, A.B. Kolstø, H. Prydz, Cell biology of tissue factor, the principal initiator of blood coagulation., Thromb. Res. 81 (1996) 1–41. http://www.ncbi.nlm.nih.gov/pubmed/8747518.

[53] P. Carmeliet, N. Mackman, L. Moons, T. Luther, P. Gressens, L. Van Vlaenderen, H. Demunck, M. Kasper, G. Breier, P. Evrard, M. Müller, W. Risau, T. Edgington, D. Collen, Role of tissue factor in embryonic blood vessel development, Nature. 383 (1996) 73–75. doi:10.1038/383073a0.

[54] Y. Förster, A. Meye, S. Albrecht, B. Schwenzer, Tissue factor and tumor: Clinical and laboratory aspects, Clin. Chim. Acta. 364 (2006) 12–21. doi:10.1016/j.cca.2005.05.018.

[55] V. Bogdanov, R.I. Kirk, C. Miller, J.J. Hathcock, S. Vele, M. Gazdoiu, Y. Nemerson, M.B. Taubman, Identification and characterization of murine alternatively spliced tissue factor., J. Thromb. Haemost. 4 (2006) 158–67. doi:10.1111/j.1538-7836.2005.01680.x.

[56] S. Butenas, Tissue Factor Structure and Function, Scientifica (Cairo). 2012 (2012) 1–15. doi:10.6064/2012/964862.

[57] W. Lösche, Platelets and tissue factor., Platelets. 16 (2005) 313–9. doi:10.1080/09537100500140190.

[58] B. Østerud, J.O. Olsen, E. Bjørklid, What is blood borne tissue factor?, Thromb. Res. 124 (2009) 640–641. doi:10.1016/j.thromres.2009.06.027.

[59] N. Mackman, Alternatively spliced tissue factor – One cut too many?, Thromb. Haemost. 97 (2007) 5–8. doi:10.1160/TH06-11-0670.

[60] B. Szotowski, S. Antoniak, U. Rauch, Alternatively Spliced Tissue Factor: A Previously Unknown Piece in the Puzzle of Hemostasis, Trends Cardiovasc. Med. 16 (2006) 177–182. doi:10.1016/j.tcm.2006.03.005.

[61] D.P. O'Brien, J.S. Anderson, D.M. a Martin, P.G.H. Byfield, E.G.D. Tuddenham, Structural requirements for the interaction between Tissue Factor and Factor VII: Characterization of chymotrypsin-derived Tissue Factor polypeptides, Biochem. J. 292 (1993) 7–12.

[62] J.E. Hobbs, A. Zakarija, D.L. Cundiff, J. a Doll, E. Hymen, M. Cornwell, S.E. Crawford, N. Liu, M. Signaevsky, G. a Soff, Alternatively spliced human tissue factor promotes tumor growth and angiogenesis in a pancreatic cancer tumor model, Thromb. Res. 120 (2007) S13–S21. doi:10.1016/S0049-3848(07)70126-3.

[63] H.F. Rønning, U.C. Risøen, L. Örning, K. Sletten, K.S. Sakariassen, SYNTHETIC PEPTIDE ANALOGS OF TISSUE FACTOR AND FACTOR VII WHICH INHIBIT FACTOR Xa FORMATION BY THE TISSUE FACTOR/FACTOR VIIa COMPLEX, Thromb. Res. 84 (1996) 73–81. doi:10.1016/0049-3848(96)00163-6.

[64] S. Butenas, Tissue factor activity in whole blood, Blood. 105 (2005) 2764–2770. doi:10.1182/blood-2004-09-3567.

[65] L.W. Brüggemann, J.W. Drijfhout, P.H. Reitsma, C.A. Spek, Alternatively spliced tissue factor in mice: Induction by Streptococcus pneumoniae [15], J. Thromb. Haemost. 4 (2006) 918–920. doi:10.1111/j.1538-7836.2006.01870.x.

[66] J. Morrissey, N. Mackman, Tissue factor and factor VIIa: understanding the molecular mechanism., Thromb. Res. 122 Suppl (2008) S1-2. doi:10.1016/S0049-3848(08)70008-2.

[67] É. Biró, K.N. Sturk-Maquelin, G.M.T. Vogel, D.G. Meuleman, M.J. Smit, C.E. Hack, A. Sturk, R. Nieuwland,

Human cell-derived microparticles promote thrombus formation in vivo in a tissue factor-dependent manner, J. Thromb. Haemost. 1 (2003) 2561–2568. doi:10.1046/j.1538-7836.2003.00456.x.

[68] U. Rauch, Tissue Factor, the Blood, and the Arterial Wall, Trends Cardiovasc. Med. 10 (2000) 139–143. doi:10.1016/S1050-1738(00)00049-9.

[69] P. Censarek, A. Bobbe, M. Grandoch, K. Schrör, A.-A. Weber, Alternatively spliced human tissue factor (asHTF) is not pro-coagulant., Thromb. Haemost. 97 (2007) 11–4. doi:10.1160/TH06.

[70] D.M. Martin, C.W. Boys, W. Ruf, Tissue factor: molecular recognition and cofactor function., FASEB J. 9 (1995) 852–9. http://www.ncbi.nlm.nih.gov/pubmed/7615155.

[71] Y. a. Muller, M.H. Ultsch, M. de Vos, Abraham, The crystal structure of the extracellular domain of human tissue factor refined to 1.7 A resolution., J. Mol. Biol. 256 (1996) 144–159. doi:10.1006/jmbi.1996.0073.

[72] I.D. Campbell, C. Spitzfaden, Building proteins with fibronectin type III modules, Structure. 2 (1994) 333–337. doi:10.1016/S0969-2126(00)00034-4.

[73] E.M. Egorina, M.A. Sovershaev, B. Østerud, Regulation of tissue factor procoagulant activity by post-translational modifications, Thromb. Res. 122 (2008) 831–837. doi:10.1016/j.thromres.2007.11.004.

[74] a C. Pike, a M. Brzozowski, S.M. Roberts, O.H. Olsen, E. Persson, Structure of human factor VIIa and its implications for the triggering of blood coagulation., Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 8925–8930. doi:10.1073/pnas.96.16.8925.

[75] E. Zhang, R. St Charles, a Tulinsky, Structure of extracellular tissue factor complexed with factor VIIa inhibited with a BPTI mutant., J. Mol. Biol. 285 (1999) 2089–104. doi:10.1006/jmbi.1998.2452.

[76] C. Eigenbrot, New Insight into How Tissue Factor Allosterically Regulates Factor VIIa, Trends Cardiovasc. Med. 12 (2002) 19–26. doi:10.1016/S1050-1738(01)00139-6.

[77] G. Petrillo, P. Cirillo, G.- Luana D'Ascoli, F. Maresca, F. Ziviello, M. Chiariello, Tissue Factor/Factor FVII Complex Inhibitors in Cardiovascular Disease. Are Things Going Well?, Curr. Cardiol. Rev. 6 (2010) 325–332. doi:10.2174/157340310793566190.

[78] B. V. Norledge, R.J. Petrovan, W. Ruf, A.J. Olson, The tissue factor/factor VIIa/factor Xa complex: A model built by docking and site-directed mutagenesis, Proteins Struct. Funct. Genet. 53 (2003) 640–648. doi:10.1002/prot.10445.

[79] K. Harlos, D.M.A. Martin, D.P. O'Brien, E.Y. Jones, D.I. Stuart, I. Polikarpov, A. Miller, E.G.D. Tuddenham, C.W.G. Boys, Crystal structure of the extracellular region of human tissue factor, Nature. 370 (1994) 662–666. doi:10.1038/370662a0.

[80] M. Huang, W. Ruf, T.. Edgington, I.A. Wilson, Ligand Induced Conformational Transitions of Tissue Factor. Crystal Structure of the Tissue Factor:Factor VIIa Complex., To Be Publ. (n.d.). doi:10.2210/pdb1j9c/pdb.

[81] F.D. George, Microparticles in vascular diseases, Thromb. Res. 122 (2008). doi:10.1016/S0049-3848(08)70020-3.

[82] D. Lechner, A. Weltermann, Circulating tissue factor-exposing microparticles, Thromb. Res. 122 (2008) S47–S54. doi:10.1016/S0049-3848(08)70019-7.

[83] R.R. Bach, Initiation of coagulation by tissue factor., CRC Crit. Rev. Biochem. 23 (1988) 339–68.

doi:10.3109/10409238809082548.

[84] D. Mezzano, V. Matus, C.G. Sáez, J. Pereira, O. Panes, Tissue factor storage, synthesis and function in normal and activated human platelets, Thromb. Res. 122 (2008) S31–S36. doi:10.1016/S0049-3848(08)70016-1.

[85] S. Nomura, Y. Ozaki, Y. Ikeda, Function and role of microparticles in various clinical settings, Thromb. Res. 123 (2008) 8–23. doi:10.1016/j.thromres.2008.06.006.

[86] J.I. Zwicker, C.C. Trenor, B.C. Furie, B. Furie, Tissue factor-bearing microparticles and thrombus formation, Arterioscler. Thromb. Vasc. Biol. 31 (2011) 728–733. doi:10.1161/ATVBAHA.109.200964.

[87] S. Butenas, J. Krudysz-Amblo, Decryption of tissue factor, Thromb. Res. 129 (2012) S18–S20. doi:10.1016/j.thromres.2012.02.022.

[88] L.V.M. Rao, H. Kothari, U.R. Pendurthi, Tissue Factor encryption and decryption: Facts and controversies, Thromb. Res. 129 (2012) S13–S17. doi:10.1016/j.thromres.2012.02.021.

[89] A.P. Owens, N. MacKman, Sources of tissue factor that contribute to thrombosis after rupture of an atherosclerotic plaque, Thromb. Res. 129 (2012) S30–S33. doi:10.1016/j.thromres.2012.02.026.

[90] P. a Kern, S. Ranganathan, C. Li, L. Wood, G. Ranganathan, Adipose tissue tumor necrosis factor and interleukin-6 expression in human obesity and insulin resistance., Am. J. Physiol. Endocrinol. Metab. 280 (2001) E745-51. http://www.ncbi.nlm.nih.gov/pubmed/11287357.

[91] M. Diamant, R. Nieuwland, R.F. Pablo, A. Sturk, J.W.A. Smit, J.K. Radder, Elevated numbers of tissue-factor exposing microparticles correlate with components of the metabolic syndrome in uncomplicated type 2 diabetes mellitus, Circulation. 106 (2002) 2442–2447. doi:10.1161/01.CIR.0000036596.59665.C6.

[92] H.S. Lim, Soluble CD40 Ligand, Soluble P-Selectin, Interleukin-6, and Tissue Factor in Diabetes Mellitus: Relationships to Cardiovascular Disease and Risk Factor Intervention, Circulation. 109 (2004) 2524–2528. doi:10.1161/01.CIR.0000129773.70647.94.

[93] D.W. Sommeijer, H.R. Hansen, R. Van Oerle, K. Hamulyak, A.P. Van Zanten, E. Meesters, H.M.H. Spronk, H. ten Cate, Soluble tissue factor is a candidate marker for progression of microvascular disease in patients with Type 2 diabetes, J. Thromb. Haemost. 4 (2006) 574–580. doi:10.1111/j.1538-7836.2005.01763.x.

[94] G.S. Hotamisligil, P. Arner, J.F. Caro, R.L. Atkinson, B.M. Spiegelman, Increased adipose tissue expression of tumor necrosis factor-alpha in human obesity and insulin resistance., J. Clin. Invest. 95 (1995) 2409–15. doi:10.1172/JCI117936.

[95] F. Samad, Regulation of tissue factor gene expression in obesity, Blood. 98 (2001) 3353–3358. doi:10.1182/blood.V98.12.3353.

[96] J. Nofer, B. Kehrel, M. Fobker, B. Levkau, G. Assmann, A. von Eckardstein, HDL and arteriosclerosis: beyond reverse cholesterol transport, Atherosclerosis. 161 (2002) 1–16. doi:10.1016/S0021-9150(01)00651-7.

[97] V. Llorente-Cortés, M. Otero-Viñas, S. Camino-López, O. Llampayas, L. Badimon, Aggregated low-density lipoprotein uptake induces membrane tissue factor procoagulant activity and microparticle release in human vascular smooth muscle cells, Circulation. 110 (2004) 452–459.

doi:10.1161/01.CIR.0000136032.40666.3D.

[98] P.L.A. Giesen, U. Rauch, B. Bohrmann, D. Kling, M. Roque, J.T. Fallon, J.J. Badimon, J. Himber, M.A. Riederer, Y. Nemerson, Blood-borne tissue factor: Another view of thrombosis, Proc. Natl. Acad. Sci. 96 (1999) 2311–2315. doi:10.1073/pnas.96.5.2311.

[99] C. Marsik, P. Quehenberger, N. Mackman, B. Osterud, T. Luther, B. Jilma, Validation of a novel tissue factor assay in experimental human endotoxemia, Thromb. Res. 111 (2003) 311–315. doi:10.1016/j.thromres.2003.09.017.

[100] J.F. Viles-Gonzalez, J.J. Badimon, Atherothrombosis: the role of tissue factor, Int. J. Biochem. Cell Biol. 36 (2004) 25–30. doi:10.1016/S1357-2725(03)00240-1.

[101] a. Migdalski, M. Kotschy, a. Jawien, Tissue Factor, Tissue Factor Pathway Inhibitor and Vascular Endothelial Growth Factor-A in Carotid Atherosclerotic Plaques, Eur. J. Vasc. Endovasc. Surg. 30 (2005) 41–47. doi:10.1016/j.ejvs.2005.02.055.

[102] J.H. Morrissey, Tissue factor: in at the start...and the finish?, J. Thromb. Haemost. 1 (2003) 878–80. doi:10.1046/j.1538-7836.2003.00219.x.

[103] M.N. Levine, A.Y. Lee, A.K. Kakkar, From Trousseau to targeted therapy: New insights and innovations in thrombosis and cancer, J. Thromb. Haemost. 1 (2003) 1456–1463. doi:10.1046/j.1538-7836.2003.00275.x.

[104] T. a Hembrough, G.M. Swartz, A. Papathanassiu, G.P. Vlasuk, W.E. Rote, S.J. Green, V.S. Pribluda, Tissue Factor / Factor VIIa Inhibitors Block Angiogenesis and Tumor Growth Through a Nonhemostatic Mechanism Tissue Factor / Factor VIIa Inhibitors Block Angiogenesis and Tumor Growth Through a Nonhemostatic Mechanism, Cancer Res. 63 (2003) 2997–3000.

[105] Y. Nadir, B. Brenner, A. Zetser, N. Ilan, I. Shafat, E. Zcharia, O. Goldshmidt, I. Vlodavsky, Heparanase induces tissue factor expression in vascular endothelial and cancer cells., J. Thromb. Haemost. 4 (2006) 2443–51. doi:10.1111/j.1538-7836.2006.02212.x.

[106] J.I. Zwicker, B.C. Furie, B. Furie, Cancer-associated thrombosis, Crit. Rev. Oncol. Hematol. 62 (2007) 126–136. doi:10.1016/j.critrevonc.2007.01.001.

[107] R.E. Tilley, T. Holscher, R. Belani, J. Nieva, N. Mackman, Tissue factor activity is increased in a combined platelet and microparticle sample from cancer patients, Thromb. Res. 122 (2008) 604–609. doi:10.1016/j.thromres.2007.12.023.

[108] A.K. Kakkar, Venous Thrombosis in Cancer Patients: Insights from the FRONTLINE Survey, Oncologist. 8 (2003) 381–388. doi:10.1634/theoncologist.8-4-381.

[109] J.E. Bluff, M. Amarzguioui, J. Slattery, M.W.R. Reed, N.J. Brown, C.A. Staton, Anti-tissue factor short hairpin RNA inhibits breast cancer growth in vivo, Breast Cancer Res. Treat. 128 (2011) 691–701. doi:10.1007/s10549-010-1149-8.

[110] K.W. Lee, A.D. Blann, G.Y.H. Lip, Inter-relationships 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, Int. J. Cardiol. 111 (2006) 302–308.

doi:10.1016/j.ijcard.2005.10.014.

[111] E. Arbit, M. Goldberg, I. Gomez-Orellana, S. Majuru, Oral heparin: status review, Thromb. J. 4 (2006) 6. doi:10.1186/1477-9560-4-6.

[112] C.E. Copeland, C.K. Six, A Tale of Two Anticoagulants: Warfarin and Heparin, J. Surg. Educ. 66 (2009) 176–181. doi:10.1016/j.jsurg.2009.03.035.

[113] D. Keeling, Duration of anticoagulation: decision making based on absolute risk, Blood Rev. 20 (2006) 173–178. doi:10.1016/j.blre.2005.09.001.

[114] R.L. Mueller, First-generation agents: Aspirin, heparin and coumarins, Best Pract. Res. Clin. Haematol. 17 (2004) 23–53. doi:10.1016/j.beha.2004.03.003.

[115] I. Seljeflot, M. Hurlen, H. Arnesen, Increased levels of soluble tissue factor during long-term treatment with warfarin in patients after an acute myocardial infarction, J Thromb Haemost. 2 (2004) 726–730. doi:10.1111/j.1538-7836.2004.00676.x\rJTH676 [pii].

[116] J.I. Weitz, Unanswered questions in venous thromboembolism, Thromb. Res. 123 (2009) S2–S10. doi:10.1016/j.thromres.2008.08.003.

[117] C.W. Francis, Ximelagatran: a new oral anticoagulant, Best Pract. Res. Clin. Haematol. 17 (2004) 139–152. doi:10.1016/j.beha.2004.03.005.

[118] M. Franchini, B. Chong, I.-K. Jang, H. Hursting, T. Warkentin, T. Warkentin, D. Cines, J. Bussel, R. McMillan, J. Zehnder, T. Warkentin, A. Greinacher, R. Strauss, M. Wehler, K. Mehler, D. Kreutzer, C. Koebnick, E. Hahn, T. Warkentin, L. Rice, T. Warkentin, M. Levine, J. Hirsh, J. Kelton, T. Baglin, M. Comunale, E. van Cott, S. Nand, W. Wong, B. Yuen, A. Yetter, E. Schmulbach, S.G. Fisher, B. Girolami, P. Prandoni, P. Stefani, E. Lindhoff-Last, P. Eichler, M. Stein, C. Locke, J. Dooley, J. Gerber, R. Reilly, G. Arepally, D. Cines, J. Kelton, J. Smith, T. Warkentin, C. Hayward, G. Denomme, P. Horsewood, D. Cines, A. Tomaski, S. Tannenbaum, G. Visentin, S. Ford, J. Scott, R. Aster, T. Warkentin, J. Kelton, T. Warkentin, J. Kelton, W. Bell, T. Warkentin, J. Kelton, A. Greinacher, T. Warkentin, A. Meyer-Lindenberg, E. Quenzel, E. Bierhoff, H. Wolff, E. Schindler, R. Biniek, L. Boshkov, T. Warkentin, C. Hayward, M. Andrew, J. Kelton, T. Warkentin, H. Klein, W. Bell, B. Alving, D. Wallis, D. Workman, B. Lewis, L. Steen, R. Pifarre, J. Moran, B. Chong, H. Magnani, A. Greinacher, U. Janssens, G. Berg, A. Greinacher, H. Volpel, U. Janssens, A. Greinacher, P. Eichler, N. Lubenow, H. Kwasny, M. Luz, B. Lewis, D. Wallis, S. Berkowitz, B. Chong, A. Gallus, J. Cade, T. Warkentin, L. Elavathil, C. Hayward, M. Johnston, J. Russett, J. Kelton, T. Ortel, J. Gockerman, R. Califf, T. Ortel, B. Chong, N. Lubenow, P. Eichler, T. Lietz, B. Farner, A. Greinacher, B. Farner, P. Eichler, H. Kroll, A. Greinacher, J. Hirsh, N. Heddle, J. Kelton, R. Gosselin, W. Dager, J. King, S. Sheth, R. DiCicco, M. Hursting, T. Montague, D. Jorkasky, M. Hursting, B. Lewis, D. Macfarlane, P. Savi, B. Chong, A. Greinacher, J. Harenberg, I. Jorg, T. Fenyvesi, K. Kuo, M. Kovacs, Heparin-induced thrombocytopenia: an update, Thromb. J. 3 (2005) 14. doi:10.1186/1477-9560-3-14.

[119] J.I. Weitz, M. Crowther, Direct thrombin inhibitors, Thromb. Res. 106 (2002). doi:10.1016/S0049-3848(02)00093-2.

[120] J.J. Van Veen, A. Gatt, M. Makris, Thrombin generation testing in routine clinical practice: Are we there yet?, Br. J. Haematol. 142 (2008) 889–903. doi:10.1111/j.1365-2141.2008.07267.x.

[121] J.I. Weitz, Factor Xa and thrombin as targets for new oral anticoagulants, Thromb. Res. 127 (2011) S5–S12. doi:10.1016/S0049-3848(10)70147-X.

[122] M.L. van Montfoort, J.C.M. Meijers, Anticoagulation beyond direct thrombin and factor Xa inhibitors: Indications for targeting the intrinsic pathway?, Thromb. Haemost. 110 (2013) 223–232. doi:10.1160/TH12-11-0803.

[123] T. Furugohri, T. Fukuda, N. Tsuji, A. Kita, Y. Morishima, T. Shibano, Melagatran, a direct thrombin inhibitor, but not edoxaban, a direct factor Xa inhibitor, nor heparin aggravates tissue factor-induced hypercoagulation in rats, Eur. J. Pharmacol. 686 (2012) 74–80. doi:10.1016/j.ejphar.2012.04.031.

[124] I. Ott, Tissue factor inhibition: Another approach reducing thrombosis after vascular injury, Thromb. Haemost. 103 (2009) 7–8. doi:10.1160/TH09-10-0712.

[125] W. Liang, W. Zhang, S. Zhao, Q. Li, Y. Yang, H. Liang, R. Ceng, A study of the ultrasound-targeted microbubble destruction based triplex-forming oligodexinucleotide delivery system to inhibit tissue factor expression, Mol. Med. Rep. 11 (2015) 903–909. doi:10.3892/mmr.2014.2822.

[126] Y. Yang, Q. Li, D. Ying, Z. Gong, R. Cheng, M. Lü, Y. Liu, Z. Zhou, J. Zheng, [Triplex-forming oligonucleotide inhibits the expression of tissue factor gene in endothelial cells induced by the blood flow shear stress in rats]., Yao Xue Xue Bao. 41 (2006) 808–13. http://linkinghub.elsevier.com/retrieve/pii/S1000194808600374.

[127] Y. Zhang, Z. Wang, R.A. Gemeinhart, Progress in microRNA delivery, J. Control. Release. 172 (2013) 962–974. doi:10.1016/j.jconrel.2013.09.015.

[128] A.C. Stephens, N.F. Ranlall, R.P.A. Rivers, Suppression of HUVEC tissue factor synthesis by antisense oligodeoxynucleotide, Thromb. Res. 122 (2008) 99–107. doi:10.1016/j.thromres.2007.08.021.

[129] J. Steffel, A. Akhmedov, C. Fähndrich, F. Ruschitzka, T.F. L??scher, F.C. Tanner, Differential effect of celecoxib on tissue factor expression in human endothelial and vascular smooth muscle cells, Biochem. Biophys. Res. Commun. 349 (2006) 597–603. doi:10.1016/j.bbrc.2006.08.075.

[130] T. Holen, M. Amarzguioui, M.T. Wiiger, E. Babaie, H. Prydz, Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor., Nucleic Acids Res. 30 (2002) 1757–1766. doi:10.1093/nar/30.8.1757.

[131] A. Eisenreich, U. Leppert, The impact of microRNAs on the regulation of tissue factor biology, Trends Cardiovasc. Med. 24 (2014) 128–132. doi:10.1016/j.tcm.2013.09.005.

[132] M.M. Fiore, P.F. Neuenschwander, J.H. Morrissey, An unusual antibody that blocks tissue factor/factor VIIa function by inhibiting cleavage only of macromolecular substrates., Blood. 80 (1992) 3127–34. http://www.ncbi.nlm.nih.gov/pubmed/1467519.

[133] E.S. Priestley, Tissue factor–fVIIa inhibition: update on an unfinished quest for a novel oral antithrombotic, Drug Discov. Today. 19 (2014) 1440–1444. doi:10.1016/j.drudis.2014.05.023.

[134] M. Jacquemin, J.M. Saint-Remy, The use of antibodies to coagulation factors for anticoagulant therapy.,

Curr. Med. Chem. 11 (2004) 2291–6. http://www.ncbi.nlm.nih.gov/pubmed/15379713.

[135] K. Carlsson, E. Persson, U. Carlsson, M. Svensson, Inhibitors of factor VIIa affect the interface between the protease domain and tissue factor, Biochem. Biophys. Res. Commun. 349 (2006) 1111–1116. doi:10.1016/j.bbrc.2006.08.148.

[136] N. Alizadeh, B. Pittet, X. Tenorio, C. Pyke, D. Baetens, K.U. Schlaudraff, D. Montandon, M. Ezban, M.S. Pepper, Active-site inactivated FVIIa decreases thrombosis and necrosis in a random skin flap model of acute ischemia, J. Surg. Res. 122 (2004) 263–273. doi:10.1016/j.jss.2004.06.014.

[137] J.I. Weitz, J. Hirsh, M.M. Samama, New antithrombotic drugs: American College of Chest physicians evidence-based clinical practice guidelines (8th edition), Chest. 133 (2008). doi:10.1378/chest.08-0673.

[138] A.C.J.M. De Pont, a H.M. Moons, E. De Jonge, J.C.M. Meijers, G.P. Vlasuk, W.E. Rote, H.R. Buller, T. Van Der Poll, M. Levi, Recombinant nematode anticoagulant protein c2, an inhibitor of tissue factor/factor VIIa, attenuates coagulation and the interleukin-10 response in human endotoxemia, J. Thromb. Haemost. 2 (2004) 65–70. doi:10.1111/j.1538-7836.2004.00526.x.

[139] P. Morawitz, Die Chemie der Blutgerinnung, Ergebnisse Der Physiol. 4 (1905) 307–422.

[140] E. Bachli, MD, Historical review, Br. J. Haematol. 110 (2000) 248–255. doi:10.1046/j.1365-2141.2000.02055.x.

[141] Q. Wu, X. Wang, Y. Gu, X. Zhang, Y. Qin, H. Chen, X. Xu, T. Yang, M. Zhang, Screening and identification of human ZnT8-specific single-chain variable fragment (scFv) from type 1 diabetes phage display library, Sci. China Life Sci. 59 (2016) 686–693. doi:10.1007/s11427-016-5077-7.

[142] Z.A. Ahmad, S.K. Yeap, A.M. Ali, W.Y. Ho, N.B.M. Alitheen, M. Hamid, scFv Antibody: Principles and Clinical Application, Clin. Dev. Immunol. 2012 (2012) 1–15. doi:10.1155/2012/980250.

Chapter3: Structural characterisation of TFI-scFv: a combined homology modelling, molecular docking, and molecular dynamic approach.

List of Abbreviations

Complementarity Determining Regions	CDR
Disulfide- stabilized Fv	dsFv
factor VII	FVII
Heavy chain variable domains	H_v
Immunoglobulin G	IgG
Light chain variable domains	V_L
magnetic resonance spectroscopy	MRS
molecular dynamic simulations	MDS
Periodic Boundary Conditions	PBC
Permutated Fv	pFv
Pico seconds	ps
Position Specific Scoring Matrix	PSSM
Protein Date Bank	PDB
Root Mean Square Deviation	RMSD
Single-chain Variable Fragment	scFv
Three-dimensional	3D
Tissue factor	TF
Yet Another Scientific Artificial Reality Application	YASARA

Chapter 3: Structural characterisation of TFI-scFv: a combined homology modelling, molecular docking, and molecular dynamic approach.

3.1 Abstract

The interaction of a single chain variable fragment (scFv) directed against human tissue factor was predicted using an *in sillico* approach with the aim to establish a most likely mechanism of inhibition. The structure of the tissue factor inhibiting antibody (TFI-scFv) was predicted using homology modelling and position of the complementarity determining regions (CDR) identified. CDR regions were utilised to direct molecular docking between the homology model of TFI-scFv and the crystal structure of the extracellular domains of tissue factor (TF). The rigid-body docking model was refined by means of molecular dynamic (MD) simulations and the most prevalent cluster was identified. MD simulations predicted improved interaction between TFI-scFv and TF and propose the formation of stable complex for duration of the 100 nanoseconds simulation. Analysis of the refined docking model suggest the interactions between TFI-scFv would sterically interfere with the association of coagulation factor VII (FVII) with TF. This interaction would prevent the formation of the TF:VIIa complex and in so doing interfere with blood coagulation.

3.2 Key-Words

Single Chain variable fragment, Homology modelling, Protein-protein docking, Molecular dynamics (MD), Tissue Factor, Factor VII

3.3 Introduction

The use of recombinant antibodies has become ubiquitous with several fields of medical sciences, be it for therapeutic or analytical purposes [1–3]. The search for smaller more effective antibodies with a reduced immune response has led to the development of several alternative sub-antibody fragments including, but not limited to, single-chain variable fragments (scFv) [4], disulfide- stabilized variable fragments (dsFv) [5], diabodies [6], and permutated variable fragments (pFv) [7,8]. Therapeutic antibodies therapy is currently the fastest growing sector of pharmaceutical biotechnology [9]. Utilising phage display technology, an inhibitory single chain variable fragment (scFv) against human tissue factor was isolated form the Human Single Fold scFv Libraries I + J in a previous study [10]. TF plays a key role within normal haemostasis, however abnormal expression has been linked to various thrombotic disorders, cardiovascular diseases and have been shown to play a key role in insulin- disorders and as well as some cancers [11–14].

Single chain variable fragments represent the smallest stable antibody fragments still capable of specifically recognising and binding an antigen [9]. The scFvs in the Tomlinson I+J libraries are structured as a single polypeptide consisting of the Heavy- (H_V) and Light chain (V_L) variable domains of human immunoglobulins (IgG) [15]. The diversity of the libraries are generated through random V-gene chain shuffling techniques [16]. Each of

these domains are structurally stabilised by intra-domain cysteine disulfide bridges [17]. The V_H and L_H domains are connected by means of a highly synthetic flexible linker peptide $(Gly_4Ser)_3$ in order to stabilise the synthetic construct [18–20]. The scFv at 27–32 kDa is only 1/6 of the molecular weight of the complete antibody. The smaller size of the scFv allows for the improved penetration of blood vessels and tissue barriers [21]. In addition, scFv have more stable structures, shorter half-life, and reduced immunogenicity that their larger counterparts [22].

Knowledge of antibody three-dimensional (3D) structure can provide deeper insights into the structural stability and functionality of proteins. It allows the identification of specific residues that are involved in the functionality of the antibody and as a result facilitate advanced techniques such as site directed mutagenesis [23]. In addition, it allows for the prediction of protein-protein interactions based on structural interaction which may identify specific inhibition mechanisms [24] . Ever since the early 1920s, X-ray crystallography and later-on nuclear magnetic resonance spectroscopy (NMR) have been at the forefront of experimental protein structure solution [25,26]. Despite many advances in the last years, these methods are still a time-consuming and require highly specialised equipment and unfortunately do not guarantee success [27]. In some cases, suitable crystals were only obtained after several years. Homology modelling provides a powerful *in silico* alternative method for the prediction of protein structure based on databases of proteins with known X-ray crystallography or NMR structures [28,29]. As a result, it provides a quick and cost effective method to generate credible protein structures [30–32].

In this study, a homology model of the antibody fragment was generated in order to gain specific insights into the structural composition, stability, and function of the anti-tissue factor scFv. The homology model was utilised to perform protein-protein docking simulations against the extra cellular domains of tissue factor based on the CDR composition to establish a possible mechanism of inhibition. The initial rigid protein-protein docking model was refined by molecular dynamic analysis in order to establish a possible mechanism of inhibition.

3.4 Materials and Methods

3.4.1 Sequencing of the scFv gene fragment

All DNA sequencing reactions were performed using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA) in conjunction with the 3130xl Genetic Analyzer HITACHI (Applied Biosystems, USA) according to manufacturer's specifications. The pET22 TFI-scFv expression vector was sequenced using T7 Promoter Primer: 5' TAATACGACTCACTATAGGG 3' and T7 Terminator Primer: 5' GCTAGTTATTGCTCAGCGG 3' (Integrated DNA technologies, Belgium).

3.4.1 *In silico* Homology modelling of TFI-scFv

Due to the novel nature of TFI-scFv, no previous structure was available and as a result a homology model was generated. The DNA sequence of TFI-scFv was translated to the final amino acid sequence using online translations tool ExPASy: translate from the Swiss Institute of Bioinformatics and the His-TAG removed for modelling purposes. The amino acid sequence was used to model a three-dimensional structure of TFI-scFv. Homology modelling of TFI-scFv was performed using Yet Another Scientific Artificial Reality Application (YASARA) version 16.4.6.L.64 available at http://www.yasara.org/ with the following set parameters; Modelling speed: Slow, Number of PSI-BLAST iterations in template search (PsiBLASTs): 3, Maximum allowed (PSI-) BLAST E-value to consider template (EValue Max): 0.5, Maximum number of templates to be used (Templates Total): 5, Maximum number of templates with same sequence (Templates SameSeq): 1, Maximum oligomerization state (OligoState): 4 (tetrameric), Maximum number of alignment variations per template: (Alignments): 5, Maximum number of conformations tried per loop (LoopSamples): 50 and Maximum number of residues added to the termini (TermExtension): 10 [33]. The Z-score are calculated from molecular dynamic force field energies and is an indication standard deviations in model quality with regards to an average high-resolution quality x-ray crystal structure of globular proteins. Z-scores were also calculated as the weighted averages of individual Z-scores using the formula Overall = 0.145 * Dihedrals + 0.390 * Packing1D + 0.465 * Packing3D.

3.4.3 *In silico* protein-protein docking predictions

Initially, the extracellular domain of tissue factor PDB no: 1boy (protein receptor) and homology model of TFI-scFv (protein ligand) were preserved as rigid bodies and the docking carried out using online docking software suit, ZDOCK v3.0.2 available at http://zdock.umassmed.edu/. The docking region of the ligand was restricted to the surfaced exposed compliment determining regions of the antibody fragment as were identified by previous *in silico* homology modelling [34]. Structural visualisation and analysis of the rigid body docking models was preformed using YASARA version 16.4.6.L.64.

3.4.4 Molecular dynamic refinement

The top rigid body docking model, generated by ZDOCK, served as starting structures for MD simulations. GROMACS utilities were used to prepare the structures for MD simulations [35–40]. Hydrogen atom and water molecule motions were constrained using the LINCS and SETTLE algorithms respectively for equilibration and production simulations [41,42]. Periodic Boundary Conditions (PBC) was used and long-range electrostatic interactions was resolved using particle mesh Ewald (PME) in conjunction with the Verlet cut-off scheme with a cut-off of 0.8 nm [43]. Energy minimization was performed using a maximum number of 50 000 of steepest descent minimization algorithm steps. Next, a constant number, volume, and temperature (NVT) equilibration was performed for 2 ns with the protein positionally restrained to reach a target temperature of 300K using a modified Berendsen thermostat. This was followed by a constant number, volume, pressure (NPT) equilibration run for 20 ns, again with the protein positionally restrained, to reach a target pressure of 1 bar using the Parrinello-Rahman

barostat. The simulation was performed for 100 ns at 300 K using the modified Berendsen thermostat, 1 bar pressure using the Parinello-Rahman barostat, and with the positional restraints on the protein removed to allow more free movement. Snapshots were saved every 200 Pico seconds (ps) to yield a trajectory of 5000 frames. GROMACS's utility: Trajconv 0.4 was used and the PBC artifacts were removed from the trajectories. The trajectory was visually inspected for artefacts using Visual Molecular Dynamics (VMD) version 1.9.2 software suite [44]. All post MDS structures, contact analysis, surface model alignments and root mean square deviations (RMSD) were calculated using YASARA. All MDS were performed using GROMACS version 5.04 on the High-Performance Computing Cluster of the University of the Free State.

3.5 Results

3.5.1 Structure of pET22 TFI-scFv.

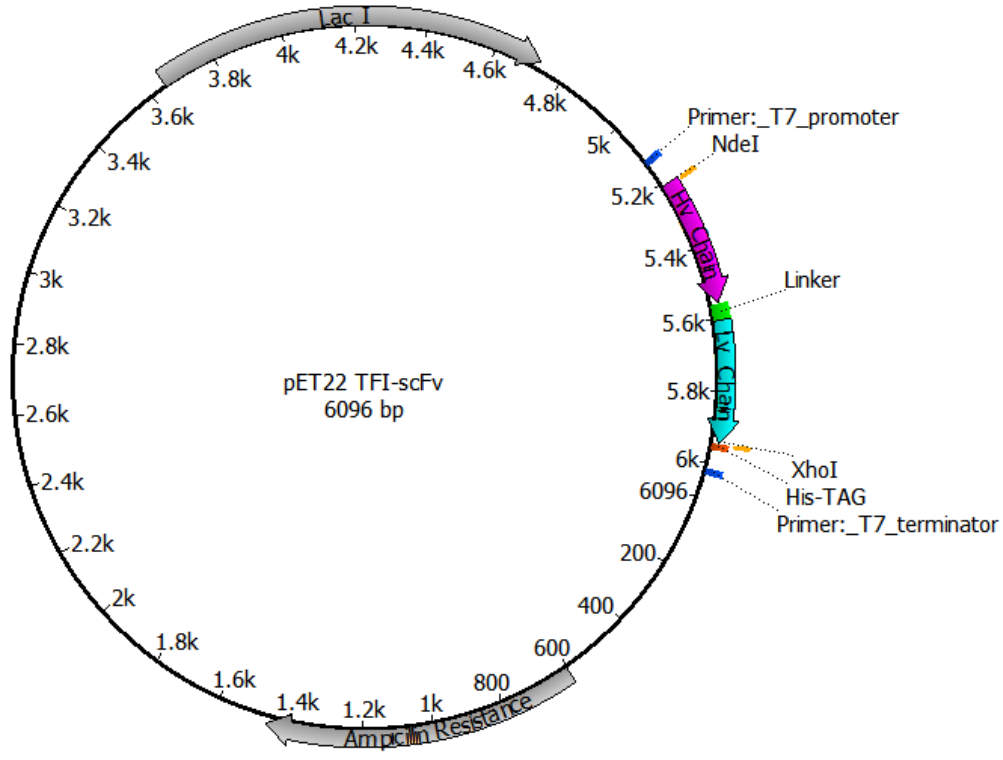


Figure 3.1 Structure pET TFI-scFv plasmid.

*The structure of the pET TFI-scFv expression vector. The TFI-scFv sequece in nested between NdeI and XhoI restriction sites and flanked by the primer binding sites (**Blue**). The N-terminal heavy chain variable region (**Magenta**), connecting linker region (**Green**), C-terminal light chain variable region (**Cyan**), and C-terminal His-TAG (**Orange**) is indicated. The lacI- and ampicillin resistance genes are indicated in gray.*

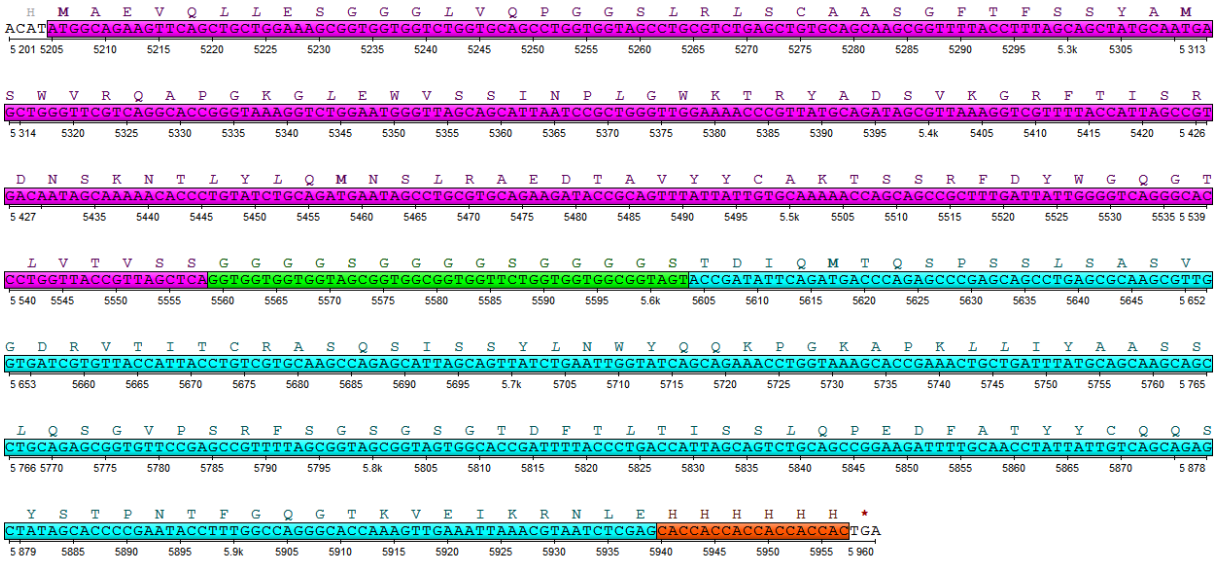


Figure 3.2 Sequence of TFI-scFv.

*The DNA- and transcribed amino acid sequence of TFI-scFv is given. The N-terminal heavy chain variable region (**Magenta**), connecting linker region (**Green**), C-terminal light chain variable region (**Cyan**), and C-terminal His-TAG (**Orange**) is indicated. TGA-stop codon is indicated with *****.*

3.5.2 *In silico* homology modelling of TFI-scFv

In short, homology modelling of the TFI-scFv was performed by screening for possible homology templated using PSI-BLAST to extract a position specific scoring matrix (PSSM) from the UniProt Reference Clusters (UniRef90). This scoring matrix was then utilized to search the protein data bank (PDB). A total of 2958 potential modelling templates were identified and ranked according structural quality by WHAT_CHECK [35] obtained from the PDBFinder2 database [36]. The PDB-redo database was scanned for refined templates of the initially identified templates. Low quality templates were filter out based on the WHAT_CHECK quality score. Secondary structure were predicted using the PSIPRED secondary structure prediction algorithm [37]. The target secondary structure profile was refined by means multiple alignments against related UniRef90 sequences. A total of 7267 alignments were made and initial models generated for each possible alignment. The modelling process firstly generates the carbon backbone followed by the generation of loops. Loops are modelled by screening multiple conformations and re-optimizing of the side chains. The model's hydrogen bonding network was then optimized. The top scoring PDB templates utilized for homology modelling of TFI-scFv are list below according to quality rank (Table 3.1).

Table 3.1: Top ranked PDB templated utilized for homology modelling.

Rank	*Structure*	*Resolution (Å)*	*Description*
1	*2YC1*	*1.9*	*Human derived scFv in complex with scorpion toxin* [38]
2	*1FGV*	*1.9*	*Humanized anti-CD18 antibody* [39]
3	*2UZI (Alignment var 1)*	*2*	*Anti-RAS Single Antibody Domain* [40]
4	*2UZI (Alignment var 2)*	*-*	*-*
5	*4D1V*	*1.7*	*Metallo-ß-lactamase* [41]
6	*2D7T*	*1.7*	*Anti polyhydroxybutyrate antibody Fv, heavy chain* [42]

Finally, an improved hybrid model is generated through the combination of the highest-ranking sections of top ranked homology models as templates. The top scoring homology models and Z-scores are summarised in Figure 3.1.

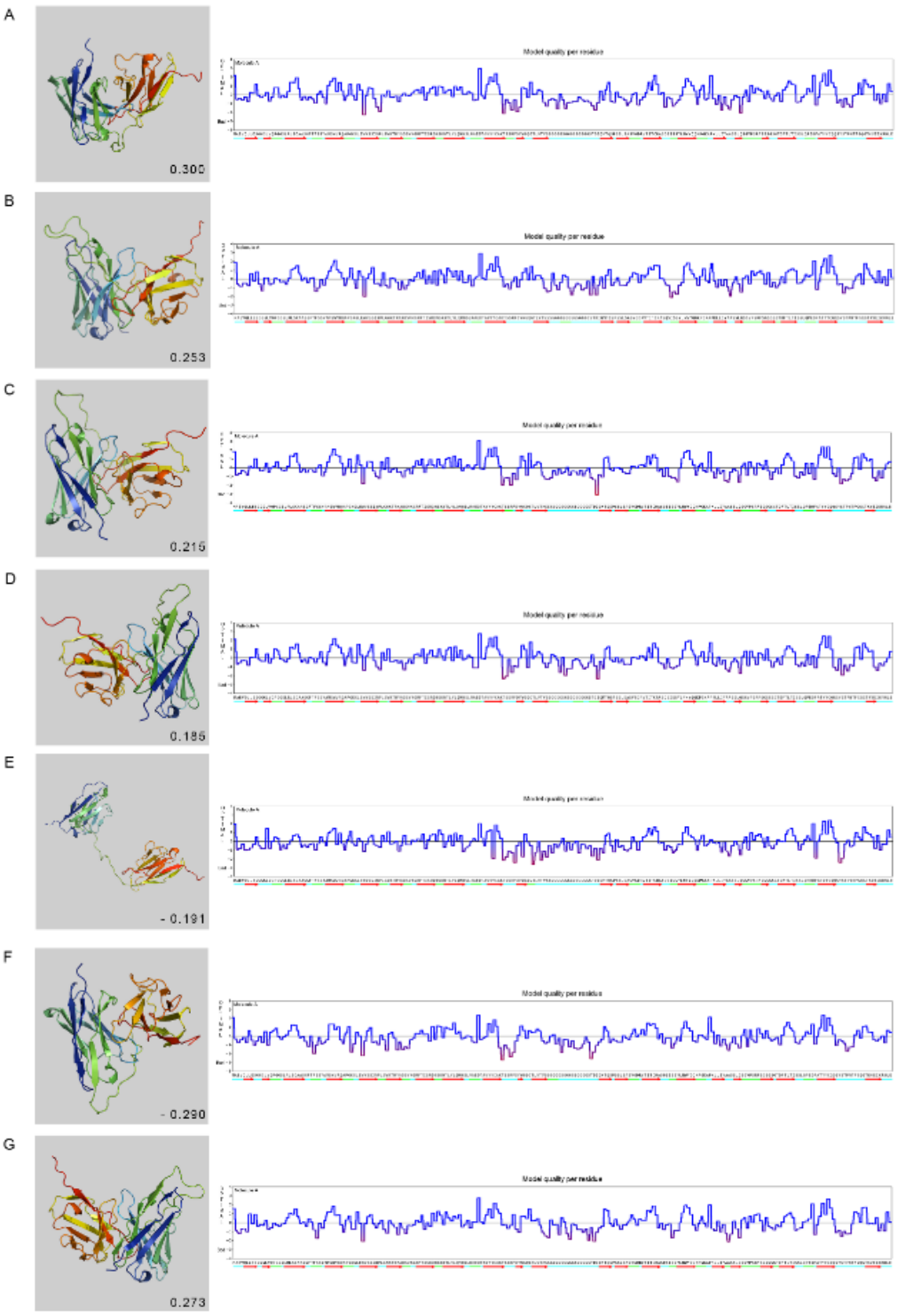


Figure 3.3: Homology modelling of TFI-scFv based on to scoring templates.

*Homology models based on top scoring templates A) 2YC1 B) 1FGV C) 2UZI alignment variant 2 D) 2UZI alignment variant 1 E) 4V1D F) 2D7T and G) TFI-scFv Hybrid model are indicated. Cartoon models are presented in colour gradient ranging from N-terminal (**Blue**) to C-Terminal (**Red**). The quality Z-score plotted per residue for each of each of the predicted models is presented to the right. The overall Z-scores are indicated in the lower right of the predicted cartoon ribbon structures.*

3.5.2 Structural analysis of TFI-scFv hybrid model

The predicted TFI-scFv structure consists of Hv- and Lv chain connected by means of a flexible linker. The secondary structure content of TFI-scFv model consists of 0.0% helix, 51.0% sheet, 19.6% turn, 29.4% coil, 0.0% 3-10 helix and 0.0% pi-helix. Energy minimizations was performed using self-parameterizing knowledge-based YASARA force field and was calculated at -459.144 kJ/mol. The total solvent accessible area was calculated at 10515.12 $Å^2$. The Hv- and Lv-chain domains are perpendicularly orientated and consists of a barrel-like structures. This orientation is facilitated by a series of hydrogen bonds interactions, formed at the interface (1118.47 $Å^2$) between Hv- and Lv-chain (Figure 3.2 A and Table 3.2). The interface is also rich in hydrophobic residues (Hv-chain: Val^{39}, Leu^{46} and Ala^{99} and Lv-Chain: Ala^{177}, Leu^{180} and Phe^{232}) which facilitates this alignment of the Hv- and Lv-chain at the interface. Each of these barrel-like structures consist of 9 antiparallel β-sheets with a hydrophobic inner region. These structures are stabilised by a single intra-domain disulfide bond that spans across the hydrophobic inner region, which is the hallmark of the immunoglobulin fold. The disulfide linkages are formed between Cys^{24}-Cys^{98} and Cys^{157}-Cys^{222} for the Hv- and Lv-chain respectively (Figure 3.2 ii B). The linker consists of three GGGGS repeats and connects the Hv- chain C-terminal (Ser^{118}) with the Lv chain N-terminal (Thr^{134}). The resulting artificial construct mimics the variable regions of a complete IgG. The linker region is designed to retain a high level of flexibility [20]. This is also evident with in the predicted models where the residue quality of the linker regions tends to be lower than the average trend due to the high flexibility.

Table 3.2: Characteristics of hydrogen bond interaction between Hv- and Lv-Chain.

Hv-Chain variable region	*Lv-chain variable region*	*Bond distance in Å*	*Bond Energy kcal/mol*	*Relative H-bond strength**
Ser103	*Asn168*	*2.14*	*17.87*	*moderate*
Ser103	*Ser225*	*1.804*	*24.43*	*moderate*
Phe105	*Typ170*	*2.045*	*20.33*	*moderate*

** Relative hydrogen bonds strength according to Jeffrey's classification.* [43]

Figure 3.4: Structure and key features of TFI-scFv hybrid model.

*i) The cartoon structure of the 27 kDa TFI-scFv hybrid model as predicted by homology modelling. The hybrid model consists of the N-terminal heavy chain variable region (**Magenta**), connecting linker region (**Green**) and C-terminal light chain variable region (**Cyan**). The model comprises of 51 % beta-sheets (Arrow), 24 % Coils (Ribbon) and 19.6 % turns (Ribbon). All C-terminal affinity tags are omitted form the homology model. ii) The ribbon diagram of TFI-scFv hybrid model show at 180° rotated positions. The key stabilising residues are highlighted and numbered. Intra-domain stabilising disulfide bonds (**Gray**) are present between Cys^{24}-Cys^{98} and Cys^{157}-Cys^{222} in the Heavy- and Light chain respectively. The residues involved in hydrogen interactions (**Yellow**) between the Heavy- and Light chains are indicated and hydrogen bonds represented by dotted lines.*

3.5.3 Complementarity Determining regions of TFI-scFv

The variable regions of the light and heavy chains are well defined in antibodies and are referred to as the CDR. Structural as well as functional antigen binding studies have shown that the residues located in the hypervariable regions of the CDRs in both chains are responsible for the specificities of the anti-body [44,45]. CDRs located on the Hv- and Lv chains were identified using Kabat numbering rule [46]. The Kabat definition is based on sequence variability and is the most commonly used. According to the TFI-scFv hybrid model, the various CDRs are clustered together to form the binding interface toward one polar region of the antibody structure (Table 3.3).

Table 3.3: The complementarity determining regions of TFI-scFv as per Kabat numbering scheme.

Hv-Chain	*CDR*	*Sequence*
	1	*Phe^{31}, Ser,^{32} Ser^{33}, Tyr^{34}, Ala^{35}*
	2	*Val^{50}, Ser^{51}, Ser^{52}, Ile^{53}, Asn^{54}, Pro^{55}, Leu^{56}, Gly^{57}, Trp^{58}, Lys^{59}, Thr^{60}, Arg^{61}, Tyr^{62}, Ala^{63}, Asp^{64}, Ser^{65}*
	3	*Val^{95}, Tyr^{96}, Tyr^{97}, Cys^{98}, Ala^{99}, Lys^{100}, Thr^{101}, Ser^{102}*
Lv -Chain	*1*	*Tyr^{166}, Lue^{167}, Asn^{168}, Trp^{169}, Tyr^{170}, Gln^{171}, Gln^{172}*
	2	*Lue^{188}, Gln^{189}, Ser^{190}, Gly^{191}, Val^{192}, Pro^{193}, Ser^{194}*
	3	*Ser^{227}, Thr^{228}, Pro^{229}, Asn^{230}, Thr^{231}, Phe^{232}, Gly^{233}, Gln^{234}, Gly^{235}*

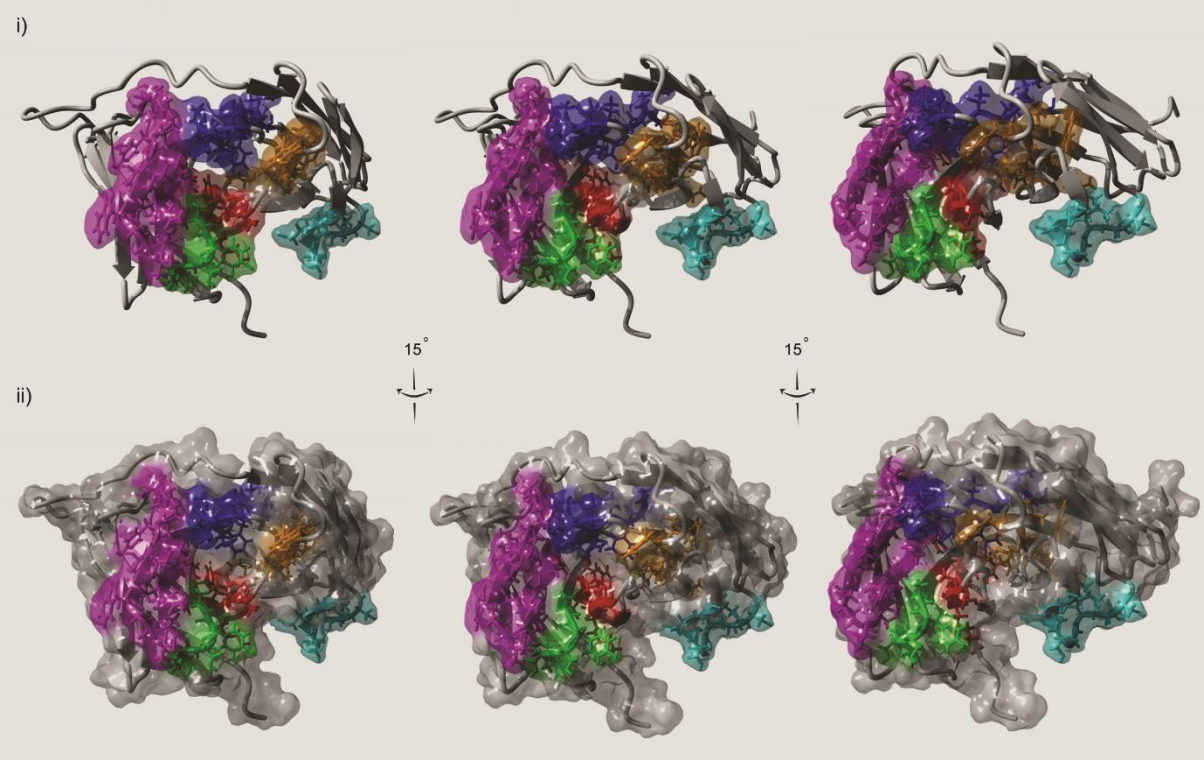


Figure 3.5: Position and orientation of the complimentary determining regions of TFI-scFv.

*Ribbon diagrams of TFI-scFv hybrid model (**Grey**) shown at 15 ° rotations with the of CDR forming residues indicated in colour: Hv-chain CDR1 (**Green**), CDR2 (**Magenta**) and CDR 3 (**Red**) and Lv-chain CDR1 (**Orange**), CDR2(**Cyan**) and CDR3 (**Blue**). i)The transparent surface areas of the CDR forming residues are indicated with the rest of the TFI-scFv surface area omitted. CDRs are located in close proximity to the ligand binding interface towards one pole of the antibody structure to form the binding interface. ii) The total transparent surface area of the TFI-scFv (**Grey**) with CDR in colour. Several the CDR residues are located within the structure of TFI-scFv structure and not exposed to the surrounding solvent.*

3.5.4 Rigid Protein-Protein docking modelling

The hybrid model of TFI-scFv was used for initial rigid body docking using ZDock online server. Rigid protein-protein interactions between the surface exposed CDR region of TFI-scFv and the extracellular domains of TF was performed. Only the top 10 most prevalent docking models were retained after the docking simulations. Of these predictions, only the top 3 most prevalent docking models were retained following the docking simulation and

visually inspection. The 3 top scoring docking models display a high level of similarity in position and orientation of the TFI-scFv. Top predictions were then again visually inspected using YASARA to verify the positioning of the CDRs towards the TF. All the docking predictions suggest that TFI-scFv is in association with the extracellular type III fibronectin domain of TF (TF domain 1) (Figure 3.4).

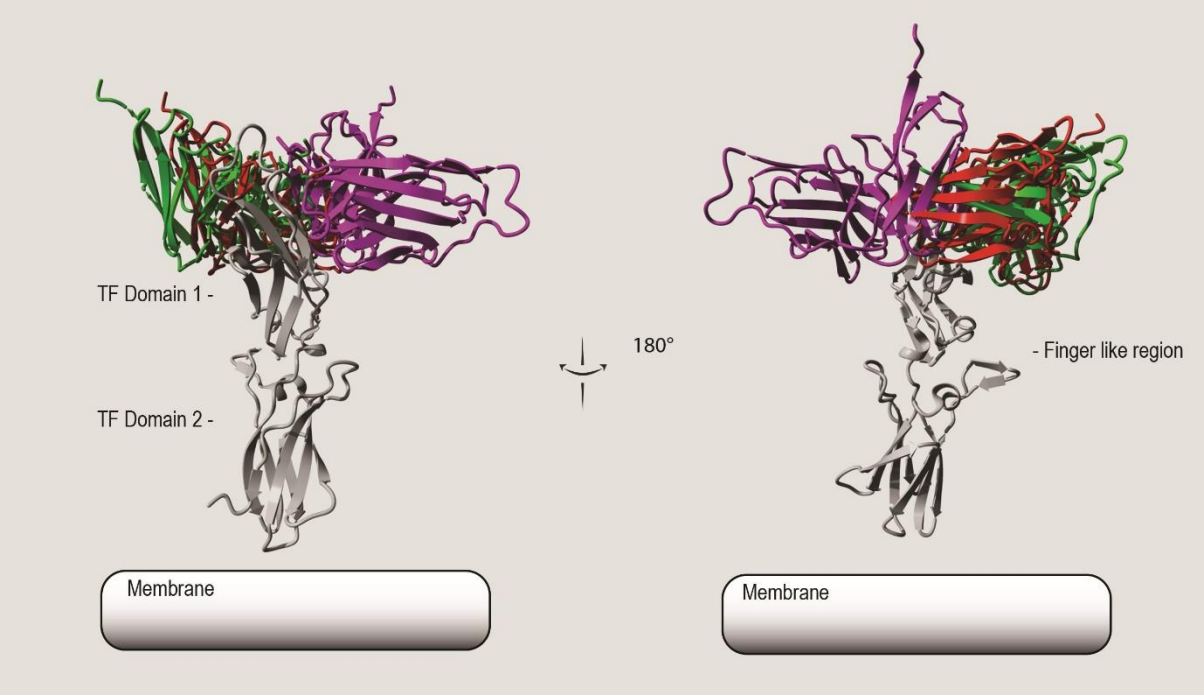


Figure 3.6: Top three rigid model docking prediction based on ZDock simulations.

*The ribbon diagram of the extracellular domains of tissue factor (PDB no: 1boy) in complex with top 3 predicted docking models of TFI-scFv relative to an illustrated lipid membrane. The transmembrane- and intra cellular regions are omitted for the tissue factor structure as per 1boy PDB. The top three predicted docking models based on the TFI-scFv CDR regions of are indicated with the highest scoring prediction (**Red**), the second highest (**Green**) and third highest (**Magenta**).*

The highest- and second highest scoring docking models differ on a few degrees (< 5 °) in orientation. The third highest scoring docking model, is slightly in a more rotated orientation with respect to models 1 and 2 but is still docked in a similar region of tissue factor. The docking position of the top rigid model predictions were similarly positioned. Due to the high level of similarity between the top docking models, only the highest scoring model was retained for final molecular dynamic refinement of interaction between TFI-scFv and tissue factor. Energy minimizations were performed using self-parameterizing knowledge-based YASARA force field and was calculated at -12820.934 kJ/mol for the highest scoring model. The total solvent accessible area of the extracellular domain of tissue factor in complex with TFI-scFv was calculated at 19646.77 $Å^2$ with an inter-molecule interface of 1865.81 $Å^2$ for the rigid docking model. Contact analysis identified only two weak hydrogen bond interactions between TFI-scFv and TF (Table 3.4).

Table 3.4: Rigid body hydrogen bond interaction between TFI-scFv and Tissue factor.

TFI-scFv	*TF*	*Bond distance in Angstrom (Å)*	*Bond Energy kJ/mol*	*Relative H-bond strength**
TRP45	*SER33*	*1.86*	*10.55*	*weak*
TYR51	*TYR183*	*2.00*	*6.85*	*weak*

3.5.5 Molecular dynamic refinement of the Protein-Protein Docking model

Molecular dynamic refinement of the protein-protein docking model was performed using GROMACS over 100 ns simulation. The simulation was analysed and the most prevalent clusters were identified (Table 3.5). Cluster 18 is the most prevailing cluster, spanning approximately 31 % of the total MD simulation. The second and third highest clusters (Cluster 4 and 10) only existed for 16 % and 14% the duration of the simulation respectively. This proportionally large clustering suggests a stable interaction between TFI-scFv and TF for the latter part of the simulations. A representative structure is defined as the specific conformation of the docking model at the middle point (time) of a specific cluster. Such a representative structure of the most prevalent cluster (Cluster 18) was isolated after 84.68 ns at frame number 4234 (Figure 3.5 and Figure 3.6 i).

Table 3.5: Clustering of most prevalent structures.

Rank	*Cluster Number*	*Number of Members*	*Middle Cluster(ns)*	*Middle Cluster Frame Number*
1	*18*	*1552*	*84.68*	*4234*
2	*4*	*799*	*13.48*	*674*
3	*10*	*702*	*38.38*	*1919*
4	*13*	*651*	*56.94*	*2847*
5	*9*	*433*	*25.56*	*1278*

3.5.6 Comparison between the rigid body - and refined model structures.

When comparing structural models of the rigid docking with the MD refined model, it is clear the general structure and interface area was retained following the MDS (Figure 3.5). When comparing the structures of the extra cellular domains of TF in isolation, it is evident that the structure remained largely unchanged with a root mean square deviation (RMSD) of only 1.559 Å using superpose and MUSTANG alignment modules in YASARA. The structure of TFI-scFv showed more of variation with a RMSD of 4.045 Å when comparing structures in isolations. The root mean square deviation (RMSD) between the two complete structures is calculated at 10.6523 Å. The total solvent accessible area of the extracellular domain of tissue factor in complex with TFI-scFv was calculated at 22662.89 $Å^2$ with an inter-molecule interface of 1865.81 $Å^2$ for the refined docking model. Energy minimizations were performed using self-parameterizing knowledge-based YASARA force field and was calculated at -13816.397 kJ/mol for the highest scoring model. This is 995.463 $kJ.mol^1$ lower than that of the rigid body docking models as predicted by ZDOCK, indicating a slight improvement in the overall structural stability.

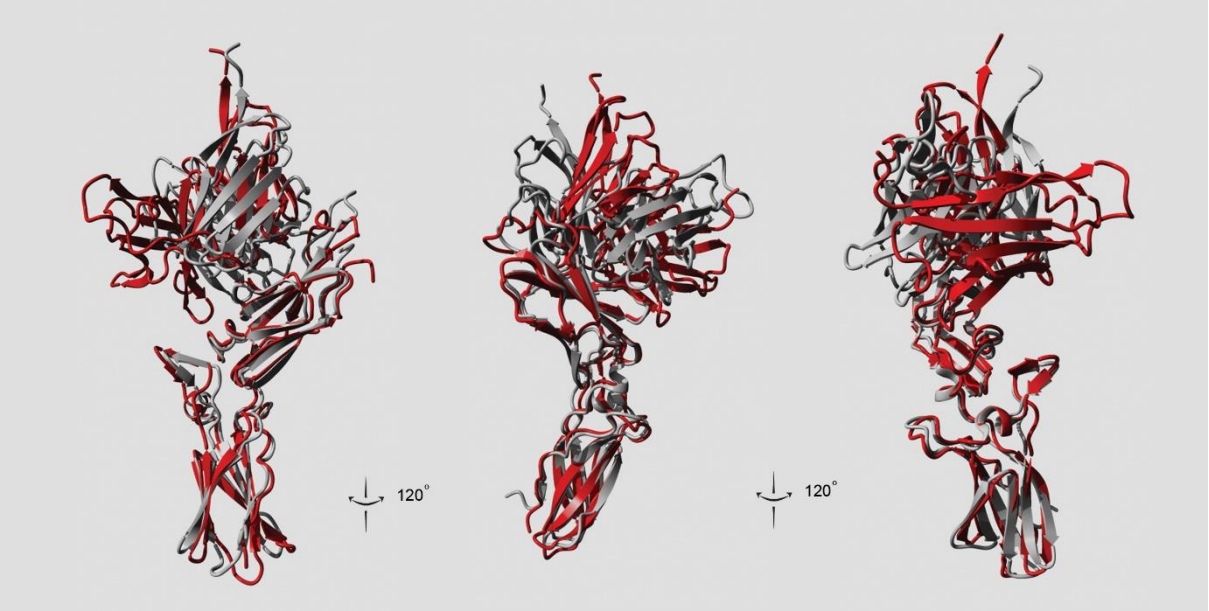


Figure 3.7: Aligned structures of the rigid body- and MD refined docking models.

*The aligned and superimposed ribbon diagrams of the rigid body- (**Grey**) and refined docking models (**Red**) shown at 120 degree rotations. The membrane is presumed to be at the bottom of this view.*

A minimal distance analysis was performed using the MDS to evaluate the extent of movement (disassociation) between TFI-scFv and tissue factor. The analysis was performed for the full 100ns simulation (Figure 3.6). The analysis suggests that the refined docking model is stable during the MD simulation and confirms that dissociation between TFI-scFv and tissue factor does not occur.

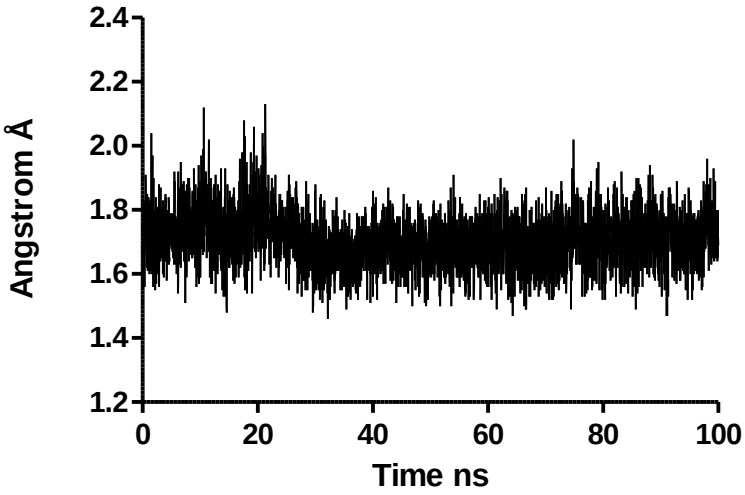


Figure 3.8: Minimal distance analysis between TFI-scFv and Tissue factor.

The distance between the structures fluctuated between 1.460 Å (min) and 2.130 Å (max). The means distance between the two structures is 1.708 Å with a standard deviation of 0.07928 Å.

3.5.7 Contact analysis of the interaction between TFI-scFv and tissue factor.

Molecular dynamic refinement of the rigid body docking model resulted in the optimisation the interactions between TFI-scFv and TF as confirmed by the contact analysis. A series of hydrogen bond, hydrophobic and PiPi interactions are determined by YASARA (Tables 3.6, 3.8, and 3.9). The amino acids forming part of the CDR are indicated in **BOLD**. The position of these interactions is visually summarised in figure 3.7.

Table 3.6: Hydrogen bond interaction between TFI-scFv and Tissue factor.

TFI-scFv	*TF*	*Bond distance in Angstrom (Å)*	*Bond Energy kJ/mol*	*Relative H-bond strength**
TYR226	*GLN37*	*1.86*	*17.60*	*moderate*
ARG61	*ASP44*	*1.75*	*19.55*	*moderate*
THR228	*ASP44*	*1.78*	*25.00*	*moderate*
ARG61	*ASP44*	*1.88*	*23.85*	*moderate*
THR228	*TRP45*	*2.05*	*22.88*	*moderate*
SER227	*SER47*	*1.88*	*25.00*	*moderate*
LYS59	*GLN69*	*1.80*	*25.00*	*moderate*
LEU56	*GLN69*	*2.16*	*17.05*	*moderate*
SER201	*GLY87*	*2.41*	*8.10*	*weak*
SER164	*GLY90*	*1.79*	*25.00*	*moderate*

** Relative hydrogen bonds strength according to Jeffrey's classification.*

Table 3.7: Hydrophobic interactions between TFI-scFv and Tissue factor.

TFI-scFv	*TF*	*Distance in Angstrom (Å)*	*Interaction strength**
TYR226	*PHE76*	*3.41*	*0.030*
SER162	*TYR78*	*3.65*	*0.293*
THR134	*PHE50*	*3.71*	*0.354*
SER162	*TYR51*	*3.73*	*0.419*
GLN161	*TYR51*	*3.73*	*0.420*
ILE136	*PHE50*	*4.98*	*0.685*
SER164	*PRO92*	*4.73*	*0.755*
TYR166	*PRO92*	*3.94*	*0.775*
LEU56	*SER42*	*4.09*	*0.795*
TYR226	*THR35*	*4.62*	*0.808*
LEU56	*GLY43*	*4.61*	*0.835*
LEU56	*SER42*	*4.48*	*0.853*
SER164	*TYR78*	*4.37*	*0.885*
TYR166	*PHE76*	*4.21*	*0.898*
TYR166	*PRO92*	*4.03*	*0.935*
LEU56	*THR40*	*4.33*	*0.949*
LEU56	*THR40*	*4.28*	*0.966*
LEU56	*SER42*	*4.23*	*0.987*
SER201	*GLY90*	*4.14*	*0.989*
SER227	*SER47*	*4.12*	*0.995*
TYR166	*PHE76*	*3.92*	*0.996*

Interaction strength based on YASARA's classification of 0 to 1 with 1 indicating an optimal strength.

Table 3.8 PiPi interaction between TFI-scFv and Tissue factor.

TFI-scFv	*TF*	*Distance in Angstrom (Å)*	*Interaction strength**
TYR226	*PHE76*	*3.41*	*1.000*
TYR166	*PHE76*	*4.26*	*1.000*

Interaction strength based on YASARA's classification of 0 to 1 with 1 indicating an optimal strength.

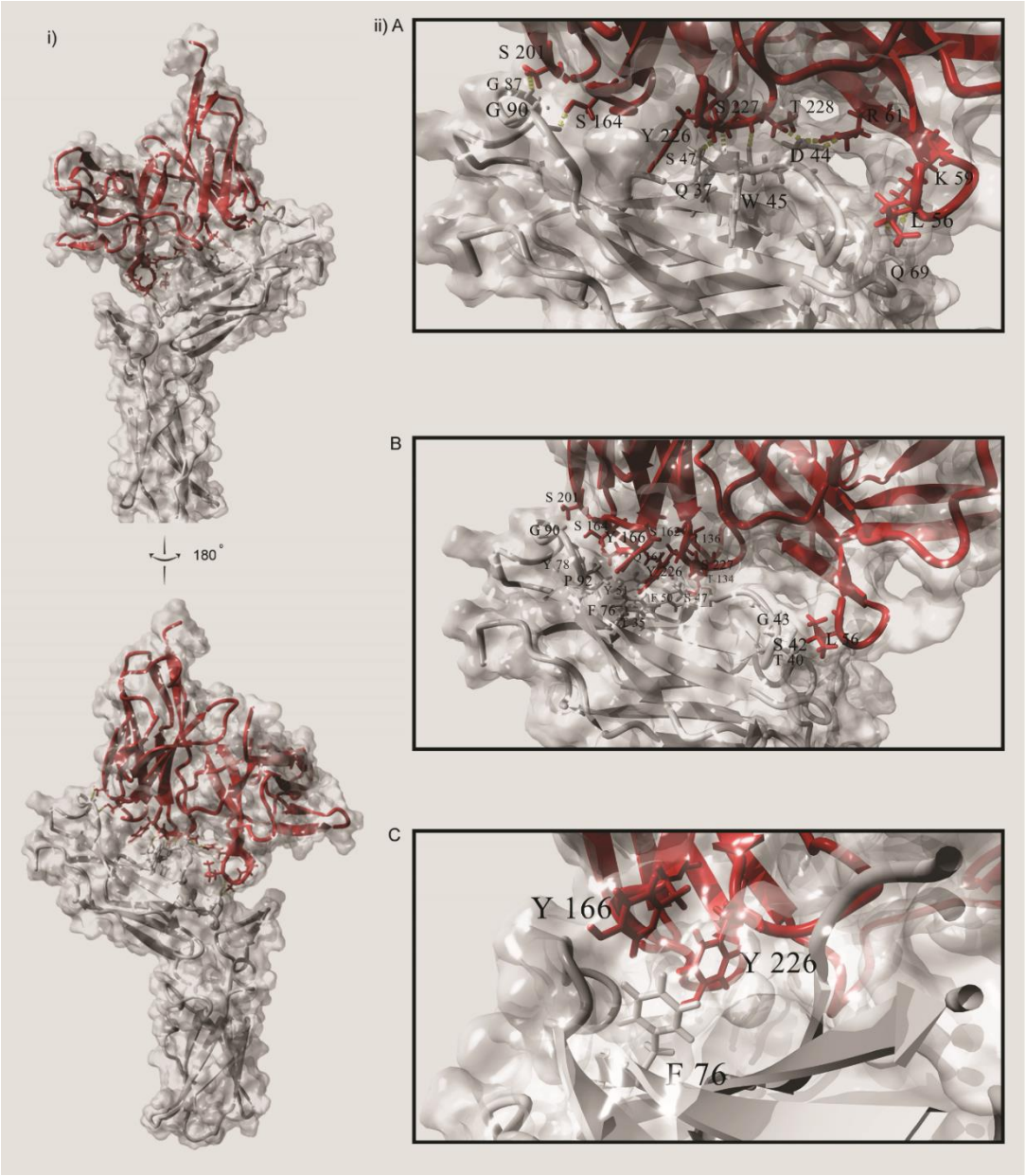


Figure 3.7: Binding interactions between TFI-scFv and tissue factor.

*i) The ribbon diagram of TFI-scFv (**Red**) in association with tissue factor (**Grey**) with total transparent surface area of the TFI-scFv (**Grey**) visualised. The refined model is shown at 180 ° rotations. The membrane is presumed to be at the bottom of this view. ii) Visual summary of the predicted A) Hydrogen bond-, B) Hydrophobic- and C) PiPi interaction present at the interface between TFI-scFv (**Red**) and tissue factor (**Grey**).*

3.5.8 Mechanism of tissue factor inhibition.

Factor VII (FVII), and Factor X (FX) have large extensive interface regions with TF [47]. The residues that are important for interaction between TF and FVII are distributed throughout both of the extracellular domains of TF [48]. In contrast, residues that are important for interaction between TF and FX are only located on the TF2 domain near the membrane surface [49] (Figure 3.8 i). According to the refined docking model, several of the key TF residues (Gln^{37}, Asp44, Lys^{48}, Phe^{76} and Tyr^{78}) responsible for interaction with FVII are obstructed by TFI-scFv and thus not exposed to the solvent accessible area (Figure 3.8).

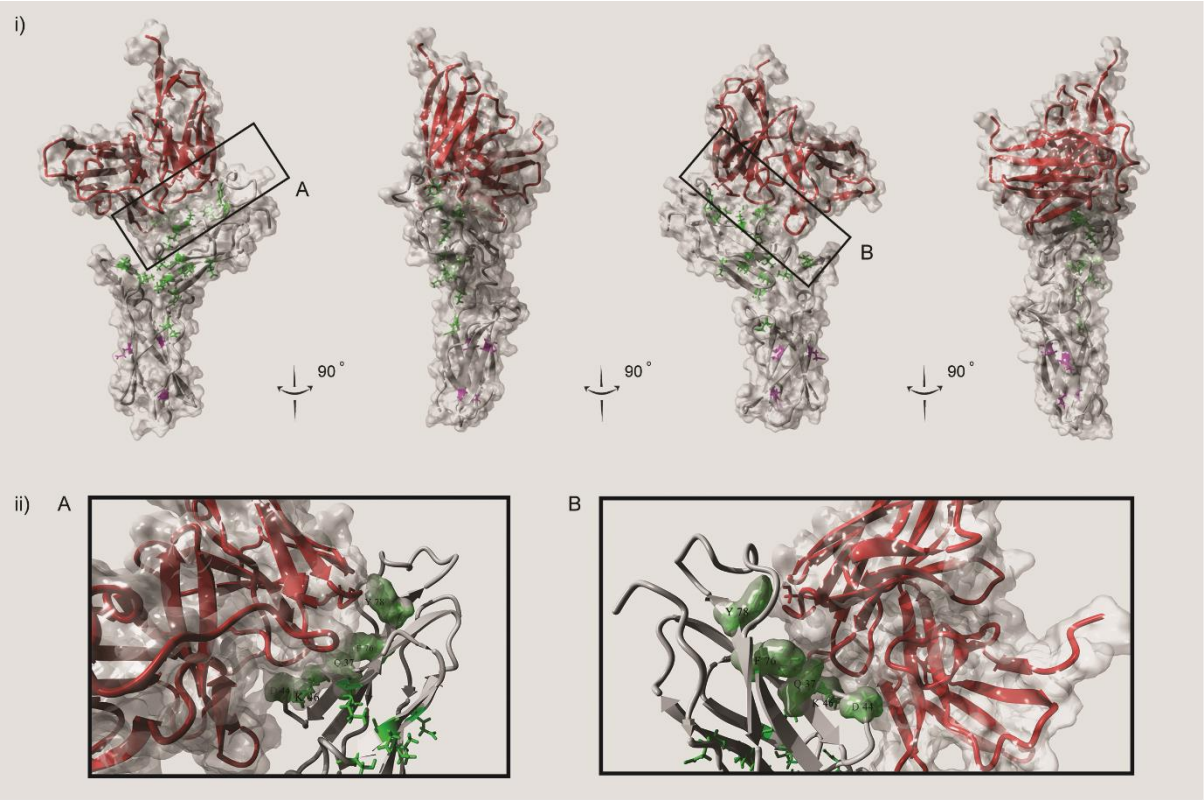


Figure 3.9: TFI-scFv inhibition mechanism.

*i) Rotated ribbon diagram of the docking model between extracellular domain of TF (**Grey**) in association with TFI-scFv (**Red**) with the total transparent surface area (**Grey**) included (See model no: TopMD and supplementary video). The residues important for the interaction between tissue factor and factor VIIa (Thr^{17}, Lys^{20}, Ile^{22}, Glu^{24}, Gln^{37}, Asp^{44}, Lys^{46}, Lys^{48}, Asp^{58}, Thr^{60}, Phe^{76}, Tyr^{78}, Gln^{110}, Leu^{133}, Arg^{135}, Phe^{140}, and Val^{207}) are indicated (**Green**). Residues important for the interaction between tissue factor and factor X (Thr^{154}, Glu^{174} and* Tyr^{185}*) (**Magenta**) The membrane is presumed to be at the bottom of this view. ii) Focused view of the interface area between TFI-scFv and tissue factor as highlighted by the frames A and B. TFI-scFv (**Red**) is show with transparent surface area (**Gray**). The key tissue factor residues (Gln^{37}, Asp44, Lys^{48}, Phe^{76} and Tyr^{78}) that are obstructed by TFI-scFv are indicated with transparent surface area (**Green**).*

The x-ray crystal structure of FVII bound to TF (PDB no: 1J9C) was aligned and super imposed with the refined docking model using superpose and MUSTANG alignment modules in YASARA. It is evident from the predicted alignment model, that TFI-scFv positioned in such a manner that it would sterically interfere with the ability of FVII to associate with TF (Figure 3.9).

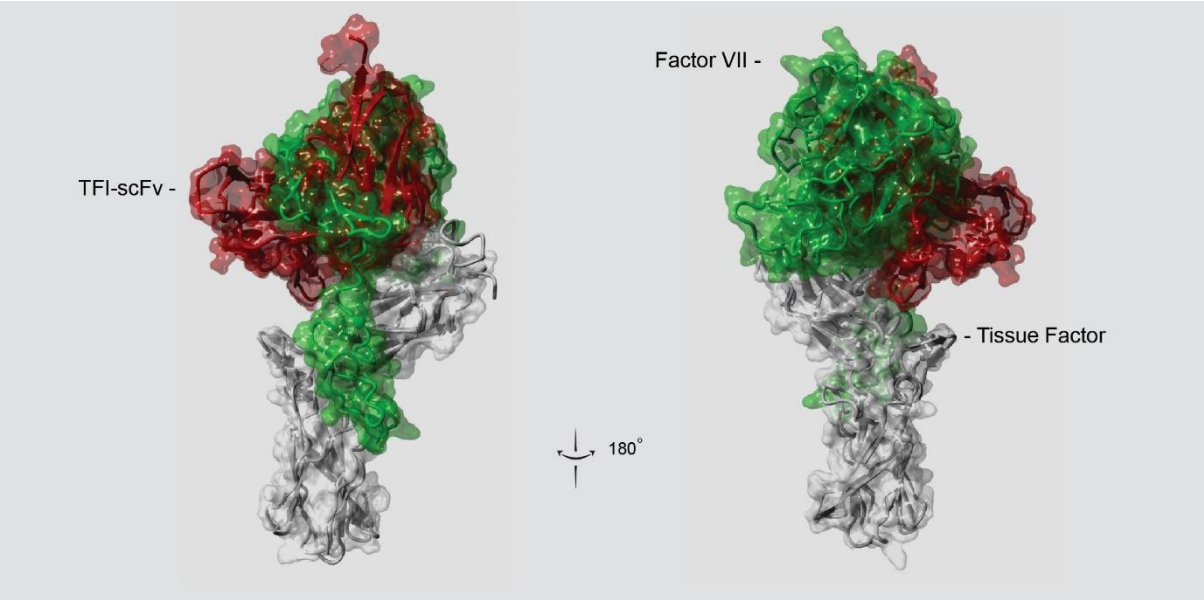


Figure 3.10 Steric hindrance of Factor VII by TFI-scFv.

*The superimposed ribbon diagram of the docking model between extracellular domain of TF (**Grey**) in association with TFI-scFv (**Red**) and FVII (PDB id: 1J9C) (**Green**). The total transparent surface area each molecule included in their respective colours. The membrane is presumed to be at the bottom of this view. Overlapping position of TFI-scFv and FVII suggests that inhibition effect of TFI-scFv is due to steric interference that prevent the activation of Factor VII*

3.6 Discussion

Information regarding the structure of a protein is important as it provides deeper insights into the structural stability of the protein itself as well as the mechanism involved in protein-protein interactions. However, the generations of these structures are often difficult to preform despite significant advances in X-ray crystallography and MNR spectroscopy [27]. The reality is that structures of many therapeutic relevant receptors are still unavailable [32]. Practically, the generation of an X-ray structure of TFI-scFv itself or a structure of the interaction between TFI-scFv would have been extremely difficult if not impossible. The generation of a TFI-scFv crystal would have simply been hampered by low expression yields of TFI-scFv as a considerable amount of target protein is required for the screening process. Even if a sufficient amount of TFI-scFv was available to perform the screening, there is no guarantee that a crystallisation process would be successful [29]. The vast majority of proteins do not yield sufficient quality crystals for structure determination [50]. A further complication would be the generation of a crystal structure interaction between TFI-scFv and TF due to the transmembrane nature of TF. Membrane proteins are notoriously difficult to study as they are highly hydrophobic and often unstable once removed from the membrane [51]. It is therefore not surprising that a complete full-length TF crystal structure has not been published.

Thus, as an alternative, homology modelling was utilised to predict the three-dimensional protein structure of TFI-scFv. Homology modelling involves the generation of protein models based on templates with known experimental X-ray crystallography or NMR. The generally conserved structure shaded between scFv and antibodies in general, is fortuitous with regards to homology modelling and the identification of high quality structural templates [22,52]. It is therefore unsurprising that the top structural template (2YC4) utilised by YASARA was a Human derived scFv and that the other templates were all antibodies of nature apart from Mettalo-β-Lactamasa (4D1V). All of the utilised templates were of high quality and have resolutions below 2 Å. The accuracy of the homology model is highly dependent on the quality of the alignment between the target (unknown) and the template(s) [32]. Using YASARA a total of 7267 alignments were made and the top models generated. The generated models all displayed the general scFv structural conformation of: Heavy chain variable region, linker and Light chain variable region. The general quality of these models were reduced in the region of the linker as indicated by the per residue quality score. This was expected due to the high flexibility of the (Glu4Ser)3 linkers utilised in most scFv [53]. The high level of flexibility is evident in model E where the Hv- and Lv chains are separated form each other with an outstretched stretched out linker connecting the two domains. The lower scoring ("bad regions") in the top ranked homology model were iteratively replaced with better scoring corresponding fragments from the other models. Finally, an improved hybrid model (Model G) was then generated as confirmed by the quality Z-score per residue. The final quality sore of the TFI-scFv hybrid model was calculated as 0.273 which falls within the "optimal" quality range for homology models. Although hybrid model Z-score is slightly lower (0.027) than the score of model A based on the 2YC1 crystal structure (0.300). It however does not infer a lower quality model as the hybrid models covers more residues that the original templates.

Homology modelling of TFI-scFv predicted a heterodimer structure, consisting of tightly packed β-sheet barrel-like structures of the Hv- and Lv chain, connected via a flexible linker. The contact analysis suggests that interaction between the Hv- and Lv chains are maintained by hydrogen bond- and hydrophobic interactions. The Hv- and Lv domains are stabilised by disulfide bridges which span the hydrophobic inner region of the β-sheet barrel-like structures. It has been found that high confidence homology models of antibodies has been generated due to the large collection of antibody crystal templates, the highly conserved nature of antibody structure and tendency of the CDR loops to adopt a few mainchain conformations [54]. ScFv and antibodies in general possess several highly conserved framework and highly variable regions. The role of conserved framework is to orientate the CDRs for interaction with the antigen while variable regions are responsible for ligand recognition [18,45,54,55]. This conformation of framework- and highly variable regions is also evident in the model of TFI-scFv in figure 3.2. The CDR loops, as identified by the Kabat numbering scheme, are oriented towards one pole of the antibody fragment to form the binding interface. The identification of the CDRs and the subsequent binding interface was an important initial step in the rigid body docking between TFI-scFv and TF. The rigid body protein-protein docking using the surface exposed CDR residues was performed using ZDOCK docking online server. ZDOCK utilises specialised algorithms and approaches to perform complex protein–protein docking simulations [34]. These algorithms are

based on spherical harmonics and an energy-based scoring function which include van der Waals-, electrostatics and desolvation terms [56]. The possible interactions in translational and rotational space between the two proteins were screened and evaluated using an energy-based scoring function [56]. Out of the best scoring predicted docking models, only the 3 top scoring docking models were initially retained. These models were visually inspected to confirm the orientation of the CDRs towards TF in such a manner as to not overlap with the omitted membrane region. Ultimately only the highest scoring model was retained for further simulation and analysis due to the high level of similarity between top 3 models. According to the highest scoring docking model, the interaction between antibody and ligand occurs on the extracellular domain of TF, located the furthers from the lipid membrane. Only two weak hydrogen bonds were predicted to facilitate the interaction between molecules. Interestingly, somewhat similar results were obtained when docking was performed using GRAMM-X Protein-Protein Docking Web Server (Data no shown).

The purpose of preforming the ZDOCK modelling was to obtain the best possible starting structure for the MD refinement of the protein-protein interactions. In terms of computational complexity, rigid body docking are much faster to perform than more complex modelling simulations [57]. The modelling performed by ZDOCK servers were generally completed within two to three hours depending on server load as opposed to the MD simulation which took just under a month of continuous computing. However the major drawback to rigid body docking with regard to protein-protein interactions, is that it does not account for the inherent flexibility of the proteins [58]. In other words, the proteins are treated as solid objects and conformational changes that occur during protein-protein interactions are not accounted for. This could result in the generation of low quality or false models. A relevant example, are the conformational changes involved in factor VII activation during its interaction with TF [59]. Rigid body docking would not be capable of accurately predicting the final structural interactions between Factor VII and TF due to the high level of conformational changes that FVII needs to undergo. As a results, rigid body docking modelling simulations are utilised for the generations of initial protein-protein models which are then refined by more complex dynamic methods [60]

Classical- and quantum MD simulations enables the study of biological systems at the atomic level on timescales from femto- to milliseconds [61]. The simulation resolve atomic movement over a time frame to simulate structural function and interactions [62]. MD simulations have provided detailed information on the fluctuations and conformational changes of peptides, proteins, and nucleic acids [63,64] They are also used in the determination of structures from x-ray crystallography and from NMR experiments [65]. The top scoring rigid body docking model, as predicted by ZDOCK, was utilised as the starting position for a 100 ns classical MD simulation. The most frequent conformations were identified though clustering of the various structures that occurs during the simulation. This most prominent cluster occurring towards the end (last 30 ns) of the simulations which suggest a stabilisation of the interaction as the simulation proceeded. The final refined structure of the interaction between TFI-scFv and TF showed a large degree of similarity to the rigid body docking model utilised as the starting position as confirmed by the relatively small RMSD. However more importantly, is that TFI-scFv remained docked to TF throughout the 100 ns simulation as indicated by the minimal distance deviation analysis. If the interactions between TFI-scFv and

TF were not sufficiently robust during the MD simulations it would have resulted in the dissociation of the two molecules. In this case however, the MD simulations resulted in a much-improved interaction between target and ligand as confirmed by the contact analysis. The quantity, quality, and distribution interactions through-out the binding interface were substantially improved. The improvement of the hydrogen bond interaction is especially promising with regards to the stability of the interactions as they are considered the strongest electrostatic dipole-dipole interactions [66]..

The interaction between TF and FVII is predominantly mediated by protein-protein interactions over both the extracellular domains of TF as indicated in figure 3.9 [59]. Large structural differences are seen between the active- and inactive zymogen FVII protease domains [47]. TF serves as a cofactor of FVII by stabilising the active site of FVII once binding takes place [59]. The association of FVII to TF results in allosteric changes to enzyme which results in the dramatic enhancement of catalytic function [49,67]

A series of key residues involved in the interactions have been identified though alanine scanning mutagenesis [49]. Alanine substitution of these residues resulted in 5- to 100 fold reduction in TF:FVII affinity [48]. This suggest that each of these interactions are critical for the allosteric changes required for activation of FVII. According to the refined docking model, TFI-scFv is positioned in such a manner that 5 of these key residues (Gln^{37}, Asp44, Lys^{48}, Phe^{76} and Tyr^{78}) are rendered inaccessible as indicated in figure 3.8. As a result, the interaction between FVII and these key TF residues would be hindered. Furthermore, it is clear from the superimposed structures of TFI-scFv and FVII in complex with tissue factor that, TFI-scFv would sterically interfere with FVII ability to associated with TF. It can therefore be concluded that TFI-scFv functions as an anti-thrombotic, by preventing the association between FVII and TF and the formation of the TF:FVII a complex in the initiation stage of coagulation. Although, the results presented here provide much needed insight in the mechanism of inhibition, it is important to remain mindful that these findings are computerized predictions and might not reflect the *in vivo* interactions correctly.

3.7 Conclusion

According to computerized docking simulations and molecular dynamic refinement of the interactions between the CCR of TFI-scFv and TF, TFI-scFv interfere with the association of FVII with TF. This interference would hinder the activation of FVII and result in the inhibition of the first stage of the coagulation cascade.

3.8 Acknowledgements

This work has been funded by the National Research Foundation (NRF) of South Africa.

3.9 References

[1] J.M. Reichert, Marketed therapeutic antibodies compendium, MAbs. 4 (2012) 413–415. doi:10.4161/mabs.19931.

[2] P. Chames, M. Van Regenmortel, E. Weiss, D. Baty, Therapeutic antibodies: successes, limitations and hopes for the future, Br. J. Pharmacol. 157 (2009) 220–233. doi:10.1111/j.1476-5381.2009.00190.x.

[3] N.C. Nicolaides, P.M. Sass, L. Grasso, Monoclonal antibodies: a morphing landscape for therapeutics, Drug Dev. Res. 67 (2006) 781–789. doi:10.1002/ddr.20149.

[4] K. Tsumoto, K. Shinoki, H. Kondo, M. Uchikawa, T. Juji, I. Kumagai, 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, J. Immunol. Methods. 219 (1998) 119–129. doi:10.1016/S0022-1759(98)00127-6.

[5] Y. Reiter, U. Brinkmann, K.O. Webber, S.-H. Jung, B. Lee, I. Pastan, Engineering interchain disulfide bonds into conserved framework regions of Fv fragments: improved biochemical characteristics of recombinant immunotoxins containing disulfide-stabilized Fv, Protein Eng. Des. Sel. 7 (1994) 697–704. doi:10.1093/protein/7.5.697.

[6] L.J. Lawrence, A.A. Kortt, P. Iliades, P.A. Tulloch, P.J. Hudson, Orientation of antigen binding sites in dimeric and trimeric single chain Fv antibody fragments, FEBS Lett. 425 (1998) 479–484. doi:10.1016/S0014-5793(98)00292-0.

[7] U. Brinkmann, a Di Carlo, G. Vasmatzis, N. Kurochkina, R. Beers, B. Lee, I. Pastan, Stabilization of a recombinant Fv fragment by base-loop interconnection and V(H)-V(L) permutation., J. Mol. Biol. 268 (1997) 107–17. doi:10.1006/jmbi.1996.0850.

[8] V. Crivianu-Gaita, M. Thompson, Aptamers, antibody scFv, and antibody Fab' fragments: An overview and comparison of three of the most versatile biosensor biorecognition elements, Biosens. Bioelectron. 85 (2016) 32–45. doi:10.1016/j.bios.2016.04.091.

[9] D. Moricoli, W.A. Muller, D.C. Carbonella, M.C. Balducci, S. Dominici, R. Watson, V. Fiori, E. Weber, M. Cianfriglia, K. Scotlandi, M. Magnani, Blocking monocyte transmigration in in vitro system by a human antibody scFv anti-CD99. Efficient large scale purification from periplasmic inclusion bodies in E. coli expression system, J. Immunol. Methods. 408 (2014) 35–45. doi:10.1016/j.jim.2014.04.012.

[10] S.M. Meiring, J. Vermeulen, P.N. Badenhorst, Development of an inhibitory antibody fragment to human tissue factor using phage display technology, Drug Dev. Res. 70 (2009) 199–205. doi:10.1002/ddr.20295.

[11] A. Breitenstein, F.C. Tanner, T.F. Lüscher, Tissue Factor and Cardiovascular Disease:, Circ. J. 74 (2010) 3–12. doi:10.1253/circj.CJ-09-0818.

[12] S. Butenas, T. Orfeo, K.E. Brummel-Ziedins, K.G. Mann, Tissue factor in thrombosis and hemorrhage., Surgery. 142 (2007) 2–14. doi:10.1016/j.surg.2007.06.032.

[13] F. Samad, Regulation of tissue factor gene expression in obesity, Blood. 98 (2001) 3353–3358. doi:10.1182/blood.V98.12.3353.

[14] N. MACKMAN, The many faces of tissue factor, J. Thromb. Haemost. 7 (2009) 136–139. doi:10.1111/j.1538-7836.2009.03368.x.

[15] E. Garet, A.G. Cabado, J.M. Vieites, A. González-Fernández, Rapid isolation of single-chain antibodies by phage display technology directed against one of the most potent marine toxins: Palytoxin, Toxicon. 55 (2010) 1519–1526. doi:10.1016/j.toxicon.2010.03.005.

[16] J.P. Yan, J.H. Ko, Y.P. Qi, Generation and characterization of a novel single-chain antibody fragment specific against human fibrin clots from phage display antibody library, Thromb. Res. 114 (2004) 205–211. doi:10.1016/j.thromres.2004.06.013.

[17] K. Proba, A. Wörn, A. Honegger, A. Plückthun, Antibody scFv fragments without disulfide bonds, made by molecular evolution, J. Mol. Biol. 275 (1998) 245–253. doi:10.1006/jmbi.1997.1457.

[18] N. a Watkins, W.H. Ouwehand, Introduction to antibody engineering and phage display., Vox Sang. 78 (2000) 72–9. doi:31154.

[19] T. Arakawa, D. Ejima, Refolding Technologies for Antibody Fragments, Antibodies. 3 (2014) 232–241. doi:10.3390/antib3020232.

[20] X. Chen, J.L. Zaro, W.-C. Shen, Fusion protein linkers: Property, design and functionality, Adv. Drug Deliv. Rev. 65 (2013) 1357–1369. doi:10.1016/j.addr.2012.09.039.

[21] Q. Wu, X. Wang, Y. Gu, X. Zhang, Y. Qin, H. Chen, X. Xu, T. Yang, M. Zhang, Screening and identification of human ZnT8-specific single-chain variable fragment (scFv) from type 1 diabetes phage display library, Sci. China Life Sci. 59 (2016) 686–693. doi:10.1007/s11427-016-5077-7.

[22] Z.A. Ahmad, S.K. Yeap, A.M. Ali, W.Y. Ho, N.B.M. Alitheen, M. Hamid, scFv Antibody: Principles and Clinical Application, Clin. Dev. Immunol. 2012 (2012) 1–15. doi:10.1155/2012/980250.

[23] F. Sefid, I. Rasooli, Z. Payandeh, Homology modeling of a Camelid antibody fragment against a conserved region of Acinetobacter baumannii biofilm associated protein (Bap), J. Theor. Biol. 397 (2016) 43–51. doi:10.1016/j.jtbi.2016.02.015.

[24] D. Yan, W. Han, Z. Dong, Q. Liu, Z. Jin, D. Chu, Y. Tian, J. Zhang, D. Song, D. Wang, X. Zhu, Homology modeling and docking studies of ENPP4: a BCG activated tumoricidal macrophage protein, Lipids Health Dis. 15 (2016) 19. doi:10.1186/s12944-016-0189-4.

[25] Z. Sayers, B. Avşar, E. Cholak, I. Karmous, Application of Advanced X-ray Methods in Life Sciences, Biochim. Biophys. Acta - Gen. Subj. (2016). doi:10.1016/j.bbagen.2016.05.008.

[26] R. Nussinov, Introduction to Protein Ensembles and Allostery, Chem. Rev. 116 (2016) 6263–6266. doi:10.1021/acs.chemrev.6b00283.

[27] T. Schwede, J. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: An automated protein homology-modeling server, Nucleic Acids Res. 31 (2003) 3381–3385. doi:10.1093/nar/gkg520.

[28] Y. Sefidbakht, O. Ranaei Siadat, F. Taheri, Homology Modeling and Molecular Dynamics Study on Schwanniomyces Occidentalis Alpha Amylase., J. Biomol. Struct. Dyn. 1102 (2016) 1–27. doi:10.1080/07391102.2016.1154892.

[29] C.N. Cavasotto, S.S. Phatak, Homology modeling in drug discovery: current trends and applications, Drug Discov. Today. 14 (2009) 676–683. doi:10.1016/j.drudis.2009.04.006.

[30] Y. Zhang, Progress and challenges in protein structure prediction, Curr. Opin. Struct. Biol. 18 (2008) 342–348. doi:10.1016/j.sbi.2008.02.004.

[31] Y. Zhang, Protein structure prediction: when is it useful?, Curr. Opin. Struct. Biol. 19 (2009) 145–155. doi:10.1016/j.sbi.2009.02.005.

[32] A. Nayeem, D. Sitkoff, S. Krystek, A comparative study of available software for high-accuracy homology modeling: from sequence alignments to structural models., Protein Sci. 15 (2006) 808–824. doi:10.1110/ps.051892906.

[33] E. Krieger, G. Vriend, YASARA View--molecular graphics for all devices--from smartphones to workstations, Bioinformatics. 30 (2014) 2981–2982. doi:10.1093/bioinformatics/btu426.

[34] B.G. Pierce, K. Wiehe, H. Hwang, B.H. Kim, T. Vreven, Z. Weng, ZDOCK server: Interactive docking prediction of protein-protein complexes and symmetric multimers, Bioinformatics. 30 (2014) 1771–1773. doi:10.1093/bioinformatics/btu097.

[35] R.W.W. Hooft, G. Vriend, C. Sander, E.E. Abola, Errors in protein structures, Nature. 381 (1996) 272–272. doi:10.1038/381272a0.

[36] R.W. Hooft, C. Sander, M. Scharf, G. Vriend, The PDBFINDER database: a summary of PDB, DSSP and HSSP information with added value., Comput. Appl. Biosci. 12 (1996) 525–9. doi:10.1093/bioinformatics/12.6.525.

[37] D.T. Jones, Protein secondary structure prediction based on position-specific scoring matrices, J. Mol. Biol. 292 (1999) 195–202. doi:10.1006/jmbi.1999.3091.

[38] J.C. Canul-Tec, L. Riaño-Umbarila, E. Rudiño-Piñera, B. Becerril, L.D. Possani, A. Torres-Larios, Structural basis of neutralization of the major toxic component from the scorpion Centruroides noxius Hoffmann by a human-derived single-chain antibody fragment, J. Biol. Chem. 286 (2011) 20892–20900. doi:10.1074/jbc.M111.238410.

[39] C. Eigenbrot, T. Gonzalez, J. Mayeda, P. Carter, W. Werther, T. Hotaling, J. Fox, J. Kessler, X-ray structures of fragments from binding and nonbinding versions of a humanized anti-CD18 antibody: structural indications of the key role of VH residues 59 to 65., Proteins. 18 (1994) 49–62. doi:10.1002/prot.340180107.

[40] T. Tanaka, R.L. Williams, T.H. Rabbitts, Tumour prevention by a single antibody domain targeting the interaction of signal transduction proteins with RAS, EMBO J. 26 (2007) 3250–3259. doi:10.1038/sj.emboj.7601744.

[41] L. Riaño-Umbarila, L.M. Ledezma-Candanoza, H. Serrano-Posada, G. Fernández-Taboada, T. Olamendi-Portugal, S. Rojas-Trejo, I. V. Gómez-Ramírez, E. Rudiño-Piñera, L.D. Possani, B. Becerril, Optimal neutralization of Centruroides noxius venom is understood through a structural complex between two antibody fragments and the Cn2 toxin, J. Biol. Chem. 291 (2016) 1619–1630. doi:10.1074/jbc.M115.685297.

[42] H. Watanabe, K. Tsumoto, S. Taguchi, K. Yamashita, Y. Doi, Y. Nishimiya, H. Kondo, M. Umetsu, I. Kumagai, A human antibody fragment with high affinity for biodegradable polymer film, Bioconjug. Chem. 18 (2007) 645–651. doi:10.1021/bc060203y.

[43] G. A. Jeffrey, An Introduction to Hydrogen Bonding, Oxford University Press, Oxford, 1997.

[44] E.A. Kabat, T. Wut, Identical V regions amino acid sequences and segments of sequences in antibodies of different specificities. Relative Contributions of VH and VL Genes, Minigenes, and Complementarity-Determining Regions to Binding of Antibody-Combining Sites, J Immunol. 147 (1991) 1709–1719. http://www.jimmunol.org/content/147/5/1709.

[45] J. Davies, L. Riechmann, Affinity improvement of single antibody VH domains: residues in all three hypervariable regions affect antigen binding, Immunotechnology. 2 (1996) 169–179. doi:10.1016/S1380-2933(96)00045-0.

[46] E.A. Kabat, T.T. Wu, H.. Perry, G. K.S, F. C, Sequences of Proteins of Immunological Interst, fifth ed, US Department of Health and Human Services, Bethesda, MD, 1991.

[47] B. V. Norledge, R.J. Petrovan, W. Ruf, A.J. Olson, The tissue factor/factor VIIa/factor Xa complex: A model built by docking and site-directed mutagenesis, Proteins Struct. Funct. Genet. 53 (2003) 640–648. doi:10.1002/prot.10445.

[48] D.M. Martin, C.W. Boys, W. Ruf, Tissue factor: molecular recognition and cofactor function., FASEB J. 9 (1995) 852–9. http://www.ncbi.nlm.nih.gov/pubmed/7615155.

[49] S. Butenas, Tissue Factor Structure and Function, Scientifica (Cairo). 2012 (2012) 1–15. doi:10.6064/2012/964862.

[50] W.N. Price II, Y. Chen, S.K. Handelman, H. Neely, P. Manor, R. Karlin, R. Nair, J. Liu, M. Baran, J. Everett, S.N. Tong, F. Forouhar, S.S. Swaminathan, T. Acton, R. Xiao, J.R. Luft, A. Lauricella, G.T. DeTitta, B. Rost, G.T. Montelione, J.F. Hunt, Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data, Nat. Biotechnol. 27 (2009) 51–57. doi:10.1038/nbt.1514.

[51] E.P. Carpenter, K. Beis, A.D. Cameron, S. Iwata, Overcoming the challenges of membrane protein crystallography, Curr. Opin. Struct. Biol. 18 (2008) 581–586. doi:10.1016/j.sbi.2008.07.001.

[52] W. Wang, S. Singh, D.L. Zeng, K. King, S. Nema, Antibody Structure, Instability, and Formulation, J. Pharm. Sci. 96 (2007) 1–26. doi:10.1002/jps.20727.

[53] X. Gu, X. Jia, J. Feng, B. Shen, Y. Huang, S. Geng, Y. Sun, Y. Wang, Y. Li, M. Long, Molecular modeling and affinity determination of scFv antibody: Proper linker peptide enhances its activity, Ann. Biomed. Eng. 38 (2010) 537–549. doi:10.1007/s10439-009-9810-2.

[54] K. He, X. Zhang, L. Wang, X. Du, D. Wei, Production of a soluble single-chain variable fragment antibody against okadaic acid and exploration of its speci fi c binding, Anal. Biochem. 503 (2016) 21–27. doi:10.1016/j.ab.2015.12.020.

[55] R. Barderas, J. Desmet, P. Timmerman, R. Meloen, J.I. Casal, Affinity maturation of antibodies assisted by in silico modeling, Proc. Natl. Acad. Sci. 105 (2008) 9029–9034. doi:10.1073/pnas.0801221105.

[56] J. Yu, M. Vavrusa, J. Andreani, J. Rey, P. Tufféry, R. Guerois, InterEvDock: a docking server to predict

the structure of protein–protein interactions using evolutionary information, Nucleic Acids Res. 44 (2016) W542–W549. doi:10.1093/nar/gkw340.

[57] V. Mohan, A. Gibbs, M. Cummings, E. Jaeger, R. DesJarlais, Docking: Successes and Challenges, Curr. Pharm. Des. 11 (2005) 323–333. doi:10.2174/1381612053382106.

[58] C.J. Camacho, D.W. Gatchell, S.R. Kimura, S. Vajda, Scoring docked conformations generated by rigid-body protein-protein docking, Proteins Struct. Funct. Genet. 40 (2000) 525–537. doi:10.1002/1097-0134(20000815)40:3<525::AID-PROT190>3.0.CO;2-F.

[59] P. Carmeliet, D. Collen, Molecules in focus tissue factor, Int. J. Biochem. Cell Biol. 30 (1998) 661–667. doi:10.1016/S1357-2725(97)00121-0.

[60] I. Halperin, B. Ma, H. Wolfson, R. Nussinov, Principles of docking: An overview of search algorithms and a guide to scoring functions, Proteins Struct. Funct. Genet. 47 (2002) 409–443. doi:10.1002/prot.10115.

[61] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the Amber biomolecular simulation package, Wiley Interdiscip. Rev. Comput. Mol. Sci. 3 (2013) 198–210. doi:10.1002/wcms.1121.

[62] G. Ciccotti, M. Ferrario, C. Schuette, Molecular Dynamics Simulation, 2014.

[63] K.P.D. Thayer, B. Slaw, I. Han, T. Quinn, Structural Bioinformatics and Molecular Dynamics of p53 Tumor Suppressor Protein, FASEB J. 29 (2015) 712.12-. http://www.fasebj.org/cgi/content/long/29/1_Supplement/712.12 (accessed December 27, 2016).

[64] M. Pasi, J.H. Maddocks, R. Lavery, Analyzing ion distributions around DNA: Sequence-dependence of potassium ion distributions from microsecond molecular dynamics, Nucleic Acids Res. 43 (2015) 2412–2423. doi:10.1093/nar/gkv080.

[65] M. Satoh, H. Saburi, T. Tanaka, Y. Matsuura, H. Naitow, R. Shimozono, N. Yamamoto, H. Inoue, N. Nakamura, Y. Yoshizawa, T. Aoki, R. Tanimura, N. Kunishima, Multiple binding modes of a small molecule to human Keap1 revealed by X-ray crystallography and molecular dynamics simulation, FEBS Open Bio. 5 (2015) 557–570. doi:10.1016/j.fob.2015.06.011.

[66] G.A. Jeffrey, W. Saenger, Hydrogen Bonding in Biological Structures, Springer Berlin Heidelberg, Berlin, Heidelberg, 1991. doi:10.1007/978-3-642-85135-3.

[67] D.P. O'Brien, J.S. Anderson, D.M. a Martin, P.G.H. Byfield, E.G.D. Tuddenham, Structural requirements for the interaction between Tissue Factor and Factor VII: Characterization of chymotrypsin-derived Tissue Factor polypeptides, Biochem. J. 292 (1993) 7–12.

Chapter 4: Improved production of a single chain variable fragment against human tissue factor using rare codon optimization and *in vitro* protein refolding techniques.

List of Abbreviations

Codon Adaptation Index	CAI
Dimethiel Sulfoxide	DMSO
*E.*coli Codon Usage Analyzer	*E*CUA
Escherichia coli	E.coli
Guanidinium-HCl	GdnHCl
Heavy chain variable domains	H_v
Immobilised Metal Affinity Chromatography	IMAC
Isopropyl β-D-1-thiogalactopyranoside	IPTG
Light chain variable domains	V_L
Medical Research Council	MRC
Oxidized glutathione	GSSG
Prothombin Time	PT
Reduced glutathione	GSH
Single chain variable fragment	scFv
Tissue factor	TF
Tris Buffered Saline	TBS

Chapter 4: Improved production of a single chain variable fragment against human tissue factor using rare codon optimization and *in vitro* protein refolding techniques.

4.1 Abstract

The development of therapeutic antibodies in their various forms has been a constant challenge since the development of the first monoclonal antibodies in 1975. We previously described the selection of a single chain variable fraction isolated from the Thomlinson I&J phage libraries. In this chapter we aim to, optimization of the expression of this previously low yielding antibody in *Escherichia coli*. The single chain variable fragment (TFI-scFv) was codon optimized and expressed in *Escherichia coli*. Modifications to the gene and expression system redirected the expression of TFI-scFv to the cytoplasm as inclusion bodies. The production of inclusion bodies resulted in an 80-fold improvement in yield. The inclusion bodies were isolated, purified and subjected to varying *in vitro* refolding conditions in order to identify optimum conditions that would yield functional TFI-scFv. Prothrombin time analyses were used to assess the biological activity of renatured antibody fragments. Results revealed that the scFv could be expressed at higher levels in *E. coli* cytoplasm, isolated by means of immobilized metal affinity chromatography (IMAC) under denaturing conditions and refolded to produce functional TFI-scFv. This strategy resulted in a 6-fold improvement in the production of final scFv yield. This procedure could be followed to improve the yield of similar scFv isolated from phage display repositories libraries.

4.2 Key-Words

Single chain variable fragments, *In Vitro* protein refolding, Tissue factor inhibitor, immobilized metal affinity chromatography, Codon optimization

4.3 Introduction

The development of antibodies as therapeutic agents has always been an attractive prospect due to their high specificity, affinity to a wide variety of molecules and stability [1]. The use of antibody fragments is an especially attractive strategy for the development of therapeutic and diagnostic applications due to their relative small size (27 kDa), low immunogenicity and relative ease of selection by means of phage display [2]. Previously we reported on the utilization of phage display technology to select a single chain antibody fragment (scFv) from the Tomlinson I + J Human Single Fold Phage Libraries that functionally inhibits human tissue factor [3]. Tissue factor (TF) plays a vital role in the haemostatic response to damage [4]. Excessive or aberrant TF expression has been linked to inflammation and thrombosis which is a major contributor to global disease and mortality [5,6]. The inhibition of TF has great potential for the development of novel anti-thrombotic agents. Phage display technology is well established and over the years has had many diverse applications such as: affinity selection, epitope mapping, identifying of receptors and ligands as well as drug discovery [7]. The scFvs produced by phage display are currently the leading source of human antibodies, and have led to development of therapeutic antibodies including

adalimumab (Humira™) and belimumab (Benlysta™) [8]. Tomlinson I + J libraries comprise of small monovalent fragments consisting of the V_H and V_L domains from human immunoglobulin (IgG). The domains are connected via a flexible linker $(Gly_4Ser)_3$ to maintain relative orientation and are stabilized by intra-domain disulfide bridges [9,10]. Despite phage display's exceptional ability to generate a vast variety of antibodies against practically any target, the utilization of scFv is limited due to complications that arise during protein production. Expression in *E. coli* systems are often preferred for the production of therapeutics, on laboratory as well as industral scale, due to low cost and simplicity of cultivation. As a result, a constant challenge in the field of protein engineering is to overcome the lack of post-translational modification mechanisms inherent to *E. coli* [11].

However, the situation is exceedingly more complex. Heterologous protein expression of recombinant proteins is an process that involves multiple synthetic pathways and is regulated at both transcriptional and translational levels [12]. The expression of foreign proteins in *E. coli* is frequently complicated due to differences in codon usage, a phenomenon known as codon bias [13]. Due to the degeneracy of the genetic code, several codons translate for a particular amino acid [14]. The frequency of usage of a specific codon for a specific amino acid is not random and differs dramatically between species [15]. In other words, a specific codon and its corresponding tRNA may be abundant in one species but rare in another [16]. These low frequency usage codons are known as rare codons [17]. The presence of rare codons leads to translational errors such as stalling, termination and amino acid substitutions [18,19]. The discrepancy in the availability of a specific codon between the original host of a gene and the heterologous expression host has been shown to have a detrimental effect on protein production and yield [20]. These complications can be circumvented by substitution of rare codons with synonymous codons of a higher frequency which contribute to higher expression levels [19]. Conversely, the higher expression rate achieved due to synonymous codon substitution frequently results in the formation of insoluble protein aggregates, known as inclusion bodies [21,22]. Bacterial inclusion bodies are dense spherical particles of 0.5–1.3 μm in size of aggregated protein that accumulate in the cell cytoplasm [23]. The major advantages of the expression of proteins as inclusion bodies are (i) very high protein yield up to 15- 25% of total protein (ii) simplistic isolation due to their high density and relative size in comparison with other cellular components, (iii) lower rate of degradation, (iv) resistance to proteolytic cleavage, and (v) high level of purity as the inclusion bodies contain little contaminants [12,23–25]. The challenge, however, lies in the generation of native proteins from isolated inclusion bodies by means of *in vitro* refolding techniques. Various refolding techniques have been developed to refold scFv, that include direct dilution, dialysis, diafilteration, chromatographic methods, matrix-assisted refolding, and chemical-assisted refolding with varying levels of success [26]. In this study, we report on the process used to (i) perform molecular modifications to the scFv gene and expression systems to improve protein yield (ii) the production, isolation and purification of scFv inclusion bodies from *E. coli* cytoplasm and (iii) refolding of inclusion bodies *in vitro* to produce functional TFI-scFv.

4.4 Materials and Methods

4.4.1 Bacterial cell lines and plasmids

E.coli strain TG1 (MRC Centre for Protein Engineering, Cambridge, UK) was used for phagemid production, Top10 (Invitrogen, USA) was used for the in vitro amplification of pSMART cloning vectors (Lucigen, USA) and BL21(DE3) (Invitrogen, USA) was use for scFv expression.

4.4.2 Rare Codon Visualization

Visualisation of the distribution of rare codons with in the JTC5-scFv was preformed using web-based application Protter: interactive protein feature visualization and integration with experimental proteomic data v1.0 [27]

4.4.3 Construction of TFI-scFv Expression Vector

The codon optimised gene was synthesised for expression in *E. coli* by GeneART (Regensburg, Germany). Rare codons were identified using *E*.coli Codon Usage Analyzer v2.1 (*E*CUA) (http://www.faculty.ucr.edu/). The frequency and distribution of codons pre- and post-optimization was analysed using Graphical codon usage analyser v2.0 utilizing codon usage table for *E. coli*. [28]. The Codon Adaptation Index (CAI) value was calculated using the CAIcal server [29]. The codon optimized TFI-scFv sequence was synthesised to incorporate 5' *NdeI* and 3' *XhoI* restriction sites for downstream cloning. The optimized scFv gene was amplified using sense primer 5'-GAAGGAAGGCCGTCAAGG-3', anti-sense primer: 5'-CTCGAGATTACGTTTAATTTCAACTTTGG-3' and sub cloned into *NdeI/XhoI* linearized expression vector pET-22b(+). The new construct was designated pET22 TFI-scFv and verified by sequencing.

4.4.4 Sequencing of the scFv gene fragment

All DNA sequencing reactions were performed using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA) in conjunction with the 3130xl Genetic Analyzer HITACHI (Applied Biosystems, USA) according to manufacturer's specifications. The original pIT2 plasmid containing the TFI-scFv gene was sequenced using LBR (sense) 5'-CAGGAAACAGCTATGAC-3' and pHEN (Anti-sense) 5'-CTATGCGGCCCCATTCA-'3 primers. The pET22-TFI expression vector was sequenced using T7 Promoter Primer: 5'TAATACGACTCACTATAGGG 3' and T7 Terminator Primer: 5'GCTAGTTATTGCTCAGCG G 3' (Integrated DNA technologies, Belgium). The DNA- and amino acid multiple sequence alignment of the unoptimized (JTC5-scFv) and optimized (pET22 TFI-scFV) was performed using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/).

4.4.5 Expression of pET22 TFI-scFv in E. coli BL21(DE3)

E. coli BL21(DE3) cells transformed with pET22 TFI-scFv in parallel with a control harbouring empty pET22b(+) plasmid were incubated in LB-media (5 $g.L^{-1}$ NaCl, 5 $g.L^{-1}$ Yeast Extract, 10 $g.L^{-1}$ Bacto-Tryptone) containing 100 $\mu g.mL^{-1}$ Ampicillin for 16 hours at 37 °C, 160 rpm. The overnight cultures were used at 1:100 dilution to inoculate 5L LB Media containing 100 $\mu g.mL^{-1}$ Ampicillin. Cultures were incubated to mid-log phase and induced by 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) for 5 hours at 37 °C, with shaking160 rpm. Cultures were harvested by centrifugation at 15000 x g for 10 minutes at 4 ˚C, the supernatant discarded and cell pellet stored at -20 °C. Harvested cells were disrupted using the One-Shot Homogenizer (Constant Systems, UK) at 30 000 KPSI. Soluble and insoluble fractions were separated by centrifugation at 15000 x g for 10 minutes at 4 ˚C. The soluble fraction was further fractionated using ultra-centrifugation at 100000 x g for 20 min at 4˚C. All fractions were evaluated by 10 % SDS-PAGE as described by Laemmli (1970) and stained as described by Fairbanks and co-workers (1971).

4.4.6 Isolation of inclusion bodies

Cell pellets were thawed on ice and washed twice using disruption buffer (50 mM tris buffer containing 500 mM NaCl, 10 mM EDTA, 5mM DTT and 2% Triton X 100 at pH 7.5). Cells were disrupted using the One-Shot Homogenizer (Constant Systems, UK) at 30 000 KPSI. Soluble and insoluble fractions were separated by centrifugation at 15000 x g for 10 minutes at 4 ˚C. The insoluble fraction was again washed twice using disruption buffer and a third time using a chelating buffer [50 mM Tris and 10 mM EDTA at pH 7.5]. The pellets were suspended in a minimal volume of Dimetiel Sulfoxide (DMSO) and kept at room temperature for 30 minutes. Pellets were solubilized in denaturing buffer (50 mM tris at pH 7.5 containing 6M guanidinium-HCl (GdnHCl) and 5 mM dithiothreitol) at room temperature for 3 hours with slow magnetic stirring. The soluble and insoluble fractions were separated again using ultra-centrifugation at 100000 x g for 20 min at 4˚C. Protein concentrations of the solubilized fractions were determined using the Q-Bit Illuminometer and Q-Bit Protein assay (ThermoFisher scientific, USA) according to manufacturer's specifications.

4.4.7 Purification of TFI-scFv using IMAC under denaturing conditions

TFI-scFv inclusion bodies were purified under denaturing conditions using a 5 mL His Trap FF column (GE Healthcare, USA) in conjunction with the ÄKTAprime plus (Amersham Biosciences, UK) chromatography system according to manufacturer's specification. The column and system were first equilibrated with 100 mL binding buffer at 5 $mL.min^{-1}$ (50 mM tris containing 6 M Gdn-HCl, 0.5 M NaCl and 20 mM Imidazole at pH 7.5). The denatured sample was diluted (1:10) with binding buffer and loaded at 1 $mL.Min^{-1}$. The column was then washed with 50 mL wash buffer at 1 $mL.min^{-1}$. Elution was performed in binding buffer using a 100 mL linear imidazole gradient (20-300 mM) at 1 $mL.min^{-1}$. Protein concentration of the solubilized fractions were determined using the Q-Bit Illuminometer and Q-Bit Protein assay. Purified TFI-scFv was resolved by means of 10% SDS-PAGE.

4.4.8 Screening of refolding conditions

Optimum refolding conditions were determined using the Pierce Protein Refolding Kit (Thermo scientific, USA) according to manufacturer's specifications. Batch-wise refolding was performed using 500 µg denatured TFI-scFv overnight at 4 °C. Due to the presence of intra-domain disulfide bonds, redox conditions were investigated using oxidized (GSSG) and reduced (GSH) glutathione. Buffer exchange was performed by overnight dialysis at 4 °C against Tris-buffered saline (50 mM tris containing 150 mM NaCl, pH7.5) using 10 000 MWCO Snake Skin Pleated Dialysis tubing (Thermo scientific, USA) as described by Stanton and co-workers (2011). The soluble fraction was isolated by centrifugation at 15000 x g for 10 minutes at 4 ˚C and evaluated by 10% SDS-PAGE. Densitometric analysis was performed using ImageJ v 1.51d (http://imagej.nih.gov/ij/) image analysis software.

4.4.9 Prothrombin Times

Diluted prothrombin time tests were performed in triplicate using Dade C Control platelet poor plasma (Siemens Healthcare, Germany) on a STAR-4 coagulation instrument (Diagnostica Stago, France). Refolded TFI-scFv fractions were individually incubated with tissue factor (Dade Tromborel S, Siemens Healthcare, Germany) for 10 min prior to the test.

4.5 Results

4.5.1 Rare Codon optimization

Multiple alignment of the unoptimized- and optimized DNA sequence reveals the extend of modification to the original sequence (Table 4.1) while the final amino acid sequences have remained unchanged (Table 4.2). Clustal Omega utilize the following notation system to indicate conservation of nucleotides between sequences; gaps are indicated by " **-** ", fully conserved nucleotides are indicated by " ***** ", conservation of all pyrimidines or all purines are indicated by " **:** ", some conservation is indicated by " **.** ", and no conservation is indicated by empty block. The plasmid backbone region of the unoptimized sequence (**Gray**) containing the C-terminal His-tag (**Orange**) and myc-tag (**Blue**) was excluded during codon optimization.

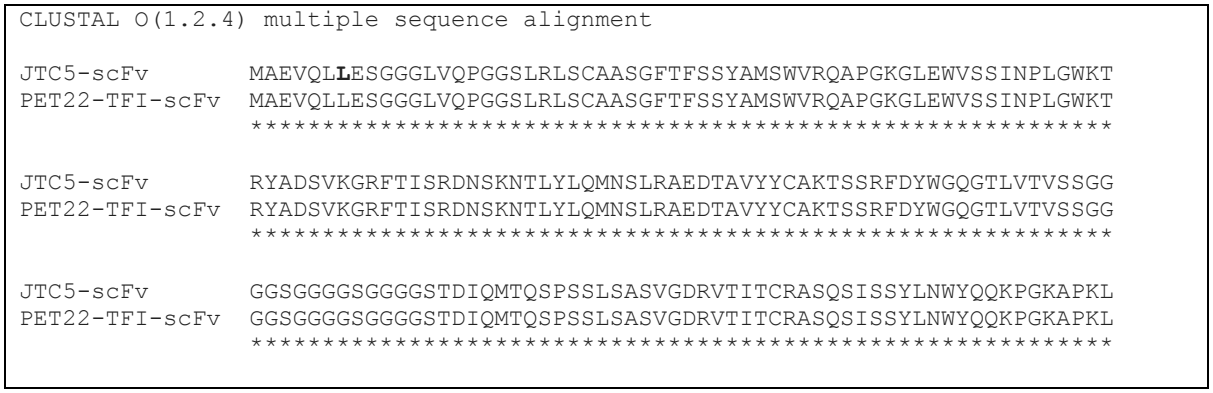
Table 4.1 DNA sequence alignment of unoptimized- and optimize sequence.

```
CLUSTAL O(1.2.4) multiple sequence alignment

Unopt       ATGGCCGAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTG
Opt         ATGGCAGAAGTTCAGCTGCTGGAAAGCGGTGGTGGTCTGGTGCAGCCTGGTGGTAGCCTG
            *****.**.** ****** ****.:  ** **:**  ****.******** ** : ****

Unopt       AGACTCTCCTGTGCAGCCTCTGGATTCACCTTTAGCAGCTATGCCATGAGCTGGGTCCGC
Opt         CGTCTGAGCTGTGCAGCAAGCGGTTTTACCTTTAGCAGCTATGCAATGAGCTGGGTTCGT
            .*:** : *********.:  **:** ******************.********** **
```

Table 4.2 Amino acid alignment of the JTC5-scFv (unoptimized) and pET22 TFI-scFv (optimized).



```
JTC5-scFv        LIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTPNTFGQGTKVEI
PET22-TFI-scFv   LIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTPNTFGQGTKVEI
                 ************************************************************

JTC5-scFv        KRAAAHHHHHHGAAEQKLISEEDLNGAA
PET22-TFI-scFv   KRNLEHHHHHH----------------
                 **   *****
```

Analysis of the TFI-scFv gene was performed using *E*CUA software with a 10 % codon frequency threshold based on highly and continuously expressed genes during exponential growth phase (Class II genes). The analysis of the gene identified a total of 46 rare (approximately 20 %) codons distributed throughout the Heavy- and Light chains of the scFv. The majority of rare codons coded for glycine (14 of 29), serine (12 of 38), arginine (9 of 11), leucine (8 of 18) and threonine (2 of 9). Rare codon pairs are present in positions Gly^{10}-Gly^{11}, Gly^{17}-Gly^{18}, Arg^{21}-Lue ^{22}, Lue^{187} -Ser^{188}, Ser^{190}-Gly^{191}, Ser^{194}-Arg^{195}, Ser^{199}-Gly^{200} and Gly^{202}-Thr^{203}. Rare codon clusters are present at positions Leu^{7} to Gly^{28} (heavy chain) and Ser^{187} to Leu^{207} (Light Chain) (Figure: 4.1). To overcome rare codon limitation, the TFI-scFv gene was codon optimized for expression in *E. coli* and cloned into pET-22b(+).

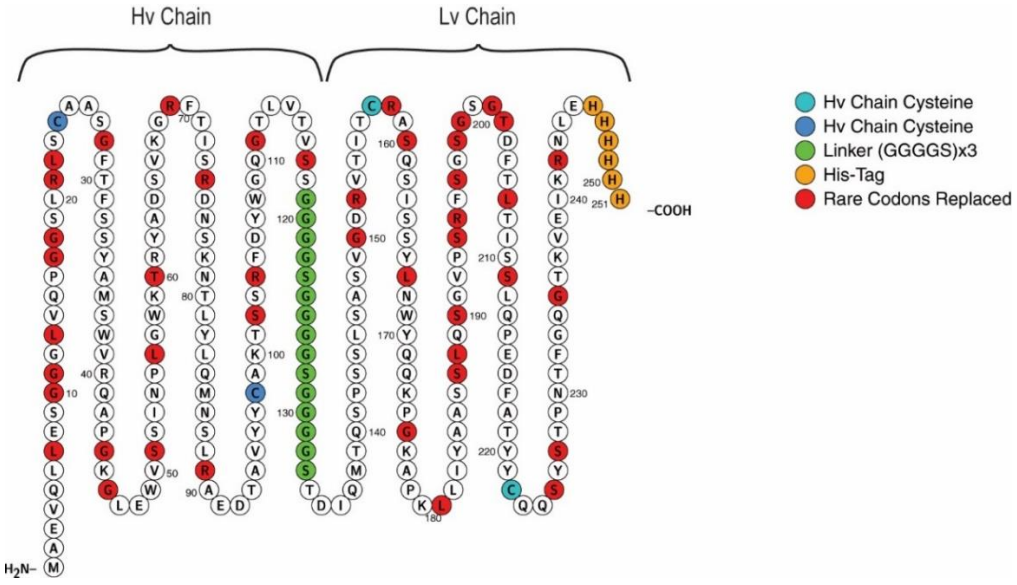


Figure 4.1: Distribution of substituted rare codons and other regions of interest in the TFI-scFv

The distribution and positions of substituted rare codons in the TFI-scFv. The scFv is represented as a single continuous ribbon for simplicity. The amino acid positions of the C-terminal His-tag, flexible (GGGGS)x3 linker, Hv and Lv Cysteine disulfide bridges and rare codons are indicated.

Analysis of the codon frequency distributions of the unoptimized and optimised TFI-scFv genes was performed using Graphical codon usage analyzer v2.0. All the rare codons and less preferred codons of the TFI-scFv (unoptimized) gene were replaced with higher frequency synonymous codons in the optimised TFI scFv gene during the optimization process (Figure 4.2). A total of 57.44 % of the codons in the un-optimized sequence had a

utilization frequency lower than 30 (less utilized codon) compared to only 30.58 % for the TFI sequence. A total of 150 out of 242 codons (approximately 62%) were substituted in the optimized gene. The Codon adaption index (CAI) is "universal' indication of synonymous codon usage. The CAI is calculated from codon frequencies of highly expressed genes for a specific species. It is an useful indicator of the level of expression of a gene between two different organisms [33]. It provides an approximate indication of the likeliness of successful heterologous gene expression. The CAI index value for the unoptimized gene was improved from 0.61 to 0.91 for expression in *E. coli.*

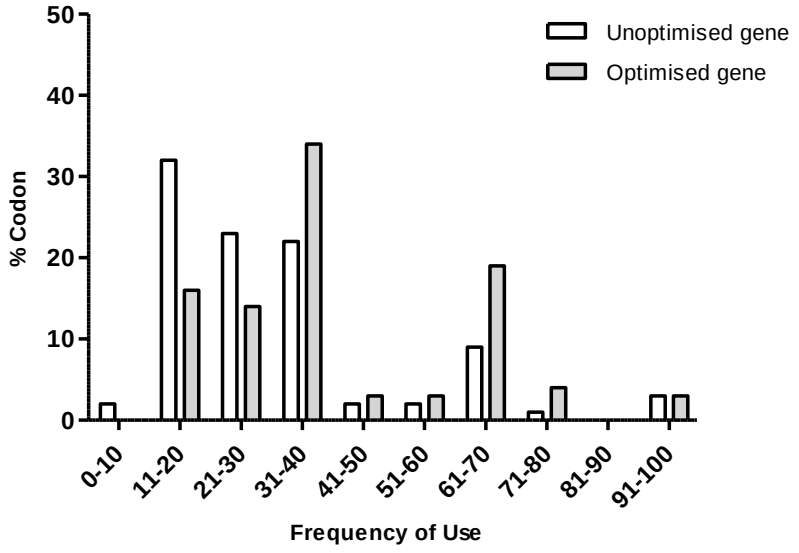


Figure 4.2: Codon frequency of usage profile.

The comparative summary of codon frequency between the unoptimized TFI-scFv gene (□) and optimized TFI-scFv gene (□). A value of 0 to 100 was assigned to synonymous codons for each amino acid based on their frequency of utilization according to the E. coli codon usage database. Codons with low relative frequency values indicate less preferred or rare codons. Low frequency codons have lower expression efficiency and as with the case of rare codons may lead to the truncation of proteins. Codons with high relative frequency (highly utilized) generally improve expression efficiency.

4.5.2 Cytoplasmic expression of codon-optimized TFI-scFv

The codon-optimized TFI-scFv gene was specifically cloned into the pET22b(+) expression vector to exclude the N-terminal *pelB* signal peptide, thus preventing the transport of target protein from the cytoplasm to the periplasm. A C-terminal His x6 tag was included for the detection and purification of expressed scFv. The *E. coli BL21 (DE3)* cells were transformed with pET22 TFI-scFv and induced by 0.1mM IPTG at 37 °C for 5h. Expression was performed in parallel with a negative control containing empty pET22b(+) plasmid. Harvested cells were fractionated and analysed by SDS-Page (Figure 4.3). A protein with a molecular weight of approximately 27 kDa was highly expressed. This correlates to the theoretical molecular weight of TFI-scFv. The majority of expressed

TFI-scFv accumulated in the insoluble cytoplasmic fraction (Lane 4) while the soluble fraction (Lane 8) only contained minimal of the 27 kDa protein TFI-scFv. The accumulation of TFI-scFv in the insoluble cytoplasmic fraction is indicative of incorrect protein folding [34]. The formation of insoluble aggregate folding intermediated (known as inclusion bodies) during high-level protein in the cytoplasm production is a well observed occurrence especially with scFv [35,36].

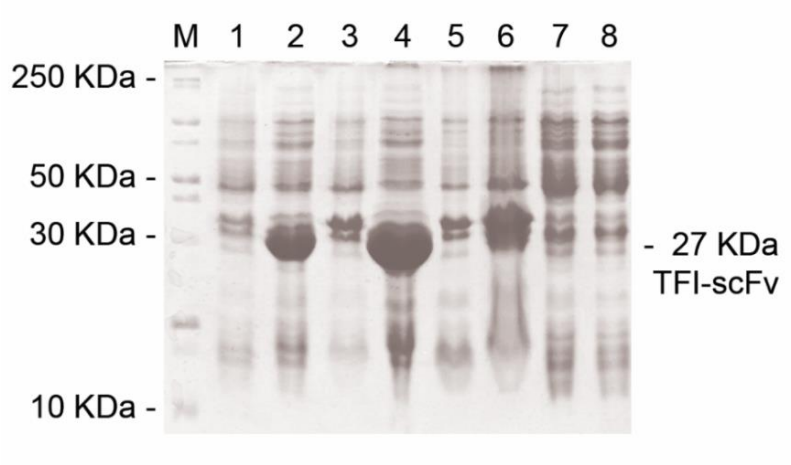


Figure 4.3. Cytoplasmic expression of TFI-scFv.

SDS-PAGE analysis of codon optimized TFI-scFv expression. The expected molecular weight of TFI-scFv is 27 kDa. Lane M, protein marker. Lane 1: Whole cell fraction of Bl21 transformed with empty pET22b(+) plasmid (negative control). Lane 2: Whole cell fraction of Bl21 transformed with pET22 TFI-scFv. Lane 3: Insoluble cytoplasmic fraction (negative control). Lane 4: Insoluble cytoplasmic fraction (pET22 TFI-scFv). Lane 5: Membrane fraction (negative control). Lane 6: Membrane fraction (pET22 TFI-scFv). Lane 7: Ultra centrifuge soluble fraction (negative control). Lane 8: Ultra centrifuge soluble fraction (pET22 TFI-scFv).

4.5.3 Isolation and purifications of TFI-scFv

The insoluble fraction was subjected to a series of wash steps in order to remove remaining membrane fractions and cellular debris. The washing steps (Figure 4.4 A: Lane 2 to 4) removed the majority of the unwanted proteins. The remaining cytoplasmic fraction was denatured using a high concentration of guanidium–chloride buffer in order to solubilize the scFv inclusion bodies. In order to visualize the scFvs, a dilution of the solubilized fraction was prepared due to the high salt concentrations of the denaturing buffer (Figure 4.4: Lanes 6- 7.) The SDS-page analysis of the cytoplasmic fraction after washing and denaturation shows the target scFv with greater that 95% purity (Figure 4.4 A: Lane 7) based on densitometry analysis (Figure 4.5). This correlates to production of approximately 40 $mg.L^{-1}$ of TFI-scFv. The solubilized inclusion bodies containing a C-terminal His-tag were purified by means of nickel based Immobilized Metal Ion Affinity chromatography (IMAC) under denaturing conditions using HisTrap Fast Flow columns. The protein absorbance peak corresponding to 70 - 100 mL was collected (Figure 4.4 B) and was analysed by means of SDS-PAGE. The denatured 27 kDa TFI-scFv was thus isolated with a high purity (Figure 4.4 C: Lane 1).

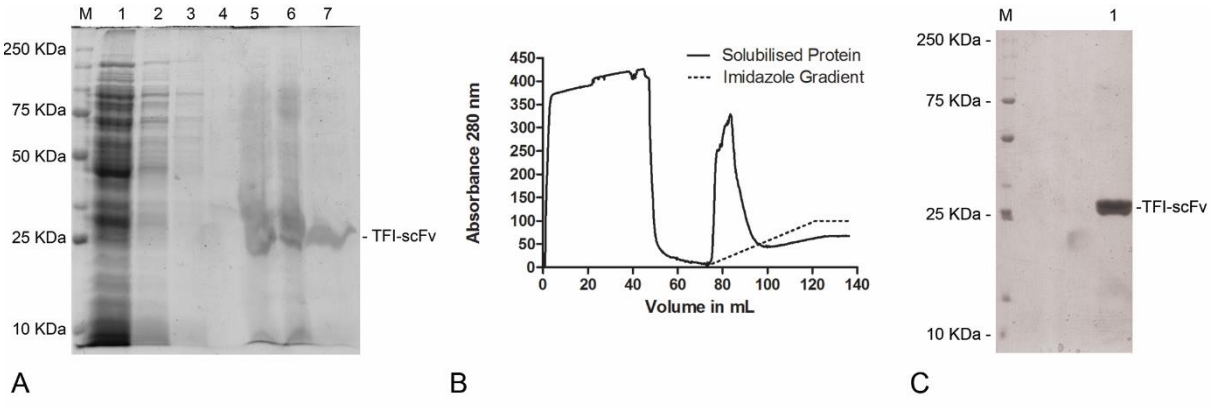


Figure 4.4: Isolation and purification of scFv from insoluble cytoplasmic fraction

A) SDS-PAGE analysis for insoluble cytoplasmic fraction. Lane M: Protein marker. Lane 1: Whole cell fraction. Lane 2: Disruption buffer Wash I. Lane 3: Disruption buffer Wash II. Lane 4: Chelating buffer. Wash Lane. 5: Solubilized inclusion bodies 1 in 12 Dilution. Lane 6: Solubilized inclusion bodies 1 in 6 Dilution. Lane 7: Solubilized inclusion bodies 1 in 30 Dilution. B) AKTÄ Purification Profile of TFI-scFv under denaturing conditions. C) SDS-PAGE analysis of purified denatured TFI-scFv. Lane M: Protein Marker. Lane 1: Purified TFI-scFv.

High purity of protein is extremely important for *in vitro* downstream protein refolding. As a result, the final purified protein fraction was analysed by means of densitometry in order to confirm the correct protein size and purity (Figure :4.5). The denatured 27 kDa TFI-scFv was successfully isolated and purified with a purity greater than 99 %. The protein yield was approximately 23 $mg.^{L-1}$ of *E. coli* culture.

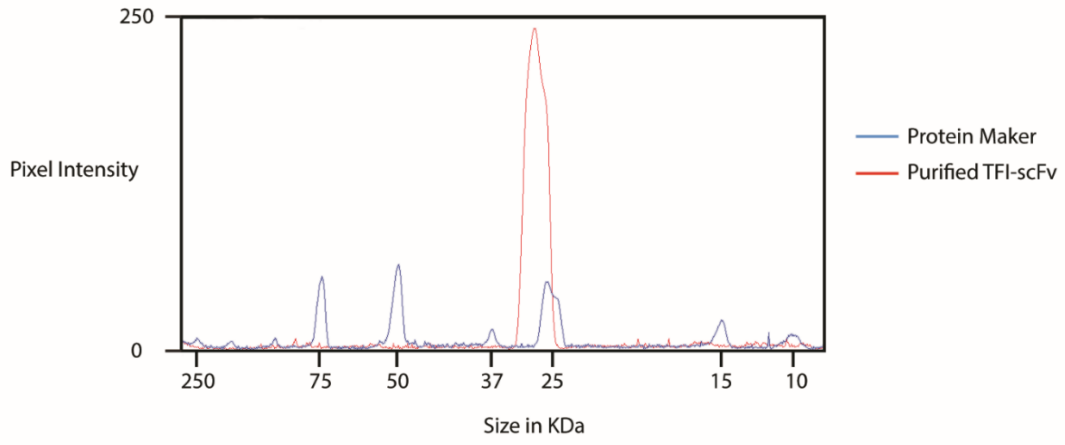


Figure 4.5: Desitometric analysis of scFv isolated from insoluble cytoplasmic fraction

*Desitometric plot of the purified TFI-scFv fraction (**red**) and the protein marker (**blue**). Pixel intensity rage is defined as as 0 = Absolute white and 255 = Absolue Black.*

4.5.4 The *in vitro r*efolding of TFI-scFV

All refolding reactions were performed at 4 °C in 1 mL batch reactions. The refolding matrix explores three stages of escalating concentrations of denaturant (Guanidine) in combination with varying L-Arginine concentrations while maintaining the pH at 8.0. The other buffer components are maintained at a constant level (Table 4.3).

Table 4.3: Refolding Buffer Matrix

Refolding Buffer	Denaturant	Guanidium	L-arginine	Reduced GSH	Oxidised GS-SG	TFI-scFv (ug/ml)
1	Range 1	300	0	2	0,2	500
2		300	400	2	0,4	500
3		300	800	1	1	500
4	Range 2	800	0	2	0,4	500
5		800	400	1	1	500
6		800	800	2	0,2	500
7	Range 3	1300	0	1	1	500
8		1300	400	2	0,2	500
9		1300	800	2	0,4	500

**All concentrations are indicated as mM unless indicated otherwise.*

Renaturing of the solubilized inclusion bodies was initialized by the removal of the refolding buffers by dialysis against Tris Buffered Saline (TBS). Soluble and insoluble fractions were separated by means of centrifugation and analysed by SDS-PAGE (Figure 4.6) and densitometry. Varying levels of protein aggregations were observed in all the samples. Refolding buffer 4 produced the lowest level of solubilisation (7.53%) while refolding buffer 7 produced the highest level of solubility (60.45 %). Solubilisation data is summarized in table 4.4.

The addition of arginine in refolding buffers in order to suppress protein aggregation is a well-established practice in protein refolding. Is has been proposed that the guanidinium group of arginine and tryptophan side chains may be responsible for suppression of protein aggregation by arginine [37]. The addition of arginine does not, however, suppress aggregation in some disulfide-containing proteins [38]. Due to this unpredictable effect of arginine on disulfide bond formation, the effect of arginine was examined across a large concentration range (0 – 800 mM). In general, the refolding buffers containing the lower levels the guanidine (range 1) had the higher levels of aggregation. Interestingly, the highest level of solubilisation was observed in sample 7 which contains no L-arginine, while in contrast samples 1 and 4 (also not containing L-arginine) produced the lowest levels of solubilisation for their respective ranges.

Table 4.4 Summary of solubilisation efficacy.

Refolding Buffer	Denaturant	Aggregated TFI-scFv (µg)	Soluble TFI-scFv (µg)	% Solubility
1	Range 1	438,66	61,34	12,27
2		429,40	70,60	14,12
3		374,26	125,74	25,15
4	Range 2	462,36	37,64	7,53
5		300,12	199,88	39,98
6		371,76	128,24	26,14
7	Range 3	197,76	302,24	60,45
8		260,64	239,36	47,87
9		219,45	280,55	56,11

Due to the presence of two intra-domain disulfide bonds in scFvs it was expected that the redox ratio would have substantial influence on the solubilisation of TFI-scFv. The redox ratio of 1 to 1 oxidized/reduced glutathione (1 mM) (sample 3, 5 and 7) produced the highest level of solubility of each of the three redox ratios examined across all three denaturant ranges. The lowest levels of solubilisation was achieved with oxidized/reduced glutathione ratio of 1 to 10 in range 1 (sample 1), a ratio 1 to 5 in range 2 (sample 4) and a ratio of 1 to 10 in range 3 (sample 8) (Figure: 4.6).

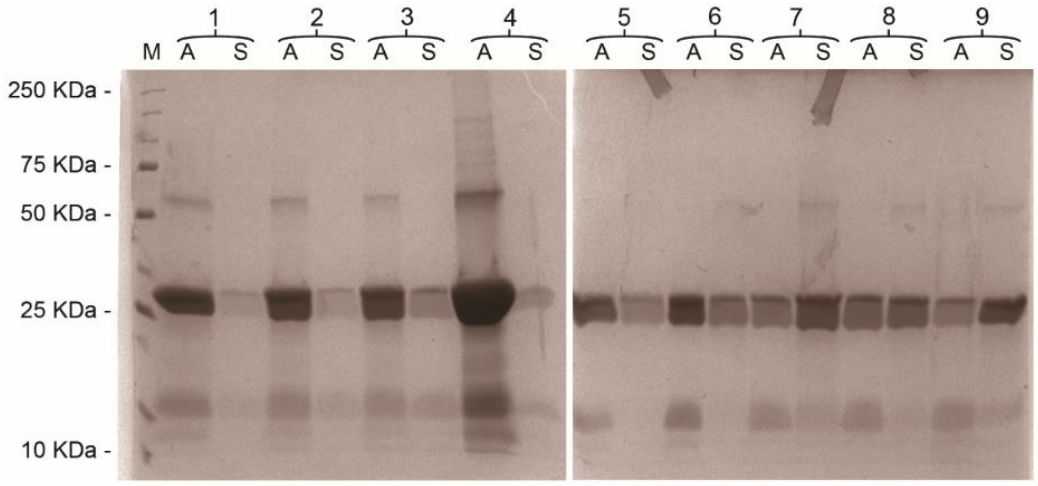


Figure 4.6 Solubility of refolded TFI-scFv

The SDS-Page analysis of the aggregated- (A) and soluble (S) fraction of each refolding buffer respectively. Refolding buffer 1 to 9 is indicated. Lane M: Protein marker. The marker was omitted from the 2nd gel due to a limited number of wells.

4.5.5 Prothrombin times

Diluted Prothombin Time (PT) assays were carried out in order to assess the efficiency of the refolding reactions to produce functional TFI-scFv. PTs were performed in triplicate and the average PT was calculated (Table 4.5). The inhibition effect of each refolded TFI-scFv was calculated as the difference between the average PT and the blank (negative control). All refolded samples were capable of extending the baseline PT value. This indicates that all of the refolding reactions were capable of producing functional TFI-scFv with varying levels of success. The highest inhibition effect was observed with sample 6 (17.2 s) which is an approximate 84 % extension of the baseline PT value. As a result, inhibition effectiveness of the refolded TFI-scFv was expressed as a percentage value relative to sample 6. Refolding buffer 2, 5 and 8 were capable of relative inhibition of approx. 25 % while refolding buffers 4, 7 and 9 were had a relative effectiveness of less than 10% relative to refolding buffer 6. The specific activity was calculated as the time extended in seconds per mg of TFI-scFv added to each reaction. The Specific activity was further compared to the level of solubilisation (Figure 4.7).

Table 4.5: Analysis of prothrombin times

Refolding Buffer	Average PT	Time extended	Final Concentration $mg.mL^{-1}$	Relative Inhibition Effectiveness (%)	Specific activity Time extended / Concentration $s / mg.mL^{-1}$
Blank	20,6	0	0	0	0
1	23,1	2,4	0.0245	14,2	9.78
2	24,8	4,1	0.0282	24,1	14.52
3	23,7	3,1	0.0503	17,9	6.16
4	22,3	1,6	0.0151	9,5	10.63
5	25,1	4,5	0.08	26	5.63
6	37,8	15,2	0.0513	100	29.63
7	21,2	0,6	0.1209	3,5	0.5
8	24,8	4,1	0.0957	24,1	4.28
9	20,9	0,3	0.1122	1,7	0.27

The specific activity analysis provides a good indication of which buffer produced the most functional TFI-scFv capable of recognizing the tissue factor epitope oppose to producing soluble but non-functional TFI-scFv. Refolding Buffer 6 (29.63 $s/mg.mL^{-1}$) and refolding buffer 2 (14.52 $s/mg.mL^{-1}$) had the highest specific activity. Interestingly, refolding buffer 5, 7, 8 and 9 which all produced a higher level of solubility than refolding buffer 6, had a much lower specific activity. This indicates that although these refolding buffers were more effective at producing soluble TFI-scFv, the antibodies were folded in incorrect intermediates that are not capable of recognizing the tissue factor epitope.

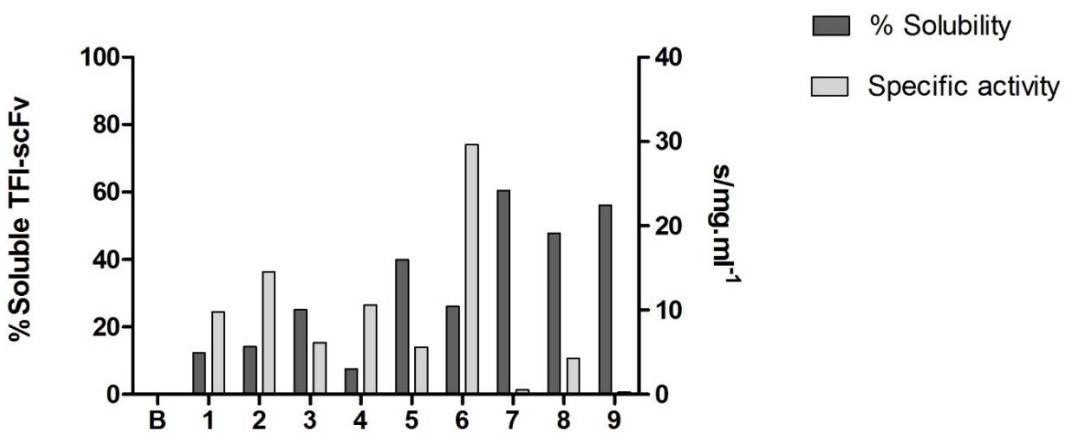


Figure 4.7: Relationship between specific activity and solubility

The specific activity ( ) (s/mg.mL^{-1}) relative to the level of solubility (% soluble TFI-scFv) ( ) achieved by each refolding buffer. Refolding buffer 6 produced the highest specific activity despite having solubility lower that buffers 5, 7, 8 and 9.

4.6 Discussion

The TFI-scFv was initially selected from the Tomlinson I & J phage display library [3]. This phage library consists of scFv that mimic the heavy- and light chain variable regions of human IgG. The TFI-scFv was originally expressed in the periplasm using the pIT2 expression vector in *E. coli* HB2151. However, the in-depth characterization of the scFv was compromised by low protein yields as well as expensive and time-consuming Protein A based purifications methods.

Codon usage has been identified as the single most important factor in prokaryotic gene expression [39]. We suspected that due to heterologous nature of the scFv gene that the low levels of expression attained was a result of codon bias. In terms of a native gene expression, rare codons are not randomly distributed throughout the genes but are rather found in large clusters [40]. These rare codon clusters lead to translational pausing of nascent polypeptide production, which facilitates the slower folding of secondary structure [15,40]. It has been suggested that rare codon clusters are associated with β-sheets and coil regions due to the slower folding rate of these regions [41,42]. However, during heterologous expression the rare codon distribution becomes randomly distributed due to difference in codon usage bias between two organisms [14]. The presence of these distributed rare codons and especially rare codon pairs has detrimental effects on gene expression efficiency as they limit the ribosome association with mRNA [43]. Expression of heterologous genes containing rare codons can lead to translational errors, as a result of ribosomal stalling at irregular positions requiring incorporation of amino acids coupled to minor codon tRNAs [18]. It has also been shown that a single rare codon can have a deleterious effect on heterologous protein expression [44]. The rare codon analysis of TFI-scFv revealed that 45 codons (approximately 20 % in total) distributed throughout the gene. Of these rare codons, there are 3 consecutive repeat

and 2 single rare codons near the N-terminal of the scFv gene. Low expression levels are exacerbated when rare codons appear in clusters or in the N-terminal region of the protein [39]. The low expression levels of TFI-scFv previously attained, and most likely the other scFvs isolated by our group, can be accredited to the excessive presence of rare codons throughout the gene.

Although the detrimental effects of codon bias on heterologous protein expression levels is well established, the influence that the replacement of these codons will exert on protein folding is not always clear. The removal of these "throttling steps" through preferred codon substitution in favour of higher translation efficiency and increased translational speed may be responsible for the incorrect folding of target protein structure [26]. In addition, scFvs are notorious inclusion body producers when expressed in the cytoplasm of *E. coli* due to the presence of intra domain disulfide bonds [45]. As a result, synthetic constructs like scFvs are traditionally expressed with a bacterial N-terminal leader peptide such as *pelB, ribose binding protein, ompA, ompT, dsbA, torA, torT,* and *tolT.* [46]. These leader peptides enable the transport of the protein product to the bacterial periplasm in order to facilitate correct disulfide bridge formation [21]. Correct folding is also facilitated by the oxidative environment and the presence of a variety of chaperone proteins present with in the periplasm [46,47]. Despite these advantages, over production of scFv in the periplasm frequently results in the production of inclusions bodies [48]. Due to these factors the scFv expression system was modified to remove the 5' *pelB* leader peptide in order to prevent the transport of the protein from the bacterial cytoplasm. Expression to the cytoplasm is generally associated with higher expression levels and simpler purification procedures [49]. The majority of TFI-scFv was expressed as insoluble inclusion bodies as confirmed by SDS-PAGE and densitometry. The use of commercial expression vectors and cell lines, synonymous substitution of rare codons together with the removal of the signal peptide from the *JTC5-scFv* gene, culminated in an 80-fold improvement in TFI-scFv expression.

Despite years of collective knowledge, the unpredictable nature of the *in vitro* protein refolding still makes it difficult to predict conditions what would be most effective at renaturing the protein aggregates [10,12]. The TFI-scFv aggregate isolated by centrifugation were solubilized using 6M of Guanidium–chloride buffer to yield approximately 80 $mg.L^{-1}$ denatured protein. The step-wise washing procedure was effective in removing the majority (> 80%) of the background protein. The solubilized TFI-scFv was purified by means of IMAC under denaturing conditions in order to obtain high purity (>90) of TFI-scFv required for downstream *in vitro* refolding. Approximately 55 $mg.L^{-1}$ of denatured and reduced TFI-scFv was purified. In order to identify the specific conditions that would facilitate the correct refolding of TFI-scFv, the denatured scFvs were subjected to a refolding matrix consisting of buffers refolding buffers. A redox potential gradient was included in the refolding due to the presence of intra-disulfide bridges in the scFv molecular structure [50]. Disulfide bond formation plays an important role in the stabilization of folding intermediates as the protein advances to maturity [24,47]. Unfortunately, due to sensitivity of batch refolding procedures to target protein concentration, TFI-scFV concentration was restricted to 500 $ug.mL^{-1}$ for screening reactions.

Although there is a clear increase in solubility with an escalation in denaturant, a high level of solubilisation did not correlate with functionality of the antibody. The oxidation-reduction ratio of 10:1 oxidized/reduced glutathione had the greatest effect on solubilisation for each stage of the escalating denaturant. This is expected as disulfide bridges play a significant role in the stabilization of many proteins as well as antibodies like scFv [51]. Refolding buffer 6 produced a substantial higher specific activity at a TFI-scFv concentration considerably lower than most of the refolding buffers. It is interesting to observe the sensitivity and cumulative effect of the *in vitro* refolding buffers. Of the nine relatively similar conditions tested, only one was capable of producing functional TFI-scFv. It would have been preferable to normalize the concentrations of the each of the nine refolding reactions prior to the PT assays, to better compare the specific activity of each refolded product at a specific concentration. However, this would have been practically impossible to achieve due to the small reaction volumes as well as the very low levels of solubilisation achieved by some refolding reactions. Instead a comprehensive approach to the prothrombin times was followed as an indication of refolding efficiency. The relative inhibition effectiveness of each the soluble fractions is therefore a good indication of the overall refolding process. It correlated the functionality of the refolded protein with total yield. In other words, the buffer capable of producing the most soluble protein, correctly folded and capable of achieving the best inhibition would be the optimum buffer to utilize for TFI-scFv refolding on a larger scale.

Based on the capacity of refolding buffer 6 to solubilized 26 % of the denatured TFI-scFv one could extrapolate that approximately 6 mg of active TFI-scFv can be produced from one litre of *E. coli* culture under the current expression levels using up-scale settings. The production of TFI-scFv at these levels would enables the additional characterization of the TFI-scFv under various *in vitro* as well as *in vivo* settings that was previous impossible due to low yields. The procedure utilized here for the improved production of TFI-scFv can be followed to improve the production of other low yield scFv isolated from phage display libraries.

4.7 Conclusion

The expression strategy followed resulted in a 80-fold increase in TFI-scfv expression level in *E. coli* cytoplasm compared to that previously achieved with periplasmic expression. The isolations and *in vitro* refolding of denatured scFv resulted in a 6-fold improvement in the final antibody yield. This improvement in the yield of functional TFI-scFv enables the additional characterization of the TFI-scFv in *in vitro* as well as *in vivo* settings that was previously impossible.

4.8 Acknowledgements

This work has been funded by the National Research Foundation (NRF) of South Africa.

4.9 References

[1] P. Chames, M. Van Regenmortel, E. Weiss, D. Baty, Therapeutic antibodies: successes, limitations and hopes for the future, Br. J. Pharmacol. 157 (2009) 220–233. doi:10.1111/j.1476-5381.2009.00190.x.

[2] X. Yang, W. Hu, F. Li, H. Xia, Z. Zhang, Gene cloning, bacterial expression, in vitro refolding, and characterization of a single-chain Fv antibody against PreS1(21-47) fragment of HBsAg, Protein Expr. Purif. 41 (2005) 341–348. doi:10.1016/j.pep.2005.02.005.

[3] S.M. Meiring, J. Vermeulen, P.N. Badenhorst, Development of an inhibitory antibody fragment to human tissue factor using phage display technology, Drug Dev. Res. 70 (2009) 199–205. doi:10.1002/ddr.20295.

[4] S.A. Smith, R.J. Travers, J.H. Morrissey, How it all starts: Initiation of the clotting cascade, Crit. Rev. Biochem. Mol. Biol. 50 (2015) 326–336. doi:10.3109/10409238.2015.1050550.

[5] A.P. Owens, N. Mackman, Tissue factor and thrombosis: The clot starts here, Thromb. Haemost. 104 (2010) 432–439. doi:10.1160/TH09-11-0771.

[6] G.E. Raskob, P. Angchaisuksiri, A.N. Blanco, H. Buller, A. Gallus, B.J. Hunt, E.M. Hylek, A. Kakkar, S. V. Konstantinides, M. McCumber, Y. Ozaki, A. Wendelboe, J.I. Weitz, Thrombosis: A Major Contributor to Global Disease Burden, Arterioscler. Thromb. Vasc. Biol. 34 (2014) 2363–2371. doi:10.1161/ATVBAHA.114.304488.

[7] D.R. Burton, Phage display, Immunotechnology. 1 (1995) 87–94. doi:10.1016/1380-2933(95)00013-5.

[8] K. Li, K.A. Zettlitz, J. Lipianskaya, Y. Zhou, J.D. Marks, P. Mallick, R.E. Reiter, A.M. Wu, A fully human scFv phage display library for rapid antibody fragment reformatting, Protein Eng. Des. Sel. 28 (2015) 307–315. doi:10.1093/protein/gzv024.

[9] N. a Watkins, W.H. Ouwehand, Introduction to antibody engineering and phage display., Vox Sang. 78 (2000) 72–9. doi:31154.

[10] T. Arakawa, D. Ejima, Refolding Technologies for Antibody Fragments, Antibodies. 3 (2014) 232–241. doi:10.3390/antib3020232.

[11] M. Kamionka, Engineering of therapeutic proteins production in Escherichia coli., Curr. Pharm. Biotechnol. 12 (2011) 268–74. doi:10.2174/138920111794295693.

[12] A. Singh, V. Upadhyay, A.K. Upadhyay, S.M. Singh, A.K. Panda, Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process., Microb. Cell Fact. 14 (2015) 41. doi:10.1186/s12934-015-0222-8.

[13] K. Terpe, Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems, Appl. Microbiol. Biotechnol. 72 (2006) 211–222. doi:10.1007/s00253-006-0465-8.

[14] E.M. Novoa, L. Ribas de Pouplana, Speeding with control: Codon usage, tRNAs, and ribosomes, Trends Genet. 28 (2012) 574–581. doi:10.1016/j.tig.2012.07.006.

[15] M. Widmann, M. Clairo, J. Dippon, J. Pleiss, Analysis of the distribution of functionally relevant rare codons., BMC Genomics. 9 (2008) 207. doi:10.1186/1471-2164-9-207.

[16] P.N. Desai, N. Shrivastava, H. Padh, Production of heterologous proteins in plants: Strategies for optimal expression, Biotechnol. Adv. 28 (2010) 427–435. doi:10.1016/j.biotechadv.2010.01.005.

[17] E. Angov, Codon usage: Nature's roadmap to expression and folding of proteins, Biotechnol. J. 6 (2011) 650–659. doi:10.1002/biot.201000332.

[18] H.P. Sørensen, K.K. Mortensen, Advanced genetic strategies for recombinant protein expression in Escherichia coli, J. Biotechnol. 115 (2005) 113–128. doi:10.1016/j.jbiotec.2004.08.004.

[19] N.A. Burgess-Brown, S. Sharma, F. Sobott, C. Loenarz, U. Oppermann, O. Gileadi, Codon optimization can improve expression of human genes in Escherichia coli: A multi-gene study, Protein Expr. Purif. 59 (2008) 94–102. doi:10.1016/j.pep.2008.01.008.

[20] M. Welch, a. Villalobos, C. Gustafsson, J. Minshull, You're one in a googol: optimizing genes for protein expression, J. R. Soc. Interface. 6 (2009) S467–S476. doi:10.1098/rsif.2008.0520.focus.

[21] H. Sonoda, Y. Kumada, T. Katsuda, H. Yamaji, Effects of cytoplasmic and periplasmic chaperones on secretory production of single-chain Fv antibody in Escherichia coli, J. Biosci. Bioeng. 111 (2011) 465–470. doi:10.1016/j.jbiosc.2010.12.015.

[22] H. Yang, H. Wang, T. Xue, X. Xue, T. Huyan, W. Wang, K. Song, Single-chain variable fragment antibody against human aspartyl/asparaginyl beta-hydroxylase expressed in recombinant Escherichia coli, Hybridoma. 30 (2011) 69–79. doi:10.1089/hyb.2010.0070.

[23] A. Singh, V. Upadhyay, A.K. Panda, Solubilization and Refolding of Inclusion Body Proteins, in: Insoluble Proteins Methods Protoc., 2015: pp. 283–291. doi:10.1007/978-1-4939-2205-5_15.

[24] R. Rudolph, H. Lilie, In vitro folding of inlusion body proteins, FASEB J. 10 (1996) 49–56. doi:10.1021/bi2012965.

[25] R. Gutiérrez, E.M. Martín del Valle, M. a. Galán, Immobilized Metal-Ion Affinity Chromatography: Status and Trends, Sep. Purif. Rev. 36 (2007) 71–111. doi:10.1080/15422110601166007.

[26] K. Sushma, C.J. Bilgimol, M.A. Vijayalakshmi, P.K. Satheeshkumar, Recovery of active anti TNF-?? ScFv through matrix-assisted refolding of bacterial inclusion bodies using CIM monolithic support, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 891–892 (2012) 90–93. doi:10.1016/j.jchromb.2012.02.011.

[27] U. Omasits, C.H. Ahrens, S. M??ller, B. Wollscheid, Protter: Interactive protein feature visualization and integration with experimental proteomic data, Bioinformatics. 30 (2014) 884–886. doi:10.1093/bioinformatics/btt607.

[28] M. Fuhrmann, A. Hausherr, L. Ferbitz, T. Schödl, P. Hegemann, Monitoring dynamic expression of nuclear genes in Chlamydomonas reinhardtii by using a synthetic luciferase reporter gene, Plant Mol. Biol. 55 (2004) 869–881. doi:10.1007/s11103-005-2150-1.

[29] P. Puigbò, I.G. Bravo, S. Garcia-Vallve, CAIcal: a combined set of tools to assess codon usage adaptation., Biol. Direct. 3 (2008) 38. doi:10.1186/1745-6150-3-38.

[30] U.K. LAEMMLI, Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4, Nature. 227 (1970) 680–685. doi:10.1038/227680a0.

[31] G. Fairbanks, T.L. Steck, D.F. Wallach, Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane., Biochemistry. 10 (1971) 2606–2617. doi:10.1021/bi00789a030.

[32] H. Stanton, S.B. Golub, F.M. Rogerson, K. Last, C.B. Little, A.J. Fosang, Investigating ADAMTS-mediated aggrecanolysis in mouse cartilage, Nat. Protoc. 6 (2011) 388–404. doi:10.1038/nprot.2010.179.

[33] A. Tiwari, A. Sankhyan, N. Khanna, S. Sinha, Enhanced periplasmic expression of high affinity humanized scFv against Hepatitis B surface antigen by codon optimization, Protein Expr. Purif. 74 (2010) 272–279. doi:10.1016/j.pep.2010.06.006.

[34] G.P. Subedi, T. Satoh, S. Hanashima, A. Ikeda, H. Nakada, R. Sato, M. Mizuno, N. Yuasa, Y. Fujita-Yamaguchi, Y. Yamaguchi, Overproduction of anti-Tn antibody MLS128 single-chain Fv fragment in Escherichia coli cytoplasm using a novel pCold-PDI vector, Protein Expr. Purif. 82 (2012) 197–204. doi:10.1016/j.pep.2011.12.010.

[35] G. Hannig, S.C. Makrides, Strategies for optimizing heterologous protein expression in Escherichia coli, Trends Biotechnol. 16 (1998) 54–60. doi:10.1016/S0167-7799(97)01155-4.

[36] M. Kudou, D. Ejima, H. Sato, R. Yumioka, T. Arakawa, K. Tsumoto, Refolding single-chain antibody (scFv) using lauroyl-L-glutamate as a solubilization detergent and arginine as a refolding additive, Protein Expr. Purif. 77 (2011) 68–74. doi:10.1016/j.pep.2010.12.007.

[37] K. Tsumoto, M. Umetsu, I. Kumagai, D. Ejima, J.S. Philo, T. Arakawa, Role of arginine in protein refolding, solubilization, and purification., Biotechnol. Prog. 20 (2004) 1301–8. doi:10.1021/bp0498793.

[38] J. Chen, Y. Liu, Y. Wang, H. Ding, Z. Su, Different Effects of L -Arginine on Protein Refolding: Suppressing Aggregates of Hydrophobic Interaction, Not Covalent Binding, Biotechnol. Prog. 24 (2008) 1365–1372. doi:10.1021/bp.93.

[39] C. Gustafsson, S. Govindarajan, J. Minshull, Codon bias and heterologous protein expression, Trends Biotechnol. 22 (2004) 346–353. doi:10.1016/j.tibtech.2004.04.006.

[40] T.F. Clarke, P.L. Clark, Increased incidence of rare codon clusters at 5' and 3' gene termini: implications for function., BMC Genomics. 11 (2010) 118. doi:10.1186/1471-2164-11-118.

[41] J.L. Chaney, P.L. Clark, Roles for Synonymous Codon Usage in Protein Biogenesis., Annu. Rev. Biophys. 44 (2015) 143–66. doi:10.1146/annurev-biophys-060414-034333.

[42] S. Zhang, E. Goldman, G. Zubay, Clustering of Low Usage Codons and Ribosome Movement, J. Theor. Biol. 170 (1994) 339–354. doi:10.1006/jtbi.1994.1196.

[43] R. Trinh, B. Gurbaxani, S.L. Morrison, M. Seyfzadeh, Optimization of codon pair use within the (GGGGS)3 linker sequence results in enhanced protein expression., Mol. Immunol. 40 (2004) 717–22. doi:10.1016/j.molimm.2003.08.006.

[44] Y. Ioannou, I. Giles, A. Lambrianides, C. Richardson, L. Pearl, D. Latchman, D. Isenberg, A. Rahman, A novel expression system of domain I of human beta2 glycoprotein I in Escherichia coli, BMC Biotechnol. 6 (2006) 8. doi:10.1186/1472-6750-6-8.

[45] I. Kumagai, R. Asano, T. Nakanishi, K. Hashikami, S. Tanaka, A. Badran, H. Sanada, M. Umetsu, Integration of PEGylation and refolding for renaturation of recombinant proteins from insoluble aggregates produced in bacteria-Application to a single-chain Fv fragment, J. Biosci. Bioeng. 109 (2010) 447–452. doi:10.1016/j.jbiosc.2009.10.016.

[46] A. de Marco, Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli., Microb. Cell Fact. 8 (2009) 26. doi:10.1186/1475-2859-8-26.

[47] S.C. Makrides, Strategies for achieving high-level expression of genes in Escherichia coli., Microbiol. Rev. 60 (1996) 512–538. http://www.ncbi.nlm.nih.gov/pubmed/8840785.

[48] D. Moricoli, W.A. Muller, D.C. Carbonella, M.C. Balducci, S. Dominici, R. Watson, V. Fiori, E. Weber, M. Cianfriglia, K. Scotlandi, M. Magnani, Blocking monocyte transmigration in in vitro system by a human antibody scFv anti-CD99. Efficient large scale purification from periplasmic inclusion bodies in E. coli expression system, J. Immunol. Methods. 408 (2014) 35–45. doi:10.1016/j.jim.2014.04.012.

[49] A.M. Latifi, K. Khajeh, G. Farnoosh, K. Hassanpour, S. Khodi, The cytoplasmic and periplasmic expression levels and folding of organophosphorus hydrolase enzyme in Escherichia coli, Jundishapur J. Microbiol. 8 (2015) 10–14. doi:10.5812/jjm.17790.

[50] W. Swietnicki, Folding aggregated proteins into functionally active forms, Curr. Opin. Biotechnol. 17 (2006) 367–372. doi:10.1016/j.copbio.2006.05.011.

[51] K. Proba, A. Wörn, A. Honegger, A. Plückthun, Antibody scFv fragments without disulfide bonds, made by molecular evolution, J. Mol. Biol. 275 (1998) 245–253. doi:10.1006/jmbi.1997.1457.

Chapter 5: Solubilisation of an anti-tissue factor single chain variable fragment (TFI-scFv) using cold shock in conjunction with E.coli SHuffle.

List of Abbreviations

Cold shock protein A	cspA
Complementarity Determining Regions	CDR
Escherichia Coli:	*E.coli*
Glutathione S-transferase:	GST
Green Fluorescent Protein:	GFP
Heavy Chain variable region	Hv
Immunoglobulin:	IgG
Light Chain Variable region	Lv
Litre per minute:	lpm
Maltose Binding Protein:	MBP
N-utilization substance:	NusA
Single chain variable fragment	scFv
Thioredoxin:	TrxA
Trigger Factor:	TrF
Transcription Enhancement Elements	TEE

Chapter 5: Solubilisation of an anti-tissue factor single chain variable fragment (TFI-scFv) using cold shock in conjunction with E.coli SHuffle.

5.1 Abstract

The development of therapeutic single chain variable fragments in *Escherichia coli* has been a constant challenge in the field of drug development. This is largely due to their tendency to accumulate within inclusion bodies. The production of functional TFI-scFv from insoluble aggregated protein through the use of *in vitro* refolding techniques was reported in chapter 4. In this chapter, the feasibility of producing a functional TFI-scFv in the cytoplasm of *E.coli.* as an alternative to the *in vitro* refolding approach was explored. Solubilisation of the scFv was achieved though commercial cold shock expression vectors (pCOLD DNA II), cold shock fusion partner expression vectors (pCOLD Trigger factor) in conjunction with *E.coli BL21* (DE3) or specialised disulfide bond promoting *E.coli SHuffle*. The level of TFI-scFv solubilisation ranged from 20 % for the non-fusion constructs up to 40 % for the trigger factor: TFI-scFv fusion constructs. The higher solubilisation achieved with the fusion construct was further examined though the generation of structural homology models. Production of soluble TFI-scFv fractions was up-scaled using 5L Bioreactors. All of the expression strategies were capable of producing soluble TFI-scFv within the cytoplasm of *E.coli.* The antibody fragments were then isolated by means of immobilized metal affinity chromatography (IMAC). Prothrombin time (PT) analyses were performed with the purified TFI-scFvs to determine biological activity. The TFI-scFv extended PT in a dose dependant manner with a specific activity of 83.8 s/mg.mL^{-1}. Functional fragments were produced through a combination of cold shock, pCOLD DNA II expression vector in conjunction with *E.coli SHuffle*. This expression systems provides a simpler, less time consuming and more cost effect alternative for the production of functional TFI-scFv than previously possible.

5.2 Key-Words

Single Chain Variable Fragment (scFv), Solubilisation Fusion Partner, Homology modelling

5.3 Introduction

In response to vascular damage, blood coagulation is initiated by a 263 residue transmembrane glycoprotein know as tissue factor [1,2]. Tissue factor is highly constitutively expressed in vascularized organs such as the brain, placenta, lungs, heart kidney, intestine, testes and uterus [3]. This distribution of tissue factor expression in vital organs provide a protective haemostatic barrier in order to minimise blood loss during injury [4]. Despite this critical role in haemostasis, excessive or aberrant expression of tissue factor has been linked to inflammatory and thrombotic disorders, which are major contributors to global disease and mortality [5–7]. Due to tissue factor's role as principle initiator of the coagulation cascade it provides an ideal target for the development of novel anti-coagulants with which to treat thrombosis [8]. As reported previously, phage display technology was used to select a single chain variable fragment (scFv) from the Tomlinson I + J Human Single Fold Phage Libraries to inhibit the

haemostatic function of human tissue factor [9]. Codon optimization of the TFI-scFv gene resulted in a substantial increase in the initial expression levels (80 fold). However, despite the increased expression the majority (>80%) of antibody was produced as insoluble aggregates known as inclusion bodies. Single chain variable fragments are notorious for their tendency to produce insoluble aggregates when expressed in Escherichia Coli (*E.col)i* [10]. The regeneration of functional scFv from insoluble aggregates through *in vitro* refolding techniques has become standard practice [10–13]. Although there are some advantages to the expression of target protein as inclusion bodies, the challenge lies in the renaturing of the native protein structure [14]. The refolding techniques commonly used are often complex and require intensive optimization of multiple operational steps in order to obtain functional protein [15]. Target protein recovery yields are also highly variable due to the protein specific nature of *in* vitro refolding techniques [16]. This is especially true for scFvs owing to the artificial nature in which the Hv- and Lv domains are connected and the presence of intra-chain disulfide bonds [17,18].

The many contributing factors that lead to protein aggregation during heterologous- as well as native protein expression are not completely understood. It is believed that high expression rates, the reducing environment of the bacterial cytosol and its influence on disulfide bond formation, insufficient chaperone assistance and the lack of post-translation modification mechanisms contribute to protein aggregation observed during bacterial expression [19–21]. Thus, a constant challenge in the field of protein engineering is to overcome these native molecular mechanisms through alternative expression strategies. This could alleviate the need for complex post-expression *in vitro* refolding techniques [15,22,23].

Cold shock expression, is a popular method used to overcome protein aggregation. It utilizes the native response of *E.coli* to sudden temperature reductions to overcome protein aggregation [24]. These sudden fluctuations in temperature strongly influences gene regulation in mesophilic bacteria such as *E.coli* [25,26]. Cold shock effect the cell at various levels (i) reduction in membrane associated functions such as active transport and protein secretion, (ii) reduced efficiency of mRNA translation and transcription, (iii) reduction in protein folding rate and (iv) the need of ribosomes to cold-adapt in order to function properly at low temperature [27]. Cold shock expression systems utilises specific cold shock protein A (cspA) promotors and transcription enhancement elements (TEE) to facilitate the over expression of the target protein at lower growth temperatures [28,29]. It also has the combined advantages of increasing solubility of the target protein while limiting unwanted degradation produced by heat shock induced proteases [16]. Additionally, a number of chaperones are up-regulated with increased activity at lower temperatures which in turn facilitates appropriate protein folding [30]. These expression mechanisms have also been successfully utilised to facilitate the solubilisation of proteins containing multiple disulfide bonds [31].

Additionally, the utilisation of fusion partner expression systems has also been shown to facilitate the production of soluble recombinant proteins in *E. coli* [32]. Fusion expression systems make use of a highly soluble partner that is fused to the N-terminal of the target protein in order to improve the solubility of the target protein [33]. Upon the completion of expression, the target protein is proteolytically cleaved from the fusion construct by specific

proteases to yield the native target protein [34]. Trigger Factor (TrF) is a 48 kDa chaperone native to *E.coli* that is often utilized as a fusion partner during recombinant protein expression [35]. In contrast to most other heat-shock induced chaperones, TrF has increased expression at low temperatures [36]. This makes TrF an ideal target for fusion partner expression in conjunction with cold shock.

An additional complication to consider is the presence of inter domain stabilising disulfide bonds. All antibodies have highly conserved disulfide bonds and scFv contains one intra chain disulfide bond for each of the variable regions [37]. The reducing conditions of the *E.coli* cytoplasm hampers the formation of disulfide bonds that are necessary for stable tertiary structure formation [38,39]. One of the approaches used to circumvent the difficulties of disulfide bonds formation in *E.coli* cytoplasm has been to develop cysteine-free scFvs [40]. However, most antibodies will not tolerate the loss of the stabilising disulfide bonds [41]. Alternatively, and perhaps a simpler approach, is the modification of the expression host itself in order to promote disulfide bond formation in the cytoplasm [42]. Today, commercial strains (*E.coli* Origami and FA113) are available that contain mutated thioredoxin reductase (trxB) and glutathione reductase (gor) genes in order to provide an oxidizing environment [43]. Other commercial strains (*E.coli SHuffle*®) expands on the trxB$^-$/Gro$^-$ mutations with the constitutive expression of a chromosomal copy of cytoplasmic DsbC [38]. DsbC is a isomerase chaperone that specifically facilitates disulfide bond formation within the cytoplasm [44]. The enhanced oxidizing cytoplasmic conditions in combination with improved isomerization capability should, in theory, improve target protein solubilisation thought improved disulfide bond formation.

The work presented in this chapter reports on: (i) The degree of solubilisation of TFI-scFv using a commercial (TAKARA) cold shock pCOLD DNA II and pCOLD Trigger Factor expression vectors in conjunction with *E. coli BL21 (DE3)* and *E. coli SHuffle*® strains; (ii) The solubilisation of TFI-scFv when expressed as a fusion construct by examining the molecular structure of the Trigger factor: TFI-scFv (TrF:TFI-scFv) fusion construct based on *in silico* homology modelling; and (iii) The up-scale production of TFI-scFv using 5 L bioreactor incubation systems.

5.4 Materials and methods

5.4.1 Bacterial cell lines and plasmids

Cloning host, *E.coli* Top10 F– *mcr*A Δ(*mrr-hsd* RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G (Invitrogen, USA) was used for the *in vitro* amplification pSMART (Lucigen, USA) cloning vectors. All scFvs were expressed using pCOLD DNA II and pCOLD TF expression vectors (TAKARA Bio Inc, Japan) utilising *E.coli BL21* (DE3) *B F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)* (Invitrogen, USA) over expression cultures as well as *E.coli SHuffle® T7 fhuA2 lacZ::T7 gene1 [lon] ompT ahpC gal λatt::pNEB3-r1-cDsbC (SpecR, lacIq) ΔtrxB sulA11 R(mcr-73:miniTn10--TetS)2 [dcm] R(zgb-210::Tn10 --TetS) endA1 Δgor Δ(mcrC-mrr)114::IS10* (New England Biolabs, USA).

5.4.2 Sequencing of the scFv gene fragment

All DNA sequencing reactions were performed using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA) in conjunction with the 3130xl Genetic Analyzer HITACHI (Applied Biosystems, USA) according to manufacturer's specifications. The pCOLD DNA II expression vector constructs were sequenced using sense primer 5'-ACGCCATATCGCCGAAAGG-3' and universal anti-sense primer 5'-GGCAGGGATCTTAGATTCTG-3'. The pCOLD TF expression vector constructs were sequence using 5'-CCACTTTCAACGAGCTGATGA-3' in conjunction with universal anti-sense primer 5'-GGCAGGGATCTTAGATTCTG-3' (Integrated DNA technologies, Belgium).

5.4.3 Construction of the pCOLD TFI-scFv Expression Vector

TFI-scFv gene was amplified to incorporate 5' *NdeI* and 3' *XhoI* restriction sites for downstream cloning using sense primer 5'-GAAGGAAGGCCGTCAAGG-3' and anti-sense primer 5'-TCTAGAATTACGTTTAATTTCAACTTTGG-3' (Integrated DNA technologies, Belgium). The TFI-scFv amplicons were sub cloned into *NdeI/XbaI* linearized expression vector pCOLD DNA II and pCOLD TF expression vectors. The new constructs were designated as pCOLD DNA II TFI-scFv and pCOLD TrF to avoid confusion. TFI-scFv respectively in order to avaind confusion. Both constructs were verified by Sanger sequencing as described previously.

5.4.4 Analysis of TFI-scFv expression levels.

Chemically competed cells *E.coli BL21*(DE3) and *E.coli SHuffle* were transformed using pCOLD DNA II TFI-scFv, pCOLD TrF TFI-scFv, pCOLD DNA II (Empty Control) and pCOLD TrF (Empty Control). Transformants were incubated in 5 mL Luria-bertani (LB) -media (5 $g.L^{-1}$ NaCl, 5 $g.L^{-1}$ Yeast Extract, 10 $g.L^{-1}$ Bacto-Tryptone) containing 100 $\mu g.mL^{-1}$ ampicillin for 16 hours at 37 °C, 160 rpm. The overnight cultures were used at 1:100 dilution to inoculate a 100 mL LB media containing 100 $\mu g.mL^{-1}$ ampicillin. The following day, cultures were incubated at 30 °C until an OD_{600} of 0.6 was reached. At this point, cold shock was initiated through the rapid reduction of incubation temperature. In short, cultures were submerged in ice slurry for approximately 20 min. The cultures were then removed from the ice-slurry and incubated for an additional 30 minutes at 15 °C with no shaking. Expression of the antibodies, were induced by the addition of 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). The cultures where then incubated for 24 hours at 15 °C with 300 rpm agitation. The following day the cultures were harvested by centrifugation at 15000 x g for 10 minutes at 4 ˚C. Whole cell fractions were normalised at OD_{600} =1.0 and evaluated by 10 % SDS-PAGE as described by Laemmli (1970) and stained as described by Fairbanks and co-workers (1971). Total protein production was visualised by SDS-PAGE and analysed by densitometry using JAVA based ImageJ v 1.8.0_101 image analysis software.

5.4.5 Analysis of TFI-scFv solubilisation levels.

Harvested cells were disrupted using the One-Shot Homogenizer (Constant Systems, UK) at 30 000 KPSI. Soluble- and insoluble fractions were separated by centrifugation at 15000 x g for 10 minutes at 4 ˚C. All fractions were evaluated by 10 % SDS-PAGE and densitometry as previously described.

5.4.6 Upscaled production of TFI-scFv using 5L bioreactors

Chemically competent *E.coli BL21*(DE3) and *E.coli SHuffle* cultures were transformed using pCOLD DNA II TFI-scFv, pCOLD TrF TFI-scFv, pCOLD DNA II (Empty Control) and pCOLD TrF (Empty Control). The Transformants were incubated in LB-media (5 $g.L^{-1}$ NaCl, 5 $g.L^{-1}$ Yeast Extract, 10 $g.L^{-1}$ Bacto-Tryptone) containing 100 $\mu g.mL^{-1}$ ampicillin for 16 hours at 37 °C, 160 rpm. The overnight cultures were used at 1:100 dilution to inoculate a 5L Sartorius bioreactor filled with 4L LB media containing 100 $\mu g.mL^{-1}$ ampicillin and controlled by BIOSTAT B Plus universal controller unit (Sartorius, Germany). The cultures were incubated at 37 °C, 300 rpm agitation and 2 litre per minute (lpm) aeration up to mid-log phase (OD_{600} = 0.6). At this point, cold shock was initiated. The incubation chamber was submerged in ice slurry to enable rapid temperature drop required for cold shock. Once the cultures reached 4 °C the bioreactor was removed from the ice-slurry and incubated for an additional 30 minutes at 15 °C. Expression was induced by adding 1 mM IPTG and incubation for 24 hours at 15 °C with 300 rpm agitation. The following day cultures were harvested by centrifugation at 15000 x g for 10 minutes at 4 ˚C stored at -20 °C. Harvested cells were disrupted using the One-Shot Homogenizer (Constant Systems, UK) at 30 000 KPSI. The soluble- and insoluble fractions were separated by means of centrifugation at 15000 x g for 10 minutes at 4 ˚C. The soluble fractions were then further fractionated by means of ultra-centrifugation at 100000 x g for 90 min at 4˚C. All fractions were evaluated by 10 % SDS-PAGE and densitometry as previously described.

5.4.7 Purification of TFI-scFv by Immobilised Metal Affinity Chromatography.

Fractions containing the soluble TFI-scFv and pCOLD TF and fusion construct Trigger factor -TFI-scFv (TrF:TFI-scFv) were purified using a 5 mL His Trap FF column (GE Healthcare, USA) in conjunction with the ÄKTAprime plus (Amersham Biosciences, UK) chromatography system according to manufacture's specifications. The column and system was first equilibrated with 100 mL binding buffer (50 mM tris, 150 mM NaCl containing 20 mM imidazole at pH 7.5) at 5 mL.min-1. The TFI-scFv and TrF-TFI-scFv sample was diluted (1:10) with binding buffer and loaded at 1 $mL.Min^{-1}$. The column was then washed with 10 column volumes binding buffer at 1 $mL.min^{-1}$. Target protein was eluted from the column using a 100 mL linear imidazole gradient (20-500 mM) at 1 $mL.min^{-1}$. The TrF-TFI-scFv fusion construct was separated using the Thrombin Cleancleave™ Kit (Sigma-Aldrich, USA) per manufactures instructions. The cleaved trigger factor and TFI-scFv were separated by means of a secondary purification using a 5 mL His Trap FF column in conjunction with the ÄKTAprime as described previously. The purified TFI-scFv fractions were pooled together and concentrated using 10 000 MWCO VIVASPIN 20 columns

(Sartorius, Germany). The protein concentrations of the solubilized fractions were determined using the Q-Bit Illuminometer and Q-Bit Protein assay. All fractions were evaluated by 10 % SDS-PAGE and densitometry as previously described.

5.4.8 Prothrombin Times

Diluted prothrombin time tests (PTs) were performed in triplicate using Dade C Control platelet poor plasma (Siemens Healthcare, Germany) in conjunction with the STAR-4 coagulation instrument (Diagnostica Stago). Tissue factor (Dade Tromborel S, Siemens Healthcare, Germany) was individually incubated with escalating dosages (0.0 $mg.ml^{-1}$ to 0.2 $mg.ml^{-1}$) of TFI-scFv at 37 °C for 10 min prior to testing. PTs were performed in triplicate and the average clotting time was calculated. The inhibition effect (extended clotting time) of the TFI-scFv fractions was calculated as the difference between the average clotting time and the blank (negative control).

5.4.9 Homology modelling of TrF-TFI-scFv fusion construct

DNA sequence of the TrF-TFI-scFv fusion construct was translated to the final amino acid sequence using online translations tool ExPASy: Translate from Swiss Institute of Bioinformatics. The amino acid sequence was used to model a three-dimensional structure of TrF-TFI-scFv. The three dimensional model of TrF-TFI-scFv construct was predicted using on-line protein structure prediction platform Iterative Threading ASSEmbly Refinement (I-TASSER) version 5.0 [47–49]. In summary, I-TASSER utilised PSI-BLAST to identify possible homology templates [50]. Secondary structure prediction was performed using PSIPRED using the multiple alignment regiments of identified templates [51]. The secondary structure was used to screen for possible homology templates against PDB libraries by LOMETS [52]. Identified models were ranked based on statistical significance of the best threading alignment. The identified continuous segments in the alignments templates were then excised from the template structures and full length models were reassembled by replica-exchange Monte Carlo simulations. Low temperature replicas (decoys) generated during the simulation were clustered by SPICKER [53]. By default, only the top five cluster centroids were considered for generating full atomic models. The lowest energy structures are utilised as input for REMO that generates the final models through the optimisation of the Hydrogen binding networks [54]. Analysis of the predicted TrF-TFI-scFv structure was performed using Yet Another Scientific Artificial Reality Application (YASARA) version 16.4.6.L.64 available at http://www.yasara.org. [55].

5.5 Results

5.5.1 Expression TFI-scFv and TrF-TFI-scFv

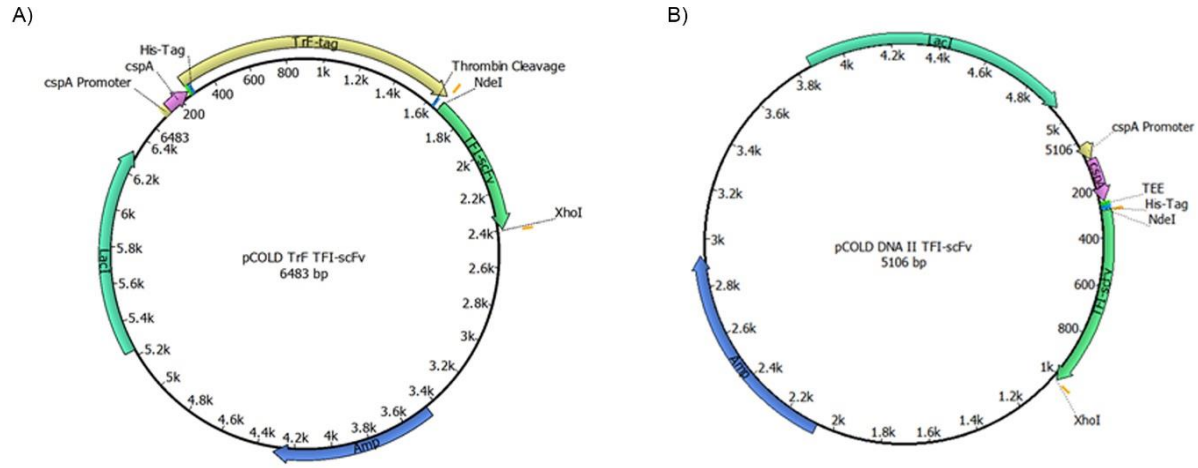


Figure 5.1 Vector maps of pCOLD TrF TFI-scFv and pCOLD DNA II FTI-scFv.

*The schematic representation of the (A) pCOLD TrF TFI-scFv and (B) pCOLD DNA II TFI-scFv expression vectors. The TFI-scFv gene was incorporated using a 5' NdeI- and 3' XhoI restriction sites. The location of the ampicillin resistance genes and LacI operons, cspA cold shock promoter (**Magenta**) and N-terminal His-Tags (**Blue**) are indicated in both vector maps. A) The TFI-scFv gene (**Green**) is incorporated downstream of the TrF fusion partner (**Yellow**) and thrombin cleavage site.*

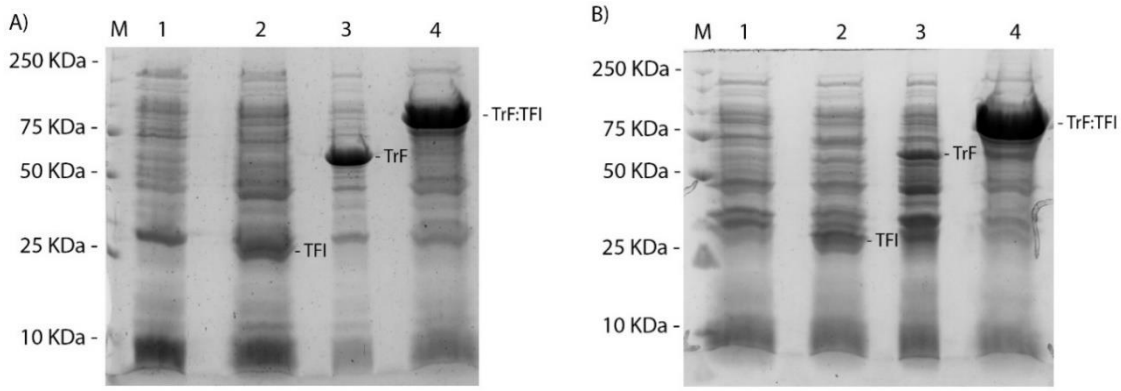


Figure 5.2 Expression Levels of TFI-scFv and TrF-TFI-scFv

SDS-Page analyses of pCOLD DNA II and pCOLD TrF expression profiles in A) E.coli BL21 (DE3) and B) E.coli SHuffle. Lane M: Protein Marker, Lane 1: pCOLD DNA II (Empty Control), Lane 2: pCOLD DNA II TFI-scFv, Lane 3: pCOLD TrF (Empty Control) and lane 4: pCOLD TrF TF-scFv. The 27 kDa TFI-scFv is indicated as TFI, the 52 kDa modified trigger factor is indicated as TrF and the 79 kDa fusion construct is indicated as TrF:TFI.

Protein expression levels were analysed by means of whole cell SDS-PAGE and densitometry. The scFv antibody was successfully expressed by *E.coli BL21*(DE3) using both pCOLD DNA II and pCOLD TrF vectors (Figure 5.1 A). Expression of the pCOLD DNA II TFI-scFv construct (lane 2) resulted in the production of TFI-scFv (27 kDa)

which consisted of approximately 15 % of the total protein. The over expression of modified fusion partner (Trigger factor 52 kDa) by pCOLD TrF (Empty Control) is clearly visible in lane 3. The Trigger factor constituted approximately 25 % of the total protein. The 79 kDa TrF-TFI-scFv fusion construct (Trigger factor (52 kDa) + TFI-scFv (27 kDa)) is also clearly visible in lane 4. Similarly, the fusion construct constituted approximately 25 % of total protein production. Both of the constructs were also successfully expressed by *E.coli SHuffle* (Figure 5.1 B). The 27 kDa antibody expressed from the pCOLD DNA II TFI-scFv vector was approximately 10% of total protein (lane 2). Interestingly, the production levels of the trigger factor fusion partner (lane 3) is slightly reduced to 15 % of total protein in comparison to the 25% of total protein attained when expressed by *E.coli BL21* in Figure 5.1 A lane 3. The TrF-TFI-scFv fusion construct (lane 4) constituted approximately 25 % of total protein.

5.5.2 Solubilisation levels of TFI-scFv and TrF:TFI-scFv

The extent of TFI-scFv solubilisation was measured by means of whole cell SDS-PAGE and densitometry. The expression of the pCOLD DNA II TFI-scFv vector by *E coli Bl21* (Figure 5.2 A Lane 1 and 2) resulted in production of approximately 19% soluble TFI-scFv with the majority (> 80%) of the protein produced as insoluble inclusion bodies. The expression of the pCOLD TrF-TFI-scFv vector by *E.coli BL21* resulted in the solubilisation of approximately 38% of the total TrF-TFI-scFv produced. Expression of the two vectors by *E.coli SHuffle* resulted in the solubilisation of approximately 20% of the TFI-scFv produced by pCOLD DNA II and 37% of the fusion constructs produced by pCOLD TrF TFI-scFv (Figure 5.2 B).

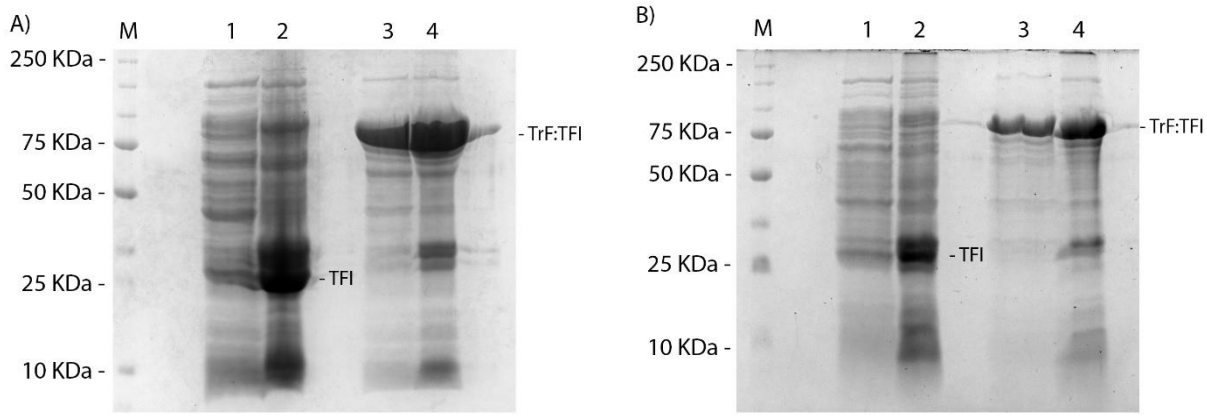


Figure 5.3 Solubilisation levels of TFI-scFv and TrF:TFI-scFv

SDS-Page analyses of pCOLD DNA II TFI-scFv and pCOLD TrF TFI-scFv expression profiles A) E.coli BL21 (DE3) and B) E.coli SHuffle. Lane M: Protein Marker, Lane 1: pCOLD DNA II TFI-scFV (Soluble Fraction), Lane 2: pCOLD DNA II TFI-scFv (Insoluble Fraction), Lane 3: pCOLD TrF TFI-scFV (Soluble Fraction) and Lane 4: pCOLD TF TF-scFv (Insoluble Fraction). The 27 kDa TFI-scFv is indicated as TFI and the 79 kDa fusion construct is indicated as TrF:TFI.

5.5.3 Large scale expression and purification.

The use of the bioreactors resulted in a final yield of roughly 7 to 9 grams of *E.coli BL21* per litre and 3 to 5 grams of *E.coli SHuffle* per litre. The final protein production ranged from 2.0 $mg.L^{-1}$ to 5 $mg.L^{-1}$ for *E.coli BL21* and 1.0 .$mg.L^{-1}$ to 3.5. $mg.L^{-1}$ for *E.coli SHuffle*. Overall the *E.coli BL21* (DE3) strains are capable of producing much higher biomasses that the *E.coli SHuffle* strains when grown under cold shock conditions. The final biomass production and protein purification details are summarised in Table 5.1. below.

Table 5.1 Summary of large scale expression and protein yeilds

Culture	*Plasmids*	*Biomass per Litre (Wet weight in g)*	*Total Protein Yield (mg per Litre)*
E.coli BL21 (DE3)	*pCOLD DNA II*	*7.3*	*2.13*
	pCOLD TrF	*8.2*	*4.83*
E.coli SHuffle	*pCOLD DNA II*	*3.6*	*1.32*
	pCOLD TrF	*4.2*	*3.26*

The antibody fragments were purified using nickel based-IMAC in conjunction with the AKTAprime Plus chromatography system. The final purified protein fractions were normalised to 0.5 $mg.L^{-1}$ and analysed by means of SDS-PAGE and densitometry to confirm correct protein size and purity (Figure 5.3). The 27 kDa TFI-scFv was successfully isolated and purified form the cytoplasm of both *E.coli* strains. The purity of the isolated protein fractions was determined by densitometry. The non-fusion construct was isolated with a purity greater than 99 %, while the purity of the antibody purified from the fusion construct was higher than 95 % for both strains utilised.

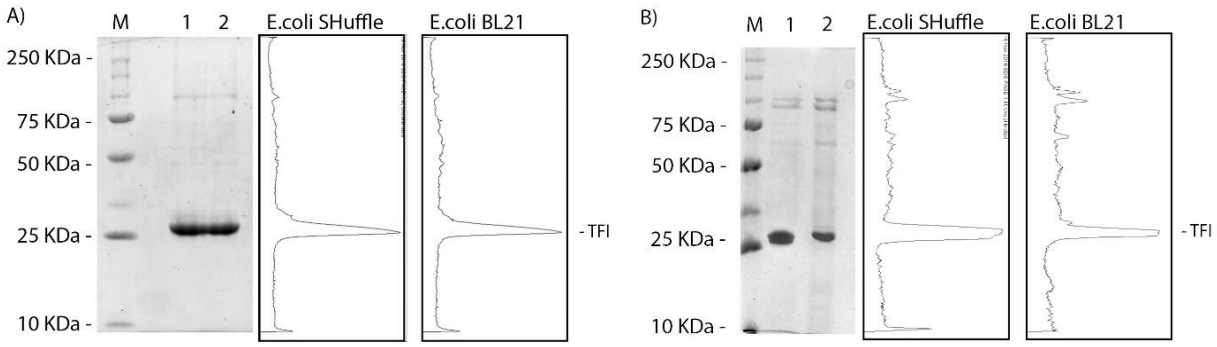


Figure 5.3 Final purified fractions.

A) The SDS-PAGE analysis of the purified TFI-scFv fractions. Lane M: Protein Marker, Lane 1: TFI-scFv produced by E.coli SHuffle using pCOLD DNA II expression vector, Lane 2 TFI-scFv :produced by E.coli BL21 using pCOLD DNA II expression vector. The respective densitometry profiles are presented to the right.

B) The SDS-PAGE analysis of the purified TFI-scFv fractions produced as a TrF-TFI-scFv fusion construct. Lane M: Protein Marker, Lane 1: TFI-scFv produced by E.coli SHuffle using pCOLD TrF expression vector, Lane 2 TFI-scFv :produced by E.coli BL21 using pCOLD TrF expression vector. The respective densitometry profiles are presented to the right.

5.5.4 Prothrombin times

Diluted PT assays were carried out in order to assess the inhibition efficiency of the various purified TFI-scFv fractions. Of the four different purified fractions that were tested only the non-fusion construct produced by *E.coli SHuffle* pCOLD DNA II was capable of notably extending the prothrombin times (Figure 5.4). The prothrombin time at the highest concentration (0.2 mg.ml^{-1}) of TFI-scFv was extended by 16.67 s which is an approximate 77% extension of the baseline prothrombin time with a specific activity of 83.8 s/mg.mL^{-1}. The remaining isolates of TFI-scFv fractions had every little inhibition effect, with all fractions only extending the prothrombin times less than 3s.

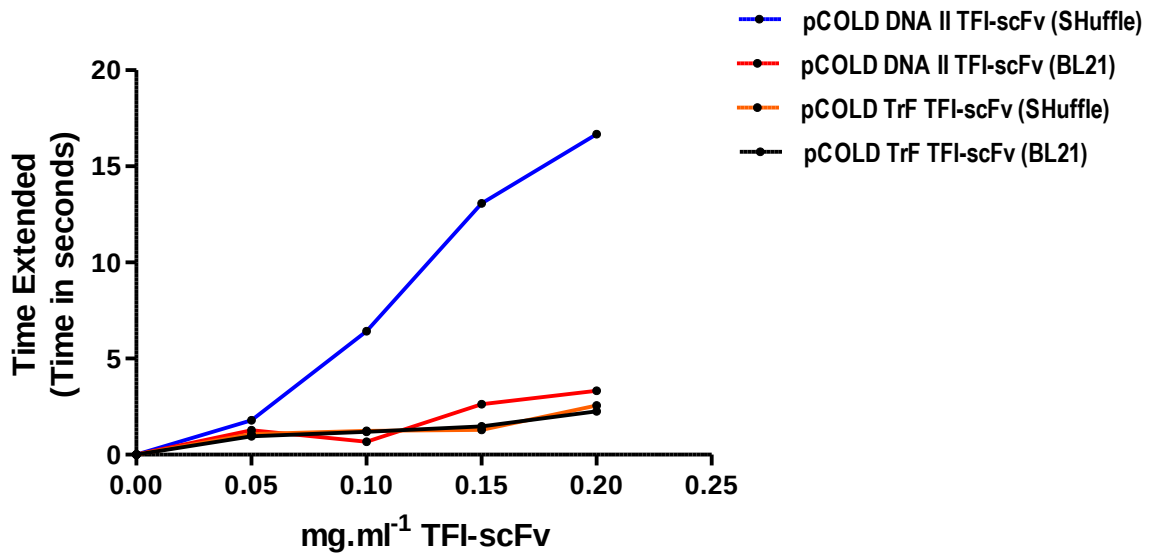


Figure 5.5 Elongation of Prothrombin Times

The inhibition efficacy of the various isolated fraction of TFI-scFv was tested by means of prothrombin times. Only the non-fusion construct produced by E.coli SHuffle using pCOLD DNA II expression vector was capable of extending the prothrombin times in a dose dependant manner. The remaining isolates were incapable of notably extending the prothrombin times.

5.5.5 Homology modelling of the TF-TFI-scFv fusion construct.

The top ranked PDB templates utilised during the homology modelling of the TrF-TFI-scFv fusion construct are listed in table 5.2. The top scoring homology model as well as estimated residue accuracy (predicted by I-TASSER) are presented in figure 5.5.

Table 5.2: Top ranked PDB templates utilized for homology modelling.

Rank	*Structure*	*Resolution (Å)*	*Description*
1	*1W26*	*2.7*	*Trigger Factor in complex with ribosome* [56]
2	*3GTY*	*3.4*	*Trigger Factor* [57]
3	*4ZHJ*	*2.5*	*Catalytic Subunit of Magnesium Chelatase* [58]
4	*5A22*	*3.8*	*L protein of vesicular stomatitis virus* [59]
5	*2PO4*	*2.0*	*Polymerase domain of the bacteriophage N4 virion* [60]

The C-Score (Range -5 to 2) is an indication of the quality of the predicted structure. The calculation is based on the significance of threading template alignments and the convergence parameters of the structures assembly simulations. Cluster density is defined as the number of structure decoys clustered SPICKER. A higher cluster density is indicative of a higher incidence of a specific structure in the simulation.

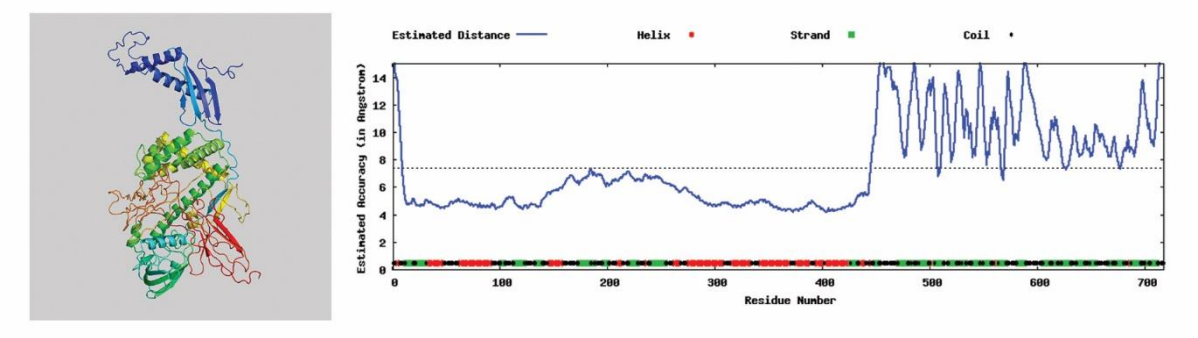


Figure 5.6: Homology modelling of TrF-TFI-scFv.

*The top homology model as generated by I-TASSER. Cartoon model is presented in colour gradient ranging from N-terminal (**blue**) towards the C-Terminal (**red**). The estimated residue accuracy of each model is plotted to the right, with the average indicated by the dotted line.*

Table 5.3: I-TASSER homology model quality score.

Rank	*C-score*	*Number decoys*	*Cluster density*	*Avg. Est. Accuracy*
A	*-2.12*	*336*	*0.0243*	*7.5*

A higher C-score, number of decoys and cluster density is indicative of a better-quality model. As the native structure is unknown, accuracy plots are estimated by ResQ using support vector regressions that makes use of the coverage of threading alignment, divergence decoys, and secondary structure and solvent accessibility predictions. The estimated accuracy values are a good indication overall model quality as well as of model quality per individual residue. [61]

5.5.6 Structure of TrF-TFI-scFv fusion construct

E.coli trigger factor is a highly flexible protein that consists of a ribosome-binding domain (RBD), a peptidyl prolyl isomerase domain (PPD) and a discontinuous C-terminal domain, referred to as the substrate binding domain (SBD), that is situated between the RBD and PPD [62]. According to the I-TASSER model, the native structure of trigger factor has remained intact within the fusion construct, with minimal distortion due inherent flexibility of the protein (Figure 5.6) [63].

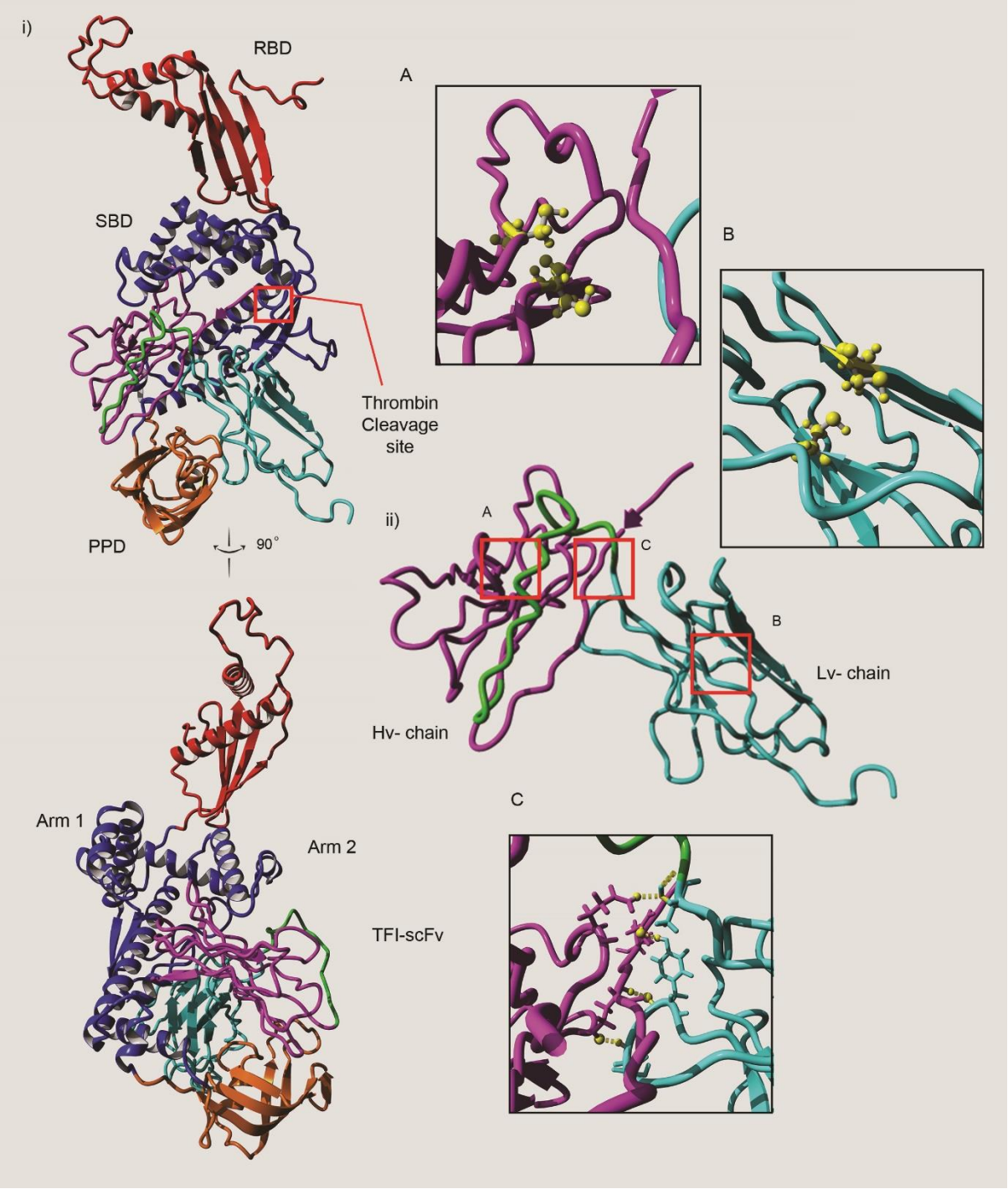


Figure 5.7: Structure of the Trigger Factor-TFI-scFv homology model.

*Two cartoon diagrams of the models are shown in i) TrF-TFI-scFv fusion construct and ii) TFI-scFv homology model with the trigger factor domain hidden. Trigger factor domains are indicated as ribosome binding domain (RBD) (**red**), the peptidyl prolyl isomerase domain (PPD) (**orange**) and substrate binding domain (SBD) (**blue**). The 90°rotation of the TrF-TFI-scFv displays the so called "Arms" of trigger factor. The heavy chain variable region is indicated (**magenta**), as well as the linker region (**green**) and the light chain variable region (**cyan**). The Thrombin cleavage site is indicated by the red square. Areas of interest are highlighted in red boxes on the main structures are defined as A) Absent Hv-chain interdomain disulfide bonds B) Absent Lv-chain disulfide bonds C) Hydrogen bond formation between the Hv-chain and Lv-chain interface. Hydrogen bonds of interest are indicated (**yellow**) within the highlighted boxes.*

According to the model, the TFI-scFv is folded in such a way that the antibody structure is nestled between the SBD and the PPD with Complementarity Determining Regions (CDR) oriented towards trigger factor and the linker region exposed to the solvent. The total solvent accessible area of the fusion construct was calculated as 10515.12 Å^2, with an interface area of 4723.32 Å^2 between trigger factor and the TFI-scFv (partially folded). This is 36.36 % of the total surface area of the partially folded TFI-scFv (Figure 5.7). Energy minimizations were performed using self-parameterizing knowledge-based YASARA force field and was calculated at -5052.36 kcal/mol. *(See model no TrF-TFI-scFv and supplementary video).*

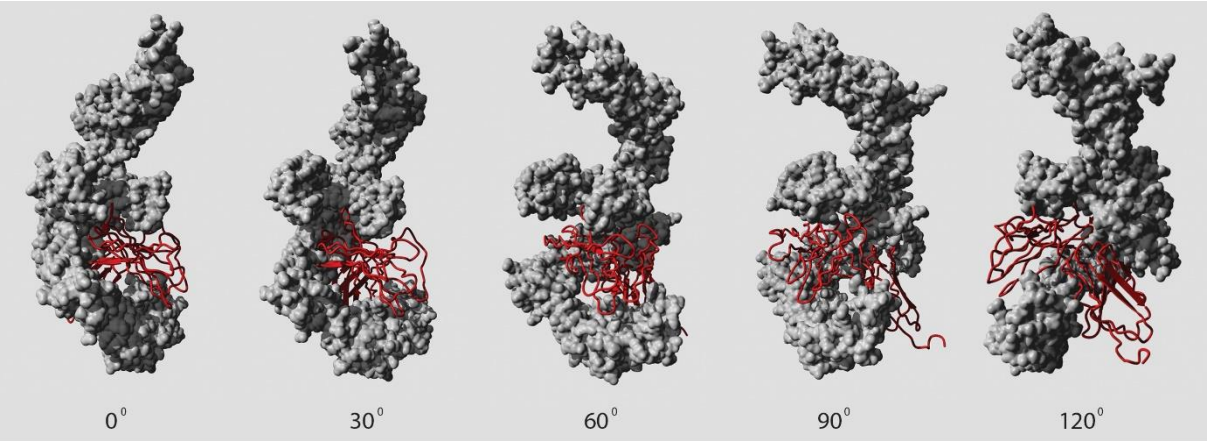


Figure 5.8: The imbedded position of TFI-scFv with in the Fusion construct.

*The surface area (van der Waal forces) of trigger factor is indicated (**Grey**) in association the cartoon diagram or the partially folded TFI-scFv (**RED**) presented at 30° degree intervals. The TFI-scFv structure is imbedded within the larger structure of trigger factor which limits the solvent exposed surface area.*

When comparing structures of the native TFI-scFv model with that of the partially folded TFI-scFv , the perpendicular orientation of the Hv- and Lv-chain barrel-like structure of the partially folded structure are more exaggerated. The formation of a cleft that is due to the occupation by the PPD of trigger factor. The perpendicular orientation is facilitated by a series of moderate strength hydrogen bonds interactions, formed at the interface between Hv- and Lv-chain (See: Figure 5.6 C and Table 5.4).

Table 5.4 Hydrogen bond interaction between Hv and Lv Chain

Heavy Chain variable region	*Light chain variable region*	*Distance in Å*	*Bond Energy kcal/mol*	*Relative H-bond strength*
Gln 475	*Tyr 696*	*1.93*	*5.98*	*Moderate*
Glu 478	*Thy 696*	*1.93*	*5.98*	*Moderate*
Glu 478	*Thr 698*	*1.89*	*5.04*	*moderate*
Gln 580	*Thr 604*	*2.13*	*2.57*	*moderate*

Although the general structure of the Hv- and Lv-chains are loosely retained, in comparison to the native TFI-scFv model much of the structure is distorted (Figure 5.8). The root-mean-square deviation (RMSD) of atomic positions for the superimposed structures was calculated as 15.4187 Å using superpose and MUSTANG alignment module in YASARA indication that the deviations are fairly substantial. The dissimilarities in the two structures are also evident when comparing the secondary structure composition (Table 5.5). The partially folded structure is missing much of the Beta-sheet structures present in the native model. This large deviation in structure is partially due to the absence of the stabilising disulfide linkages spanning the hydrophobic regions (Figure 5.6 A and B), which contribute to the less defined barrel-like structures observed in the partially folded constructs. Thus, the partially folded TFI-scFv was predicted to consist mostly (75 %) of coil-like structures.

Table 5.5 Summary of Secondary structure content of native- and partially folded TFI-scFv.

TFI-scFv Structure	*Helix (%)*	*Sheet (%)*	*Turn (%)*	*Coil (%)*
Partially folded	*0.0*	*12.6*	*11.7*	*75.7*
Native	*0.0*	*51.0*	*19.6*	*29.4*

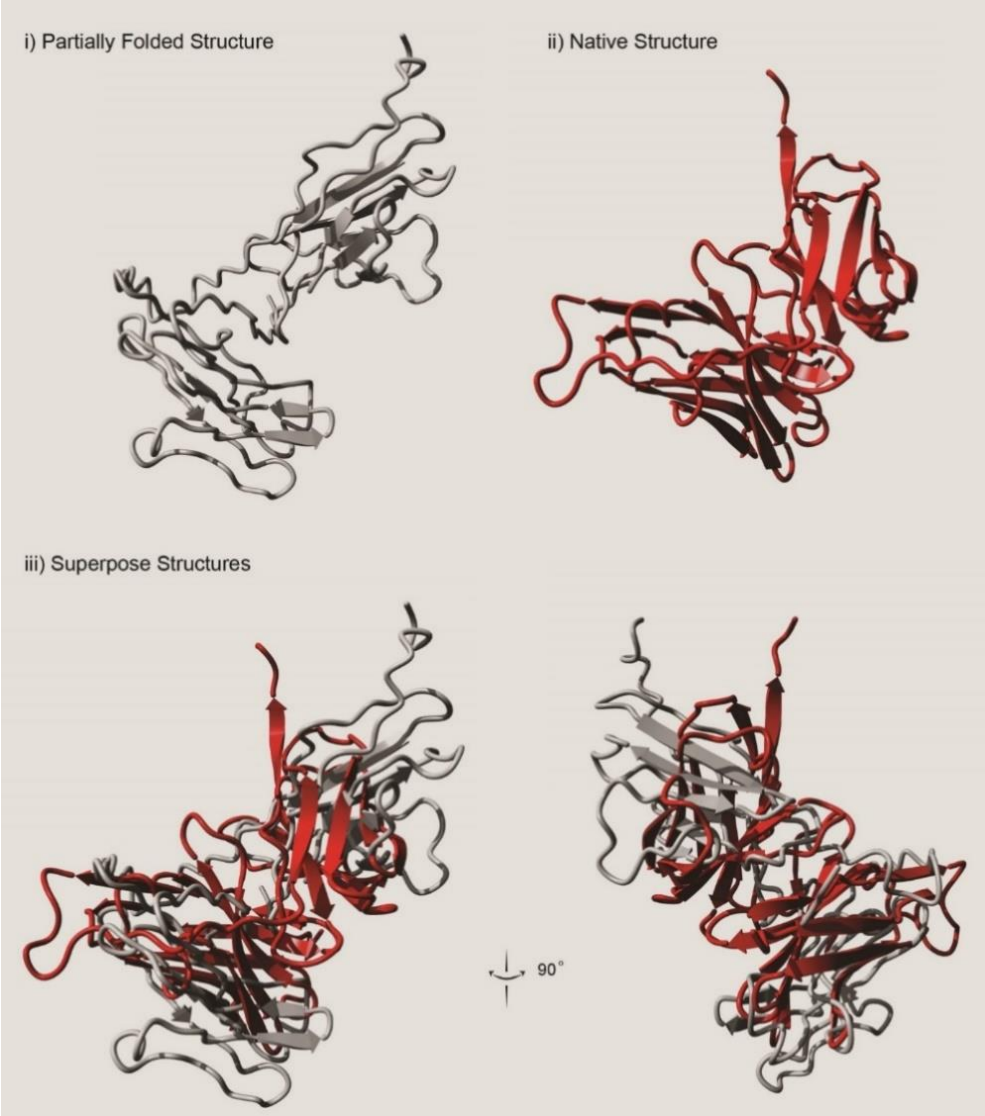


Figure 5.9: Structural difference between partially folded and native TFI-scFv

*i) The cartoon diagram of partially folded TFI-scFv structure following thrombin cleavage form the trigger factor fusion partner (**Grey**). ii) The cartoon diagram of the native TFI-scFv structure as predicted by homology modelling (**Red**). iii) The superimposed views of TFI-scFv cartoon structures at 90° rotation showing the deviation between the native- and partially folded structures.*

5.6 Discussion

Soluble TFI-scFv was produced in the cytoplasm of *E.coli* using a combination of cold shock, solubilising fusion partner and specialised disulfide bond promoting *E.coli* strains. When comparing the expression of TFI-scFv using the pCOLD DNA II cold shock vector by *E.coli BL21 (DE3)*- and *SHuffle* strains, it is interesting to note that the expression level by *E.coli BL21* was only marginally higher (5%), while the level of solubilisation between the two strains remained similar at 20% of the target protein. By comparison, the expression levels of the TrF-TFI-scFv fusion constructs using pCOLD TF were roughly 10% higher. This increase in protein expression is largely due to

the high expression levels of the fusion partner itself. Trigger factor constitutes 15 to 25 % of total protein yield in the respective controls. Trigger factor is a Class II cold shock protein that is naturally induced upon cold shock. The expression is further enhanced by incorporation of cspA (cold shock protein A) promoter elements [70]. The TrF-TFI-scFv fusion constructs were expressed at similar levels (25% of the total protein). The solubilisation levels were also higher than that of the pCOLD DNA II constructs. A total of 37-38 % of the expressed fusion construct was produced as soluble TFI-scFv. These similar expression levels are in contrast to the findings of Fati-Roudsari *et al (*2016*)* where they investigated the ability of *E.coli BL21* (DE3), Rosetta-gami and *E.coli SHuffle* to express a multi disulfide bond containing reteplase using the pET-21a vector [69][71]. Although the conflicting results are interesting, one must be careful when comparing results. Factors such as protein yield, biomass production and level of solubilisation are largely dependent on the specific nature of the target protein itself, and a large degree of variation is expected.

It is important to note that, despite the specialised mutations to *E.coli SHuffle* to reduce aggregation, the level of solubilisation was not higher than that of *E.coli BL21*. These similarities in the level of solubilisation were unexpected as the *E.coli SHuffle* strain is specially designed to facilitate disulfide bond formation and in so doing, should in theory, reduce target protein aggregation [72]. The cytoplasm of *E.coli SHuffle* is rendered slightly oxidising due to the gor^{-}/ trxB^{-} mutations. These modifications to the redox state of the host cytoplasm, in conjunction with the constitutively co-expressed DsbC isomerase, should allow for more efficient disulfide bond formation and result in improved production of soluble scFv [23]. Furthermore, the strain is also modified to be deficient in protease Lon and OmpT to prevent degradation of target protein.

The similarities in protein solubilisation when comparing the two cell lines might be attributed to the comprehensive effect of cold shock. The sudden drop in temperature causes physical stresses that influence all physiochemical parameters of the cell. It influences solute diffusion rates, enzyme kinetics, membrane fluidity, flexibility, topology and the interactions of DNA, RNA as well as proteins [26]. Protein expression at low temperature results in improved solubilisation of the target protein likely due to the reduction in translation rate of mRNA to the polypeptides [23,73,74] The activity and expression of a number of chaperones are also increased at lower temperatures [24]. Furthermore, activity of endogenous proteases is suppressed at low temperature [74]. These factors in conjunction with the physical state of the host cell all contribute to the stabilisation the target protein.

However, a major drawback to cold shock expression systems is the reduction in target protein-as well as final biomass yields, caused by to the lower growth- and protein synthesis rates [75]. This issue becomes further compounded by the slower growth rate of *E.coli SHuffle.* To compensate for the low expression and growth rates, cold shock expression are generally performed over an extended period (24h). The use of bioreactors provided a higher level of control over the induction as well as the maintenance of the cold shock conditions. Additionally, they are substantially less labour intensive than shake-flask incubation methods. The *E.coli Bl21* (DE3) strains yielded

roughly double the biomass compared to *E.coli SHuffle* when grown in the 5L bioreactors. A higher biomass yield is advantageous as it assures a higher final protein yield

The various expressed antibody fragments were purified using IMAC. The isolated TF-TFI-scFv fusion constructs were separated by means of thrombin cleavage and purified by IMAC. The final isolated TFI-scFv fractions were > 95% pure once separated from the fusion partner and the non-fusion constructs > 99 % pure. Of the four isolated fractions, only the TFI-scFv, produced via the expression of pCOLD DNA II plasmid in *E.coli Shuffle,* was capable of inhibition as demonstrated using PT assay. The remaining isolates exhibited very little to no inhibition. Prothombin Time test are especially well suited for the testing of functionality of the TFI-scFv as the reaction is only initiated once tissue factor is added. The extension of the prothrombin times is therefore directly correlated to TFI-scFv's ability to recognise the tissue factor ligand and prevent the formation of the tissue factor-factor VII complex. It is this action of the TFI-scFv that is responsible for the extension of the prothrombin times. PTs also make use of platelet poor plasma which makes the PT an excellent tool for obtaining an overview of the tissue factor (Extrinsic) coagulation pathway in semi-isolation. Thus, it is not surprising that the PT has been the gold standard for monitoring INR rations for patients receiving warfarin treatment.

It is important to note that solubilisation does not necessarily translate into functionality, as was observed for the *in vitro* refolded TFI-scFv. This is likely due to the formation of folding intermediaries that are soluble but do not have the correct conformation of CDR structure to recognise the target ligand. Antibody fragments, such as scFv, are highly dependent on the formation of stabilising disulfide bonds for functionality. ScFv have two intra-domain stabilising disulfide bond (Cys^{24}-Cys^{98} and Cys^{157}-Cys^{222}) that span the Hv- and Lv chains respectively which facilitate the correct orientation of the CDR. If the TFI-scFv is improperly folded the CDR will not be able to recognition of the target ligand.

When comparing the TFI-scFv produced by the pCOLD DNA II vectors the major difference between the production processes of these two TFI-scFv fractions are the enhanced disulfide bond forming capabilities of the *E.coli SHuffle* strain. The expression of proteins with disulfide bonds in the cytoplasm of *E.coli* is often a substantial challenge in protein engineering as the environment is not entirely conducive to it. Although there was no notable difference in the level of expression- and solubilisation levels, it is fair to assume that the *E.coli SHuffle strain* did promote the production of functional TFI-scFv likely through the promotion of proper disulfide bond formation.
In light of the inability of the of TFI-scFv to inhibit the action of tissue factor once cleaved from the fusion construct, despite a relatively higher level of expression and solubilisation, a homology model of the fusion construct was generated in order to gain insight in to the structural interactions between TFI-scFv and trigger factor. The highest scoring predicted model suggested that much of the native trigger factor structure is retained in the fusion construct while the structure of the TFI-scFv located toward the C-terminus of the fusion construct is distorted. According to the homology model the TFI-scFv folded in such a way that the antibody structure is nestled between the SBD and the PPD with CDR determining region of TFI-scFv oriented towards trigger factor and the linker region exposed

towards the solvent. This nestled orientation of TFI-scFv would help limit interaction between individual TFI-scFv due to the large contact area of 4723.32 $Å^2$ between TFI-scFv and trigger factor. This nestled orientation of TFI-scFv within the fusion construct is possibly responsible for the improved solubilisation of the TrF:TFI-scFv fusion constructs. However, the majority of the expressed TrF-TFI-scFv fraction remained insoluble. One could speculate that is most likely due to the association of various folding intermediates with one another that result in the formation of the inclusion bodies.

Although the solubilisation was improved through the use of trigger factor as a solubilisation partner, the structure of TFI-scFv was distorted when orientated into this nestled position. This was also confirmed by the relatively large RMSD of 15.4187 Å when comparing the superimposed structures of the partially folded- and native TFI-scFv. Thus, TFI-scFv would have undergone a substantial amount of additional folding in order to attain the native structure once cleaved from the fusion partner. It is also important to remember that once the partially folded TFI-scFv is cleaved from the trigger factor fusion partner that the target protein was collected in TBS buffer solution. TBS is specifically utilised as it does not interfere with downstream functional assays. The buffer does not contain any additives that would modify the redox potential to such extent that if would facilitate the *in vitro* formation of disulfide bonds. It is therefore highly unlikely that the partially folded TFI-scFv would undergo the required conformational changes to A) form the required stabilising disulfide bridges and B) fold the CDR into the correct conformation to be able to recognise the target ligand.

The production of the TFI-scFv could be further improved upon using specialised growth medias such as: Auto-induction medias [76], Terrific Broth [43] or EnPresso medium [72]. Zarschler et al (2013) reported that EnPresso medium resulted in a 10 fold increase in *E.coli SHuffle* biomass when expressing antibodies. In addition, the bioreactors can be further optimised and up-scaled to improve biomass production and thus increase final antibody yield. The expression systems described here can be upscaled and potentially utilised on an industrial scale to produce functional TFI-scFv for the future characterisation of the antibody fragment.

5.7 Conclusion

It is feasible to produce functional TFI-scFv in the cytoplasm of *E.coli.* as an alternative approach to the *in vitro* refolding of insoluble aggregates**.** The combination effect of three solubilisation strategies; cold shock, the use of trigger factor as a solubilisation partner as well as the use of specialised *E.coli* strain that promote the formation of disulfide bonds was investigated. The combined use of these strategies resulted in the production of soluble TFI-scFv within the cytoplasm of *E. coli* to varying degrees. Although all the strategies resulted in the improved solubilisation of TFI-scFv, only one of the expression strategies followed resulted in the production of a functional antibody as demonstrated. The utilisation of the TAKARA pCOLD DNA II expression vector in conjunction with *E.coli SHuffle* strain under cold shock conditions resulted in the successful expression of functional TFI-scFv. Functional TFI-scFv was successfully expressed in the bacterial cytoplasm and purified by means of IMAC at a final yield of 1.32 $mg.L^{-1}$. The expression systems described here can be upscaled and utilised to produce functional TFI-scFv for the future characterisation of the antibody fragment.

5.8 Acknowledgements

This work has been funded by the National Research Foundation (NRF) of South Africa.

5.9 References

[1] E.K. Spicer, R. Horton, L. Bloem, R. Bach, K.R. Williams, A. Guha, J. Kraus, T.C. Lin, Y. Nemerson, W.H. Konigsberg, Isolation of cDNA clones coding for human tissue factor: primary structure of the protein and cDNA., Proc. Natl. Acad. Sci. U. S. A. 84 (1987) 5148–52. http://www.ncbi.nlm.nih.gov/pubmed/3037536.

[2] S.A. Smith, R.J. Travers, J.H. Morrissey, How it all starts: Initiation of the clotting cascade, Crit. Rev. Biochem. Mol. Biol. 50 (2015) 326–336. doi:10.3109/10409238.2015.1050550.

[3] A.J. Chu, Tissue Factor, Blood Coagulation, and Beyond: An Overview, Int. J. Inflam. 2011 (2011) 1–30. doi:10.4061/2011/367284.

[4] N. Mackman, Role of tissue factor in hemostasis and thrombosis, Blood Cells, Mol. Dis. 36 (2006) 104–107. doi:10.1016/j.bcmd.2005.12.008.

[5] A.P. Owens, N. Mackman, Tissue factor and thrombosis: The clot starts here, Thromb. Haemost. 104 (2010) 432–439. doi:10.1160/TH09-11-0771.

[6] G.E. Raskob, P. Angchaisuksiri, A.N. Blanco, H. Buller, A. Gallus, B.J. Hunt, E.M. Hylek, A. Kakkar, S. V. Konstantinides, M. McCumber, Y. Ozaki, A. Wendelboe, J.I. Weitz, Thrombosis: A Major Contributor to Global Disease Burden, Arterioscler. Thromb. Vasc. Biol. 34 (2014) 2363–2371. doi:10.1161/ATVBAHA.114.304488.

[7] G.E. Raskob, P. Angchaisuksiri, A.N. Blanco, H. Buller, A. Gallus, B.J. Hunt, E.M. Hylek, A. Kakkar, S. V Konstantinides, M. McCumber, Y. Ozaki, A. Wendelboe, J.I. Weitz, Thrombosis: A Major Contributor to Global Disease Burden, Arterioscler. Thromb. Vasc. Biol. 34 (2014) 2363–2371. doi:10.1161/ATVBAHA.114.304488.

[8] I. Ott, Tissue factor inhibition: Another approach reducing thrombosis after vascular injury, Thromb. Haemost. 103 (2009) 7–8. doi:10.1160/TH09-10-0712.

[9] S.M. Meiring, J. Vermeulen, P.N. Badenhorst, Development of an inhibitory antibody fragment to human tissue factor using phage display technology, Drug Dev. Res. 70 (2009) 199–205. doi:10.1002/ddr.20295.

[10] Z.A. Ahmad, S.K. Yeap, A.M. Ali, W.Y. Ho, N.B.M. Alitheen, M. Hamid, scFv Antibody: Principles and Clinical Application, Clin. Dev. Immunol. 2012 (2012) 1–15. doi:10.1155/2012/980250.

[11] L. Sánchez, M. Ayala, F. Freyre, I. Pedroso, H. Bell, V. Falcón, J. V. Gavilondo, High cytoplasmic expression in E. coli, purification, and in vitro refolding of a single chain Fv antibody fragment against the hepatitis B surface antigen, J. Biotechnol. 72 (1999) 13–20. doi:10.1016/S0168-1656(99)00036-X.

[12] S. Tapryal, L. Krishnan, J.K. Batra, K.J. Kaur, D.M. Salunke, Cloning, expression and efficient refolding of carbohydrate–peptide mimicry recognizing single chain antibody 2D10, Protein Expr. Purif. 72 (2010) 162–168. doi:10.1016/j.pep.2010.03.024.

[13] L. Vallejo, U. Rinas, Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins, Microb. Cell Fact. 3 (2004) 11. doi:10.1186/1475-2859-3-11.

[14] A. Singh, V. Upadhyay, A.K. Panda, Solubilization and Refolding of Inclusion Body Proteins, in: Insoluble Proteins Methods Protoc., 2015: pp. 283–291. doi:10.1007/978-1-4939-2205-5_15.

[15] J.A. Vasina, F. Baneyx, Recombinant protein expression at low temperatures under the transcriptional control of the major Escherichia coli cold shock promotor cspA., Appl. Environ. Microbiol. 62 (1996) 1444–1447.

[16] E.D.B. Clark, Protein refolding for industrial processes, Curr. Opin. Biotechnol. 12 (2001) 202–207. doi:10.1016/S0958-1669(00)00200-7.

[17] T. Arakawa, D. Ejima, Refolding Technologies for Antibody Fragments, Antibodies. 3 (2014) 232–241. doi:10.3390/antib3020232.

[18] A. Mitraki, J. King, Protein Folding Intermediates and Inclusion Body Formation., Bio/Technology. 7 (1989) 690–697. doi:10.1038/nbt0789-690.

[19] J. Song, Why do proteins aggregate? “Intrinsically insoluble proteins” and “dark mediators” revealed by studies on “insoluble proteins” solubilized in pure water, F1000Research. 2 (2013) 94. doi:10.12688/f1000research.2-94.v1.

[20] A. Singh, V. Upadhyay, A.K. Upadhyay, S.M. Singh, A.K. Panda, Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process., Microb. Cell Fact. 14 (2015) 41. doi:10.1186/s12934-015-0222-8.

[21] K. Sushma, C.J. Bilgimol, M.A. Vijayalakshmi, P.K. Satheeshkumar, Recovery of active anti TNF-?? ScFv through matrix-assisted refolding of bacterial inclusion bodies using CIM monolithic support, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 891–892 (2012) 90–93. doi:10.1016/j.jchromb.2012.02.011.

[22] G.P. Subedi, T. Satoh, S. Hanashima, A. Ikeda, H. Nakada, R. Sato, M. Mizuno, N. Yuasa, Y. Fujita-Yamaguchi, Y. Yamaguchi, Overproduction of anti-Tn antibody MLS128 single-chain Fv fragment in Escherichia coli cytoplasm using a novel pCold-PDI vector, Protein Expr. Purif. 82 (2012) 197–204. doi:10.1016/j.pep.2011.12.010.

[23] H.P. Sørensen, K.K. Mortensen, Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli., Microb. Cell Fact. 4 (2005) 1. doi:10.1186/1475-2859-4-1.

[24] C.A. White-Ziegler, S. Um, N.M. Perez, A.L. Berns, A.J. Malhowski, S. Young, Low temperature (23 C) increases expression of biofilm-, cold-shock- and RpoS-dependent genes in Escherichia coli K-12, Microbiology. 154 (2008) 148–166. doi:10.1099/mic.0.2007/012021-0.

[25] C.O. Gualerzi, A. Maria Giuliodori, C.L. Pon, Transcriptional and Post-transcriptional Control of Cold-shock Genes, J. Mol. Biol. 331 (2003) 527–539. doi:10.1016/S0022-2836(03)00732-0.

[26] S. Phadtare, Recent Developments in Bacterial Cold-Shock Response, Curr. Issues Mol. Biol. (2004) 125–136.

[27] W. Bae, P.G. Jones, M. Inouye, CspA, the major cold shock protein of Escherichia coli, negatively regulates its own gene expression., J. Bacteriol. 179 (1997) 7081–8. http://www.ncbi.nlm.nih.gov/pubmed/9371456.

[28] X. Hu, L. O'Hara, S. White, E. Magner, M. Kane, J. Gerard Wall, 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 Expr. Purif. 52 (2007) 194–201. doi:10.1016/j.pep.2006.08.009.

[29] A. Mogk, M.P. Mayer, E. Deuerling, Mechanisms of protein folding: Molecular chaperones and their application in biotechnology, ChemBioChem. 3 (2002) 807–814. doi:10.1002/1439-7633(20020902)3:9<807::AID-CBIC807>3.0.CO;2-A.

[30] T. Rathnayaka, M. Tawa, S. Sohya, M. Yohda, Y. Kuroda, Biophysical characterization of highly active recombinant Gaussia luciferase expressed in Escherichia coli, Biochim. Biophys. Acta - Proteins Proteomics. 1804 (2010) 1902–1907. doi:10.1016/j.bbapap.2010.04.014.

[31] G. Hannig, S.C. Makrides, Strategies for optimizing heterologous protein expression in Escherichia coli., Trends Biotechnol. 16 (1998) 54–60. doi:10.1016/S0167-7799(97)01155-4.

[32] K. Terpe, Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, Appl. Microbiol. Biotechnol. 60 (2003) 523–533. doi:10.1007/s00253-002-1158-6.

[33] M.H. Hefti, C.J.G. Van Vugt-Van der Toorn, R. Dixon, J. Vervoort, A Novel Purification Method for Histidine-Tagged Proteins Containing a Thrombin Cleavage Site, Anal. Biochem. 295 (2001) 180–185. doi:10.1006/abio.2001.5214.

[34] A. Basters, L. Ketscher, E. Deuerling, C. Arkona, J. Rademann, K.-P. Knobeloch, G. Fritz, High yield expression of catalytically active USP18 (UBP43) using a Trigger Factor fusion system, BMC Biotechnol. 12 (2012) 56. doi:10.1186/1472-6750-12-56.

[35] O. Kandror, a L. Goldberg, Trigger factor is induced upon cold shock and enhances viability of *Escherichia coli* at low temperatures., Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 4978–4981. doi:10.1073/pnas.94.10.4978.

[36] R. Glockshuber, T. Schmidt, A. Plueckthun, The disulfide bonds in antibody variable domains: effects on stability, folding in vitro, and functional expression in Escherichia coli, Biochemistry. 31 (1992) 1270–1279. doi:10.1021/bi00120a002.

[37] A. de Marco, Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli., Microb. Cell Fact. 8 (2009) 26. doi:10.1186/1475-2859-8-26.

[38] I. Kumagai, R. Asano, T. Nakanishi, K. Hashikami, S. Tanaka, A. Badran, H. Sanada, M. Umetsu, Integration of PEGylation and refolding for renaturation of recombinant proteins from insoluble aggregates produced in bacteria-Application to a single-chain Fv fragment, J. Biosci. Bioeng. 109 (2010) 447–452. doi:10.1016/j.jbiosc.2009.10.016.

[39] K. Proba, A. Wörn, A. Honegger, A. Plückthun, Antibody scFv fragments without disulfide bonds, made by molecular evolution, J. Mol. Biol. 275 (1998) 245–253. doi:10.1006/jmbi.1997.1457.

[40] M. Berkmen, Production of disulfide-bonded proteins in Escherichia coli, Protein Expr. Purif. 82 (2012) 240–251. doi:10.1016/j.pep.2011.10.009.

[41] H. Sonoda, Y. Kumada, T. Katsuda, H. Yamaji, Functional expression of single-chain Fv antibody in the cytoplasm of Escherichia coli by thioredoxin fusion and co-expression of molecular chaperones, Protein Expr. Purif. 70 (2010) 248–253. doi:10.1016/j.pep.2009.11.003.

[42] J.G. Thomas, A. Ayling, F. Baneyx, Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from E. coli - To fold or to refold, Appl. Biochem. Biotechnol. 66 (1997) 197–238. doi:10.1007/BF02785589.

[43] U.K. LAEMMLI, Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4, Nature. 227 (1970) 680–685. doi:10.1038/227680a0.

[44] G. Fairbanks, T.L. Steck, D.F.H. Wallach, Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane., Biochemistry. 10 (1971) 2606–2617. doi:10.1021/bi00789a030.

[45] Y. Zhang, I-TASSER server for protein 3D structure prediction., BMC Bioinformatics. 9 (2008) 40. doi:10.1186/1471-2105-9-40.

[46] A. Roy, A. Kucukural, Y. Zhang, I-TASSER: a unified platform for automated protein structure and function prediction., Nat. Protoc. 5 (2010) 725–38. doi:10.1038/nprot.2010.5.

[47] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The I-TASSER Suite: protein structure and function prediction, Nat. Methods. 12 (2014) 7–8. doi:10.1038/nmeth.3213.

[48] S. Altschul, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402. doi:10.1093/nar/25.17.3389.

[49] D.T. Jones, Protein secondary structure prediction based on position-specific scoring matrices, J. Mol. Biol. 292 (1999) 195–202. doi:10.1006/jmbi.1999.3091.

[50] S. Wu, Y. Zhang, LOMETS: A local meta-threading-server for protein structure prediction, Nucleic Acids Res. 35 (2007) 3375–3382. doi:10.1093/nar/gkm251.

[51] Y. Zhang, J. Skolnick, <I>SPICKER</I>: A clustering approach to identify near-native protein folds, 25 (2004) 865–871 %U http://dx.doi.org/10.1002/jcc.20011.

[52] Y. Li, Y. Zhang, REMO: A new protocol to refine full atomic protein models from C-alpha traces by optimizing hydrogen-bonding networks, Proteins Struct. Funct. Bioinforma. 76 (2009) 665–676. doi:10.1002/prot.22380.

[53] E. Krieger, G. Vriend, YASARA View--molecular graphics for all devices--from smartphones to workstations, Bioinformatics. 30 (2014) 2981–2982. doi:10.1093/bioinformatics/btu426.

[54] L. Ferbitz, T. Maier, H. Patzelt, B. Bukau, E. Deuerling, N. Ban, Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins, Nature. 431 (2004) 590–596. doi:10.1038/nature02899.

[55] E. Martinez-Hackert, W.A. Hendrickson, Promiscuous Substrate Recognition in Folding and Assembly Activities of the Trigger Factor Chaperone, Cell. 138 (2009) 923–934. doi:10.1016/j.cell.2009.07.044.

[56] X. Chen, H. Pu, Y. Fang, X. Wang, S. Zhao, Y. Lin, M. Zhang, H.-E. Dai, W. Gong, L. Liu, Crystal structure of the catalytic subunit of magnesium chelatase, Nat. Plants. 1 (2015) 15125. doi:10.1038/nplants.2015.125.

[57] B. Liang, Z. Li, S. Jenni, A.A. Rahmeh, B.M. Morin, T. Grant, N. Grigorieff, S.C. Harrison, S.P.J. Whelan, Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy, Cell. 162 (2015) 314–327. doi:10.1016/j.cell.2015.06.018.

[58] K.S. Murakami, E.K. Davydova, L.B. Rothman-Denes, X-ray crystal structure of the polymerase domain of the bacteriophage N4 virion RNA polymerase, Proc. Natl. Acad. Sci. 105 (2008) 5046–5051. doi:10.1073/pnas.0712325105.

[59] J. Yang, Y. Wang, Y. Zhang, ResQ: An Approach to Unified Estimation of B-Factor and Residue-Specific Error in Protein Structure Prediction, J. Mol. Biol. 428 (2016) 693–701. doi:10.1016/j.jmb.2015.09.024.

[60] M. Gamerdinger, E. Deuerling, Trigger Factor Flexibility, Science (80-.). 344 (2014) 590–591. doi:10.1126/science.1254064.

[61] A. Hoffmann, B. Bukau, G. Kramer, Structure and function of the molecular chaperone Trigger Factor., Biochim. Biophys. Acta. 1803 (2010) 650–61. doi:10.1016/j.bbamcr.2010.01.017.

[62] D. Esposito, D.K. Chatterjee, Enhancement of soluble protein expression through the use of fusion tags, Curr. Opin. Biotechnol. 17 (2006) 353–358. doi:10.1016/j.copbio.2006.06.003.

[63] H.P. Sørensen, K.K. Mortensen, Advanced genetic strategies for recombinant protein expression in Escherichia coli, J. Biotechnol. 115 (2005) 113–128. doi:10.1016/j.jbiotec.2004.08.004.

[64] H. Sonoda, Y. Kumada, T. Katsuda, H. Yamaji, Effects of cytoplasmic and periplasmic chaperones on secretory production of single-chain Fv antibody in Escherichia coli, J. Biosci. Bioeng. 111 (2011) 465–470. doi:10.1016/j.jbiosc.2010.12.015.

[65] J.X. Zhao, L. Yang, Z.N. Gu, H.Q. Chen, F.W. Tian, Y.Q. Chen, H. Zhang, W. Chen, Stabilization of the single-chain fragment variable by an interdomain disulfide bond and its effect on antibody affinity, Int. J. Mol. Sci. 12 (2011) 1–11. doi:10.3390/ijms12010001.

[66] L. Zheng, U. Baumann, J.L. Reymond, Production of a functional catalytic antibody ScFv-NusA fusion protein in bacterial cytoplasm, J. Biochem. 133 (2003) 577–581. doi:10.1093/jb/mvg074.

[67] H. Xiong, S. Li, Z. Yang, R.R. Burgess, W.S. Dynan, E. coli expression of a soluble, active single-chain antibody variable fragment containing a nuclear localization signal, Protein Expr. Purif. 66 (2009) 172–180. doi:10.1016/j.pep.2009.03.002.

[68] K. Yamanaka, Cold shock response in Escherichia coli., J. Mol. Microbiol. Biotechnol. 1 (1999) 193–202.

[69] M. Fathi-Roudsari, A. Akhavian-Tehrani, N. Maghsoudi, Comparison of Three Escherichia coli Strains in Recombinant Production of Reteplase., Avicenna J. Med. Biotechnol. 8 (2016) 16–22. http://www.ncbi.nlm.nih.gov/pubmed/26855731.

[70] K. Zarschler, S. Witecy, F. Kapplusch, C. Foerster, H. Stephan, High-yield production of functional soluble single-domain antibodies in the cytoplasm of Escherichia coli, Microb. Cell Fact. 12 (2013) 97. doi:10.1186/1475-2859-12-97.

[71] S.C. Makrides, Strategies for achieving high-level expression of genes in Escherichia coli., Microbiol. Rev. 60 (1996) 512–538. http://www.ncbi.nlm.nih.gov/pubmed/8840785.

[72] S. Costa, A. Almeida, A. Castro, L. Domingues, Fusion tags for protein solubility, purification, and

immunogenicity in Escherichia coli: The novel Fh8 system, Front. Microbiol. 5 (2014) 1–20. doi:10.3389/fmicb.2014.00063.

[73] G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in Escherichia coli: advances and challenges, Front. Microbiol. 5 (2014) 1–17. doi:10.3389/fmicb.2014.00172.

[74] F.W. Studier, Protein production by auto-induction in high-density shaking cultures, Protein Expr. Purif. 41 (2005) 207–234. doi:10.1016/j.pep.2005.01.016.

Chapter 6: General discussion and conclusion

List of abbreviations

Circular Dichroism	CD
Complementarity Determining regions	CDR
Deep Vein Thrombosis	DVT
Global Burden of Diseases	GBD
Human Immunodeficiency Virus	HIV-AIDS
Human Single Chain Antibody Fragment	scFv
Molecular Dynamic	MD
Prothrombin Times	PT
Tissue Factor Inhibitor scFv	TFI-scFv
Tissue Factor: Factor VIIa Complex	TF:FVIIa
Venous Thromboembolism	VTE

Chapter 6: General discussion and conclusion

6.1 Initial characterisation of TFI-scFv

Thrombosis and thrombosis related disorders are responsible for millions of deaths worldwide each year. The Global Burden of Diseases, Injuries, and Risk Factors (GBD) Study 2010 attributed one in four deaths to ischemic heart disease and stroke collectively [1]. Thrombosis includes disorders such as during deep vein thrombosis (DVT), venous thromboembolism (VTE), sepsis, ischemic heart disease, ischemic stroke, and atherosclerosis. The list of thrombotic-related disorders include diseases such as obesity, diabetes, cancer and HIV-AIDS [2–10]. 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 [11]. Tissue factor functions as the initiator of blood coagulation plays a key role in the pathology of various thrombosis disorders. Consequently, the development of safe and effective anti-tissue factor agents has great potential to supplement and replace current commercial anticoagulants. Thus, a human single chain antibody fragment (scFv) was isolated from the Tomlinson I + J Human Single Fold Phage Libraries that binds to and inhibit the actions of tissue factor. Initial characterisation of Tissue Factor Inhibitor scFv (TFI-scFv) have shown a dose dependent extension of prothrombin times (PT) and a reduction in peak thrombin generation [12]. As described in chapter 1, the initial findings were promising but the characterisation TFI-scFv was hampered by low protein yields. It was therefore primary the goal of this study to develop methods for the improved production of TFI-scFv at sufficient levels to allow the continued characterisation. The secondary goal was to identify the mechanism involved in the inhibition of the coagulation cascade.

6.2 Production of functional TFI-scFv

The production of scFv in *Escherichia coli* is notoriously difficult [24]. Heterologous protein expression of recombinant proteins is a process that involves multiple synthetic pathways and is regulated at both transcriptional and translational levels [25]. Differences in codon usage between species and the presence of rare codons can have detrimental effects on scFv production in *E.coli* [26]. In addition, one of the main problems regarding recombinant protein production is that heterologous proteins often result in the formation of insoluble protein aggregates known as inclusion bodies [25]. In other words, the main challenge in recombinant protein expression is to find the balance between quantity and quality of target protein. Faced with these obstacles, the production of functional TFI-scFv was achieved through two methods: i) *in vitro* protein refolding of inclusion bodies described in chapter 4 ii) the use of specialised cold shock expression systems described in chapter 5. It is however important to provide a clear explanation of the logical progression of expression strategies that lead the subsequent production of functional TFI-scFv.

When this project was started, the Tomlinson expression system was utilised in an upscale setting to produce as much as scFv as possible given the available equipment. The expression method was modified through the

addition of osmotic shock steps to improve the yield of TFI-scFv isolated from the periplasm. This however was not especially effective as the increase in antibody yield was minimal. The problem was that the Tomlinson system is primarily designed for the effective identification and isolation of desired phagemids. It is not designed to be use as an over-expression system for the large-scale production of the soluble scFv. The original unoptimized gene was then expressed using commercial pET22 expression vectors in *E.coli BL21*. The expression system was modified to redirect the expressed antibody to the cytoplasm, rather than the periplasm, with the hope to improve antibody retention. This unfortunately did not lead to the improved production of the antibody. It did however prompt the initial investigation to the presence of rare codons. A brief scan of the unoptimized gene sequence revealed the presence of rare codons. Consequently, the unoptimized gene was then co-expressed with pRARE plasmid. The pRARE plasmid harbours a set of tRNA genes that are co-expressed to compensate for the presence of rare codons. This also did not result in any improvement in TFI-scFv yield. Rare codon analysis of the original unoptimized gene identified at total of 46 rare (approximately 20 %) codons distributed throughout TFI-scFv. The low expression levels were attributed to the excessive presence of rare codons throughout the gene. It was at this stage that the TFI-scFv gene was codon optimised for expression in *E.coli*. Expression of the codon optimised gene resulted significant increase (80-fold increase) in protein expression levels but did result in the production of inclusion bodies. Initial attempts to alleviate the production of the inclusion bodies included expression of TFI-scFv at lower temperatures ranges (30 °C, 25 °C and 20 °C), the co-expression of TAKARA chaperones in *E.coli -Bl21* and *Origami* strains, the combined effect of the co-expression of TAKARA chaperones at lower temperatures, and the use of auto-induction medias over a wide range of incubation temperatures (37°C – 20 °C). Although some of these strategies did result in a minor improvement in solubilisation levels of TFI-scFv, they were not capable of producing functional antibody.

At this point in the study it became clear that solubilisation of the antibody, with focus on the formation disulfide bonds, would be the key to the production of functional TFI-scFv. Thus, the *in vitro* refolding of TFI-scFv from the inclusion bodies was the next logical step. As discussed in chapter 3, the challenge was to convert the inactive and insoluble scFv aggregates into soluble, correctly folded biologically active products [27]. In reality, the generation of biologically active antibodies from insoluble aggregates has become common practice when working with scFvs [28]. The expression of scFv as inclusion bodies in *E.coli* is such a common occurrence, that various methods have been proposed for the generation of scFv [29]. These methods vary in efficacy as well as complexity but are generally divided into i) Dilution methods ii) Solvent exchange methods, and iii) Solid state support or on-column refolding methods [27,30,31]. For the screening of the refolding conditions suitable for generation of functional TFI-scFv, the dilution method was used due to low cost, availability of equipment, and relative ease of use in comparison to the more complex methods. During a single round of expression, approximately 55 $mg.L^{-1}$ denatured TFI-scFv was isolated, but batch refolding procedures were restricted to only 500 ug.mL because of the screening reactions' limitations. Of the nine refolding conditions utilised during the screening only one (buffer 6) yielded functional TFI-scFv. The top performing buffer had refolding efficiency of approximately 26 % of the denatured inclusion bodies. It was found that the redox protentional had a significant effect on the production of

functional antibody due to the presence of intra-chain disulfide bonds. [32,33]. Although the refolding techniques utilised here were relatively simple in comparison to other advanced methods, it did require intensive optimization of multiple operational steps. A high level of accuracy with regards to protein concentration, buffer composition and temperature control was required in order to obtain functional protein. In short, the regeneration of functional TFI-scFv from insoluble aggregates was possible but the refolding process is an intricate and labour intensive method.

Due to the drawbacks of *in vitro* refolding, the feasibility of producing a functional TFI-scFv in the cytoplasm of *E.coli* was explored as an alternative approach in chapter 5. Solubilisation of the scFv was achieved though commercial cold shock expression vectors (pCOLD DNA II and pCOLD Trigger factor) in conjunction with *E.coli BL21* (DE3) and specialised disulfide bond promoting strain (*E.coli SHuffle)*. The level of TFI-scFv solubilisation ranged from 20 % for the non-fusion constructs up to 40 % for the trigger factor-TFI-scFv fusion constructs. Molecular modelling suggests that the higher level of solubilisation was due to nestled position of TFI-scFv within the fusion construct. Although the solubilisation of the fusion constructs were higher, the TFI-scFv was not biologically active once separated from the fusion partner. Functional antibody fragments were produced through a combination of cold shock, pCOLD DNA II expression vector in conjunction with *E.coli SHuffle*. The specific mutation to *E.coli SHuffle* which facilitated the formation of disulfide bonds in the cytoplasm seemed to be the key to producing functional TFI-scFv. The cold shock expression systems thus provide a simpler, less time consuming, and more cost-effective approach to produce functional TFI-scFv.

When comparing the two expression strategies, the TFI-scFv produced by cold shock had a specific activity of 83.8 s/mg.ml^{-1} opposed to 29.63 s/mg.mL^{-}1 of the refolded TFI-scFv. This difference suggests that although functional scFv was produced by *in vitro* refolding, much of the refolded fraction was soluble but inactive. One of the major challenges of working with a novel antibody is the fact that there is no commercially available counterpart to serve as reference point. As an example, it is impossible to confirm if the specific activities reported here are at maximum. By comparing the two methods, deduction can be made as to which production strategy is more efficient but if these strategies are fully optimised remains unknown. As a second, example in terms of the TFI-scFv structure during refolding, comparative circular dichroism (CD) spectra between the reference and that of the refolded fraction would have indicated whether the scFv was in one, two or multiple conformations [34]. This would have provided much needed insight into the effectiveness of the *in vitro* refolding reactions. It is complications such as these that makes working with novel heterologous proteins a complex and challenging task. Nevertheless, functional TFI-scFv was finally expressed in the cytoplasm of *E.coli* at sufficient levels enable the further characterisation of the antibody.

6.3 Mechanism of inhibition

The TFI-scFv was isolated from single fold phage libraries based on its ability to inhibit the actions of tissue factor as confirmed by tissue factor activity-, prothrombin times, and eventually thrombin generation assays [12]. Although

it was clear that TFI-scFv inhibited blood coagulation by associating with tissue factor, the mechanism involved in the inhibition was unknown. Upon vessel damage, constitutively expressed tissue factor is exposed to circulating zymogen factor VII. The association between tissue factor and factor VII results in the formation of the tissue factor: - factor VIIa complex (TF-FVIIa Complex). This association results in the allosterical activation of factor VII. The TF-FVIIa complex in turn is responsible for the activations of factor X [13]. As both coagulation factors VII and X associate with tissue factor during the initiation of coagulation, it was unclear whether TFI-scFv interfere with the association of factor VII, factor X or both.

Homology modelling provides a powerful *in silico* method for the prediction of protein structure based on the databases of proteins with known X-ray crystallography or NMR structures [14,15]. In chapter 3, the structure of TFI-scFv was modelled and the position and orientation of the Complementarity Determining Regions (CDR) identified. The CDRs in both chains of the scFv are responsible for the specificities of the antibody [16,17]. The CDR regions were utilised to direct molecular docking between the homology model of TFI-scFv and the crystal structure of the extracellular domains of tissue factor. Although tissue factor is a transmembrane glycoprotein, no crystal structure of the full length tissue factor is available as membrane proteins are notoriously difficult to crystallise [18]. However, the published structure of the extracellular domain of tissue factor has been widely utilised to study the interactions between tissue factor and factors VII and X [19–22]. The rigid-body docking model was refined by means of molecular dynamic (MD) simulations and the most prevalent cluster was identified. MD simulations predicted continuous interaction between TFI-scFv and tissue factor for duration of the 100 nanoseconds simulation. According to the refined docking model, TFI-scFv is positioned in such a manner that it sterically interferes with the association between tissue actor as factor VII. It was therefore concluded that TFI-scFv functions as an anti-thrombotic, by preventing the formation of the TF-FVIIa complex during the initiation stage of coagulation.

It is however important to be mindful of the fact that the interactions predicted here are based on assumptions regarding the homology model of TFI-scFv, position and orientation of the CCR, and the interactions of these regions with tissue factor. Nevertheless, it has been found that high confidence homology models of antibodies have been generated due to the large collection of antibody crystal templates and the highly conserved nature of antibody structure [23]. It is therefore reasonable to assume that the structure and inhibition mechanism predicted by *in silico* modelling provide an accurate representation of the physiological state. This work represented here is the first predict the inhibition mechanism in the association between TFI-scFv and tissue factor. It provides a good basis for the future characterisation of the biochemical interaction involved in the inhibition mechanism.

6.4 Future research

The production of functional TFI-scFv in the cytoplasm of *E.coli* finally allows the further in depth characterisation of the antibody structure and inhibitor function. The expression systems utilised for the production of TFI-scFv can

also be further optimised. Zarschler *et.al* (2013) reported on the improved biomass production of specifically *E.coli SHuffle* when incubated using Terrific broth and EnPresso medium. The use of these media results in a 5 to 10 fold increase in biomass production respectively compared to the standard Lysogeny Broth (LB media) [35]. The use of these medias would be a simple alteration to the current expression method and shows great potential for the further improved of TFI-scFv production. If these modification to the expression systems were to substantially improve the yield TFI-scFv in would pave the way to produce TFI-scFv on an industrial-scale.

The improved production also make the study of the TFI-scFv structure through X-RAY crystallography possible. Currently sufficient TFI-scFv can be produce to perform the intense screening process involved in the generation protein crystals. Once a TFI-scFv structure is resolved it would be interesting to compare it to the predicted homology model. A comparison of the structures would provide much insight into the accuracy of the TFI-scFv homology model. It would also be interesting to generate a crystal structure of TFI-scFv bound to tissue factor. Although a difficult undertaking, this would provide the information to confirm, improve or disprove the molecular dynamic simulations of the inhibition mechanism.

Perhaps the most interesting would be the in depth *in vitro* and *in vitro* characterisation of TFi-scFv inhibitory effect. The manner in which TFI-scFv performed in prothrombin times, thromboelastography (TEG), and thrombin generations would provide information regarding the effectiveness of the antibody. The effectiveness of TFI-scFv as an antithrombotic can also be tested in various diseases such as diabetes, hypertension and protein S deficiencies. In addition, the antithrombotic capability of TFI-scFv can also be tested in animal thrombosis models using mice and baboons.

Ultimately, the improved expression of functional TFI-scFv opens several different pathways for the future characterisation and study of the structure and interaction of TFI-scFv with tissue factor.

6.4 References

[1] G.E. Raskob, P. Angchaisuksiri, A.N. Blanco, H. Buller, A. Gallus, B.J. Hunt, E.M. Hylek, A. Kakkar, S. V. Konstantinides, M. McCumber, Y. Ozaki, A. Wendelboe, J.I. Weitz, Thrombosis: A Major Contributor to Global Disease Burden, Arterioscler. Thromb. Vasc. Biol. 34 (2014) 2363–2371. doi:10.1161/ATVBAHA.114.304488.

[2] G. Granata, T. Izzo, P. Di Micco, B. Bonamassa, G. Castaldo, V. Viggiano, U. Picillo, G. Castaldo, A. Niglio, Thromboembolic events and haematological diseases: a case of stroke as clinical onset of a paroxysmal nocturnal haemoglobinuria, Thromb. J. 2 (2004) 10. doi:10.1186/1477-9560-2-10.

[3] C. Milsom, J. Yu, L. May, B. Meehan, N. Magnus, K. Al-Nedawi, J. Luyendyk, J. Weitz, P. Klement, G. Broze, N. Mackman, J. Rak, The role of tumor-and host-related tissue factor pools in oncogene-driven tumor progression, Thromb. Res. 120 (2007) S82–S91. doi:10.1016/S0049-3848(07)70135-4.

[4] L. Pantanowitz, B.J. Dezube, Monocytes tied to HIV-associated thrombosis, Blood. 115 (2010) 156–157. doi:10.1182/blood-2009-11-250597.

[5] J. Nofer, B. Kehrel, M. Fobker, B. Levkau, G. Assmann, A. von Eckardstein, HDL and arteriosclerosis: beyond reverse cholesterol transport, Atherosclerosis. 161 (2002) 1–16. doi:10.1016/S0021-9150(01)00651-7.

[6] A.K. Kakkar, Venous Thrombosis in Cancer Patients: Insights from the FRONTLINE Survey, Oncologist. 8 (2003) 381–388. doi:10.1634/theoncologist.8-4-381.

[7] M.N. Levine, A.Y. Lee, A.K. Kakkar, From Trousseau to targeted therapy: New insights and innovations in thrombosis and cancer, J. Thromb. Haemost. 1 (2003) 1456–1463. doi:10.1046/j.1538-7836.2003.00275.x.

[8] D. a Manly, J. Boles, N. Mackman, Role of Tissue Factor in Venous Thrombosis, Annu. Rev. Physiol. 73 (2011) 515–525. doi:10.1146/annurev-physiol-042210-121137.

[9] A.I. Schafer, Thrombotic Disorders: Diagnosis and Treatment, Hematology. 2003 (2003) 520–539. doi:10.1182/asheducation-2003.1.520.

[10] K.A. Illig, A.J. Doyle, A comprehensive review of Paget-Schroetter syndrome, J. Vasc. Surg. 51 (2010) 1538–1547. doi:10.1016/j.jvs.2009.12.022.

[11] A.D. Callow, Cardiovascular disease 2005 — the global picture, Vascul. Pharmacol. 45 (2006) 302–307. doi:10.1016/j.vph.2006.08.010.

[12] S.M. Meiring, J. Vermeulen, P.N. Badenhorst, Development of an inhibitory antibody fragment to human tissue factor using phage display technology, Drug Dev. Res. 70 (2009) 199–205. doi:10.1002/ddr.20295.

[13] K. Tsumoto, K. Shinoki, H. Kondo, M. Uchikawa, T. Juji, I. Kumagai, 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, J. Immunol. Methods. 219 (1998) 119–129. doi:10.1016/S0022-1759(98)00127-6.

[14] A. Singh, V. Upadhyay, A.K. Upadhyay, S.M. Singh, A.K. Panda, Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process., Microb. Cell Fact. 14 (2015) 41. doi:10.1186/s12934-015-0222-8.

[15] M. Widmann, M. Clairo, J. Dippon, J. Pleiss, Analysis of the distribution of functionally relevant rare codons., BMC Genomics. 9 (2008) 207. doi:10.1186/1471-2164-9-207.

[16] M. Li, Z.G. Su, J.C. Janson, In vitro protein refolding by chromatographic procedures, Protein Expr. Purif. 33 (2004) 1–10. doi:10.1016/j.pep.2003.08.023.

[17] K. Ikeda, Y. Kumada, T. Katsuda, H. Yamaji, S. Katoh, Refolding of single-chain Fv by use of an antigen-coupled column, Biochem. Eng. J. 44 (2009) 289–291. doi:10.1016/j.bej.2008.12.014.

[18] M. Liu, X. Wang, C. Yin, Z. Zhang, Q. Lin, Y. Zhen, H. Huang, One-step on-column purification and refolding of a single-chain variable fragment (scFv) antibody against tumour necrosis factor alpha., Biotechnol. Appl. Biochem. 43 (2006) 137–45. doi:10.1042/BA20050194.

[19] J.R. Sinacola, A.S. Robinson, Rapid refolding and polishing of single-chain antibodies from Escherichia coli inclusion bodies, Protein Expr. Purif. 26 (2002) 301–308. doi:10.1016/S1046-5928(02)00538-7.

[20] G. Lemercier, N. Bakalara, X. Santarelli, On-column refolding of an insoluble histidine tag recombinant exopolyphosphatase from Trypanosoma brucei overexpressed in Escherichia coli, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 786 (2003) 305–309. doi:10.1016/S1570-0232(02)00745-6.

[21] E.D.B. Clark, Protein refolding for industrial processes, Curr. Opin. Biotechnol. 12 (2001) 202–207. doi:10.1016/S0958-1669(00)00200-7.

[22] T. Arakawa, D. Ejima, Refolding Technologies for Antibody Fragments, Antibodies. 3 (2014) 232–241. doi:10.3390/antib3020232.

[23] K.K. Fursova, A.G. Laman, B.S. Melnik, G. V. Semisotnov, P.K. Kopylov, N. V. Kiseleva, V.A. Nesmeyanov, F.A. Brovko, Refolding of scFv mini-antibodies using size-exclusion chromatography via arginine solution layer, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 877 (2009) 2045–2051. doi:10.1016/j.jchromb.2009.05.038.

[24] A.J. Chu, Tissue factor mediates inflammation, Arch. Biochem. Biophys. 440 (2005) 123–132. doi:10.1016/j.abb.2005.06.005.

[25] Y. Sefidbakht, O. Ranaei Siadat, F. Taheri, Homology Modeling and Molecular Dynamics Study on Schwanniomyces Occidentalis Alpha Amylase., J. Biomol. Struct. Dyn. 1102 (2016) 1–27. doi:10.1080/07391102.2016.1154892.

[26] C.N. Cavasotto, S.S. Phatak, Homology modeling in drug discovery: current trends and applications, Drug Discov. Today. 14 (2009) 676–683. doi:10.1016/j.drudis.2009.04.006.

[27] E.A. Kabat, T. Wut, Identical V regions amino acid sequences and segments of sequences in antibodies of different specificities. Relative Contributions of VH and VL Genes, Minigenes, and Complementarity-Determining Regions to Binding of Antibody-Combining Sites, J Immunol. 147 (1991) 1709–1719. http://www.jimmunol.org/content/147/5/1709.

[28] J. Davies, L. Riechmann, Affinity improvement of single antibody VH domains: residues in all three hypervariable regions affect antigen binding, Immunotechnology. 2 (1996) 169–179. doi:10.1016/S1380-2933(96)00045-0.

[29] E.P. Carpenter, K. Beis, A.D. Cameron, S. Iwata, Overcoming the challenges of membrane protein crystallography, Curr. Opin. Struct. Biol. 18 (2008) 581–586. doi:10.1016/j.sbi.2008.07.001.

[30] K. Harlos, D.M.A. Martin, D.P. O'Brien, E.Y. Jones, D.I. Stuart, I. Polikarpov, A. Miller, E.G.D. Tuddenham, C.W.G. Boys, Crystal structure of the extracellular region of human tissue factor, Nature. 370 (1994) 662–666. doi:10.1038/370662a0.

[31] M. Huang, R. Syed, E. a Stura, M.J. Stone, R.S. Stefanko, W. Ruf, T.S. Edgington, I. a Wilson, The mechanism of an inhibitory antibody on TF-initiated blood coagulation revealed by the crystal structures of human tissue factor, Fab 5G9 and TF.G9 complex., J. Mol. Biol. 275 (1998) 873–94. doi:10.1006/jmbi.1997.1512.

[32] Y. a. Muller, M.H. Ultsch, M. de Vos, Abraham, The crystal structure of the extracellular domain of human tissue factor refined to 1.7 A resolution., J. Mol. Biol. 256 (1996) 144–159. doi:10.1006/jmbi.1996.0073.

[33] D.M. Martin, C.W. Boys, W. Ruf, Tissue factor: molecular recognition and cofactor function., FASEB J. 9 (1995) 852–9. http://www.ncbi.nlm.nih.gov/pubmed/7615155.

[34] K. He, X. Zhang, L. Wang, X. Du, D. Wei, Production of a soluble single-chain variable fragment antibody against okadaic acid and exploration of its speci fi c binding, Anal. Biochem. 503 (2016) 21–27. doi:10.1016/j.ab.2015.12.020.

[35] K. Zarschler, S. Witecy, F. Kapplusch, C. Foerster, H. Stephan, High-yield production of functional soluble single-domain antibodies in the cytoplasm of Escherichia coli, Microb. Cell Fact. 12 (2013) 97. doi:10.1186/1475-2859-12-97.

Chapter 7: Summaries

Abstract (English)

Recognising the potential in tissue factor inhibition as a novel approach to antithrombotic therapy, our laboratory utilized phage display technology to select a human single chain antibody fragment from the Tomlinson I + J Human Single Fold Phage Libraries which functionally inhibits human tissue factor. Although the initial findings were promising the further characterisation of the tissue factor inhibiting scFv was hampered by low proteins yields as well as financial complication associated with the initial purification methods. In this study, the production of functional antibody was improved through the use of *in vitro* refolding- and cold shock expression methods. The protein structure and inhibition mechanism was characterised by means of *in silico* modelling. The improved expression of functional TFI-scFv opens several different pathways for the future characterisation and study of the structure and interaction of TFI-scFv with tissue factor.

Opsomming (Afrikaans)

Ter erkenning in die potensiaal in weefsel faktor inhibisie as 'n nuwe benadering tot anti-trombotiese terapie, het ons laboratorium gebruik gemaak van faagvertooning tegnologie om 'n menslike enkele ketting teenliggaam fragment vanuit die Tomlinson I + J Human single fold library te isoleer wat menslike weefsel faktor inhibeer. Hoewel die aanvanklike bevindings belowend was, was die verdere karakterisering van die antiliggaam bemoeilik deur lae proteïen vlakke asook finansiële komplikasies wat verband hou met die aanvanklike suiwerings metodes. In hierdie studie, is die produksie van funksionele teenliggaam verbeter deur gebruik te maak van in vitro hervouing en koue skok proteïen uitdrukking metodes. Die proteïenstruktuur en inhibisie meganisme is geidentifiseer deur middel van in silico modellering. Die verbeterde uitdrukking van funksionele antiliggaam open verskillende roetes vir die toekomstige karakterisering en studie van die struktuur en interaksie van die antiliggaam met weefsel factor.