Generation of a long acting GCSF

Abdulrahman T Alshehri

Department of Oncology and Metabolism

The University of Sheffield

This thesis is submitted for the degree of

Doctor of Philosophy

Jan 2016

TABLE OF CONTENTS

List of Figures

List of Tables

Declaration

I hereby declare that this thesis has been composed by myself and has not been accepted in any previous application for a higher degree. The work reported in this thesis is novel and has been carried out by myself with all source of information being specifically acknowledged by means of references.

Abdulrahman Alshehri

January 2016

PUBLICATIONS AND PRESENTATIONS

Published Abstracts and Posters:

- School Research Meeting Conference, Sheffield University June 2012

'Generation of a long acting GCSF for treatment of neutropenia and stem cell harvest - Abdulrahman Alshehri, Richard Ross and Ian R Wilkinson'

- Society for Endocrinology BES Meeting Liverpool, March 2014

'Generation of a long acting GCSF for treatment of neutropenia and stem cell harvest - Abdulrahman Alshehri, Richard Ross and Ian R Wilkinson'

- Advanced Biomanufacturing Conference, Sheffield May 2015

'Generation of a long acting GCSF for treatment of neutropenia and stem cell harvest - Abdulrahman Alshehri, Richard Ross and Ian R Wilkinson'

- Society for Endocrinology BES Meeting Edinburgh, November 2015

(Awarded the Best junior poster prize)

'Generation of a long acting GCSF for treatment of neutropenia and stem cell harvest - Abdulrahman Alshehri, Richard Ross and Ian R Wilkinson'

Medical School Presentations and Mellanby Centre Internal Seminars:

First year PhD presentation, July 2012

Third year PhD presentation, July 2014

Mellanby Centre Internal Seminar, December 2012

Mellanby Centre Internal Seminar, December 2013

ACKNOWLEDGMENT

I would like to acknowledge the support and help of many colleagues. In particular, I would like to thank Professor Richard Ross for his continuous support and guidance during this project. His support was not only to develop me as a scientist but also as an individual. I am indebted to Dr Ian Wilkinson who provided excellent supervision and critical guidance throughout this project. Also, I would like to thank him for the difficult time of correcting this thesis. Without him the project would surely not have progressed so successfully.

I would like to thank Dr Hamid Zarkesh for his help doing the rat injection, counting WBCs using an automated coulter counter, blood smears and sample collection. Also, I would like to thank Dr Miguel Debonno for his help to calculate *in vivo* data using Win Non Lin.

I am extremely grateful to Dr Sarbandra Pradhananga, Sue Justice, Hadel Ghaban, Mahmoud Habibullah and Jude Akinwale for all their help and time during laboratory work.

I would like to thank my wife, daughters, mother, father, brothers, sisters and all friends for their endless support and believe in me.

Also, I am grateful to thank the Prince Mohammed Bin Naif (Ministry of Interior) and cultural Bureau of Saudi Arabia for their financial support.

Abstract

Rationale: Current therapies require daily injections of GCSF to treat patients with neutropenia and response to treatment is often unpredictable as GCSF is rapidly cleared. A number of approaches to reducing GCSF clearance have been tried mainly through conjugation with another moiety. The technologies already being employed, include PEGylation, immunoglobulins or albumin to increase the half - life of GCSF. However, although these approaches have reduced clearance the pharmacokinetic profile of GCSF has remained unpredictable.

Aim and Hypothesis: a glycosylated linker between two ligands could delay clearance with out blocking bioactivity.

Methodology: GCSF tandem molecules with linkers containing between 2-8 N-linked glycosylation sites (NAT motif) and their respective controls (Q replaces N in the sequence motif NAT so there is no glycosylation) were cloned, and sequenced. Following expression in CHO cells, expressed protein was quantified by ELISA and analysed by western blot to confirm molecular weights and protein integrity. *In vitro* bioactivity was tested using an AML-193 proliferation assay. IMAC was used to purify the protein. Pharmacokinetic and pharmacodynamic of GCSF tandems were measured in Sprague Dawley rats.

Results: Purified glycosylated tandem molecules showed increased molecular weight according to the number glycosylation of their sites when analysed by SDS-PAGE. All GCSF tandems showed increased *in vitro* bioactivity in comparison to rhGCSF. GCSF2NAT, GCSF4NAT and GCSF8NAT containing 2, 4 & 8 glycosylation sites respectively and GCSF8QAT displayed a three-fold increased terminal half-life compared to that published for GCSF, however there was no difference in serum half-life according to the level of glycosylation. Both GCSF2NAT and GCSF4NAT showed a higher increase in the percentage of neutrophils over controls at 12 hrs post injection only. In contrast, GCSF8NAT exhibited a higher increase in neutrophil levels over controls at 48 hrs.

Conclusion: Using glycosylated linkers in GCSF tandems results in molecules

with increased molecular weight according to the number of glycosylation sites. Tandems of GCSF have increased *in vitro* bioactivity compared to monomeric GCSF. Tandems with and without glycosylation had three-fold greater half-lives than rhGCSF. There was evidence that GCSF8NAT was biologically active *in vivo*. The results confirm the hypothesis that it is possible to predictably increase the molecular weight of GCSF tandems and retain biological activity but this was not associated with a predictable prolongation of the serum half-life.

1. Introduction

1.1 History of Granulocyte Colony Stimulating Factor

Prior to the 1960s, several studies on animal models had been performed to find the answer to how white blood cell (WBC) homeostasis can be regulated in the circulation. The specific regulator remained unknown until 1966, when two groups simultaneously performed a method for growing colonies of monocytes and granulocytes from spleen cells, and bone marrow (BM) in semi sold cultures *in vitro* (Bradley and Metcalf, 1966, Ichikawa et al., 1966). However, the growth of these colonies was dependent on the presence of unknown proteins that were given the name of colony stimulating factors (CSFs). Since the middle of the 1980s, efforts have been made by several laboratories to identify and purify these CSF proteins. These efforts revealed that there are four CSF proteins with different activities. They were named dependent on the type of colony of cells they stimulated: M-CSF stimulated macrophage colonies; GM-CSF stimulated both granulocyte and macrophage colonies; G-CSF stimulated granulocyte colony formation and multi-CSF (known as interleukin 3, IL3) stimulated multiple of hematopoietic cell colonies (Metcalf, 2010).

In 1983, Murine GCSF was first purified from mouse lung-conditioned medium by Nicola and colleagues in Melbourne, Australia (Nicola et al., 1983), while Human GCSF (hGCSF) was first purified from the human bladder carcinoma cell line 5637 in 1984 (Welte et al., 1985). The molecular cloning of complementary deoxyribonucleic acid (cDNA) for GCSF and the first expression from E. coli were attained by Souza and Boone in 1986 (Souza et al., 1986).

1.2 GCSF Structure

It has been reported that the hGCSF is encoded by a gene located on chromosome 17 (known as CSF3 gene) and due to differential splicing of GCSF, this gene encodes two different messenger ribonucleic acids (mRNA) products: isoform A contains 177 amino acids (18.8kD) and isoform B contains 174 amino acids (19.6kD). The difference between these isoforms is that isoform A contains an additional three residues (Valine-Serine-Glycine) added after Leucine35. The short isoform B (174 amino acids) contains a glycosylation site on the oxygen (O-linked glycosylation) attached to one threonine at residue 133 as the form expressed in mammalian cells. The B isoform obtains high stability and biological activity and therefore is the source for commercial pharmaceutical products of GCSF (Aapro et al., 2011).

The central structure of hGCSF contains four antiparallel, left-handed α-helical bundles in a form that two helices (A with 29 amino acids & B with 21 amino acids) extend up and two helices (C with 24 amino acids & D with 30 amino acids) extend down (Figure 1.1) (Arvedson and Giffin, 2012). Additionally, hGCSF contains five cysteine residues; Four of these cysteines form two internal disulfide bonds at positions Cys36– Cys42 and Cys64– Cys74, thus leaving one free cysteine residue at position Cys17 with a free sulfhydryl group (Werner et al., 1994).

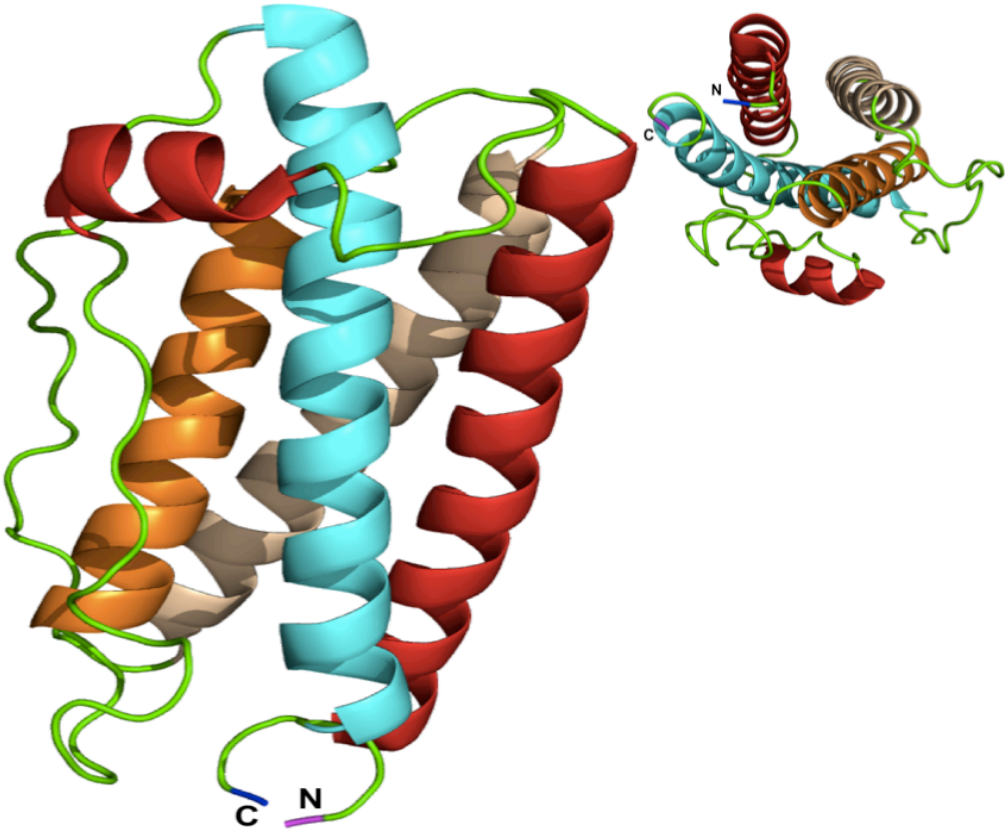


Figure 1-1: Human GCSF structure

The molecular structure of hGCSF contains four antiparallel, left-handed α-helical bundles in a form that two helices A (Red) & B (Orange) extend up and two helices C (White) & D (Cyne) extending down. C= carboxyl-terminus and N= amine-terminus (Using PyMOL molecular graphics system).

1.3 GCSF Expression and Action

Human GCSF is a glycoprotein that regulates the proliferation, differentiation and functional activation of granulopoiesis (Cox et al., 2014) since proved by the significant decrease of neutrophils in both GCSF and GCSF-R deficient mice (Lieschke et al., 1994, Liu et al., 1996). In response to several inflammatory factors such as, interleukin β (IL-1β) necrosis factor alpha (TNF-α) and lipopolysaccharide (LPS), GCSF can be produced by a variety of cells, including endothelial cells, fibroblasts, macrophages, monocytes, and bone marrow stromal cells. GCSF has recently been shown to be highly expressed on a number of cancer cell types including human gastric and colon cancers (Gascon, 2012, Morris et al., 2014), as well as acute myeloid leukemia (AML) and other carcinoma cells (Beekman et al., 2012).

In healthy individuals, GCSF is present at low levels and is rapidly increased in severe cases such as, infection (20 times increase) (Cheers et al., 1988, Kawakami et al., 1990). Therefore, the physiological role of GCSF in the body is to maintain the production of neutrophils during steady state situations and increase neutrophil production during severe inflammatory conditions such as, infection (Hartung et al., 1999).

Several mechanisms showed that GCSF can enhance and regulate the production of neutrophils from the bone marrow to the blood circulation. It enhances the proliferation of all granulocytic lineages from myloblast (stem cell) to mylocyte. It also drives neutrophil differentiation and accelerates the maturation of metamyelocytes. As results of these functions, GCSF shows rapid and continuous elevation in the number of neutrophils (Lord et al., 1991, Basu et al., 2002).

1.4 Regulation of GCSF Expression

During infection, several inflammatory factors in the extracellular microenvironment are elevated, such as IL-1β, TNF-α and LPS and thereafter act on target cells to stimulate GCSF expression by intracellular signaling transcriptional factors, such as NF-κB and C/EBP β. The GCSF promoter region contains binding sites for these factors, which in turn stimulate the GCSF production (Figure 1.2 right). The circulatory levels of GCSF enhance the production and mobilization of neutrophils from the bone marrow to the blood circulation (Panopoulos and Watowich, 2008).

In addition to this pathway, recent studies show that IL-23 produced by dendritic cells and macrophages, induces T helper 17 (Th17) to synthesis IL-17. IL-17 then drives the production of GCSF from cells contained in the stroma, such as endothelial cells, epithelial and fibroblasts through the IL-17 receptor (IL-17R) (Fossiez et al., 1996, Langrish et al., 2005). In response to bacterial pneumonia infection, deficiency of IL-17R resulted in decreased levels of GCSF and delayed neutrophils production, indicating the important role of IL-17-induced granulopoiesis *in vivo* (Figure 1.2 left) (Ye et al., 2001, Nguyen-Jackson et al., 2010).

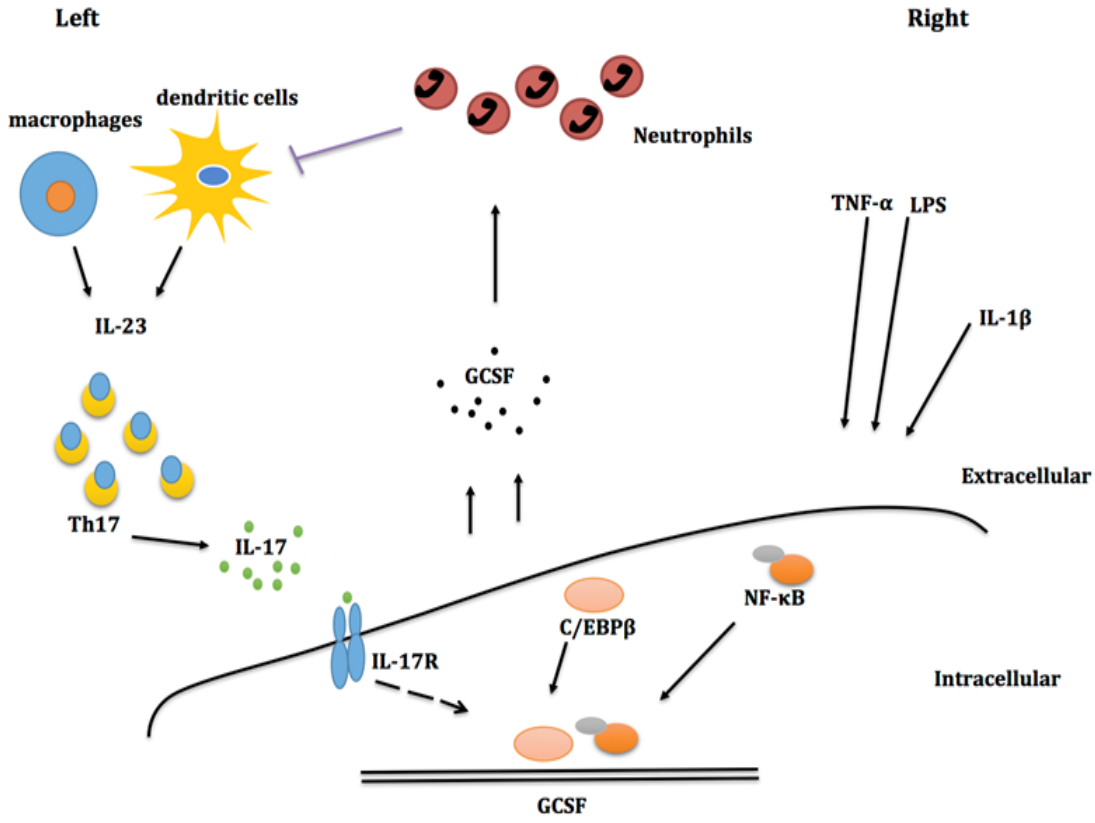


Figure 1-2: Regulation pathways of GCSF expression and production of neutrophils

1.5 GCSF-Receptor

1.5.1 Discovery, Expression and Cloning

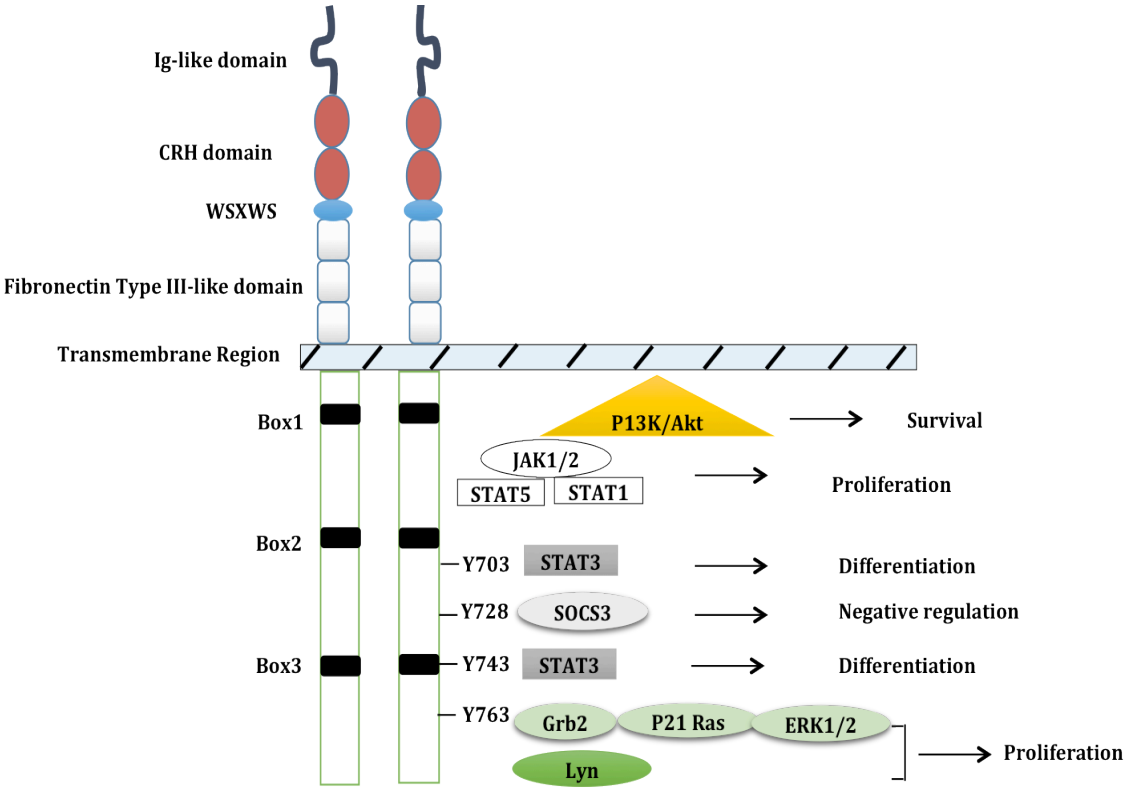
The biological action of GCSF is mediated by binding to its receptor (GCSF-R). Thus, the regulation, proliferation and differentiation of neutrophilic granulocytes are highly dependent on this binding (Gascon, 2012). GCSF-R was discovered as a membrane protein expressed in all granulocytic lineage cells, including neutrophils and their precursors, and myeloid leukemia cells (Nicola and Metcalf, 1984). Later, GCSF-R was detected on normal B & T lymphocytes, monocytes (Boneberg et al., 2000, Morikawa et al., 2002) and non- hematopoietic tissues, such as, cardiomyocytes (Harada et al., 2005), vascular endothelial cells (Bussolino et al., 1989), neural stem cells (Schneider et al., 2005), placenta (McCracken et al., 1996), many non-haematopoietic tumours cell lines (Roberts, 2005) and has recently been shown to be highly expressed on human gastric and colon cancers (Morris et al., 2014). However, GCSF-R is predominantly expressed on stem cells, common myeloid progenitors (CMP) and mature neutrophils, where by the expression of this receptor increases during maturation (Manz et al., 2002).

In 1990, granulocyte colony stimulating factor receptor (GCSF-R) was first cloned from mouse myeloid leukemia cell line (NFS-60) and shown to form homo-dimers upon binding to its ligand (GCSF), resulting in a complex 2:2 ligand receptor subunit (Fukunaga et al., 1990a).

1.5.2 Structure and Function of GCSF-R

The human GCSF receptor is a 120kDa cell surface receptor, which belongs to the hematopoietic cytokine receptor super-family, HCR. The GCSF-R is 836 amino acids in length and consists of an extracellular region with 604 amino acids, transmembrane region with 26 amino acids and a cytoplasmic (intracellular) region with 183 amino acids. The extracellular domains consist of the cytokine receptor homology (CRH) domain, N-terminal immunoglobulin (Ig)-like domain, and a Trp-Ser-X-Trp-Ser (WSXWS) motif required for ligand binding, and the rest of the extracellular region is formed by 3 fibronectin type III (FNIII) domains (Molineux et al., 2012).

The intracellular region contains 2 conserved sub-domains termed Box 1 and Box 2, and a membrane-distal domain that includes Box 3 (less conserved sequence) (Fukunaga et al., 1990b). The intracellular region of GCSF-R has also 4 tyrosine residues at locations 704, 729, 744 and 764 of the human receptor (corresponding to Y703, Y728, Y743, and Y763 in the murine receptor); these conserved residues play an important role in the induction of GCSF cell survival, proliferation and differentiation (Figure 1.3) (Molineux et al., 2012).



.

Figure 1-3: Structure and downstream signal pathways of GCSF-R

The GCSF-R contains of extracellular region and intracellular region. The extracellular region of the GCSF-R includes 3 fibronectin type III (FNIII)-like domains, WSXWS motif, a cytokine receptor homologous (CRH) domain, and immunoglobulin (Ig)-like domain. Conserved Box 1, Box 2, and Box 3 and 4 tyrosines (Y703, Y728, Y743 and Y763) mediate downstream signal transduction in the intracellular region. Binding of GCSF to its receptor induces phosphorylation of JAKs, resulting in phosphorylation of the 4 tyrosine residues located on the cytoplasmic region. Once these tyrosine residues are phosphorylated, they serve as docking sites for numerous proteins characterized by Src Homology 2 (SH2) domains resulting in activation of many signaling pathways that regulate different cell processes.

The binding of GCSF to its receptor forms homo-dimers, resulting in a complex of two GCSF molecules and two GCSF-R molecules (Larsen et al., 1990). Each GCSF interacts with the immunoglobulin (Ig)-like domain of one GCSF-R subunit and the CRH domain of the second GCSF-R subunit, resulting in a crossover configuration of the receptor subunits (Figure 1.4.B) (Tamada et al., 2006).

The binding of GCSF to its receptor activates the Janus kinase (Jak)/signal transducer and activator of transcription (STAT) signalling pathways. Molecules that get activated as part of the pathway include Jak1, Jak2, STAT1, STAT3, and STAT5. It has been reported that tryptophan residues localized between Box 1 and Box 2 in the intracellular region of GCSF-R serves as a docking site for Jaks. Because GCSF-R lacks intrinsic kinase activity; thus, it mostly relies on different non-receptor kinases, for instance, activation of JAK family, mainly through JAK 1 and 2 (Meshkibaf, 2015). Upon activation, the GCSF-R dimerizes and brings the Jaks together into proximity, resulting in their trans-phosphorylation of one another which in turn phosphorylate tyrosine (Y) resides (Y703, Y728, Y743, and Y763) located in the cytoplasmic region. Once these tyrosine residues are phosphorylated, they serve as docking sites for STAT's. STAT's are transcriptional factors found in the intracellular region (cytoplasmic region), and can interact with phosphotyrosine residues of the GCSF-R via their Src Homology 2 (SH2) domains (the function of this domain is to identify the phosphorylated state of tyrosine (Y) residues). STAT's get phosphorylated then form dimers and migrate to the nucleus, where they bind DNA and activate transcription. Although there are seven family members of STATs, activation of low level of GCSF results in phosphorylation of Y703 and Y743, resulting in strong stimulation of STAT3 with slight stimulation of STAT1 and STAT5 (Figure 1.4). *In vitro*, STAT3 activation seems to push neutrophil differentiation mediated by activation of neutrophil marker genes. Activation of GCSF-R is also appeared to activate STAT1 and STAT5, resulting in cell proliferation. In addition, it has been reported that Y728 is a docking site for Suppressor of cytokine signalling 3 (SOCS3), which is a critical feedback inhibitor of GCSF-R signalling pathway (Molineux et al., 2012).

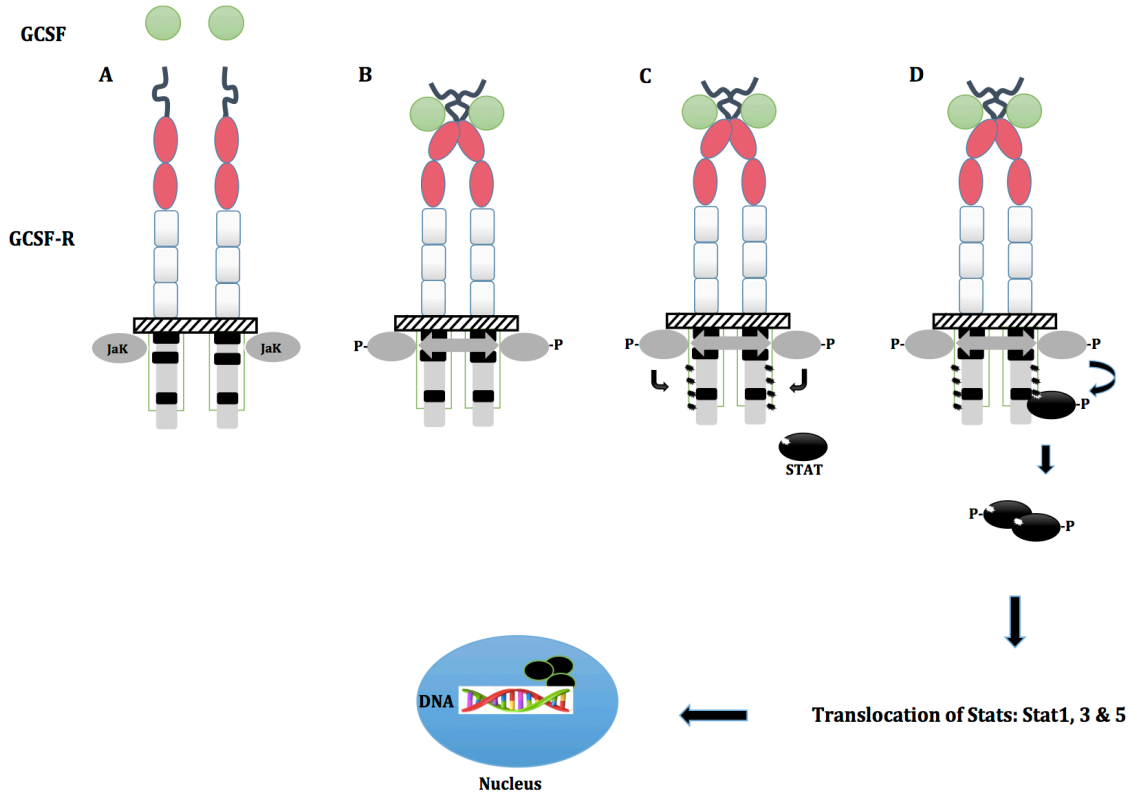


Figure 1-4: Pathway upon activation of Jak-STAT signals

[A] In the absence of the ligand, G-CSFR is associated with Janus kinases (Jaks). [B] The binding of the ligand to the receptor occurs at a 2:2 ligand:receptor subunit stoichiometry, forming a cross-over configuration between the receptor subunits bring the Jaks into proximity and enables their trans-phosphorylation and stimulation. [C] The intracellular 4-tyrosine residues of the GCSF-R (represented by stars) are phosphorylated by Jaks. [D] STAT interacts with the phosphotyrosine residues through their Src Homology 2 (SH2) domains and become phosphorylated by the Jak. Phospho-dimers of STATs accumulate in the nucleus and activate transcription factors that drive the neutrophils from the bone marrow to the blood circulation.

Although that GCSF-R is widely accepted to activate Jak/STAT pathways, it has also been reported that GCSF-R is linked to numerous components of Mitogen-activated protein (MAP) kinase and phosphoinostide-3-kinase-protein kinase B (PI-3K-PKB) and pathways resulting in activation of transcription factors and regulation. For instance, Y764 serves as docking site for Growth factor receptor-bound protein 2(Grb2) and has been linked to activation of p21 Ras pathway. A significant reduction of p21 Ras activation and neutrophil proliferation was noticed *in vitro* when Y764 was absent (Hermans et al., 2003). Extracellular signal-regulated 1/2 (Erk 1/2) MAP kinases are considered the main downstream effectors from the p21 Ras pathway resulting in signaling proliferation of myeloid progenitor cells. In neural cells, it is also reported that Erk1/2 is strongly activated upon exposure to GCSF (Hamilton, 2008, Panopoulos and Watowich, 2008, Touw and van de Geijn, 2007). In Swan 71 cells, binding of GCSF to its receptor leads to the activation of both PI3K/Akt and Erk1/2 pathways leads to the migration of NF-kB to the nucleus, stimulating an increase of matrix metalloproteinase-2 (MMP-2) activity and Vascular endothelial growth factor (VEGF) secretion (Furmento et al., 2014). It is also reported that the activation of GCSF-R triggers PI-3K-PKB pathway that is important for the stimulation of cell survival by inhibiting the apoptotic cascades (Figure 1.3) (Hunter & Avalos, 2000; Touw & van de Geijn, 2007). Another pathway that are induced by GCSF is the Tyrosine-protein kinase (Lyn), play a crucial role in GCSF mediated cell proliferation, in addition to the activation of GCSF primed pro-inflammatory responses in neutrophils (Sampson et al., 2007, Sivakumar et al., 2015).

1.5.2.1 Mobilization of Neutrophils

Previously, it was reported that STAT3 is the major transcription factor activated upon binding of GCSF to its receptor but the role of STAT3 in the mechanism of neutrophils mobilization was not clear (Panopoulos et al., 2002). The chemokines, macrophage inflammatory protein-2 (MIP-2, known as Cxcl2) and keratinocyte derived chemokine (KC, Cxcl1) and their shared receptor CXCR2 induce the mobilization of neutrophils from the BM to the circulating blood. In contrast to this the stromal cell–derived factor 1 (SDF-1, CXCL12) which is expressed in the BM and its chemokine receptor 4 (CXCR4), expressed on the surface of neutrophils, contribute to the retention of neutrophils in the BM and requiring

dawn-regulation to induce the releasing of neutrophils. Nguyen-Jackson et al. (2010) demonstrated that STAT3 controls the neutrophils migration from the BM to the circulating blood in response to GCSF treatment by binding to the chemokines MIP-2 and KC and increasing the production of these chemokines and reducing bone marrow SDF-1 expression in WT mice (Figure 1.5) (Nguyen-Jackson et al., 2010, Nguyen-Jackson et al., 2012). In summary, because GCSF is itself not chemotactic, this concept is supported by the observation that GCSF fails to induce circulating neutrophil amounts in CXCR2-knockout mice (Pelus et al., 2002). Inhibiting the SDF-1/CXCR4 interaction is sufficient to enable neutrophil release, as shown by use of the CXCR4 antagonist AMD3100 (Plerixafor) (Broxmeyer et al., 2005).

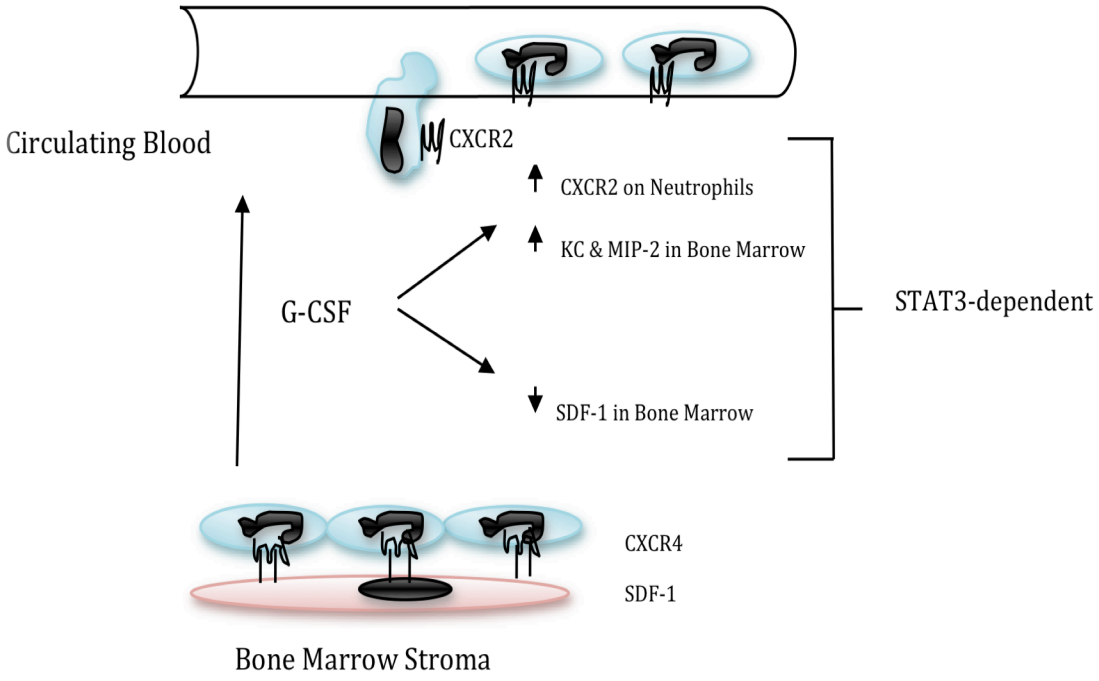


Figure 1-5: Scheme of the proposed model

How STAT3 induces the mobilization of neutrophils. Neutrophils are reserved in the BM in part throughout their CXCR4 expression, which binds to SDF-1 (stromal cells express SDF-1). Administration of GCSF leads to down-regulation and decreases of SDF-1 with its receptor CXCR4; suppression of SDF-1 needs STAT3. Intake of GCSF also stimulate the neutrophil chemo-attractants MIP-2 and KC in the BM together with up-regulation of their shared CXCR2 on the neutrophils surface, STAT3 is required in the stimulation of MIP-2, KC and CXCR2. Modified from (Nguyen-Jackson et al., 2010, Nguyen-Jackson et al., 2012).

1.6 The Major Clinical Use of GCSF

1.6.1 Febrile Neutropenia Prophylaxis

Febrile neutropenia (FN) complications (defined as development of fever >38.5^{0}C with absolute neutrophil counts < 1.0x106/L) are the main symptoms observed in patients treated with systemic cancer chemotherapy. Administration of GCSF shows benefits in reducing the risk of FN and accelerates the neutrophils number (Shah and Welsh, 2014).

1.6.2 Mobilization of Stem Cells

Administration of rhGCSF promotes and accelerates hematopoietic stem cell (HSC) secretion into peripheral blood in order to facilitate collection of large numbers of stem cells from the peripheral blood. Thus, GCSF is given to patients following chemotherapies in two processes either autologous or allogeneic stem cell transplantation. Autologous stem cell transplantation is a process that depends on the collection of a patient's own stem cells, administration of GCSF is mainly alone or in combination with chemotherapeutic drugs (Bensinger et al., 1995, Martino et al., 2014). The combination of GCSF with chemotherapeutic drugs showed higher numbers of CD34+ cells and lower levels of apheresis sessions in comparison to administration of GCSF alone (Pusic and DiPersio, 2008, Tanhehco et al., 2010). In contrast, allogeneic stem cell transplantation is a process in which the patient receives stem cells from a healthy individual who have been injected with GCSF for the purpose of donation. This process is more effective and safe than autologous stem cell transplantation (Hölig, 2013).

1.6.3 Controlling of SCN and AML

Administration of rhGCSF therapy has been shown to improve the conditions of patients with acute myeloid leukemia (AML) prior to chemotherapy in two ways; increasing both the neutrophil counts in the first 24 hours and the susceptibility of myeloid leukemia blast cells to chemotherapy (Löwenberg et al., 2003, Beekman and Touw, 2010). Also, it has shown to directly reduce the risk of several leukemia's. For instance, lymphoblastic AML patients achieved complete remission when treated with GCSF alone (Nimubona et al., 2002).

Although GCSF can increase the number of neutrophils in patients with severe congenital anaemia (SCN) and thereafter minimize this risk of recurrent infections, GCSF-R mutations have been reported in a patient with SCN who may developed secondary AML and myelodysplastic syndrome (MDS) due to administration of GCSF (Ancliff et al., 2003, Beekman et al., 2012). More details about the adverse effects of GCSF administrations will be discussed in the next section.

1.7 The Main Side Effects of GCSF Administration

To improve the safety of GCSF treatment in the clinical situation, it is important to consider the main side effects of GCSF administration. Over two decades, several distinct GCSF-R mutations, which cause intracellular receptor truncations, have been reported in a patient with SCN who developed secondary AML/MDS (Beekman et al., 2012). These truncations are considered a crucial step in the expansion of the pre-leukemic clones and the possibility that GCSF administration will give rise to these mutant clones and thereafter cause AML.

All GCSF-R mutants have no differences in their juxta-membrane and extracellular domains, but do differ in their cytoplasmic (intracellular) regions. Among all these mutations, the class IV isoform mutant (differentiation-defective) has been detectable in hematopoietic cells and is associated with administration of GCSF. This isoform maintains the membrane proximal sequence that is required for proliferative signaling but loses at position 725 the carboxy-terminal 87 amino acids, which are replaced with a unique 34 amino acid sequence (White et al., 1998, White et al., 2000). As a result, out of the 4-tyrosine residues (Y704, Y729, Y744 and Y764) in the full-length form, only Y704 is conserved (Class IV isoform) (Figure 1.6). Overexpression of isoform IV has been observed in patients with AML/MDS, potentially due to the ability of this isoform to block maturation (Liongue and Ward, 2014). In addition to this mutation, common side effects reported for patients who were treated with GCSF are summarized in Table 1.1.

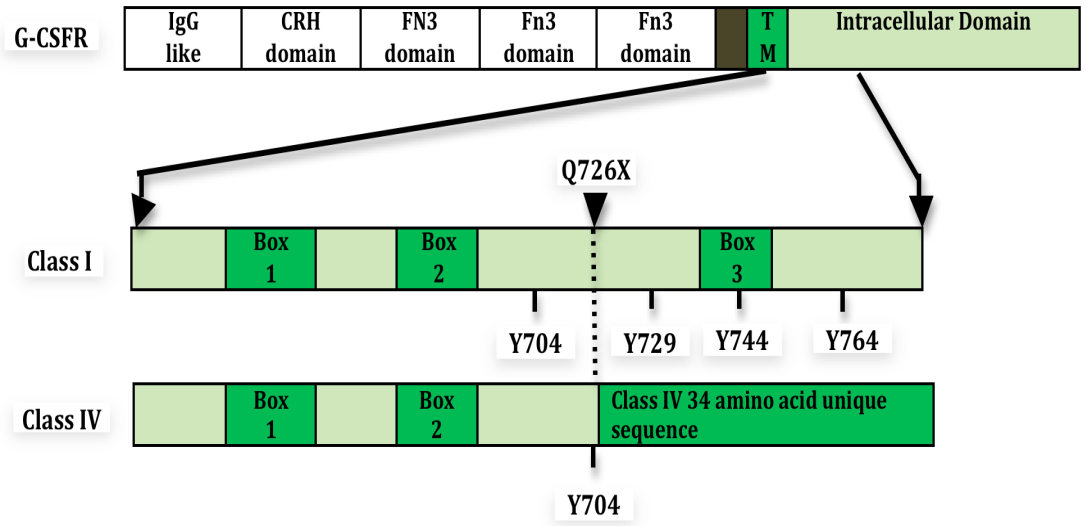


Figure 1-6: Comparison of carboxyl-terminal region of the GCSF-R in patients with AML

Class IV isoform has no differences in their juxta-membrane and extracellular domains, but differing in their cytoplasmic domains. The isoform maintains the membrane proximal sequence that required for proliferative signaling and loses at position 725 the carboxy-terminal 87 amino acids, which are replaced with a unique 34 amino acid sequence. As a result, the 4-tyrosine residues (Y704, Y729, Y744 and Y764) in the full-length form class1 (wild type) truncated and only Y704 is conserved among this isoform (Class IV). Modified from (Liongue and Ward, 2014).

Table 1-1: The main side effects reported for patients treated with GCSF

Side effect	Organ	Possible mechanism	References
Osteopenia & osteoporosis	Bone	Administration of G-CSF increases bone resorption via increasing the activity of osteoclasts leading to significant bone loss.	(D'Souza et al., 2008)
Joint pain & generalized weakness	Bone	1) Expansion of bone marrow, 2) Enhancing of GCSF-R on afferent nerve fibers lead to generate peripheral nociceptor sensitization, 3) Stimulation of inflammatory cells, such as, macrophages and monocytes that contributes to nerve remodeling, 4) Osteoblast and osteoclast activation.	(Lambertini et al., 2014)
Erythematous rash, urticarial and Sweet syndrome	Skin	Infiltration of neutrophils into dermis and epidermis.	(Nomiyama et al., 1994, Prendiville et al., 2001, Llamas-Velasco et al., 2013)
Splenomegaly & extramedullary hematopoiesis	Spleen	Stimulation of myelopoisis.	(Litam et al., 1993, O'Malley et al., 2003, Dagdas et al., 2006)
Lung cancer and acute respiratory distress syndrome	Lung	Accumulation of neutrophils due to releasing of chemoattractant molecules. As a result, these cells release a number of substances injurious, for instance, platelet, leukotrienes, proteases and, oxidants that cause damage in the alveolar endothelium and epithelium.	(Asano et al., 1977, Wada et al., 2011, Yamaguchi et al., 2012, Inokuchi et al., 2015)
Reversible renal impairments	Kidney	Stimulation of leukocytosis in the kidneys.	(Hirokawa et al., 1996)

1.8 Available Commercial GCSF Preparations and Their Limitations

Human GCSF has been cloned and is currently available as two recombinant human GCSF (rhGCSF) preparations for HPC mobilization:

1.8.1 Filgrastim (NEUPOGEN®)

Filgrastim is a non-glycosylated form (18.8kD), obtained from E. *coli* and has a methionine group at its N-terminal end. In 1991, Filgrastim was licensed and marketed as a treatment for patients with neutropenia following chemotherapeutic drugs. Since its launch, clinicians have recommended the use of Filgrastim in bone marrow transplantation procedures, aplastic anaemia, sever congenital neutropenia, and to support patients with AIDS and myelodysplastic syndromes (Molineux, 2004). However, it was reported that Filgrastim has a short half-life in the serum of 3.5 hours when injected in to healthy volunteers and patients with malignancies because E. *coli* derived Filgrastim lacks the O-linked glycosylation (Cooper et al., 2011, Hoggatt and Pelus, 2014).

1.8.2 Lenograstim (Granocyte®)

Lenograstim is an O-glycosylated form at Thr-133 position obtained from Chinese hamster ovarian (CHO) cells (Nagata et al., 1986). It has a short half-life in the serum of between3–4 hours and O-linked was shown to have an important role in providing greater stability to the GCSF by protecting the cysteine-17 sulfhydryl group from oxidation by free radicals (Hasegawa, 1993, Cooper et al., 2011). It was believed that the O-linked glycosylation might show clinical advantages over the non-glycosylated form (Filgrastim), however the *in vivo* comparative studies showed that no differences between them in all aspects (Ataergin et al., 2008). Besides, leukemic patients need daily injections to maintain its activity in the circulation, which are inconvenient, expensive and painful especially for children. As Lenograstim is similar to the natural GCSF, studies shifted to focus on this form.

1.9 Strategies Used to Delay the Clearance of GCSF

Since the main limitations of the previous forms of rhGCSF (1st generation) are short circulating half-life and daily injection, two strategies have been developed to overcome these issues. The first strategy is to increase the molecular weight of the therapeutic proteins (hydrodynamic radius) above the renal filtration threshold. Predominantly, this increase in molecular weight can be attained by conjugation with another moiety such as, PEGylation. The second strategy to extend the half-life of the therapeutic proteins offers the benefits of the neonatal Fc receptor (FcRn) recycling mechanism by forming fusion proteins with both Fc portion of immunoglobulin and albumin (Natalello et al., 2012, Cox et al., 2014, Chung et al., 2011). Additional new approaches using Asterion profuse technology will also be discussed.

1.9.1 Extension of Half-life by Increasing the Molecular Weight

1.9.1.1 PEGylation

The short circulating half-life of many recombinant proteins can be increased via conjugation with poly ethylene glycol (PEG), in a process termed PEGylation (Natalello et al., 2012). In the 1970's, It was first described by Abuchowski and Davis who found that PEG may improve the immunological properties and serum half-life of proteins such as bovine liver catalase and albumin (Abuchowski et al., 1977). Since then, widespread research has been carried out into PEG technology resulting in highly variable PEGs with several molecular weights (Jain and Jain, 2008). A range of Pegylated proteins are now clinically available, such as, Pegvisomant (Pegylated growth hormone antagonist, licensed for the treatment of acromegaly in 2003 (Trainer et al., 2000, Hamidi et al., 2006).

To generate a new 2nd generation product, a 20 kDa PEG molecule was attached covalently to recombinant methionyl (r-met) human GCSF Neupogen (Filgrastim). This new molecule was marketed as Neulasta (Pegfilgrastim). In Pegfilgrastim, each ethylene oxide unit of PEG binds to three water molecules which increases its water solubility and also hydrodynamic radius (molecule's diameter) resulting in increased size of the molecule to ~38.8kDa, thus reducing

renal clearance. PEGylation also creates a hydrophilic shield that protects the protein from proteolysis and immunologic recognition (Bailon and Won, 2009, Milla et al., 2012). The key behind the successful progress of Pegfilgrastim over Filgrastim was an understanding of the GCSF clearance processes in the body. In humans, GCSF has two clearance mechanisms: renal clearance and neutrophil-mediated clearance (Yowell and Blackwell, 2002). The presence of the PEG moiety decreases the renal clearance of Pegfilgrastim and as a result it is mainly cleared via a self–regulating neutrophil-mediated mechanism, which is dependent on the number of neutrophils. Following administration of Pegfilgrastim, concentrations remain high in patient serums during neutropenia, but are reduced when the numbers of neutrophils increase. Therefore, a single injection of Pegfilgrastim per chemotherapy cycle is as efficient as the daily administration of Filgrastim (Curran and Goa, 2002).

The PEGylation of a protein used to be one of the major limitations because the whole PEG is often processed for excretion in the human body without undergoing an initial biodegradation which could be toxic to the body (Patel et al., 2014). This possible toxicity was supported by the detection of PEG in bile (Caliceti and Veronese, 2003). Vacuole formation has also been observed in renal tubules upon administration of PEG thereby affecting the tissue distribution and in turn clearance (Zhang et al., 2014).

Modification of GCSF with PEGylation reduced *in vitro* biological activity of GCSF (2 to 3 fold), as the conjugation of PEG to the GCSF could induce structural change in the molecule that attenuates the potency (Kinstler et al., 1996, Gaertner and Offord, 1996).

The increased cost of PEGylation is considered to be one of the main limitations because Pegfilgrastim requires a post-expression chemical modification and purification processes (Pisal et al., 2010). The table below summarizes the advantages and limitations of PEGylation and other different strategies used to generate long acting GCSF (Table 1.2).

Table 1-2: Summary of pros and cons of PEGylation and other different strategies used to generate a long acting GCSF

Strategy	Pros	Cons	References
PEGylation 2nd generation Pegfilgrastim	Extended circulating half-life by reduced renal clearance. Creates a hydrophilic shield that protects the protein from proteolysis and immunologic recognition and proteolysis.	Detection of vacuole formation within the kidney and macrophages. Non-biodegradable. Reduced *in vitro* biological activity of GCSF (2 to 3 fold). High cost due to the chemical modifications and purification process	(Kinstler et al., 1996, Gaertner and Offord, 1996, Caliceti and Veronese, 2003), (Pisal et al., 2010), (Patel et al., 2014), (Zhang et al., 2014)
Fusion to Antibodies 2nd generation G-CSF/IgG-Fc & G-CSF/IgG-CH	Prolonged circulating half-lives (5 to 8-fold longer than G-CSF), accelerate number of neutrophils in vivo and decrease risk of immunogenicity.	Several proteins showing no bioactivity when attached to the Fc-IgG1 domain. Reduced *in vitro* biological activity of GCSF (3 to 4 fold) when attached to the IgG-CH domain due to significant amount of disulfide-linked aggregates/oligomers.	(Cox et al., 2004, Cox et al., 2014, Czajkowsky et al., 2012, Mitragotri et al., 2014)
Fusion to Albumin 2nd generation	Increased half-life of GCSF, increased WBC counts of neutropenia mice, decrease risk of immunogenicity, biodegradable and showed high stability.	Not offering any secondary functions, for instance, cytotoxicity. High cost product.	(Zhao et al., 2013, Schmidt, 2009, Mitragotri et al., 2014)
LR-fusion (Asterion) 3rd generation	Increased half-life of GH, reduced immunogenicity and toxicity. Naturally occurring sequences		(Wilkinson et al., 2007, Ferrandis et al., 2010)

1.9.2 Extension of Half-life Using the FcRn-Mediated Recycling Mechanism

In the last years, the FcRn recycling mechanism has been used expensively as a strategy to prolong the half-life of different proteins. Plasma proteins such as, albumin and Immunoglobulins (IgG's) are found to recycle by the neonatal Fc receptor (FcRn) pathway resulting in a longer circulating half-life.

It has been shown that the FcRn is responsible for the extraordinary long circulating half-life of albumin and IgG's (19 days for albumin and 23 days for IgG's in human (Dall'Acqua et al., 2002, Chaudhury et al., 2003, Anderson et al., 2006, Baker et al., 2009)). Studies have shown that IgG's are catabolized more quickly (Ghetie et al., 1996, Israel et al., 1996) and albumin is degraded approximately twice as fast in FcRn deficient mice than in wild type mice (Chaudhury et al., 2003, Baker et al., 2009). Fusion or non-covalent binding of small proteins (e.g. GCSF) to the Fc-part of IgG's or albumin significantly improved their pharmacokinetic properties.

Generally, FcRn can bind tightly to albumin as well as IgG's in a pH-dependent manner. Due to the presence of histidines in albumin and IgG's, the imidazole group of histidine becomes protonated at acidic pH 6.0 and interacts with the FcRn receptor. The interaction complex of FcRn-albumin or IgG's is then internalized and taken by cells to protect albumin or IgG's from lysosomal degradation (Chaudhury et al., 2006, Andersen and Sandlie, 2009, Dumont et al., 2006). FcRn-albumin or IgG's complex is released back to the cell surface membrane, where exposure to physiological pH 7.2 within the blood circulation causes the release of albumin or IgG's from the receptor (Ober et al., 2004, Andersen et al., 2006).

Extensive studies on the Human serum albumin (HAS) structure indicated that carboxy terminal domain III HAS (3DHAS) alone is sufficient for binding FcRn receptor and the histidine present in 3DHSA may dominates the binding between HSA and the FcRn receptor (Figure 1.7) (Andersen et al., 2010).

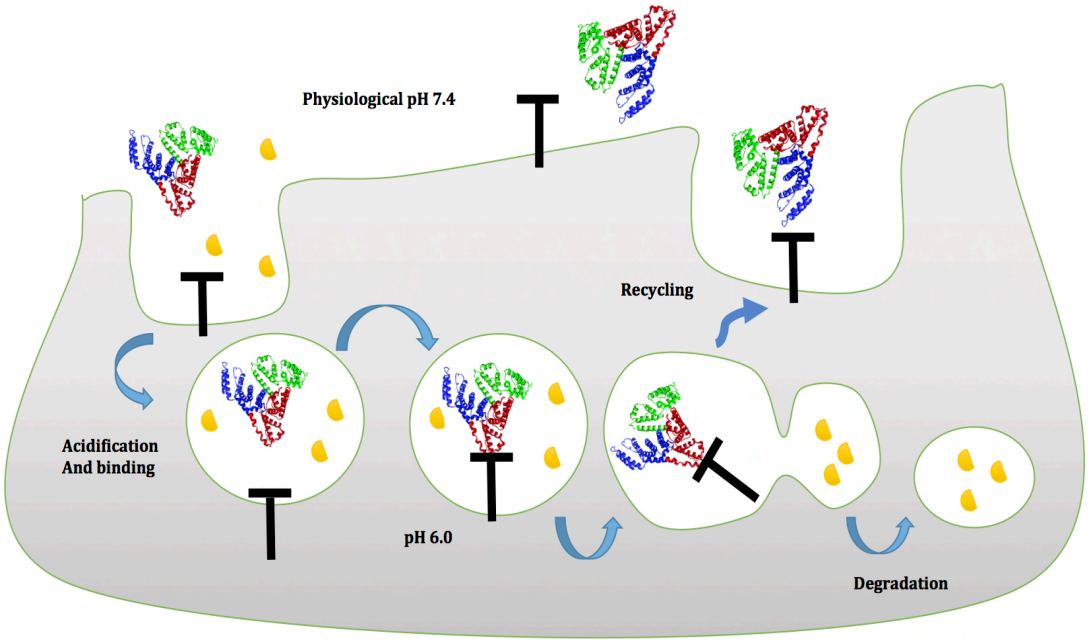


Figure 1-7: Model of the pH-dependent recycling mechanism of albumin via the FcRn receptor in serum

The three domains of albumin are marked in green (domain I), blue (domain II) and red (domain III). Domain III albumin is binding FcRn receptor at acidic pH 6.0 to protect from lysosomal degradation and recycling again to the circulation where exposure to physiological pH 7.2 causes the release of albumin.

1.9.2.1 Fusion of GCSF to Albumin

Human serum albumin (HSA) is produced in liver and has numerous physiological roles including transportation of fatty acid and metal ions as well as maintenance of plasma pH and colloid blood pressure. As albumin is a large protein (67kDa), it can be fused to recombinant proteins to extend their half-life. As a result, fused proteins will automatically attain a molecular weight too large to be filtered through the kidney and increase plasma protein residency time (Dennis et al., 2002, Andersen and Sandlie, 2009)

The advantage of carboxy terminal domain III HAS (3DHAS) has been considered widely and later 3DHAS was genetically fused to the N-terminal of GCSF. The pharmacokinetic/ pharmacodynamics (PK/PD) studies showed increased half-life and increased WBC counts of neutropenia model mice compared to native GCSF (Zhao et al., 2013).

1.9.2.2 Fusion of GCSF to IgG-Fc

In humans, the circulating half-life of IgG1 and IgG4 immunoglobulins is 23 days, and that IgG1 and IgG4 immunolglobulins have been used to form several long-acting fusion proteins (Gaberc-Porekar et al., 2008). Thus, immunoglobulins selected as the choice antibody for Fc fusion proteins.

Structurally, immunoglobulins are composed of two identical light and heavy chains connected by disulphide bonds. Both chains contain two regions: the fragment of antigen binding (Fab) (the head region of an antibody) responsible for immunogenic detection and the crystallisable fragment (Fc) (the tail region of an antibody that interacts with cell surface receptor) responsible for maintenance of IgG in the blood circulation (Gaberc-Porekar et al., 2008).

The IgG immunoglobulin has two fragments: CH (CH1-Hinge-CH2-CH3) and Fc (Hinge-CH2- CH3) domains. The Hinge domain is responsible for linking Fab and Fc regions and provides more flexibility. Many therapeutic proteins have been reported to joined via the amino-termini of CH (CH1-Hinge-CH2-

CH3) and Fc (Hinge-CH2- CH3) domains of human IgGs through their carboxy-termini (Cox et al., 2004). In mammalian cells, IgG fusion proteins are frequently expressed and secreted as disulfide-linked homodimers because of inter-chain disulfide bonds that are created between cysteine residues sited in the hinge region of the IgGs. The effective size and circulating half-life's of IgG fusion proteins are further increased by the dimeric structure of IgG fusions.

Chimeric genes have been produced encoding human GCSF that are fused through a 7-amino acid flexible linker (Ser-Gly-Gly-Ser-Gly-Gly-Ser) to the N-termini of the CH (CH1-Hinge-CH2-CH3) and Fc (Hinge-CH2- CH3) domains of human IgG4 and IgG1 immunoglobulins (Figure 1.8). Fusions of GCSF to human IgG domains were shown to form homodimers with high molecular weight, prolonged circulating half-lives (5 to 8-fold longer than GCSF) and accelerate number of neutrophils *in vivo*, without any significant effect on GCSF biological activity *in vitro* (Cox et al., 2004, Cox et al., 2014).

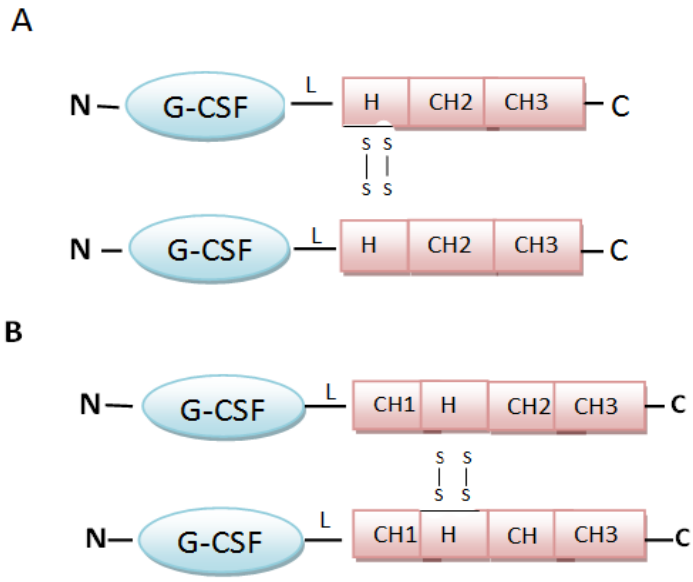


Figure 1-8: The schematic diagram shows (A) G-CSF/IgG-Fc protein and (B) G-CSF/IgG-CH fusion

The GCSF carboxy-terminus is linked through a 7 amino acid fixable linker (L) to the amino termini of the IgG-Fc and IgG-CH domains. The CH1, CH2, and CH3 regions and hinge (H) of the IgG domains are also showed. The dimeric of fusion proteins located in the IgG hinge region is due to the presence of disulfide bonds (SS) that form between cysteine residues. Modified from (Cox et al., 2004, Cox et al., 2014).

1.9.3 New Approach by Asterion

Asterion is a Sheffield University spin out company formed in 2001. Prof Richard Ross is one of the founding directors. Their strategy is to develop long acting biological using novel platform technologies, which utilises the fusion of ligand with soluble extracellular receptor, termed Profuse™ technology.

Over the last fifteen-years, plans have been developed by Asterion to focus on the utility of Profuse™ technology to produce long acting biopharmaceutical products. The current therapeutics regime for protein replacement requires daily injections, which are expensive and inconvenient. Thus, there is a need for a 3rd generation of protein therapies that are easy to administrator, acceptable and convenient to patients and also minimizes manufacturing costs. Thus, two approaches have been created by Asterion technology:

1.9.3.1 Ligand/Receptor Fusion

Using flexible linker $(Gly_4Ser)_n$ technology, Wilkinson et al. (2007) demonstrated that a fusion of Growth Hormone (GH) to its extracellular receptor (GH binding protein: GHBP) creates an effective long acting agonist with exceptional delayed clearance properties. PK analysis in rats showed that ligand-receptor growth hormone fusion molecule (LR-fusion) had a resulting 300-times reduced clearance rate when compared to native GH (Wilkinson et al., 2007).

In addition, preclinical work on a various number of long acting GCSF molecules has been performed and the preliminary data for one construct (4A1) showed that it is possible to design, clone and purify a GCSF linked to its extracellular receptor. It was also shown to reduce the rate of clearance following subcutaneous injections (up to 60 hours) in rats consistent with an increase of 15-fold over that recorded for the native GCSF and 2-fold over that reported for Pegfilgrastim.

1.9.3.2 Glycosylation

In medicine, any protein used for therapeutic purposes is not only a sequence of amino acids determined by a particular gene, but it still requires editing, altering of the amino acids or addition of carbohydrates. These modifications following the preliminary translation of the protein are called post-translational processes (Li and d'Anjou, 2009). Glycosylation refers to the post-translational process that attaches oligosaccharide to polypeptides. It is one of the most common protein modifications and more than 50% of proteins are glycosylated in the body, which are mainly secreted or part of cell membrane components (Sola et al., 2007).

Glycosylation play a fundamental role in forming or maintaining glycoprotein integrity. Generally, it can increase the molecular weight of proteins, provide a high degree of protection against proteolytic degradation and enhance thermal stability by decreasing immunogenicity due to the presence of terminal sialic acid that creates negative charge around the glycoprotein resulting in delay clearance. Given these important functions, it is now believed that glycosylations contribute to regulating protein-protein interactions, which is highly important in optimizing and developing glycoprotein drugs. In addition, an understanding of the association of the carbohydrate moieties to receptor binding can be used to enhance treatment efficacy (Li and d'Anjou, 2009).

Therefore, efforts have been made to improve a strategy to delay clearance of GCSF by the addition of natural carbohydrates (e.g. glycosylation) and avoiding limitations that were observed in section 1.8**.** For example**,** PEG molecule is often processed for excretion in the human body without undergoing an initial biodegradation which could be toxic to the body (Patel et al., 2014). Whereas, modifying proteins with glycosylation can undergo degradation within the human body with no potential toxicity of PEGylation was supported by the detection of PEG in bile (Caliceti and Veronese, 2003). For the safe conjugation of glycosylation to GCSF, it is important first to understand the forms and functions of glycosylation and also the best conjugation method to the protein.

Normally, the majority of proteins synthesized begin in the endoplasmic reticulum (ER) undergo glycosylation and are completed in the Golgi apparatus. Five classes of glycosylation are produced; N-linked glycosylation, O-linked glycosylation, Phospho-serine glycosylation, C-mannosylation and formation of GPI anchors. However, the two major forms are O-linked and N-linked glycosylation (Saint-Jore-Dupas et al., 2007).

O-linked Glycosylation

O-linked glycosylations are attached to the hydroxyl groups (-OH) of serine or threonine residues within a protein (Wongtrakul-Kish et al., 2012). However, no particular consensus sequences have been recognized for this reaction and it is still ambiguous why certain Ser/Thr residues are glycosylated as opposed to others. One theory is that it could be down to alternative structural properties of the protein that may contribute to the availability of the glycosylation site (Sinclair and Elliott, 2005).

In human GCSF, O-linked glycosylation is located at Thr-133 and is necessary for the increased stability of the GCSF molecule. The actual molecular mechanism of the glycosylated forms increased stability remains to be established, but a study carried out by Hasegawa (1993) suggested that O-linked glycosylation might be involved in the protection of the cysteine-17 sulfhydryl group by preventing oxidation by free radicals (Hasegawa, 1993).

N-linked Glycosylation

N-linked glycosylation is attached to the amide nitrogen of asparagine (Asn) residues within the common consensus sequence Asn-X-Ser/Thr where X can be any amino acid except proline (Pro) (Kornfeld and Kornfeld, 1985). In nature, N-linked is the most common form of glycosylation and is thus preferentially used in many technologies of protein modification (Spiro, 2002). Therefore, it is the focus of further discussion.

Glycosylation biosynthetic pathways start in ER (Figure 1.9). At this early stage, a 9-Mannose glycan is added to the peptide of an N-linked glycan

(known as a high mannose type). The addition of these glycans is fundamental to the control of the folding of the newly synthesized proteins. Upon successful folding of the protein, the glycoprotein migrates into the Golgi apparatus where mannosidases facilitate the removal of the mannose groups. This is then followed by the addition of different monosaccharides to the growing glycan chain by particular glycosyltransferases (This is known as a hybrid type which contains both high mannose and complex type).

At this late stage the biosynthetic process is completed in the Golgi apparatus with a fully sialylated glycan complex, which contains six sugars: N-acetylglucosamine (GlcNAc), mannose (Man), fucose (Fuc), galactose (Gal) and sialic acid (NeuAc), which are linked by different α- or β glycosidic linkages (Kim et al., 2009, Butler and Spearman, 2014).

Sialic acid (NeuAc) occupies the terminal site of the N-linked glycan and has been found to be critical in maintaining the circulating half-life of glycoproteins (Sola and Griebenow, 2009, Kim et al., 2009). It contains a large group of nine carbon-containing carbohydrates the most predominant of which is N- acetylneuraminic acid (Neu5Ac). Byrne et al., (2007) showed that many properties conferred onto proteins are due to the presence of the negative charge on C1 of sialic acid.

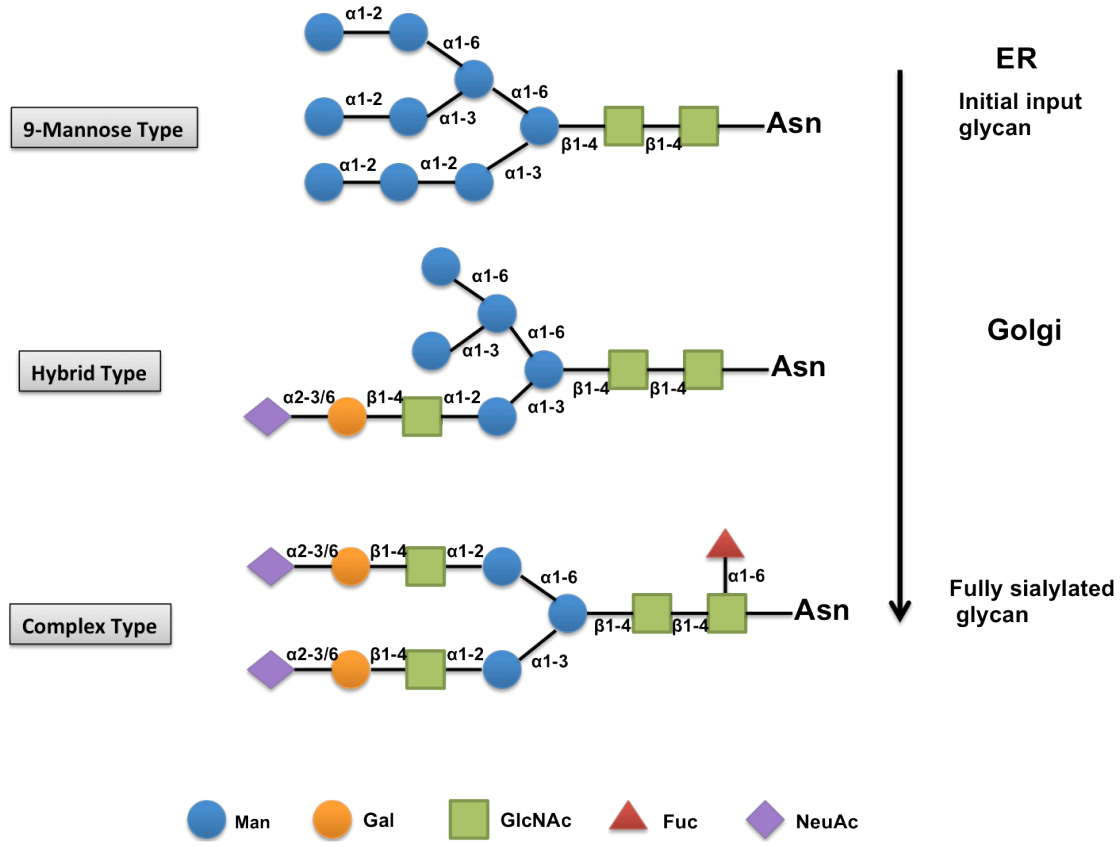


Figure 1-9: Glycan structure

Initial input glycan (9-Mannose glycan) starts biosynthetic pathway in ER (Top), the glycoprotein migrates into Golgi where the removal of mannose group and the addition of different monosaccharides in a process called hybrid type (mid). The biosynthetic processed then is completed in the Golgi as fully sialylated glycan complex (bottom). Man: mannose; Gal: galactose; GlcNAc: N-acetylglucosamine; Fuc: fucose; NeuAc: sialic acid.

Glycan and sialic acid function

In naturally-glycosylated proteins, oligosaccharides (glycans) and their terminal sialic acids fulfil a number of functions:

1. Overexpression charge repulsion between glomerular filtration barrier and glycoprotein reducing renal clearance

The presence of sialic acid on cell membrane surface proteins within the glomerular basement membrane creates over-expression of a negative charge barrier alongside the glomerular filtration barrier preventing the passage of glycoproteins through charge repulsion. This is applicable to therapeutic glycoproteins as a possible method to increase circulatory half-life (Varki, 2008).

2. Prevention of proteoylatic degradation

Sialic acid occupies a comparatively large volume in glycosylation, which protects the underlying peptides from protease detection and cleavage within the circulation (e.g. papain) (Sola and Griebenow, 2009). A study carried out by Raju and Scallon (2007) demonstrated that removal of terminal sugars from Fc antibody fragments (Antibodies are naturally glycosylated in the CH2 of the Fc fragment) resulted in increased sensitivity to papain.

3. Reduced receptor binding affinity

Elliott et al. (2004) showed that non-glycosylated erythropoietin (EPO) has seven times greater receptor binding affinity than native EPO (highly glycosylated). This is due to the presence of the negative charge created by sialic acid, which reduces binding affinity via charge repulsion. Also, *in vivo* activity of non-glycosylated analogue was significantly decreased due to increased receptor binding affinity resulting in increased internalization and protein degradation (Elliott et al., 2004).

4. Stabilization of the protein

Glycosylations offer stability for therapeutic proteins from two points of view:

Glycosylation may help decrease the immunogenicity of polypeptides in two ways: Walesh and Jefferis (2006) reported that one potential mechanism is via increased solubility of polypeptides due to interaction with H_2O and shielding of hydrophobic residues reducing the possibility of aggregate formation that could form a stationary precipitate for antibody recognition. Another possible mechanism is via shielding the polypeptide, reducing the accessible surface area exposed for antibody recognition, resulting in a masking effect provided by sialic acid. This is shown with Darbepoetin alpha (a rhEPO analogue composed of two additional N-linked glycosylations sequences) whereby ELISA tests were unable to detect the protein in the circulation from different patients (Sinclair and Elliott, 2005, Byrne et al., 2007).

From a manufacturing point of view, glycosylations offer stability for therapeutic proteins against chemical denaturation and pH by increasing the potency of internal forces and this is essential in maintaining therapeutic protein conformation. It is also important to increase the protein structural compactness, resulting in a decreased available surface area to denaturing pH or chemicals (Sola and Griebenow, 2009).

From the previous discussion, it can be shown that N-linked is the most common form of glycosylation and is thus preferentially used in many technologies of protein modification (Spiro, 2002). N-linked strategies can be divided in two classes:

1- Hyperglycosylation via Site Direct Mutagenesis

Hyperglycosylation is defined as the process of increasing glycosylation on a protein to alter its pharmacokinetic and biological activity (Sola and Griebenow, 2010). DNA mutagenesis is one of the main strategies to increase glycosylation. *In vivo*, mutagenesis of the DNA can incorporate

additional glycosylation sites. This process can be achieved by identifying Thr/Ser residues occupying the third position in the sequence of a protein and mutating the first amino acid in the sequence to Asn or by identifying Asn residues within the sequence of a protein and mutating the third amino acid to Thr/Ser. For example, Elliott et al (2003) used site directed DNA mutagenesis in the development of Darbepoetin alpha. The process started by mutating Ala-30, His-32 to Asn-30, Thr-32, and Pro-87, Trp-88, Pro-90 to Val-87, Asn-88, Thr-90. All mutations were shown to be glycosylated with an increase in molecular weight from 35kDa to approximately 43kDa, whilst maintaining biological activity.

Site direct mutagenesis has also been performed on GCSF. Hee et al. (2011) produced an N-linked glycosylation site on rhGCSF by mutating Phe140 to Asn140producing a novel form of human GCSF mutant. The new mutant rhGCSF was shown to be glycosylated and more effective than native GCSF for stimulating differentiation and proliferation of hematopoietic cells.

2- Glycosylated Linkers

The uses of glycosylated linkers are another glycan strategy that can be utilised to extend the half-life of therapeutic proteins. The glycosylated linkers can be inserted between subunits of the same ligand as observed in recombinant human Follicle-Stimulating Hormone (rhFSH) between the α- and β- subunits in which either N- linked or O-linked glycosylated linkers were used. The effect of the glycosylated linker was to increase the half-life of the molecules by as much as 2-fold compared to rhFSH (Weenen et al., 2004).

This method can also be used to insert glycosylated linkers between two ligands (tandem molecules). The insertions of glycosylated linkers between tandem molecules are preferable as they alleviate potential problems with direct glycosylation of the ligand itself, which may inhibit bioactivity and potentially introduce immunogenic sites.

The Ross group (UoS) have used the advantages of both glycosylated linker design and tandem molecules to create a long acting GH molecules. They created a tandem of two GH molecules joined by a flexible $(Gly_4Ser)_n$ linker containing variable glycosylation motifs. The results successfully showed that the use of glycosylated linkers between two GH ligands results in their glycosylation and increased molecular weight whilst maintaining biological activity and delaying clearance in a rat model. This methodology could be applied to other cytokine hormones such as GCSF.

1.10 Aim and Hypothesis

The role of recombinant human rhGCSF has been successful in the treatment of patients with neutropenia and stem cell mobilization in the circumstance of bone marrow transplantation, making it a multi-million pound market. The fast growing market size will increase appetite for generating similar rhGCSF. Therefore, it is essential to produce new GCSF compounds with improved properties over other products.

The aim of this project is to create a long acting GCSF with a predictable pharmacokinetic profile to provide a more effective treatment for generating HSCs for transplantation purposes.

We hypothesized that the incorporation of variable glycosylation motifs (2NAT, 4NAT and 8NAT) within a flexible linker $(Gly_4Ser)_n$ between two GCSF ligands will Increase the molecular weight (MW) of GCSF according to the number of glycosylation motifs with protection from proteolysis resulting in reduced clearance with out blocking bioactivity (Figure 1.10 and Table 1.3). This approach also alleviates potential problems with direct glycosylation of the ligand, which may reduce bioactivity and potentially introduce immunogenic sites.

Main Objectives

1- Design, cloning and expression of GCSF tandems.
2- Purification of GCSF tandems from a mammalian cell line.
3- Analyse and compare the biological activity of GCSF tandems with rhGCSF using an *in vitro* proliferation assay.
4- Determine the pharmacokinetic/pharmacodynamic properties of rhGCSF & GCSF tandems in a normal rat model.

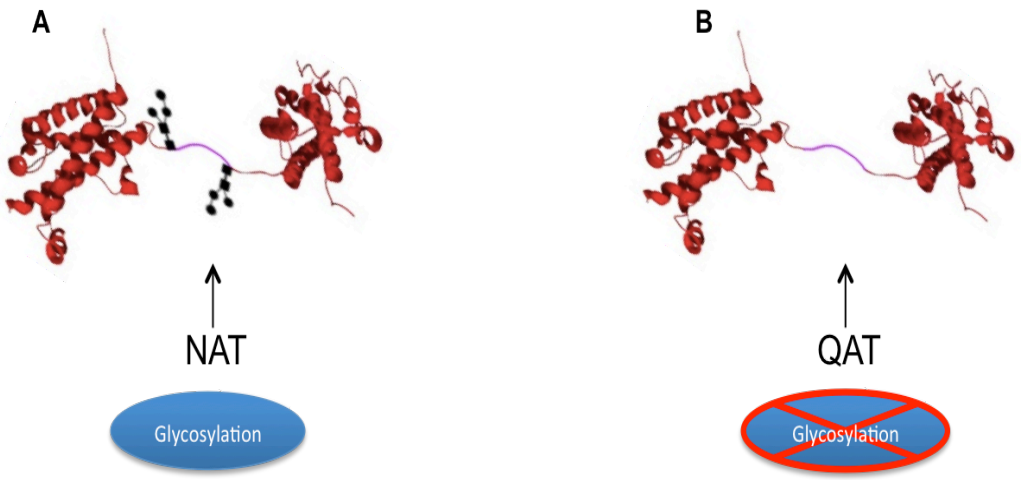


Figure 1-10: An example of 2NAT glycosylation motifs and its control 2QAT within a flexible linker $(Gly4Ser)_n$ between two GCSF ligands

(A) The glycosylation motif 2NAT inserted to the linker (glycosylated linker). (B) Non-glycosylation motif 2QAT control.

Table 1-3: List of GCSF tandems

GCSF tandems linked via a flexible linker incorporating increasing numbers of the glycosylation motif NAT (glycosylated molecules) or control motif QAT (non-glycosylated molecules).

Molecule name	Number of NAT/QAT motifs	Size of linker
GCSF8NAT_Hist	8 x NAT	282bp
GCSF8QAT_Hist	8 x QAT	282bp
GCSF4NAT_Hist	4 x NAT	147bp
GCSF4QAT_Hist	4 x QAT	147bp
GCSF2NAT_Hist	2 x NAT	147bp
GCSF2QAT_Hist	2 x QAT	147bp

2. Materials

2.1 Cell Culture

Easy filtered flasks 75cm^2	Nalgen Nunc intl
Easy filtered flasks 25cm^2	Nalgen Nunc intl
100mm tissue culture dishes	Iwaki SLS
6 well Cell Culture plates	Costar
24 well Cell Culture plates	Costar
CHO Flp-In cell lines	(Invitrogen Corp) Gibco
Cryogenic vials	Nalgen Nunc intl
Dimethylsulfoxide (DMSO)	Sigma-Aldrich
Dulbecco's Modified Eagles Medium (DMEM)/F-12	Sigma-Aldrich
Fugene-6	Roche Diagnostics
Foetal Calf serum (FCS)	Labtech
HyClone SFM4CHO Cell Culture Media	Thermo Scientific
Hygromycin B Antibiotic (50mg/ml)	(Invitrogen Corp) Gibco
Ham's F12	(Invitrogen Corp) Gibco
Labtech chambers coverglass	Nalgen Nunc intl
L-Glutamine	(Invitrogen Corp) Gibco
Phosphate Buffer 10X	(Invitrogen Corp) Gibco
Penicillin-Streptomycin (100µg/ml)	(Invitrogen Corp) Gibco
pOG44	(Invitrogen Corp) Gibco
Trypsin-EDTA	(Invitrogen Corp) Gibco
Zeocin Antibiotic (100mg/ml)	(Invitrogen Corp) Gibco
Counting chamber	Hawksley
Trypan blue	Sigma

2.2 DNA Manipulation

1 kb ladder	(Biolabs) New England
100 bp ladder	(Biolabs) New England
Agar	Sigma-Aldrich
Agarose	Sigma-Aldrich
CaCl2	Sigma-Aldrich
Dithiothreitol (DTT)	Sigma-Aldrich
dNTPs	Upjohn and Pharmacia
DNA loading dye (6X)	Promega
EDTA	BD Biosciences
Expand high fidelity PCR system	Roche Diagnostics
Ethidium bromide (0.5µg/ml)	Sigma-Aldrich
Gen Elute™ Gel Extraction Kit	Sigma-Aldrich
Glycerol	BD Biosciences
KCl	BD Biosciences
MgCl2	Sigma-Aldrich
NaCl	BD Biosciences
QIAGENprep Spin Miniprep Kit	QIAGEN
QIAGENquick PCR Purification Kit	QIAGEN
QIAGENquick Gel Extraction Kit	QIAGEN
Restriction enzymes	Promega
Synthesised DNA primers	MWG Biotech
T4 DNA ligase	Promega
Taq polymerase	Promega
Tryptone	Melford
Tris –base	Sigma-Aldrich

2.2.1 Restriction Endonucleases

2.2.1.1 Enzyme

Age1	New England Biolabs
BamH1	Promega
Nhe1	Promega
Xho1	Promega

2.2.2 Bacterial Cell Culture

2.2.2.1 Antibiotics

The table shows the antibiotics used and their final concentrations:

Antibiotics	Final concentration of antibiotic
Carbenicillin	100µg/ml
Kanamycin	10µg/ml
SURE cells	Genotype; e14-(McrA-) Δ(mcrCB-hsdSMR-mrr)171 endA1 gyrA96 thi-1 supE44 relA1 lac recB recJ sbcC umuC::Tn5 (Kanr) uvrC [F′ proAB lacIqZΔM15 Tn10)

2.2.2.2 Media

The table explains the different bacterial growth media used in this project:

Media	Formula	Supplier
SOC medium	0.5% (w/v) yeast extract 2.5mM KCl 1mM $MgSO_4 \cdot 7H_2O$ 1mM $MgCl_2 \cdot 6H_2O$ 10mM NaCl 2mM glucose 2% (w/v) tryptone	Invitrogen
Luria- Bertani (LB) medium	1% sodium chloride, 1% tryptone and 0.5% yeast	In-house

2.3 Protein Analysis

Acetic acid	Fisher chemical
Ammonium persulfate	Merck
Ammonium Sulphate	Sigma-Aldrish
Avidin-HRP	Biolegend
30 % Acrylamide/Bis solution (Geneflow)	National diagnostic
Benzamidine	Sigma-Aldrish
Roller bottle (2L)	Greiner Bio-one
Bovine gamma-globulin	Sigma
Bovine Serum Albumin (BSA)	Sigma
Bradford reagent	Bio-Rad
[BVD13-3A5] Purified anti-human GCSF	Biolegend
[BVD11-37G10] Biotin anti-human GCSF	Biolegend
Coomassie Blue reagent	Sigma-Aldrich
Chemiluminescence Blotting substrate	Roche
Cuvettes (polystyrol/polystyrene	SARSTEDT
EDTA	BDH
Ethanol	Fisher Scientific
ECL western blotting detection system	Amersham-Pharmacia
Imidazole	Sigma
2X Laemmli sample buffer	Bio-Rad
Fuji Medical X-ray Film	Fuji
Glacial acetic acid	Fisher
Glycine	Sigma-Aldrich
Glycerol	Fisher Scientific
Vortex mixer	Stuart
Traceable Nano Timer	Fisher Scientific
Methanol	Merck
Minispin centrifuge	Eppendorf
Mixing Table	Biotek instrument Inc
$NaHCO_3$	Sigma
NaN_3 [AnalR]	BDH

Nickel Chloride	Sigma
PBS Tablets	OXOID
Plate Reader (450nm/630nm)	Yellowline OS10 Basic
Polyvinylidene diflouride membranes	Amersham-Pharmacia
96 well ELISA plates	Costar
Rat Serum	Sigma
Recombinant human GCSF @ 200µg/ml	Biolegend
Sodium acetate (anhydrous)	Fisher Scientific
Sodium hydroxide	BDH
Sodium Lauryl Sulfate (SDS)	Sigma-Aldrich
Sodium phosphate monobasic	BDH
Sodium phosphate dibasic	BDH
Spectrophotometer	Unicam
Sprague male Dawley Strain rats (~300 gram)	Royan Institute (Esfahan)
Sulphuric Acid	
Tetramethylethylenediamine (TEMED)	Merck BDH
TMB Solution	Sigma
TRIS Base	Sigma
Tris-Hcl	Sigma-Aldrich
Tween 20	Sigma-Aldrich
Vivaflow 200 concentrator	Sartorius Stedin
1ml Blue Sepharose column	PIERCE
1ml IMAC HP column	GE Healthcare
3510 pH Meter	JENWAY

2.3.1 Proliferation Assay

AML-193 cell line	(ATCC; Cat. #: CRL-9589; Lot #: 3475266)
Iscove's Modified Dubellco Media (IMDM)	Gibco
L-glutamine	Gibco
Penicillin/Streptomycin	Gibco
Foetal Calf Serum	Labtech
Transferrin	Sigma
Insulin	Sigma
GMCSF (5ng/ml)	Biosource
CellTitre 96 AQueous Proliferation Assay Reagent	Promega
96-well Cell Culture Plate	Costar

3. General Methods

3.1 Preparation of Luria-Bertani (LB) Media

LB media comprised of 1% (w/v) sodium chloride, 1% (w/v) tryptone and 0.5% (w/v) yeast dissolved in MilliQ water and then autoclaved to fully dissolve these components as well as sterilize the media.

3.2 Preparation of Agar Plates

Agar plates were made up by adding granulated agar to LB media at a concentration of 1.4% (w/v) and autoclaved. Appropriate antibiotics (carbencillin at 100µg/ml) were added to the agar plates when cooled to approximately 60°C. Agar plates were allowed to set and dry.

3.3 Preparation of Chemically Competent Cells

Replicating eukaryotic DNA in E. *coli* can be difficult because eukaryotic genes might incorporate inverted repeats or secondary structures, such as Z-DNA, that can be deleted or rearranged by conventional E. *coli.* The SURE (Stop Unwanted Rearrangement Events) cell is particular strain of E. *coli* and has a unique advantage over conventional E. *coli* commonly used, as it is void of essential components of the pathways that hinders cloning of eukaryotic DNA in conventional E. *coli* (largely due to induction of DNA rearrangement and deletion of nonstandard DNA fragments). It was used in this study to allow for cloning of multiple repeat DNA sequences.

To produce competent E. *coli* cells with the ability to accept foreign DNA, the E. *coli* SURE cells were plated out on agar plate containing selective antibiotic (Kanamycin at 10µg/ml) and incubated overnight at 37°C. The next day, a single isolated colony was selected and grown in LB media containing Kanamycin 10µg/ml during the day in an orbital shaker at 37°C. A dilution of 1/1000 was seeded into 50ml LB media and grown continued overnight (37°C) at 200rpm in an orbital shake. The following day, 50ml of overnight culture was transferred into 500ml warmed LB and grown until an optical density (measured at 600nm) of approximately 0.9 was reached.

Cells were centrifuged at 2397 x g for 30 minutes at 4°C in a Beckman Avanti J201centrifuge. The supernatant was discarded and pellet resuspended in 200ml ice cold sterile 0.1M $MgCl_2$ and centrifuged again using previous parameters. Following removal of supernatant, the cell pellet was resuspended in 10ml ice cold sterile 0.1M $CaCl_2$ followed by a further addition of 90ml 0.1M $CaCl_2$ lower case and left on ice for 1 hour (hr) , then centrifuged as before. The supernatant was discarded and cells resuspended by gentle swirling or pipetting in 25ml sterile 15% glycerol/85mM ice cold $CaCl_2$. Finally, cells were aliquoted into 1.5ml pre-chilled sterile eppendorfs on ice and immediately flash frozen in liquid nitrogen then stored at 80°C.

3.4 DNA Cloning of GCSF Tandems for Expression

3.4.1 Polymerase Chain Reaction (PCR)

In this project, the PCR was used for DNA fragment amplification and to introduce restriction enzymes sites into the expression constructs for ease of cloning. The GCSF molecule was PCRed from pGCSFSecTag4A1 (GCSF fusion protein construct containing full length GCSF and its signal sequence which was available in the laboratory) using a forward primer and reverse primer giving a final amplicon of ~600bp (all primers used in this study are listed in Appendix A.1). The PCR reactions were set up as a single step reaction utilizing two-master mixes as highlighted in the Table 3.1. A negative control reaction that contained primers but no template was also included. Two GCSF molecules were PCRed in this project, GCSF and GCSF with signal sequence.

Table 3-1: PCR reaction utilizing two-master mixes

Master Mix 1	Volume	Master Mix 2	Volume
Template DNA (plasmid)	1µl (100ng)	10x polymerase buffer + $MgCl_2$ (1.5mM)	5µl
Primer 1 (10pmol/µl = 10µM)	1µl	Sterile water	19.15µl
Primer 2 (10pmol/µl = 10µM)	1µl	Expand Polymerase	0.85µl
dNTPs (10mM)	1.25µl		
Sterile water	20.75µl		

Using the TC-312 PCR thermocycler, master mix one and two were combined and cycled as follows:

Table 3-2: PCR stages

Stage	Time	Temperature
Denaturation	**2 min**	94°C
25 cycles @	**30s**	94°C
	30s	54°C
	30s	72°C
Extension	10min	72°C

Thereafter, the PCR reactions were loaded on 1% (w/v) TAE agarose gel. A 1 Kb DNA ladder runs alongside the samples during electrophoresis at 100 V for 30 minutes using Bio-Rad PowerPac™ HC and gel isolated (see section 3.1.5.2) Gene Snap software from SYGENE was used for the gel Imaging.

3.4.2 DNA Isolation from Agarose Gel Electrophoresis

A 1% (w/v) agarose gel was prepared by dissolving powdered agarose in TAE buffer (1mM EDTA, 20mM acetic acid and 40mM Tris base, pH 8.5) upon heating for 1 minute will not dissolve otherwise. 0.5µg/ml Ethidium bromide was used as staining agent. 1kb DNA ladder (0.5µg/µl) was run on each gel as standard. The agarose gel was run at 100V using the Bio-Rad Power PC.

To prevent damage to the amplified DNA for the purification step, 10% of the samples (PCR reaction) were loaded in an inner well for visualization with ultraviolet (UV) light using the Gene Snap software from SYGENE, while the remaining samples were loaded in a peripheral lane. This was done to prevent potential damage to the remaining 90% DNA samples (in the peripheral lane) by UV light during visualization. The peripheral lane of the gel was separated from the rest of the gel using a clean razor blade prior to UV exposure. After visualization and excision of the DNA fragment of interest in the main gel, the corresponding region on the peripheral lane was then cut out from the agarose gel, which was thereafter extracted using GenElute extraction kit as per manufacture's protocols in order to get rid of the PCR contaminates. DNA samples were eluted by the addition of 50µl of buffer EB (10mM Tris-HCl, pH 8.5) and quantified at a wavelength of 260nm using the Nanodrop spectrometer ND100V. Samples were then ready for ligation into plasmid or stored at -20°C to preserve DNA.

3.4.3 Single and Double Restriction Enzyme Digests

Restriction enzymes are routinely used during cloning to transfer a gene of interest into a vector plasmid. Often, both the vector plasmid and the target gene are digested with the same restriction enzymes to generate similar overhangs. The target gene to be expressed, often referred to as insert, is thereafter ligated to the vector plasmid using DNA ligase and polymerase.

Restriction Digestion of Insert and Vector DNA

In this project, single or double restriction digestions (involving use of one or two restriction enzymes respectively) were set up for both the insert (PCR product of target DNA) and the vector DNA (i.e. the expression plasmid DNA) as stated below in Table 3.3 and 3.4 respectively.

1- Digestion of insert

For the digestion, the volume of DNA used was dependent on the sample concentration. The volume was made up to 50µl with sterile MilliQ water. The reaction mixture was incubated for 2 hrs in a thermostatic water bath at 37°C (all restriction enzymes and restriction enzyme buffers used in this study are given in Appendix A.2).

Table 3-3: Single step double digestion of Insert

Undigested DNA (PCR fragment)	500ng
10x restriction Buffer	5µl
Restriction enzyme 1 (10units per 0.5-1µg of DNA)	1µl
Restriction enzyme 2 (10units per 0.5-1µg of DNA)	1µl
ac BSA (1mg/ml)	5µl
Sterile MiliQ water	Xµl
Total volume	50µl

After incubation, samples were centrifuged at 11,337 x g for 5 seconds to remove condensation and then analyzed using 1% agarose gel electrophoresis Reaction products were purified from agarose gel using QIAquick PCR purification kit as per manufacture's protocol. Purified elutions were quantified using the Nanodrop spectrometer ND 100 at 260nm.

2- Digestion of plasmid DNA

The plasmid DNA was digested in a two-step double digestion. The double digestion was carried out by setting up single restriction digestions as stated below and incubated for 2 hrs in a thermostatic water bath at 37°C. The volume was made up to 10µl with sterile MilliQ water.

Table 3-4: Two-step double digestion of Plasmid DNA

	Single digest with enzyme 1	**Single digest with enzyme 2**
Plasmid	1µl (500ng)	1µl (500ng)
10x restriction Buffer	1µl	1µl
Restriction enzyme 1 (10units per 0.5-1µg of DNA)	1µl	-
Restriction enzyme 2	-	1µl
ac BSA (1mg/ml)	1µl	1µl
Sterile MiliQ water	6µl	6µl
Total volume	10µl	10µl

After incubation, 2.5µl aliquots from each single digest were taken for further analysis to assess the integrity of the restriction enzymes used. The remainder volumes in the single digests (7.5µl) were pooled together and the following reagents were added (2µl acBSA, 2µl 10x restriction Buffer, 1µl Restriction enzyme 1, 1µl Restriction enzyme 2 and 14µl Sterile MilliQ water) to give a final total volume of 35µl. The reaction mixture containing both restriction enzymes was then incubated for 1-2 hrs in a thermostatic water bath at 37°C.

1% agarose gel electrophoresis was used to analyse samples from the single and double digestions as described previously in section 3.4.2. When required**,** reaction products were purified from agarose gel using QIAquick PCR purification kit as per manufacture's protocol. Elutions were quantified using the Nanodrop spectrometer ND 100 at 260nm.

3.4.4 General DNA Ligation

Fragments and linkers were ligated into the pSecTag plasmid to create mammalian expression vectors. Two different controls (one without insert and, another without insert and DNA ligase) were used in each reaction to ascertain the success of DNA ligation. For each ligation, the amount of insert (I) used was calculated using the formula below from the Promega website. A molar ratio of insert to plasmid (P) of 3:1 was used.

$$\text{ng of insert (I)} = \frac{(\text{ng of P} \times \text{Kb size of I})}{\text{Kb size of P}} \times \text{molar ratio of I:P (3/1)}$$

The volume of MilliQ water was added to compensate for the variation in volume of fragment (insert) to achieve a total reaction volume of 10µl. The Plasmid (P) typically used in this method was around 50ng.
Ligation reactions were set up as shown in the table below and left overnight at room temperature.

Table 3-5: Preparation of ligation reactions

	(Experimental ligation)	**Control 1**	**Control 2**
PCR fragment (insert) or linkers	Variable	-	-
Plasmid vector	X (~50ng)	Xµl	Xµl
10x ligase buffer	1µl	1.5µl	1.5µl
MilliQ water	Variable	X µl	X µl
T4 DNA ligase (3U/µl)	1µl	1µl	-
Total volume	10µl	10µl	10µl

The following day, the ligation reactions were subsequently transformed into SURE chemically competent E.*coli* cells, for vector amplification and initial assessment of ligation success.

3.4.5 Transformation of Plasmid into Chemically Competent E.*coli*

To determine if the target gene of interest or linkers has been integrated into the vector plasmid DNA, the expression plasmid constructs were transformed into competent E. *coli* cells. To transform the cells, 200µl of the chemically competent E.*coli* cells were added into each eppendorf and a suitable volume of the plasmid of interest added (no more than 10% total volume of cells), all this was carried out on ice. Cells were gently mixed and incubated on ice for 15 minutes. Heat shocking at 42°C for 1 minute was used to enhance the bacterial cell wall permeability, allowing uptake of plasmid. After that, cells were immediately chilled on ice for an additional 5 minutes. Under sterile condition, 800µl SOC media was added to each eppendorf of cells and incubated for a further 40-45 minutes at 37°C. After incubation, cells were centrifuged for 5-10 minutes at 1073 x g using the bench top centrifuge at room temperature. The supernatant was discarded and cell pellet resuspended in 100µl SOC media. Transformed cells were plated out on LB agar plates containing 100µg/ml carbenicillin at two dilutions; 10% (1 tenth cells diluted in 90µl SOC media) and 90% (9/10 of cells) to accommodate any potential overgrowth of colonies, and allow isolated colony selection. Plates were incubated overnight at 37°C. Next day, colony counts were carried out for each experimental ligation and both controls to evaluate whether DNA ligations had been successful.

3.4.6 Plasmid Preparation and Glycerol Stocks

Small-scale copies of expression plasmid are generally made to propagate transformed cells from ligation reactions and for initial analyses of expression constructs for protein expression in mammalian cells. A number of colonies from the previously transformed competent E. *coli* cells containing the expression plasmid were selected and seeded into 1ml LB broth containing 1µl of 100mg/ml carbenicillin. Colonies were then incubated during the day at 37°C in orbital shaker (200rpm). Thereafter, a 1/1000 (5µl) dilution of each was seeded into 5ml LB broth with 100µg/ml

stock carbenicillin and incubated overnight. On the following day, 200µl of 50% (v/v) glycerol (important for long-term storage of bacteria, as it prevents the formation of ice crystals that can damage the cell wall) were added to 800µl of each clone and stored at -80°C. Up to 5ml plasmid DNA from overnight cultures were purified using a Qiagen plasmid mini preparation kit as per the manufacturer's procedure. The concentrations of purified plasmid were quantified using a Nanodrop spectrometer ND100V at a wavelength of 260nm.

3.4.7 Screening of Potential Clones from Ligations

For initial screening of the clones so as to identify potential positive clones with insert prior to sending for sequencing, plasmid samples from mini preps were double-digested with the restriction enzymes that were used for the cloning. Plasmid digestion was confirmed using agarose gel electrophoresis as outlined before at section 3.4.3. Positive plasmids were sent for DNA sequencing.

3.4.8 Plasmid Sequence Analysis

All plasmid samples with an appropriate forward and reverse primers were sent to the Department of Core Genomic within the Sheffield University School of Medicine to confirm that DNA cloning had generated a vector of desired nucleotide sequence. The genomic sequence results were analyzed using SeqMan (DNASTAR Lasergene software).

3.5 Cell Culture and Protein Expression

3.5.1 Growth and General Maintenance of CHO Flp-In Cells

Mammalian CHO Flp-In cells were used for protein expression of the tandem GCSF constructs, which are the subjects of our studies. From the liquid nitrogen stock of untransfected Chinese hamster ovary cells (CHO Flp-In), a vial (at 2-3 x 10^6 cells/ml) was removed, thawed and grown in growth media (Ham's/F12 DMEM media containing 10% Fetal Calf Serum (FCS), 2mM L-glutamine, and 100µg/ml Streptomycin / Penicillin and 100µg/ml

Zeocin) and were grown in a 5% CO_2 incubator at 37°C.

For passaging, the adherent cells were removed by the addition of trypsin/EDTA and incubated 2-3 minutes. The resulting suspension was then diluted with complete medium transferred to a sterile 30ml universal and centrifuged for 5 minutes at 67 x g to remove supernatant. The cell pellet was resuspended in the appropriate media. Cells were maintained at approximately 80-90% confluence in a T-75 flask. At this density the cells are routinely passaged at 1 in 10 dilution.

3.5.2 Trypan Blue Exclusion Method

Trypan blue is used to stain cells to differentiate between live and dead cells during cell counting. The dead cells exclusively take up the blue dye and thus appear blue when visualized under a light microscope. This distinguishes them from the live cells, which appear bright. This characteristic of dead cells differentiation is particularly important when assessing the viability of the cells culture. Using heamocytometer (used to count cells number in a specific volume of fluid to give an approximate number of cells in the whole fluid), Trypan blue was applied at 1:1 dilution with media sample. The total number of cells was calculated per ml using the below equation:

$$\text{Total cell number per ml} = \frac{\text{total cell number} \times \text{Dilution factor} \times 10^4}{\text{No of squares}}$$

3.5.3 Transient Expression of GCSF Tandems in CHO Flp-In

All GCSF constructs (plasmids) were transiently transfected into CHO Flp-In cells as an initial screen for protein expression prior to development of a stable cell line.

CHO Flp-In cells were grown in antibiotic free media (Ham's/F12 DMEM media containing 10% Fetal Calf Serum (FCS), 2mM L-glutamine, and 100µg/ml Streptomycin / Penicillin). At approximately 70% confluence, cells were removed from the plate by the addition of trypsin/EDTA and

reseeded at a density of 0.2 x 10^6 cells/well into a 24 well plate on the day prior to transfection. Cells were allowed to grow overnight in antibiotic free media. The following day, media was replaced with 500µl antibiotic free growth media.

A transfection reaction mix was prepared using a Fugene-6 transfection reagent (transfection reagent designed to transfect plasmid DNA into multiple cell lines with low toxicity and high efficiency) to experimental DNA ratio of 3:2, in a volume of 100µl serum free growth media (Ham's/F12 DMEM media containing 2mM L-glutamine, and 100µg/ml Streptomycin / Penicillin). Prior to use, Fugene-6 was equilibrated to room temperature and vortexed. Then, 6µl of Fugene-6 was added to the reaction mix, being careful to avoid contact with the plastic sides of eppendorf, and followed by the addition of 4µg of plasmid. Transfection with a known plasmid that expresses well (pGCSFSecTag4A1) was used as a positive control. Non-transfected cells were used as a negative control in each experiment. The transfection mixes were gently flicked to mix and incubated at room temperature for 15 minutes. Each transfection mix was pipetted drop-wise into its corresponding well and gently rocked to ensure an even distribution. All plates were incubated at 37°C with 5% CO_2 for 72 hrs prior to harvesting the media. Finally, the media for each transfection (containing secreted product) was centrifuged and supernatant was transferred to a clean universal tube for protein expression analyses using any available convenient method (western blot or Elisa).

3.5.4 Generation of Stable CHO Flp-In Cell lines

The expression for GCSF tandems was enabled using the Invitrogen CHO Flp-In system. The CHO Flp-In system was chosen since it enabled rapid integration of a GOI into a specific site within the host genome for high expression.

The CHO Flp-In host cell line has a single Flp recombinase target (FRT) site located at a transcriptionally active genomic locus and is resistant to the antibiotic zeocin. The plasmid (A modified version of pSecTag-V5/FRT-Hist:

see Appendix C) contains the gene of interest as well as an FRT site and has the hygromycin B resistance gene.

Stable cells are generated by co-transfection of the plasmid containing the gene of interest (GOI) and another plasmid, pOG44 (a 5.8 kb Flp-recombinase expression vector responsible for Flp-recombinase expression) into the Flp-In cell line.

The expression of Flp recombinase results in the integration of the GOI into the genome via the FRT site. Stable cells are then selected using hygromycin B, thus the need for clonal selection is not necessary as integration of the DNA is directed. The process of culturing the Flp-In cell line was conducted as per manufacture's protocol using basic cell culture techniques.

The day before transfection cells were removed from a T75 flask by the addition of trypsin/EDTA and reseeded into 6 well plates at a concentration of 0.25 x 10^6 cells per well in a total volume of 2ml. Cells were left overnight to achieve about 60-70% confluency.

Transfections were carried out the next day. Briefly, a number of sterile eppendorf tubes each containing 92.5µl serum free media (without antibiotics) were set up. To each tube, 7.5µl of Fugene-6 was slowly dropped onto the surface of media and flicked to mix. In separate sterile eppendorfs, 250ng of plasmid of interest was mixed with 5µg of the pOG44 plasmid and pipetted into the tube containing Fugene-6 mix and contents mixed gently by flicking. Tubes were incubated at room temperature for 15 minutes. Fugene-6 only was used as a negative control and a known expression plasmid (pGCSFSecTag4A1) was used as a positive control. All transfection mixes were carefully pipetted drop-wise onto cells in each labelled individual well of a 6 well plate, and incubated at 37 °C, 5% CO_2 for 24-48 hrs.

At 24-48 hrs post-transfection, the culture medium in each well was removed and replaced with 2ml growth medium (Ham's/F12 medium containing 10% FCS, 100µg/ml Streptomycin /100IU/ml Penicillin, 2mM L-glutamine) containing 600µg/ml Hygromycin B antibiotic as selective reagent. Cells were then allowed to grow until desired confluency with media replacement every 2 days. Routine observations were made as to the cells appearance, with successfully transfected cells appearing fibroblastic and growing out in clumps, whereas, dead cells appeared spherical in solution. Cells were not allowed to get too confluent, as the antibiotics would become ineffective. When plates had a number of individual colonies growing out they were removed from the wells by the addition of 0.5ml Trypsin-EDTA (T/E) and centrifuged at 67 x g for 5 minutes The supernatant was discarded and cell pellet resuspended in 5ml transfected growth media and grown in T-25 flasks until cells became nearly confluent with media change every 2 days. Once confluent cells were removed and transferred to T-75 flasks with the same growth media in a total volume of 12-15ml. Cells were considered stable when all control cells were dead and cells were dividing normally. This took normally 2-4 weeks.

3.5.5 Analysis of Crude Media from Transfected Cells Lines

For protein expression analysis of stable & transiently transfected cells lines, it was required to serum starve cells prior to analysis. T-75 flasks containing stable cells were grown until an appropriate confluency was reached. Thereafter, cells were serum-starved: Media was removed from the T-75 flasks and replaced with 12-15ml serum free media and incubated for 2-3 days at 37 °C, 5% CO_2. After incubation media was transferred to clean 30ml universal tube and centrifuged to remove any cellular debris. The supernatant of each sample was then carefully removed to another sterile universal tube and stored at 4°C until required for protein expression analyses using any available convenient method (western blot or Elisa).

3.5.6 Storage of Stable Cell Lines in Liquid Nitrogen

Stable cell lines were grown to confluency in T-75 flasks at which point cells were removed using an appropriate volume of trypsin/EDTA. Once cells were detached an appropriate volume of serum growth media was added to each flask to neutralise Trypsin/EDTA. The suspended cells for all flasks were then pooled together in a clean 30ml universal tube and centrifuged at 67 x g for 5 minutes to pellet. The supernatant was discarded and cells were resuspended in a freezing mixture (foetal calf serum/DMSO mixture are ratio 9/1) to make a final concentration between 2-4 x 10^6 cells/ml. Once cells were resuspended, immediately 1ml portions were aliquoted to clean cryogenic freezing vials labelled with construct name, cell line, date and cell number. Cryogenic vials were finally placed in a polystyrene freezer box surrounded with cotton wool and stored at -80°C. This step is important to freeze cells gently and, prevent ice crystal formation within the stable cells. Later, cells were transferred to liquid nitrogen for long time storage.

3.5.7 Adaptation of Stable CHO Cells to Hyclone Media

The stable cells taken from liquid nitrogen were thawed and grown in T75 flask growth medium for 1-2 days until an appropriate confluency. Cells were resuspended in Hyclone SMF4CHO Utility media (no Hygromycin B) for adaption and media was change every 2-3 days.

3.5.8 Expression of GCSF Tandems in Roller Bottle Culture

After growing to the maximum level in T75 flask, stable cells were transferred to 2 x 1 litre roller bottles (maximum volume 500ml) at a starting density of ~0.25x10^6/ml (each roller bottle contained about 500ml). Cells were grown at 37°C 5% CO_2 with mixing (Cell Roll Cellspine Control Unit was used to mix the stable cells at 4rpm) until viability was ~30%, with samples regularly taken every 2 or 3 days to assess protein expression and cell viability. Thereafter, the cells were centrifuged at 14,981 x g (JLA 16.250) for 30 minutes at 4°C to clear cellular debris and 10mM final concentration of Benzamidine HCl (serine protease inhibitor) was

added to the 1L media sample to prevent protein degradation. The media sample was then concentrated using a vivaflow concentrator (See section 3.6) and stored at -40°C until required. Samples were analysed by western blotting and Elisa.

3.6 Vivaflow 200 Concentrator

Vivaflow 200 concentrator was used to concentrate media sample to 10x less volume (i.e. from 1L to 100ml). This was important to save time and make the process of purification easier. The Vivaflow 200 concentration system comprised of a membrane module with 10kDa molecular weight cut-off, inlet and outlet tubes, and pump (Masterflex pump). Before starting the process, the tubes were cleaned with 0.5M NaOH followed by washing with 500ml of deionized water with the filtrate going to the waste. The media sample was then circulated through the system for approximately 3-4 hrs (the media sample was placed on ice during concentration to avoid protein degradation), and the culture volume concentrated down to 10 times less volume within this period, followed by storage at -40°C until ready for purification. After each concentration cycle, deionized water was flushed through the system with the filtrate going to the waste. The system was cleaned by recirculating 250ml of 0.5M NaOH through the system at 100ml/min for 30-40 minutes. Finally, the system was drained and recirculated with 500ml deionized water for 5-10 minutes with the filtrate going to the waste. For storage, the module was filled with 20% ethanol and refrigerated at 4°C.

3.7 Protein Purification

3.7.1 Purification of GCSF Tandems Using IMAC

Tandem proteins were designed with a His-Tag at the C-terminus. The His-tag in each tandem facilitates easy purification of the desired protein when passed over a metal chelate column. The high affinity of histidine for nickel ions allows the protein to bind to the Immobilized Metal Affinity Chromatography (IMAC) column while most of other proteins remain

unbound. The column was washed with a solution of moderate ionic strength to dissociate proteins that may have bound to the column non-specifically. The target protein with His-tag is released from the IMAC by adding high concentrations of histidine analogues (e.g. imidazole) to compete with his-tag for nickel binding. As a result, the eluted target protein with his-tag could be separated from any protein contaminants (Block et al., 2009). The separated protein fraction undergoes dialysis in PBS at 4°C for 2hrs and then overnight (after PBS buffer change). The dialysed protein was stored at -80°C. Step-wise detail for the purification step is provided below.

Before applying the media sample to the 1 ml IMAC column (GE Healthcare), all tubing was cleaned with 0.2M NaOH and then rinsed with H_2O prior to use. A fresh IMAC column was charged with 4mg/ml Nickel chloride and equilibrated with equilibration buffer (20mM NaP_3, pH 7.4, 0.5M NaCl and 10% glycerol). Media sample was defrosted and centrifuged at 30,910 x g (JA-25.5) for 30 minutes at 4°C to clarify. Media sample was diluted 1:1 with 40mM NaP_3, pH 7.4, 1M NaCl, 20% glycerol and 20mM imidazole. Media sample was then filtered (using Millipore 0.22um filters) and loaded on to the IMAC column at 2ml/minute (sample contained 10mM imidazole) at room temperature or sometimes 4°C to avoid protein degradation. The unbound fraction (flow through) was collected and the column washed with 10 column volumes (CV's) equilibration buffer with the addition of 10mM imidazole, followed by 10 CV's with the addition of 10mM imidazole at 2ml/minute. These two washes were used to remove contaminants.

Different concentrations of imidazole were made up in 20mM sodium acetate, 0.5M NaCl, pH 6.0, 10% glycerol (pH 6.0 buffer) as shown in Table 3.6 were used to elute the bound protein (Target protein) from the IMAC column using a step elution method. 1ml fractions were collected (3 x 1ml per concentration). The eluted proteins were analysed by SDS-PAGE/Bradford assay and the relevant fractions were pooled and dialysed against at least 1L PBS buffer at 4°C for 1, 2 hrs and overnight (A fresh change of PBS buffer was used per dialyse) to remove all salts and

imidazole. At the end of each purification, the IMAC column was washed with 0.2M NaOH followed by 10 CV's of water and stored in 20% ethanol.

Table 3-6: Concentrations of imidazole

Imidazole concentration (mM)	Imidazole (ml)	pH 6.0 buffer (ml)
20	0.2	4.8
50	0.5	4.5
100	1	4
200	2	3
350	3.5	1.5
500	5	-

3.7.2 Purification of GCSF Using Cibacron Blue Sepharose

The stable cells of native GCSF (ligand only) were taken from liquid nitrogen, thawed and grown in T75 flasks as described in section 3.5.7 and 3.5.8 and were thereafter transferred to 4 big roller bottles (maximum volume 1 liter) at a starting density of ~$0.5x10^6$/ml (each roller bottle contained about 500ml). Cells were grown at 37°C 5% CO_2 until viability was ~30%, with samples regularly taken every 2 or 3 days to assess protein expression and cell viability. Because native GCSF has no His-tag, a Cibacron Blue Sepharose (www.gelifesciences.com) column was used to purify. This dye ligand chromatography resembles native substrate which proteins have affinity for. When proteins pass through this column, the goal of this dye is to bind the target protein and expel all unbound proteins. The bound protein could then be eluted by changing the buffer composition, often by increasing the buffer pH.

Media sample of recombinant human GCSF (rhGCSF) was placed on ice and precipitated with 35% (w/v) (20.11g/100ml) ammonium sulphate ($AmSO_4$). The solution was then incubated at 4°C with mixing for 1hr. The resulting suspension was centrifuged at 43589 x g, 4°C for 30 minutes and the supernatant was kept to check if it contains rhGCSF (1^{st} supernatant).

The pellet was solubilized in 10ml 20mM TRIS buffer pH 7.0. On ice, the pH of the media was dropped from 7.0 to 5.0 by adding 10ml 100mM sodium acetate buffer and incubating on ice at 4°C for 30 minutes. The resulting suspension was centrifuged at 43589 x g, 4°C for 30 minutes and the supernatant was kept (2nd supernatant). A 1ml Blue Sepharose column (GE Healthcare), was first equilibrated with 10 column volumes (CV's) of 20mM sodium acetate, pH 5.0 and then supernatant was loaded onto the 1ml column at room temperature and unbound fraction collected (Flow through). The column was washed with 5 CV's sodium acetate buffer, pH 5.0 buffer followed by 3 × 5ml (3 × 5 CV's) 20mM Tris buffer pH 7.0. The column was finally washed with 5 × 2 ml 20mM Tris, 1mM EDTA at pH 8.0 (the target protein was observed to elute at a high concentration in this step). Eluted proteins were stored at -80°C. Finally, the column was washed with 5 × 2ml of 20mM Tris, 1mM EDTA and 1M NaCl, pH 8.0 to clean the column and stored at 4°C. When required, samples were analysed by coomassie stain, western blotting and Elisa.

3.8 Analysis of Protein

3.8.1 Bradford Protein Assay

The Bradford protein assay was routinely used to measure the concentration of our purified tandem proteins and rhGCSF in a solution. A 10mg/ml solution of BSA was prepared in double distilled water and diluted to 1mg/ml, then to 100µg/ml by 10-fold dilution (i.e. 1ml BSA + 9ml ddH2O). The standards below were prepared from this 100µg/ml (Table 3.7).

Table 3-7: Standard curve preparation

Standard BSA (µg/ml)	Final concentration in assay (µg/ml)	Dilution	
25	20	x4	2ml 100µg/ml + 6ml ddH2O
12.5	10	x2	2.5ml 25µg/ml + 2.5ml ddH2O
6.25	5	x2	2.5ml 12.5µg/ml + 2.5ml ddH2O
2.5	2	x2.5	2ml 6.25µg/ml + 3ml ddH2O
1.25	1	x2	2ml 2.5µg/ml + 2ml ddH2O

Standards and Unknown samples were prepared in duplicate as follows:

0.8ml of each standard was pipetted into separate1.5ml eppendorf tubes and 0.2ml dye reagent was then added, mixed gently and incubated for 5mins at room temperature. Unknown samples were diluted appropriately into a final volume of 0.8ml ddH_2O. 0.2ml dye reagent was then added, mixed gently and incubated for 5mins. All standard and unknown samples were transferred into plastic disposable cuvettes before reading in spectrophotometer at 595nm.

3.8.2 Analysis of Proteins by SDS- PAGE

3.8.2.1 Preparation of SDS-PAGE Gels

Bio-Rad electrophoresis tank and gel apparatus were used for the analyses of expressed tandem proteins as per manufacture's protocol. A 1.0 mm thick discontinuous 10% gel was made by mixing 3.3ml of 30% Acrylamide mix (0.8% (w/v) bis-acrylamide stock solution (37.5:1 ratio)), 3.6ml sterile MilliQ water, 2.5ml resolving buffer (1.5 M Tris-HCl. O.4% SDS (w/v), pH 8.8), 5µl of tetramethylethylenediamine (TEMED), 100µl of 10% (w/v) ammonium persulfate (APS). The mixture was poured into the gel plate to fill about 4/5th of the plate's height while 200µl isopropanol was poured to the top of the gel to create a smooth level and bubble-free meniscus. The gel was allowed to set at room temperature for approximately 30 minutes. Afterwards, the isopropanol was poured off and the excess isopropanol was

removed with a filter paper. A 5% stacking gel was prepared for the second layer by combining 830µl of 30% Acrylamide (0.8% (w/v) bis-acrylamide stock solution (37.5:1 ratio)), 3.45ml sterile MilliQ water, 630µl of stacking buffer (0.5 M Tris-HCl, 0.4% SDS (w/v), pH 6.8), 5µl of TEMED, 50µl of 10% APS and the mixture was poured onto the top of the resolving gel up to the top of the gel plate, and a 1.0 mm well comb was immediately inserted into the top of the stacking gel, avoiding air bubble formation. The gel was allowed to set at room temperature for 15 minutes.

3.8.2.2 Preparation of Samples for SDS-PAGE

Following quantification of the proteins, 0.5ml eppendorf tubes were labeled and the final volume of each sample was added to equal volume of 2x Laemmli sample buffer. Sometimes, dithiothreitol (DDT), a reducing agent was added to the mixture to a final concentration of 25mM (to minimize dimerization of samples). The eppendorf tubes were centrifuged at 11,337g for 4 seconds and incubated at 95°C for 5 minutes using the Techne Dri-Block BD-2D. Samples were carefully pipetted into wells of the 10% gel in 1x running buffer (0.25 M Tris HCl, 1.92 M Glycine and 1% SDS (w/v), pH 8.3). Samples were run against a standard molecular weight marker. The gel was initially run for 30 minutes using PS250-2 power pack at 75v until samples had passed through the stacking gel and then turned up to 100v for 1 hour. The separated proteins were visualized by coomassie staining (0.25% Bromophenol blue R-25 (w/v), 50% methanol (v/v), 10% acetic acid (v/v)), or transferred onto a PVDF membrane for further protein analyses by western blotting.

3.8.2.3 Visualized Protein Gels with Coomassie Blue

Coomassie blue is a rapid and sensitive technique for the visualization of microgram quantities of protein using the principle of protein-dye binding between dye sulfonic acid groups and positive protein amine groups through ionic interaction. SDS-PAGE gel was incubated in the coomassie blue staining solution at room temperature with gentle shaking for 30 minutes on an orbital shaker. The stain was decanted and rinsed with

deionized water. The gel was then destained (Destain solution: 10% methanol (v/v) and 5% acetic acid (v/v)). To aid the destaining process the solution was heated in a microwave oven for 1 minute at full power and gently mixed on an orbital shaker at room temperature until the desired background was achieved.

3.8.3 Western Blotting

3.8.3.1 Transfer of Proteins to PVDF Membrane

The transfer of proteins separated by SDS-PAGE onto polyvinylidene diflouride (PVDF) membrane allows specific protein analyses by western blotting. The separated proteins were routinely transferred from the gel onto PVDF membrane using a Mini-Protean 3 blotting apparatus (Bio-Rad). As PVDF membrane is hydrophobic, 30 seconds treatment with methanol was required in order to wet the membrane before equilibrating in the transfer buffer (Transfer buffer: 2g glycine and 5.8g Tris base dissolved in 1000 ml of MilliQ water). After that, 2 gauze pads, 2 filter papers, the PVDF membrane and the gel were assembled in the blotting apparatus as shown in Figure 3.1.

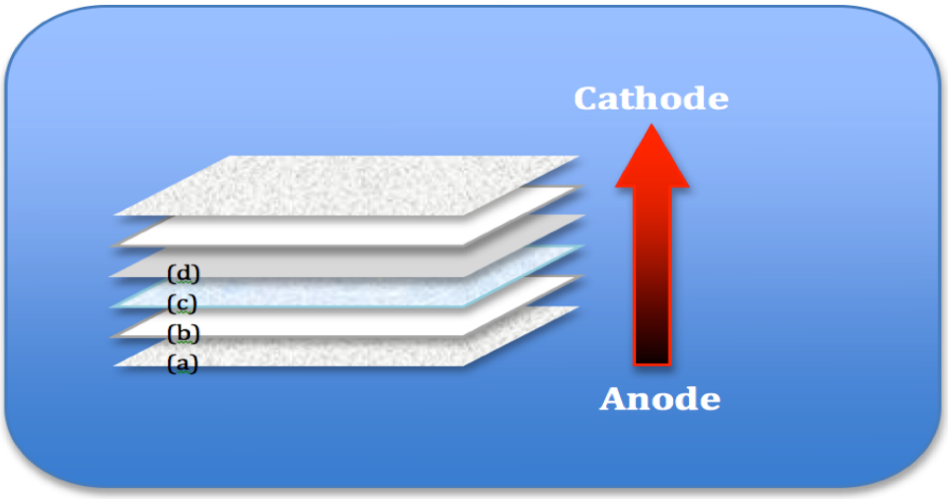


Figure 3-1: Design of transferring assembly

The scheme displays the preparation of the variant coats in the transfer buffer. In the presence of high pH buffer, the protein transfer from anode towards cathode, thus the variant coats were prearranged to permit moving of proteins into the PVDF membrane. Gauze layer (a), filter papers (b), the gel comprising the separated proteins (c) and the PVDF membrane (d).

Using the PS250-2 power pack, the system was run at 100v for 1hr. Following completion of transfer the PVDF membrane was blocked overnight at 4°C in 100 ml of 5% milk protein prepared in PBS-Tween-20 (0.05%).

3.8.3.2 Western Blotting Detection of GCSF

After transfer and blocking, the PVDF membrane was briefly washed with 50 ml PBS/Tween-20 (0.05%). The membrane was then incubated with primary rabbit anti-human GCSF antibody at a dilution of 1:5000 (2µl antibody in 10 ml of 5% blocking buffer) at room temperature on a Stuart Mini Orbital Shaker at 115rpm for 1.5 hrs. The membrane was then washed with 50 ml of PBS/T for 15 minutes. The membrane was then incubated with secondary antibody (Goat anti-Rabbit antibody linked to horseradish peroxidase (HRP) at a dilution of 1:10000 (1µl antibody into 10 ml of 5% blocking buffer)) on a Mini Orbital Shaker at 115rpm for 35 minutes. After incubation, the membrane was rinsed three times in 50 ml of PBS/T on a Mini Orbital Shaker at 65 rpm for 15 minutes per wash. Visualization of GCSF constructs was carried out using BM chemiluminescence blotting

substrate solution according to the manufacturer's instructions. Under red safety light in a dark room, a sensitive Fuji film was placed over the membrane inside a hyper cassette and exposed for approximately 10 seconds. The film was then transferred into developing solution (Kodak) until bands were detected. The film was placed into water to remove excess developing solution and then placed into fixing solution (Kodak) for approximately 1 minute. Different lengths of exposure were used to optimise the quality of the western blot. Molecular weights of the GCSF constructs were determined by comparison to loaded protein standards of known molecular weights.

3.8.4 Enzyme Linked Immunosorbent Assay (Elisa)

Elisa is a sensitive and specific method for quantification of proteins. It was used in this project to measure GCSF tandem protein in crude and purified media, and also in rat serum (*in vivo* study). The method involves the use of a specific monoclonal antibody (mAb), used to coat a microtiter plate (Capture antibody). The Ab on the plate will capture the protein of interest following the addition of the sample. The addition of secondary mAb (Detection antibody) labelled with biotin, will also bind to the protein of interest. The biotin labelled antibody allows binding of a streptavidin-conjugated enzyme. Washing was used between steps to remove any unbound proteins. A substrate is added to the reaction to produce a colour reaction that is directly proportional to the amount of bound protein. The concentration of bound protein is measured by comparison with a standard curve of known protein concentration. A standard curve ranging from 0nM to 5nM was generated from the rhGCSF working concentration (shaded in gray in Table 3.8). Similar standard curves were generated for the GCSF tandems (test proteins and controls) from purified proteins by diluting each individual stock concentrations to a starting of 5nM.

Table 3-8: Preparation of GCSF standards

Sample (µl)	LKC buffer (µl)	[GCSF] nM	Dilution
Stock @ 16000nM	-	16000	-
5 of 16000nM	495	160	100x
25 of 160nM	400	10	16x
500 of 10nM	500	5	2x
500 of 5nM	500	2.5	2x
400 of 2.5nM	600	1	2.5x
500 of 1nM	500	0.5	2x
500 of 0.5nM	500	0.25	2x
500 of 0.25nM	500	0.125	2x
500 of 0.125nM	500	0.0625	2x
500 of 0.0625nM	500	0.03125	2x
500 of 0.03125M	500	0.0156	2x
500 of 0.156M	500	0.0078	2x

Details of the initial stock concentrations for all GCSF (tandems and native) used for the standard curve are tabulated below (Table 3.9).

Table 3-9: Initial stock concentrations of rhGCSF and GCSF tandem proteins

Construct	Concentration stock (nM)
Commercial rhGCSF	500
Purified rhGCSF	16000
GCSF2NAT	10,800
GCSF2QAT	1,0000
GCSF4NAT	15,000
GCSF4QAT	5,000
GCSF8NAT	15,000
GCSF8QAT	15,100

For the ELISA assay: 96-well microtiter plates were coated with 100µl of capture antibody (BVD13-3A5) at 1µg/ml in coating buffer (0.1M $NaHCO_3$, pH 9.2) and incubated at 4°C overnight.

The next day plates were washed 3x with 230µl PBS-Tween-20 (0.05%), patted dry and blocked for 1hr at room temperature with 200µl 3% (w/v) BSA in PBS-Tween-20.

Control and unknown protein samples were prepared in LKC buffer as shown in Table 3.6. (LKC: 50ml of 0.5M Tris, 15ml of 0.5M NaCl, 50µl 0.1% Tween-20, 0.25g bovine gamma globulin, 0.25g NaN_3 & 2.5g BSA and made up to with sterile water and stored at 4°C).

Plates were washed as previously described followed by the addition of 100µl of either control or unknown protein samples to specified wells and incubated for 2 hrs at room temperature with mixing.

Plates were washed again as previously described and 100µl of detection antibody (BVD11-37G10) at 2µg/ml in LKC buffer was then added to all wells followed by incubation for 2 hrs at room temperature with mixing.

Plates were again washed as described followed by the addition of 100µl streptavidin-HRP at 1µg/ml made up in 0.5% BSA/PBS-T to all wells and incubated at room temperature for 30 minutes with mixing. Plates were finally washed 6x with 230µl PBS-T before the addition of 100µl TMB (3,3'5,5'-tetramethylbenzide liquid substrate) to all wells. Once a good colour change was observed, 100µl of 5% sulphuric acid (H_2SO_4) was added to all wells to stop the reaction. The plates were read at 450nm with background plate correction at 630nm using a Biotech FLX800 plate reader and Gen5 software. Results were analysed using Microsoft Excel and GraphPad Prism 6 used for curve fit analysis.

3.8.5 AML-193 Proliferation Assay

The biological activities of rhGCSF and GCSF tandems were evaluated using an AML-193 cell-based proliferation assay (Human acute myeloid leukemic cell line). GCSF stimulates the proliferation of the AML-193 cell line. The main objective of this method was to show that GCSF tandems could stimulate the proliferation of AML-193 cells in comparison to rhGCSF. Details of this assay method and associated cell culture development are provided below.

3.8.5.1 Growth of the AML-193 Cell Line

Cells (ATCC, Batch No. 3475266) were removed from liquid nitrogen storage and defrosted by placing into a 37°C water bath for 2 minutes. The contents of the vial were then transferred to a T-25 flask containing 4 ml of culture medium (5% FBS, 4mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, 5 µg/ml transferrin, 5 µg/ml insulin and 5 ng/ml GM-CSF in Iscove's modified Dulbecco's medium). The AML-193 cells were routinely cultured to a density of 2 x10^6 cells/ml but not exceeding 2.5 x 10^6 cells/ml (5% CO_2, 37°C). Passages were performed 2 times a week and cell density and viability was assessed by trypan blue exclusion as previously described at section 3.5.2.

3.8.5.2 AML-193 Bioassay

After a minimum of two passages, the AML-193 cells were ready for use in the bioassay. The cells were prepared for the assay by washing 3 times with 10ml PBS and the washed cells were recovered by centrifuging at 168 x g for 5 minutes. The final cell pellet was then reconstituted in the assay medium (5% FBS, 4mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, 5 µg/ml insulin, 5 µg/ml transferrin in Iscove's modified Dulbecco's medium) and cell density adjusted to 0.5 x 10^6 cells/ml.

A commercial GCSF was used to generate a standard curve for the assay. 0.2mg/ml rhGCSF (Biolegend) was reconstituted in PBS and 1% (w/v) BSA to a concentration of 10 µg/ml (500nM stock), divided into 10 µl aliquots

and stored at -80°C. On each day of assay 1 vial was removed from the frozen stock and a standard curve ranging from 0nM to 5nM was generated as shown in Table 3.10. Similar standard curves were generated for the GCSF tandems (test proteins and controls) from purified proteins by diluting each individual stock concentrations to a starting concentration of 10nM.

Table 3-10: Preparation of GCSF standard curve

Sample (µl)	Assay media (µl)	[GCSF] nM	Dilution
5ul of Stock @ 500nM	495	5	100x
220 of 5nM	220	2.5	2x
200 of 2.5nM	300	1	2.5x
250 of 1nM	250	0.5	2x
250 of 0.5nM	250	0.25	2x
250 of 0.25nM	250	0.125	2x
250 of 125nM	250	0.06	2x
250 of 0.06nM	250	0.03	2x
250 of 0.03nM	250	0.015	2x
250 of 0.015nM	250	0.008	2x
150 of 0.008nM	300	0.003	3x
150 of 0.003nM	300	0.0015	3x
150 of 0.015nM	300	0.0008	3x

50µl of each test protein was added to the wells of a 96-well microplate in triplicate. 50µl of AML-193 cells at 0.5 x 10^6 cells/ml were then added to each well with gentle agitation to mix the contents (i.e. cells suspension, standard and samples). Control wells, which contained only assay medium and cells suspension (50µl + 50µl), and blank wells, which contained only assay medium (100µl) were also set up.

For 3 days, AML-193 cells were exposed to the different concentrations of test proteins in a CO_2 incubator (5% CO_2, 37°C) and then 20µl of MTS

(Celltiter 96 Aqueous One Solution from Promega) was added to each well. Readings were taken every 40 minutes at 490nm using a Biotech FLX800 plate reader for a total of 2 hrs. The results were displayed using Microsoft Excel and analysed with Gen5 software. GraphPad Prism 6 was used for curve fit analysis.

3.8.6 Short Term Stability of GCSF Tandem Molecules

The protein stability was assessed by testing purified samples at 3 different temperatures (4°C, RmT and -80°C freeze thaw (F/T) cycles) over an 8 day period. Purified tandem proteins were taken from the -80°C and diluted to 0.8 mg/ml with filter sterile PBS and kept on ice. All manipulations were carried out under sterile conditions. Aliquots of protein at 0.8 mg/ml were placed at RmT, 4°C or -80°C in sterile 1.5 ml eppendorf tubes. Samples were taken on days 0, 1, 4 and 8 (samples taken on day zero, represent untreated sample controls) and immediately diluted with an equal volume of SDS-PAGE buffer to 0.4 mg/ml (Laemmlli buffer) and heated at 95°C for 5 minutes to denature. Samples were analysed by 10% SDS-PAGE and protein bands visualised by coomassie staining (A total of 7.5µg protein was loaded per lane) and western blotting (a total of 100ng protein was loaded per lane).

3.9 Experimental Procedure for *In vivo* Study

The aim of this protocol was to determine the pharmacokinetic (PK) and pharmacodynamics (PD) properties of rhGCSF and GCSF tandems (GCSF2NAT, 4NAT, 8NAT & 8QAT) in Sprague Dawley rats following intravenous injection and look at the effects of these constructs on the WBCs and neutrophils population.

Pre-Dose (-24 hr) samples of 300-400µl of blood sample were taken from each rat before injection with test protein. These served as control blood samples. Next day, groups of six male rats were administered a single intravenous dose of rhGCSF, GCSF tandem, or vehicle (PBS only) at 250µg per /kg protein. At the following time points of 0.5, 1, 2, 4, 8, 12, 24, 48 & 72

hrs post=injection, ~300-400µl of blood samples were collected from each rat from the tail vein under anaesthesia using isoflurane. Blood samples were centrifuged (for serum preparation), labelled and stored at -80°C. Counts of blood cells (CBCs) were performed on selected samples (-24, 12, 24, 48 and 72 hrs) using an automated coulter counter. Blood smears were fixed and stained by a routine laboratory method (H&E or Giemsa). Thereafter, all samples for analysis were transferred to the University of Sheffield for data confirmation and further analyses. Elisa was used to measure the concentration of proteins in each serum sample at 0.5, 1, 2, 4, 8, 12, 24, 48, 72 hrs post-dose.

3.9.1 Ethics

All animal experiments were approved by the local ethical committee of the University of Isfahan.

3.10 Statistical Analysis

In vitro analyses of all purified GCSF tandems were performed on GraphPad Prism 6 using a Mann-Whitney test (a non-parametric test that is often used to compare two groups that come from the same population). The statistical comparison of the tandem proteins in rat plasma for the pharmacokinetics study data were analysed using the non-compartmental method of data analysis. This involved the use of Winnonlin 6.3 PK program developed by Phoenix Certara. Significance between terminal half-life data of GCSF tandems was performed with GraphPad Prism 6 using one-way ANOVA (used to determine any significant differences between three or more groups of sample data). The pharmacodynamics studies of the GCSF tandems at selected sampling time points (-24, 12, 24, 48 and 72 hrs) were analysed with GraphPad Prism 6 using multiple T-test.

4. Results 1: Cloning and Expression of GCSF Tandems

4.1 Summary

Previous studies carried out by the Ross Group (UoS) have shown that the use of glycosylated linkers between two GH ligands to create protein tandems results in their glycosylation and an increased molecular weight whilst maintaining biological activity. This technology can be easily applied to other molecules such as GCSF. This chapter describes cloning, sequencing and expression of different GCSF constructs to produce GCSF tandems linked via glycosylated and non-glycosylated flexible linkers $(Gly_4Ser)_n$. The initial data obtained from this chapter has shown that it is possible to clone and express two GCSF molecules linked by a flexible linker $(Gly_4Ser)_n$ in mammalian cell lines. Expressed proteins were analysed by Elisa, and western blotting used to confirm the glycosylated proteins had a greater molecular weight than the non-glycosylated proteins.

4.2 Introduction

In medicine, any protein used for therapeutic purposes is not only a sequence of amino acids determined by a particular gene, but still require editing, altering of amino acids or addition of sugars. These modifications following the preliminary translation of the protein are called post-translational processes (Li and d'Anjou, 2009). Glycosylation refers to the post-translational process that covalently linking oligosaccharide to polypeptides. It is one of the most common protein modifications and more than 50% of proteins are glycosylated in the body, which are mainly secreted or part of cell membrane components (Sola et al., 2007). Normally, the attachment reaction of glycans to a protein begins in the endoplasmic reticulum (ER) and is completed in the Golgi apparatus. The two major forms are O-linked glycosylation and N-linked glycosylation (Saint-Jore-Dupas et al., 2007).

N-linked glycosylation is attached to the amide nitrogen of asparagine (Asn) residues within the common consensus sequence Asn-X-Ser/Thr where X can be any amino acid except proline (Pro) (Kornfeld and Kornfeld, 1985). No particular consensus sequences are recognized for O-linked glycosylation. Thus, N-linked glycosylation is preferentially used in many technologies of protein modification (Spiro, 2002).

Asterion have designed different strategies of protein fusion technologies with a view to generating longer acting therapies. One such strategy is based upon tandem proteins linked a via flexible linker $(Gly_4Ser)_n$ that has been designed to contain N-linked glycosylation motifs (glycosylation consensus sequences). It has been hypothesized that inserted glycosylation motifs within the linker region rather than the ligand would be recognized by mammalian cells for glycosylation. This would result in an increased molecular weight without interfering with the biological activity of the ligand.

The Linker region was designed to incorporate 2, 4 and 8 glycosylation motifs (Asn-Ala-Thr (NAT)) and their respective controls in which Glutamine (or Q) replaces N in NAT sequence motif to produce QAT which would not be recognized by mammalian cells for glycosylation. Up to eight glycosylation sites were introduced to assess whether more efficient glycosylation could increase the molecular weight and subsequently delay renal clearance.

The table below summarizes all GCSF tandem structures and their respective controls (Table 4.1). A full description of the amino acid sequences of these tandems can be seen in Appendix A.2.

Table 4-1: The structure of GCSF tandems with modified flexible linkers $(Gly4Ser)_n$

The glycosylation consensus sequences are highlighted in red (NAT: Asn-Ala-Thr). The control non-glycosylation consensus sequences are highlighted in blue (QAT: Gln-Ala-Thr). The linkers are built around a flexible glycine-serine $(Gly_4Ser)_n$ sequence (these linkers were gene synthesised by Eurofins MWG). Each tandem contains a C-terminal 6 x Histidine tag (Used to aid purification using IMAC).

Molecule name	Number of NAT/QAT motifs	Size of linker	Summary of sequence
pSecTagGCSF2NAT_Hist	2 x NAT	147bp	**GCSF**-G4S- G2NAT- G4S-G4S-G4S-G4S-G2NAT-G4S-GS-**GCSF**x6H
pSecTagGCSF2QAT_Hist	2 x QAT	147bp	**GCSF**-G4S- G2QAT- G4S-G4S-G4S-G4S-G2QAT-G4S-GS-**GCSF**x6H
pSecTagGCSF4NAT_Hist	4 x NAT	147bp	**GCSF**-G4S- G2NAT- G4S-G2-NAT-G4S-G2- NAT-G4S-G2NAT-G4S-GS-**GCSF**x6H
pSecTagGCSF4QAT_Hist	4 x QAT	147bp	**GCSF**-G4S- G2QAT- G4S-G2-QAT-G4S-G2-QAT-G4S-G2QAT-G4S-GS-**GCSF**x6H
pSecTagGCSF8NAT_Hist	8 x NAT	282bp	**GCSF**-G4S- G2-NAT- G4S-G2-NAT-G4S-G2-NAT-G4S-G2NAT- G4S- G2-NAT- G4S-G2-NAT-G4S-G2- NAT-G4S-G2NAT-G4S-GS-**GCSF**x6H
pSecTagGCSFT8QAT_Hist	8 x QAT	282bp	**GCSF**-G4S- G2QAT- G4S-G2-QAT-G4S-G2-QAT-G4S-G2-QAT- G4S- G2-QAT- G4S-G2-QAT-G4S-G2-QAT-G4S-G2-QAT-G4S-GS-**GCSF**x6H

4.2.1 Aim and Hypotheses

As mentioned previously, proteins have been directly modified by increasing the glycosylation on the proteins (Elliott et al., 2003), and this could potentially affect protein bioactivity and increase immunogenicity. Previously a tandem of two GH molecules joined by a flexible $(Gly_4Ser)_n$ peptide linker containing variable numbers of glycosylation motifs from 2 to 8 was created. This method avoided altering or modifying the protein itself directly. These molecules were shown to be glycosylated with increased MW and were biologically active. In this chapter, the aim is to replace two GH ligands in a tandem with two GCSF ligands. The initial stage will be to construct a tandem molecule containing two GCSF ligands using PCR and then ligate in relevant linker regions. As GCSF is known to bind to its receptor at 2:2 stoichiometry, we hypothesized that the resulting tandem GCSF molecule will be available to bind to the two GCSF-receptors and induce downstream signal transduction.

4.2.2 Objectives

1- Construct a tandem molecule containing two ligands of GCSF using PCR.
2- Insert variable linkers between a tandem GCSF molecule that contain increasing numbers of glycosylation motifs and control non-glycosylation motifs.
3- Express and analyse GCSF tandems from a mammalian cell line.

4.3 Construction of GCSF Tandems

To initially construct the GSCF tandem a template plasmid was used. This plasmid contained a tandem GH molecule linked via a glycosylated linker containing 2 x NAT motifs and was available in the Ross group laboratory. The method employed will be to replace each GH molecule with a GCSF molecule and thus produce a GCSF tandem containing 2 x NAT motifs with a GCSF secretion signal. Once constructed, it will be a simple matter to replace the linker region in the GCSF tandem using the restriction enzymes XhoI/BamH1 with other suitable linkers containing either NAT or control QAT motifs (Figure 4.1). A full protein sequence of GH, GCSF and linker region constructs in Appendix B.

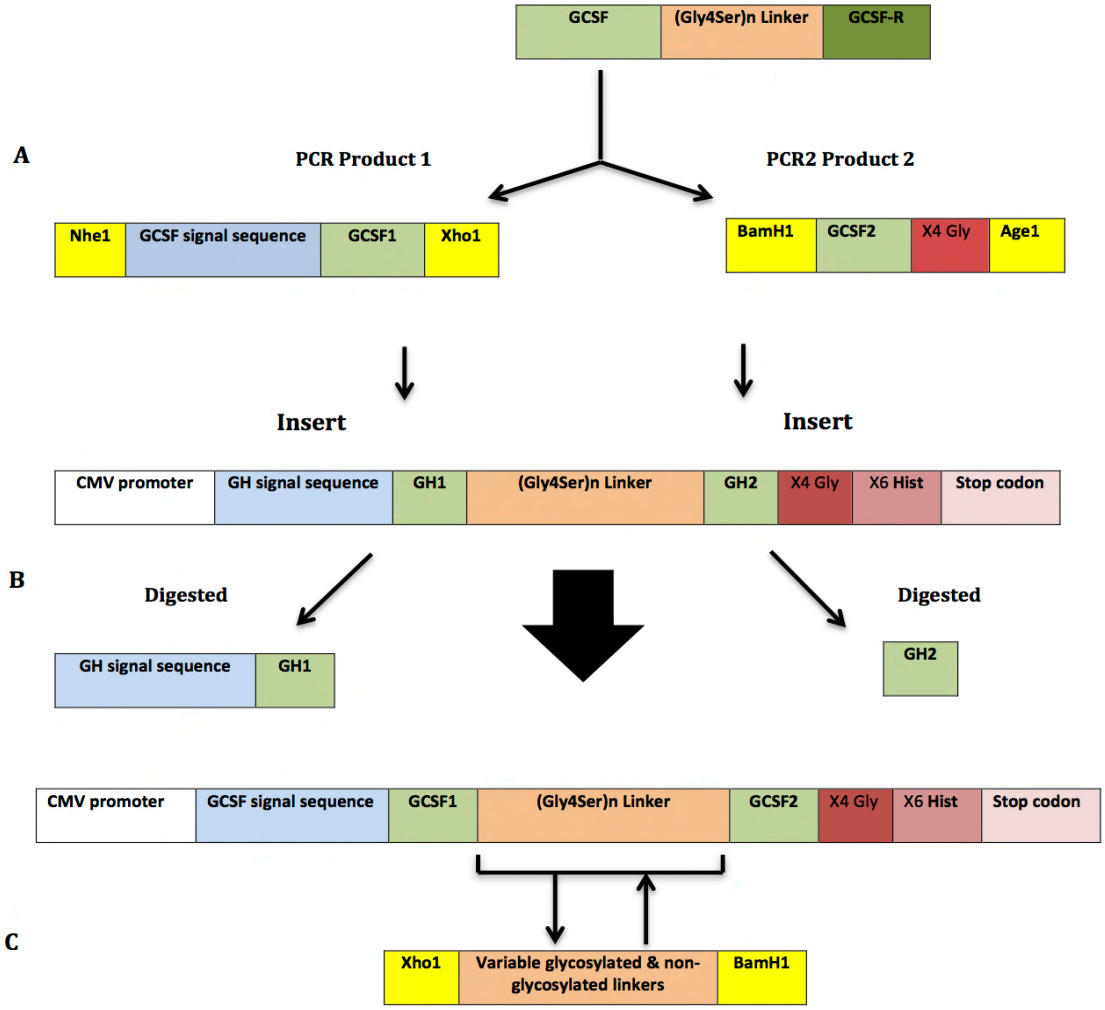


Figure 4-1: The diagram summarizes the process of producing GCSF tandems containing variable linkers

[A] Both GCSF-L1 and L2 (highlighted in green) containing restriction enzyme sites (highlighted in yellow) were PCRed from pGCSFSecTag4A1. [B] Both GH molecules (highlighted in green) are removed from pSecTagGH2NAT plasmid and replaced with the GCSF (GCSF L1 & L2). [C] The original $(Gly4Ser)_n$ linker (highlighted in orange) is replaced with variable glycosylated and non-glycosylated linkers. The plasmid contains a CMV promoter region (highlighted in white), GCSF signal sequence (highlighted in blue), 6x Hist tag (highlighted in pink) and a stop codon (highlighted in pink). A full diagram of the plasmid used in this study is provided in Appendix C.

4.3.1 Construction of pSecTag GCSF-L1_Hist

The first full-length nucleotide sequence of GCSF with signal sequence (GCSF-L1) was PCRed from the template plasmid pSecTagGCSF4A1 (See Table 4.2 and Figure 4.1.A for description) using a forward primer (GCSF Nhe1) and reverse primer (GCSF XhoI Rev) (See full nucleotide sequences for both primers in Appendix A.1) and digested with restriction enzymes NheI/XhoI. This produced a full-length GCSF molecule, GCSF-L1 (612bp) containing the signal sequence of GCSF with a 5 prime NheI site along with a 3 prime XhoI restriction site (Figure 4.2).

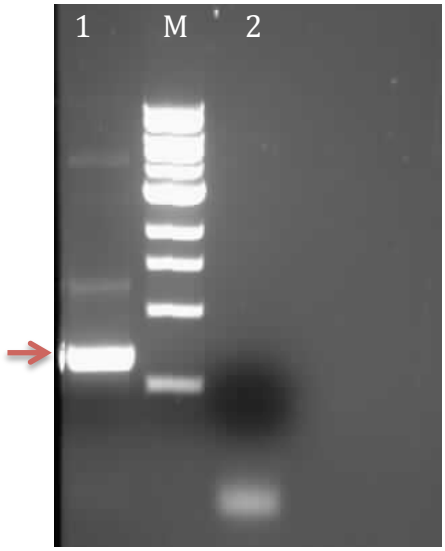
Table 4-2: Description of the two plasmids that were used to produce psecTagGCSF2NAT_Hist

Construct name	N-terminal domain	Linker	C-terminal domain	Expression vector
GCSF4A1	GCSF	$(Gly_4Ser)_{x6}$	GCSF-R	pSecTag
GH2NAT_Hist	GH	$(Gly_4Ser)_n$ with 2 glycosylated motifs	GH	pSecTag

The pSecTagGH2NAT_Hist was digested with double restriction enzymes Nhe1/Xho1 to remove GH Ligand 1 (GH-L1). Single enzyme digests were used to confirm enzyme activity (Figure 4.3).

Thereafter GCSF-L1 was ligated into pSecTagGH2NAT_Hist to produce the plasmid pSecTagGCSF2NAT_Hist-L1 (This contains GCSF-L1 and GH-L2 with a linker region containing 2 x NAT glycosylation sites). The ligation reaction was performed as described at section 3.4.4.

To confirm the construction of pSecTagGCSF2NAT_Hist-L1, plasmid mini preparations were digested using Nhe1/Xho1 to generate a GCSF-L1 0.6kb insert (Figure 4.4).



Lane 1: PCR GCSF-L1 (~0.6kb).

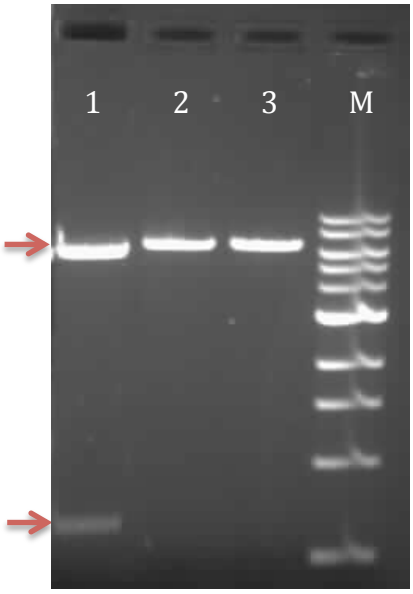
Lane M: 1kb DNA Ladder: Molecular weight marker: Bands starting at the bottom (0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0 Kb).

Lane 2: Negative control (primers only).

Figure 4-2: PCR of GCSF-L1

PCR of GCSF-L1 (DNA) using forward primer (GCSF Nhe1) and reverse primer (GCSF XhoI Rev) run on 1% agarose gel alongside 1kb ladder as standard.

The presence of GCSF-L1 at (~0.6kb) which is the expected size indicates that PCR has been successful.



Lane 1: The upper band represents the pSecTagGH2NAT_Hist (cut plasmid), lower band (~0.6kb) represents GH-L1 insert as a result of a successful double digest by both restriction enzymes.

Lane 2: NheI digest of pSecTagGH2NAT_Hist.

Lane 3: XhoI digest of pSecTagGH2NAT_Hist.

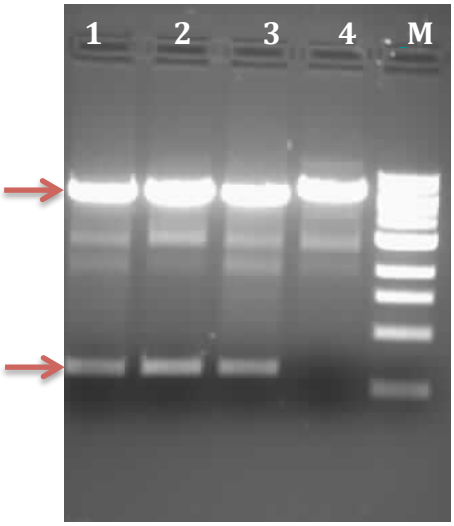
Lane M: 1kb Ladder: Molecular weight marker: Bands starting at the bottom (0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0 Kb).

Figure 4-3: Double digest of pSecTagGH2NAT_Hist

Double digest of pGH2NAT_Hist using NheI and XhoI separated on 1% agarose gel electrophoresis.

The presence of an insert at ~0.6kb in lane 1 indicates a successful double digest. A single band is observed in both lanes 2 and 3, which shows that plasmids used in the digestion reaction were cut by the individual restriction enzyme. The double digested plasmid runs at a slightly lower molecular weight compared to the single digested plasmids, which is due to the excision of the GH-L1 from the starting plasmid pSecTagGH2NAT_Hist.

GCSF-L1 was then ligated into pSecTagGH2NAT_Hist to produce the plasmid pSecTagGCSF2NAT_Hist-L1. This contains GCSF-L1 and GH-L2 with a linker region containing 2 x NAT glycosylation sites. For potential plasmid clones were generated from the ligation plate colonies and were analysed by both double digestion with NheI/XhoI (see Figure 4.4) and DNA sequencing.



Lane 1, 2 & 3: The upper band represents the cut plasmid and the small fragment band represents the GCSF-L1 (~0.6kb) as a result of a successful double digests by both NheI/XhoI restriction enzymes.

Lane 4: Unsuccessful digest.

Lane M: 1kb Ladder: Molecular weight marker: Bands starting at the bottom (0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0 Kb).

Figure 4-4: Double digest of pSecTagGCSF2NAT_Hist-L1 potential clones

Double digest of pSecTagGCSF2NAT_Hist-L1 using NheI/XhoI separated on 1% agarose gel electrophoresis.

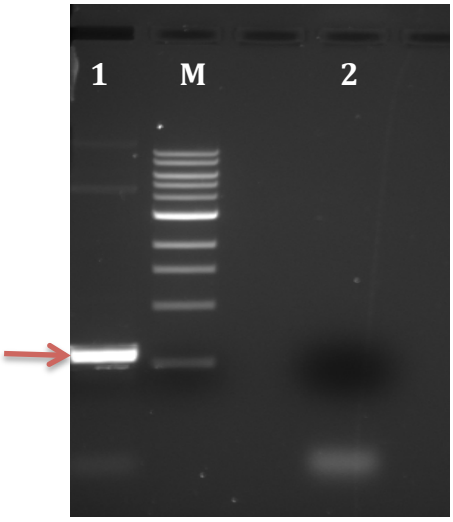
The presence of the insert (~0.6kb) in lanes 1, 2 and 3 indicates a positive ligation. Unsuccessful digest observed at lane 4 could due to missing one or both NheI/XhoI restriction enzymes or that this clone has no insert. The three positive clones were taken forward for sequencing using CMVF and GHseq2Rev primers (see full nucleotide sequences for both primers in Appendix A.1). Sequencing confirmed the ligation of GCSF L1 to pSecTagGH2NAT_Hist to form pSecTagGCSF2NAT_Hist-L1.

4.3.2 Construction of pSecTagGCSF-L2_Hist

The second full-length nucleotide sequence of GCSF-L2 was PCRed using a forward primer (GCSF BamH1) and reverse primer (GCSF Age1 Rev) (Appendix A.1) from the template plasmid pSecTagGCSF4A1 and digested with BamHI/AgeI. This produced a full length GCSF molecule containing a 5 prime BamHI site along with a 3 prime AgeI site (present in plasmid just prior to the Hist tag) (Figure 4.5.). The PCR fragment was then digested with BamH1/Age1 and gel isolated.

pSecTagGCSF2NAT_Hist-L1 was then digested with restriction enzymes BamHI/AgeI to remove GH-L2 and the digested plasmid gel isolated. Single enzyme digests were used to confirm enzyme activity (Figure 4.6). Thereafter GCSF-L2 was ligated into pSecTagGCSF2NAT_Hist-L1 to form the new plasmid, pSecTagGCSF2NAT_Hist. This contains a GCSF tandem with a linker region containing 2 x NAT glycosylation sites (refer to Figure 4.1.B). To confirm the authenticity of the new construct, plasmid mini preparations were digested using BamHI/AgeI to visualise the GCSF-L2 insert (Figure 4.7).

During sequencing, it was difficult to read the whole gene for tandem GCSF (i.e. GCSFL1-linker-GCSF-L2) using CMVF due to the presence of GCSF-L1 in the same plasmid. Therefore we decided to follow the protocol below to remove GCSF-L2 using BamHI/AgeI (Figure 4.8) and ligate to the plasmid pSecTag_link to form pSecTag_link_GCSF-L2 (Figure 4.9). Using this method the GCSF_L2 insert was successfully sequenced using CMVF.



Lane 1: PCR of GCSF-L2 (~0.6kb).

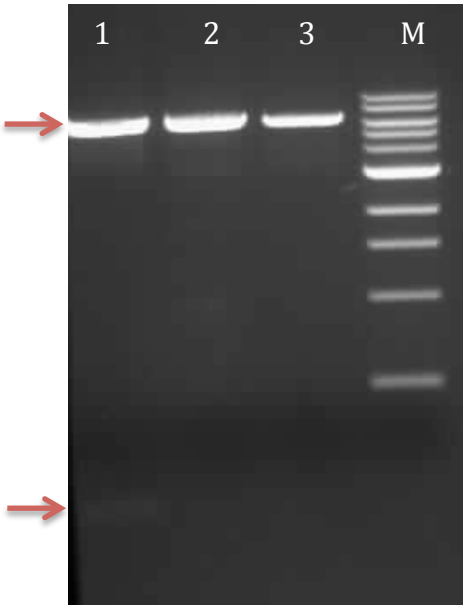
Lane M: 1kb DNA Ladder: Molecular weight marker: Bands starting at the bottom (0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0 Kb).

Lane 2: Negative control (primers only).

Figure 4-5: Generation of PCR fragment GCSF-L2

Generation of PCR fragment using a forward primer (GCSF BmaH1) and reverse primer (GCSF Age1 Rev) separated on a 1% agarose gel.

The result in lane 1 shows that GCSF-L2 is at the expected size of ~ 0.6kb, indicating a successful PCR.



Lane 1: The upper band represents the pSecTagGCSF2NAT_Hist_L1 (cut plasmid) lower band (~0.6kb) represents GH-L2 insert (~0.6kb) as a result of a successful double digest.

Lane2: BamHI digest of pSecTagGCSF2NAT_Hist_L1.

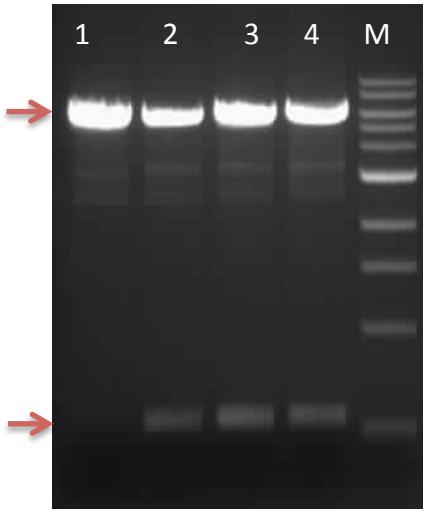
Lane3: AgeI digest of pSecTagGCSF2NAT_Hist_L1. Age1.

Lane M: 1kb Ladder. Molecular weight marker: Bands starting at the bottom (0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0Kb).

Figure 4-6: Double digestion of pSecTagGCSF2NAT_Hist_L1

Double digestion of pSecTagGCSF2NAT_Hist_L1 using BamHI and AgeI RE's separated on 1% agarose gel electrophoresis.

The presence of a fragment of ~0.6kb in lane 1 indicates a successful double digest by both restriction enzymes for the second GH-L2. A single band is observed in both 2 and 3 lanes, which shows that the plasmid used in this digestion reaction was cut by the individual restriction enzyme. The double digested plasmid fragment runs at a slightly lower molecular weight compared to the single digested plasmids, which is due to the excision of the GH-L2 fragment from pSecTagGCSF2NAT_Hist_L1.



Lane 1: Unsuccessful digest.

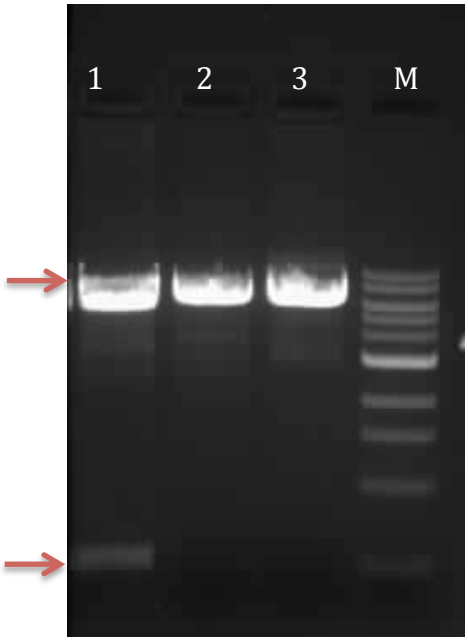
Lane 2, 3 & 4: The upper band represents the pSecTagGCSF2NAT_Hist cut plasmid and the below small fragment band represents the GCSF-L2 (~0.6kb) as a results of a successful double digests by both BamHI/AgeI restriction enzymes.

Lane M: 1kb Ladder: Molecular weight marker: Bands starting at the bottom (0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0 Kb).

Figure 4-7: Double digest of pSecTagGCSF2NAT_Hist potential clones

Double digest of pSecTagGCSF2NAT_Hist using BamHI/AgeI separated on 1% agarose gel electrophoresis.

The presence of an insert at ~0.6kb in lanes 2, 3 and 4 indicates a positive ligation. Unsuccessful digest observed at lane 1 probably due to missing one or both BamHI/AgeI restriction enzymes or that this clone has no insert. The three positive clones take forward for sequencing.



Lane 1: The upper band represents the pSecTagGCSF2NAT_Hist cut plasmid and the small fragment represents the second GCSF-L2 (~0.6kb) as a result of a successful double digests by both BamHI/AgeI restriction enzymes.

Lane2: BamHI digest of pSecTagGCSF2NAT_Hist.

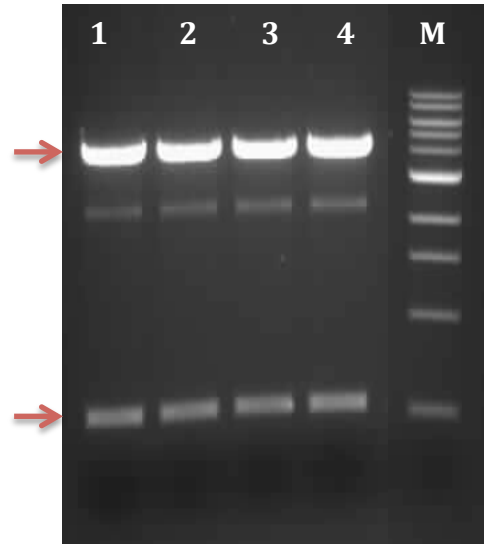
Lane3: AgeI digest of pSecTagGCSF2NAT_Hist.

Lane M: 1kb Ladder. Molecular weight marker: Bands starting at the bottom (0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0Kb).

Figure 4-8: Removal of GCSF L2 from pGCSFsecTagGCSF2NAT_Hist

Double digest of pSecTagGCSF2NAT_Hist using BamHI/AgeI separated on 1% agarose gel electrophoresis.

The presence of the below small fragment band (~0.6kb) at lane 1 indicates a successful double digest by both restriction enzymes for the second GCSF-L2. A single band is observed in both 2 and 3 lanes, which shows the plasmid used in the digestion reaction was cut by the individual restriction enzyme. Also the double digested fragment runs at a slightly lower molecular weight compared to the single digestion fragments, which is due to excision of the GCSF-L2 fragment from the pSecTagGCSF2NAT_Hist. The GCSF L2 fragment was gel isolated and ligated to pSecTag_link (Figure 4.9).



Lane 1, 2, 3 & 4: The upper band represents the pSecTag_link_GCSF-L2 cut plasmid and lower band represents the GCSF-L2 (~0.6kb) as a result of a successful double digests by both BamH1/Age1 restriction enzymes.

Lane M: 1kb Ladder: Molecular weight marker: Bands starting at the bottom (0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0 Kb).

Figure 4-9: Double digest of pSecTag_link_GCSF-L2 potential clones

Double digest of pSecTag_link_GCSF-L2 using BamHI/AgeI separated on 1% agarose gel electrophoresis.

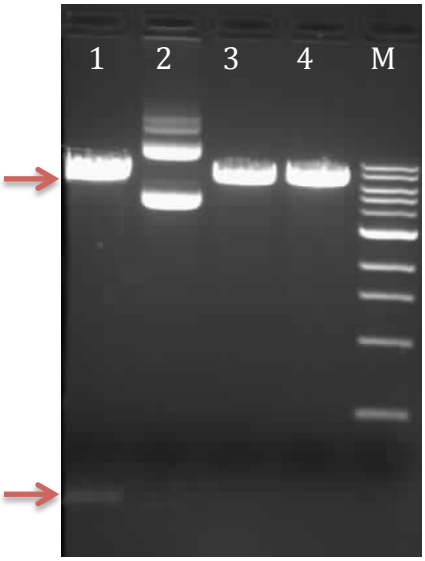
The presence of the small fragment at ~0.6kb in lanes 1, 2, 3 and 4 show successful double digests of all 4 clones. Positive clones were confirmed by sequencing using CMVF.

4.4 Generating GCSF Tandems with Variable Linkers

Successful construction of pSecTasgGCSF2NAT_Hist produced a vector containing a flexible linker with two glycosylation motifs between two GCSF ligands and read through to the X6 Hist purification tag present within this plasmid. As a result, this tandem could be used as a template for creating multiple GCSF constructs containing variable glycosylated and non-glycosylated linkers (refer to Table 4.1). A full description of the amino acid sequences of these tandems can be found in Appendix B.

The 2NAT linker fragment was digested from pSecTagGCSF2NAT_Hist using Xho1/ BamH1 restriction enzymes and replaced with other linkers containing glycosylation and non-glycosylation motifs (2QAT, 4NAT, 4QAT, 8NAT and 8QAT) (See Figure 4.3.C). The linkers were already available as BamHI/XhoI digested fragments.

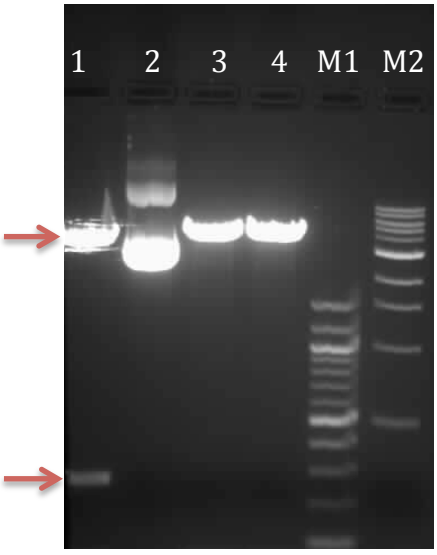
The ligation of each linker was sequenced using CMVF and/or BGHRev primers to confirm the successful ligation (A full description of the amino acid sequences of all linkers can be found in Appendix B.3) and double digestion with XhoI and BamHI. pSecTagGCSF4NAT_Hist & pSecTagGCSF8QAT_Hist are shown as examples of the restriction digests. Single enzyme digests were used to confirm enzyme activity (Figure 4.10 & 4.11).



Lane 1: Double Xho1/BamH1 digests of pSecTagGCSF4NAT_Hist to produce the linker 4NAT. The presence of fragment at 282bp indicates successful cloning.
Lane 2: Undigested pSecTagGCSF4NAT _Hist.
Lane 3: XhoI digest of pSecTagGCSF4NAT_Hist.
Lane 4: BamHI digest of pSecTagGCSF4NAT_Hist.
Lane M: 1kb Ladder: Molecular weight marker: Bands starting at the bottom (0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0 Kb).

Figure 4-10: Double digest potentially positive clone of pSecTagGCSF4NAT_Hist.

1% agarose gel electrophoresis analysis of double digest potentially positive clone of GCSF4NAT_Hist using XhoI/BamHI as a double digest for the linker (4NAT).



Lane 1: Double XhoI/BamHI digests of pSecTagGCSF8QAT_Hist to produce the linker 8QAT. The presence of fragment at 282bp indicates successful cloning.
Lane 2: Undigested pSecTagGCSF8QAT _Hist.
Lane 3: XhoI digest of pSecTagGCSF8QAT_Hist
Lane 4: BamHI digest of pSecTagGCSF8QAT_Hist
Lane M1: 100bp ladder: Molecular weight marker: Bands starting at the bottom (100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1200, 1517 bp).
Lane M2: 1kb Ladder: Molecular weight marker: Bands starting at the bottom (0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0 Kb).

Figure 4-11: Double digest potentially positive clone of pSecTag GCSF8QAT_Hist

1% agarose gel electrophoresis analysis of double digest potentially positive clone of pSecTagGCSF8QAT_Hist using XhoI/BamHI as a double digest for the linker (8QAT).

In conclusion, the presence of a fragment at 282bp in lane 1 for both figures as resulted of double digest using Xho1/BamH1 indicates successful cloning for both linker 4NAT and 8QAT. A supercoiled band is observed in lane 2 for both figures as a resulted of undigested plasmid. Also, a single band is observed in both 3 and 4 lanes for both figures, which shows that the plasmid used in this digestion reaction was cut by the individual restriction enzyme.

4.5 Expression and Analysis of GCSF Tandems

After the successful construction of GCSF2NAT_Hist, GCSF2QAT_Hist, GCSF4NAT_Hist, GCSF4QAT_Hist, GCSF8NAT_Hist and GCSF8QAT_Hist, all plasmids were transiently and stably transfected into CHO Flp-In cells, and serum free media harvested for analysis (refer to section 3.5). Media samples were analysed using Elisa and western blotting to detect expression and an increase in molecular weight of proteins due to glycosylation.

4.5.1 Transient Transfection of CHO Flp-In Cells

As an initial assessment of protein expression, all plasmid constructs were transiently transfected into CHO Flp-In cells and serum free media harvested and analysed via Elisa and western blot.

4.5.1.1 Analysis of Expression by Elisa

To verify whether or not protein tandems were successfully expressed, media of protein samples were analysed via Elisa using the procedure described in section 3.8.4. The results of Elisa are given in Table 4.3.

Table 4-3: GCSF sandwich Elisa analysis of transiently expressed tandem proteins. GCSF was used as standard

Protein	µg/ml	nM	SEM	%CV
Negative control	0.07	0.0	0.0	2.80
GCSF2QAT_Hist	2.35	51.77	0.02	3.89
GCSF2NAT_Hist	1.82	40.19	0.01	2.09
GCSF4QAT_Hist	0.94	20.72	0.01	1.95
GCSF4NAT_Hist	2.07	45.81	0.01	3.09
GCSF8QAT_Hist	3.00	61.88	0.03	6.28
GCSF8NAT_Hist	5.64	116.20	0.02	2.81

It can be seen in the table above that the Elisa has detected all tandem proteins, which indicates a successful transient expression. However, the purpose of this method is to detect expression only, as true level of expression is likely to be inaccurate since the GCSF standard curve is based upon monomeric GCSF, not specific tandem molecules (more details in discussion part).

A maximum expression of 5.64µg/ml for GCSF8NAT_Hist is observed and a minimum of 0.94µg/ml for GCSF4QAT_Hist. %CV of less than 10% can be seen for all protein samples which is considered an acceptable degree of variability between samples (between replicate values) in many commercially used Elisa's.

4.5.1.2 Analysis of Expression by Western Blotting

To assess whether protein tandems have undergone successful glycosylation, 10µl of each media sample was diluted equally with 2 x Laemmli buffer and separated on a 10% SDS-PAGE gel. The samples were then transferred to PVDF membrane and western blotted as outlined in methods (Section 3.8.3). The results of the western blot are presented in Figure 4.12.

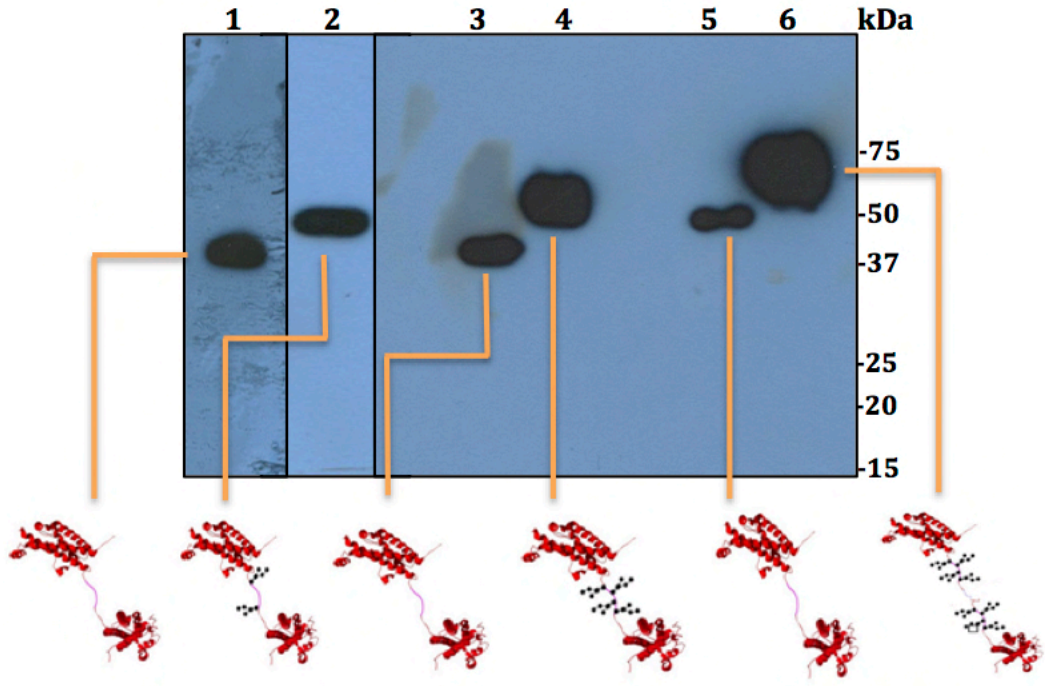


Figure 4-12: Western blot of media samples from transiently transfected CHO Flp-In cells

Lane 1*; GCSF2QAT_Hist (2 x QAT). Lane 2*; GCSF2NAT_Hist (2 x NAT). Lane 3; GCSF4QAT_Hist (4 x QAT). Lane 4; GCSF4NAT_Hist (4 x NAT). Lane 5; GCSF8QAT_Hist (8 x QAT). Lane 6; GCSF8NAT_Hist (8 x NAT). Western blot shows a definite increase in molecular weight for glycosylated molecules (GCSF2NAT_Hist, GCSF5NAT_Hist, and GCSF8NAT_Hist) compared to non-glycosylated controls (GCSF2QAT_Hist, GCSF4QAT_Hist, GCSF8QAT_Hist). *Lane 1 and 2 were taken from another gels as didn't run correctly with the above gel.

As can be observed in Figure 4.12 western blotting successfully detected all GCSF tandems. A very clear shift can be seen in molecular weight for glycosylated molecules (GCSF2NAT_Hist, GCSF5NAT_Hist, and GCSF8NAT_Hist) compared to non-glycosylated controls (GCSF2QAT_Hist, GCSF4QAT_Hist, GCSF8QAT_Hist). Glycosylated molecules show an increase in molecular weight, which is consistent with their linkers' being successful glycosylated.

In addition, it can also be seen that there is a slight difference in molecular weight between GCSF8QAT_Hist compared to GCSF4QAT_Hist and GCSF2QAT, this is due to the difference in linker length. GCSF8QATHist has a

linker length of 282bp = 94 amino acids compared to 147bp =49 amino acids for both GCSF4QAT_Hist and GCSF2QAT_Hist.

4.5.2 Stable Cell Line Development in CHO Flp-In Cell Lines

Since all tandems were shown to be expressed and intact from transient transfections, all GCSF tandems were taken forward to make stable cell lines in CHO FIp-In cells as described in section 3.5.4. Non-transfected cells were used as a negative control.

4.5.2.1 Analysis of GCSF Tandems by Elisa

Stable cells were incubated in serum free media for 2-3 days then harvested and analysed (See section 3.5.5). Elisa data is shown in Table 4.4.

Table 4-4: Results of GCSF sandwich Elisa for stably expressed GCSF tandems

Construct	µg/ml	nM	SEM	%CV
GCSF2QAT	2.82	62.04	0.038	8.11
GCSF2NAT	3.97	87.81	0.021	3.46
GCSF4QAT	4.67	102.95	0.047	6.98
GCSF4NAT	3.13	69.20	0.018	3.48
GCSF8QAT	7.69	158.62	0.032	3.29
GCSF8NAT	3.89	80.14	0.047	8.01

From the above data, all protein samples have been detected. It can be shown that GCSF8QAT_Hist is the most highly expressed protein at 7.69µg/ml and GCSF2QAT_Hist being the lower expressed at 2.82µg/ml. %CV is below 10% for all protein samples, which indicates good consistency between triplicates.

4.5.2.2 Analysis of GCSF Tandems by Western Blotting

10µl of each serum free media sample was equally diluted in 2 x Laemmli buffer and separated on a 10% SDS-PAGE gel. The samples were then transferred to PVDF membrane and analysed by western blotting. The results are shown in Figure 4.13.

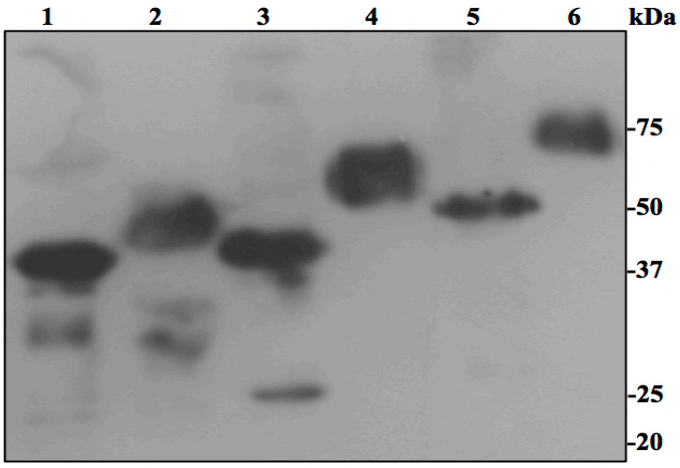


Figure 4-13: Western blot of stable CHO Flp-In cell media expressing GCSF tandems

Lane 1; GCSF2QAT_Hist (2 x QAT). Lane 2; GCSF2NAT_Hist (2 x NAT). Lane 3; GCSF4QAT_Hist (4 x QAT). Lane 4; GCSF4NAT_Hist (4 x NAT). Lane 5; GCSF8QAT_Hist (8 x QAT). Lane 6; GCSF8NAT_Hist (8 x NAT). Western bot shows an increase in molecular weight for glycosylated molecules (GCSF2NAT_Hist, GCSF4NAT_Hist, and GCSF8NAT_Hist) compared to non-glycosylated controls (GCSF2QAT_Hist, GCSF4QAT_Hist, GCSF8QAT_Hist).

Western blotting successfully detected all GCSF tandems. All glycosylated and non-glycosylated GCSF tandems are running at approximately the same molecular weights that were observed previously in transient transfections. This suggests the separation by SDS-PAGE has run appropriately and that samples have successful glycosylation at the consensus sequence within the linker region. Glycosylated constructs exhibited increased molecular weight over non-glycosylated molecules and rhGCSF (Table 4.5).

Table 4-5: Determined and observed MW's of expressed GCSF tandems

Construct	Determined MW (kDa)*	Observed MW (kDa) (approximation)
rhGCSF	18.8	_
GCSF2QAT_Hist	45.4	~ 45
GCSF2NAT_Hist	45.2	~ 52
GCSF4QAT_Hist	45.4	~ 45
GCSF4NAT_Hist	45.2	~ 60
GCSF8QAT_Hist	48.5	~ 49
GCSF8NAT_Hist	48.5	~ 70

*MW's calculated using DNASTAR Lasergene version 8 and the observed MW estimated using western blot analysis (see Figure 4-13).

4.6 *In vitro* Biological Activity of Crude Media

Biological activity of GCSF tandems was measured using the human acute myeloid leukemic cell line (AML-193 cell line), which proliferates in response to GCSF. To determine if the expressed GCSF tandem proteins are biologically active, stable clone crude media was tested for bioactivity using the AML 193 proliferation assay as described in section 3.8.5. As expected, all crude media obtained from stable cell lines are biologically active above that of the negative control, growth hormone (GH). The assay appears to reach absorbance saturation of around 0.25 for the rhGCSF and all GCSF tandems (Figure 4.14.A & B).

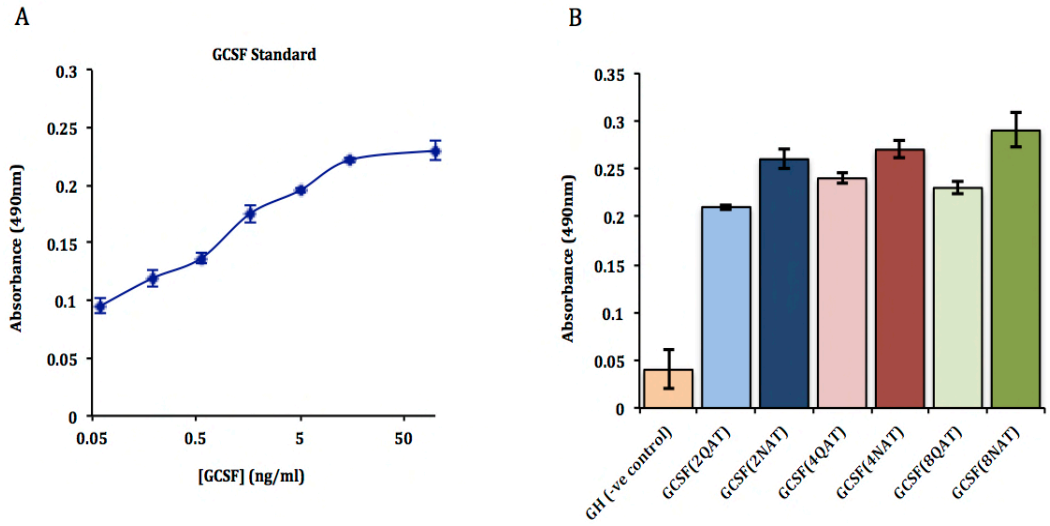


Figure 4-14: *In vitro* biological activity for rhGCSF and GCSF tandems

(A) Represents a good GCSF standard curve showing progressive increase in biological activity when tested using an in house AML-193 proliferation assay. **(B)** Represents biological activity of GCSF tandems. Results are given as standard mean of error SEM for triplicate wells.

The in house bioassay indicates a good standard curve with gradual increase in the absorbance with increasing GCSF concentrations. All constructs have biological activity with a maximum absorbance of about 0.25 observed for rhGCSF, glycosylated GCSF tandems and their respective controls (non-glycosylated tandems). Whether the glycosylations affect the biological activity is difficult to interpret since quantification was completed using monomeric GCSF and not the specific tandems: there could well be differences in detection between the tandems perhaps due to steric hindrance by glycosylations.

4.7 Discussion

In this chapter, the cloning, sequencing and expression of two GCSF ligands linked by a variably glycosylated and non-glycosylated linkers were examined.

The data obtained has shown that it is possible to incorporate 2, 4 and 8 NAT (Asn-Ala-Thr) or QAT (Glu-Ala-Thr) motifs between tandem GCSF ligands. The NAT linker containing tandems have been successfully glycosylated as shown by an increase in the molecular weight of the constructs upon expression in mammalian CHO Flp-In cells as verified by western blotting.

CHO Flp-In cell lines were used to express all GCSF constructs as these cell lines have been widely used to express recombinant proteins. The main reason for their use is the possession of glycosylation mechanism similar to those in humans. The cells are efficient and easy to culture and relatively express high level of proteins (Damiani et al., 2009). Additionally, these mammalian cell lines are preferred due to their ability to correctly fold proteins during biosynthesis (Wurm, 2004). Ordinarily, CHO cell lines were not capable of both sialyation types observed in human body (alpha 2,3 and 2,6-glycosidic linkages are participate in glycan structure) as it lacks the enzyme alpha 2,6- sialyltransferase and this could cause immunological response and thus influence the therapeutic application of the GCSF constructs. However, CHO cells that are capable of sialylation have been genetically engineered by insertion of alpha 2,6-sialylation expressing into the DNA sequence of the cells (Damiani 2009). Furthermore, the CHO cell line is the industry standard for recombinant therapeutic proteins production and many protein drugs are manufactured in this system (mostly antibodies) without problems.

The Elisa has detected all tandem proteins, indicating successful transient and stable expression. However, the purpose of this method was to detect expression only as quantifiable levels of expression are likely to be inaccurate. This may be due to overestimating of Elisa since the standard

curve is based upon monomeric GCSF and therefore is not ideal to measure these tandems. Purified GCSF tandem was not obtainable at the same time of carrying out these experiments and thus the monomeric GCSF was the best alternative. As a result, the number of GCSF molecules per mole of tandem is double the number of rhGCSF molecules. Besides, although our GCSF tandems are not directly glycosylated, it is difficult to understand how the influence of protein conformation and glycosylation of these tandem proteins affects the binding of antibody and thus influence Elisa results. Byrne et al., (2007) indicated that glycosylation could completely eradicate the polypeptide recognition site of antibody. In contrast, analysis of these tandems using western blotting may provide a truer visual of expression since proteins are in their denatured form and this may expose their recognition sites for the antibody more. Protein denaturation as a result of sample heating may also reduce the possible steric hindrance caused by glycosylation to antibody recognition sites. It is important to consider these factors since it is difficult to determine whether or not equivalent concentrations of protein sample have been tested in the GCSF bioactivity assay.

Previously we stated that, western blotting showed an increase in the molecular weight of glycosylated tandems (GCSF2NAT_Hist, GCSF4NAT_Hist and GCSF8NAT_Hist). The only different in these tandems compared to the non-glycosylated control is the presence of N-linked glycosylation motifs with in the linker, supporting successful glycosylation.

The key behind successful N-linked glycosylation is the selection of a consensus sequence (As mentioned previously, the common consensus sequence for N-linked glycosylation is Asn-X-Ser/Thr where X can be any amino acid except proline (Kornfeld and Kornfeld, 1985). Thus, efficiency of glycosylation is likely to be determined by the amino acid (Thr) that occupies the third position. It has been shown that glycosylation is more efficient when Threonine is occupying the third position as opposed to Serine, with efficiency shown to be twice as high and in other cases even forty fold (Kasturi et al., 1995, Picard et al., 1995).

The crude media of each tandem GCSF stable clones was tested using the AML 193 proliferation assay. At this stage, this assay would give a preliminary indication whether GCSF tandems could maintain biological activity since their biological activity showed similar data to rhGCSF. Whether the glycosylations affect the biological activity is difficult to interpret since quantification was completed using monomeric GCSF and not the specific tandems: there could well be differences in detection between the tandems perhaps due to steric hindrance by glycosylations.

It has been reported that the insertion of a Histidine purification tag in to recombinant proteins may potentially dislocate the three-dimensional structure of the polypeptide protein (Block et al., 2009), this may eventually impact upon receptor binding and negatively affect protein bioactivity. All tandems were Hist tagged and the initial data of AML193 assay from crude media showed that it is possible to introduce another GCSF and $_{X}$6His purification tag into the C-terminus of a GCSF (to create a tandem GCSF) and still maintain biological activity. Pure proteins would give better indication as to whether a $_{X}$6His purification tag or glycosylation could negatively affect protein bioactivity. For further analysis and characterisation, GCSF tandem proteins were purified using IMAC. This will be discussed in the next chapter.

5. Results 2: Large-scale Production and Analysis of GCSF Tandems

5.1 Summary

The previous chapter showed that it is possible to clone, sequence and express two GCSF ligands linked by variable glycosylated and non-glycosylated linkers. Therefore, for further PK/PD analysis, it was necessary to purify these GCSF tandem proteins using Nickel IMAC. This chapter shows that IMAC is an appropriate method for purifying GCSF tandems with histidine tags. IMAC permits a fast and easy way for purification due to the strong affinity of a nickel-complex (Ni^{2+}-NTA) for the histidine tag present in the GCSF tandems. All GCSF tandems produced high levels of protein (between 1 to 4mg per litre as assessed by Bradford assay) from a stable CHO cell line and were considered to be 90-95% pure as judged by SDS-PAGE. Molecular weights and the integrity of each GCSF tandem were confirmed by SDS-PAGE and western blotting. Purified protein was then passed over for PK/PD studies.

5.2 Introduction

For decades, cloning and the following recombinant expression of protein is commonly used in molecular biology. As purification of recombinant proteins is frequently a challenging step, a short affinity epitope tag was designed to aid purification process. Usually, these short amino acid sequences are added either to the N or C terminal ends of a protein. Then, these amino acid sequences can expose epitopes for specific binding partners like antibodies. The $_{X}$6 His tag (HHHHHH) is one of the most commonly used protein tags that permits a fast and easy way for purification that is based on the strong affinity of histidine for divalent cations (such as Zn, Cu, Co and Ni). The electron donor groups on the imidazole ring of histidine form coordination bonds with the immobilized metal ion matrices under native conditions in high or low salt buffers. In our purification system, immobilized nickel-complex Ni^{2+}-NTA beads were used. In the Immobilized Metal Affinity Chromatography (IMAC) purification system, the bound protein is eluted off the column using imidazole, which is a histidine analogue, as it competes with his-tagged protein for binding to the divalent-metal ions matrix. As a result, the eluted target protein can simply be separated from unwanted contaminated proteins in the elution mixture (Block et al., 2009, Kreisig et al., 2014).

5.2.1 Aim and Hypothesis

Following the successful cloning and expression of the tandem proteins in the previous chapter, it was necessary to purify these tandem proteins from mammalian cell lines (CHO FIp-In cells) stably expressing each tandem. The tandem GCSF expressing cells were grown in roller bottles and target protein harvested from the media as a secreted product, purified and passed over for *in vivo* PK/PD analysis. In order to facilitate the purification of the target proteins, the expression constructs were tagged at the C-terminal with a 6-Histidine. As Nickel IMAC column is known to have high binding affinity for histidine, we hypothesized that the histidine tag in our GCSF tandems will bind to nickel column and facilitate the purification.

5.3 Results

5.3.1 Cell Growth and Productivity

Hyclone SMF4CHO Utility media adapted stable CHO Flp-In cell lines of all GCSF tandems were thawed from liquid nitrogen stocks. After a period of growth and expansion, cells were transferred to two non-vented roller bottles containing Hyclone SFM4CHO Utility media and grown at 37°C with gentle agitation in a CO_2-free incubator at 37°C, for protein production (each roller bottle contained approximately 500ml of culture volume). Starting with a density of ~0.25 x 10^6 cells/ml, the cells were grown until viability declined below 30% (took approximately 10 - 11 days) as previously described in section 3.5.8. 100µl of media samples were routinely taken every 2 or 3 days and analysed for total cell number and viability using a trypan blue exclusion (see section 3.5.2). Elisa was also used to analyse protein productivity as described in section 3.8.4.

There were cell growth and maximal viable cell population differences observed for the GCSF tandems during the expression studies. While all GCSF tandem cells had approximately a similar % of viable cells (Figure 5.1b), they however had varied peak viable cells number, ranging from 1.25 to 1.9 million cells per ml (1.25-1.9 x 10^6/ml) with GCSF2NAT and GCSF4NAT having the least and highest peak cells densities, respectively. A steady exponential growth was observed for the GCSF2QAT and GCSF4NAT up to a maximal viable cell density at day 4 (viable densities of 1.7 x 10^6/ml & 1.14 x 10^6/ml respectively). This is in contrast to other cell lines in which maximal viable cell densities were reached on day 7 for 2NAT (1.24 x 10^6/ml), GCSF8NAT (1.68 x 10^6/ml) and GCSF8QAT (1.34 x 10^6/ml), and at day 9 for GCSF4QAT (1.62 x 10^6/ml) (Figure 5.1a). The GCSF8NAT culture produces product at an earlier time point than other cultures, peaking on day 7 but losses viability at an earlier stage. Most of the tandem GCSF cultures retained a high viability up to at least day 7 (day 6 for 8NAT culture) with viabilities at ~60% (Figure 5.1b). Consequently, 8NAT-expressing cells were harvested for protein production on day 7 of growth,

as the %viability declined rapidly beyond this period, whereas all other GCSF tandems were grown up to day 9 before culture media were harvested. Beyond day 9 there was a significant loss of cell viability (Figure 5.1b; 4QAT and 4NAT expressing cells). This was expected as nutrient depletion and host cell metabolic waste would impact cell proliferation. Therefore, at the time of harvest, majority of the tandem GCSF expressing cells were at least 60% viable. Harvesting the culture media at this stage helped prevent any potential protein degradation.

Interestingly, protein production was detectable as early as day 2 of culture growth with a consistent increase in protein productivity as the cells number increases (Figure 5.1c). There was a dramatic increase in protein productivity for most GCSF tandems up to at least day 7 with productivity peaking at >50mg/L except 2QAT and 8QAT which maintain a steady low level of expression showing a maximal peak expression of 10mg/L & 17mg/L respectively. The significant difference in protein productivity for most of the GCSF tandems expressing cells was observed to occur between day 4 and day 9, during which 2NAT and 4NAT protein yield rose from approximately 20mg/L on day 7 to about 55mg/L and 50mg/L on day 9 respectively, and 30mg/L on day 4 to 50mg/L on day 7 for 8NAT. While the underlying reason for the sudden enhanced productivity was unclear, it was inferred that the expressing cells might have become better adapted to the conditions of growth.

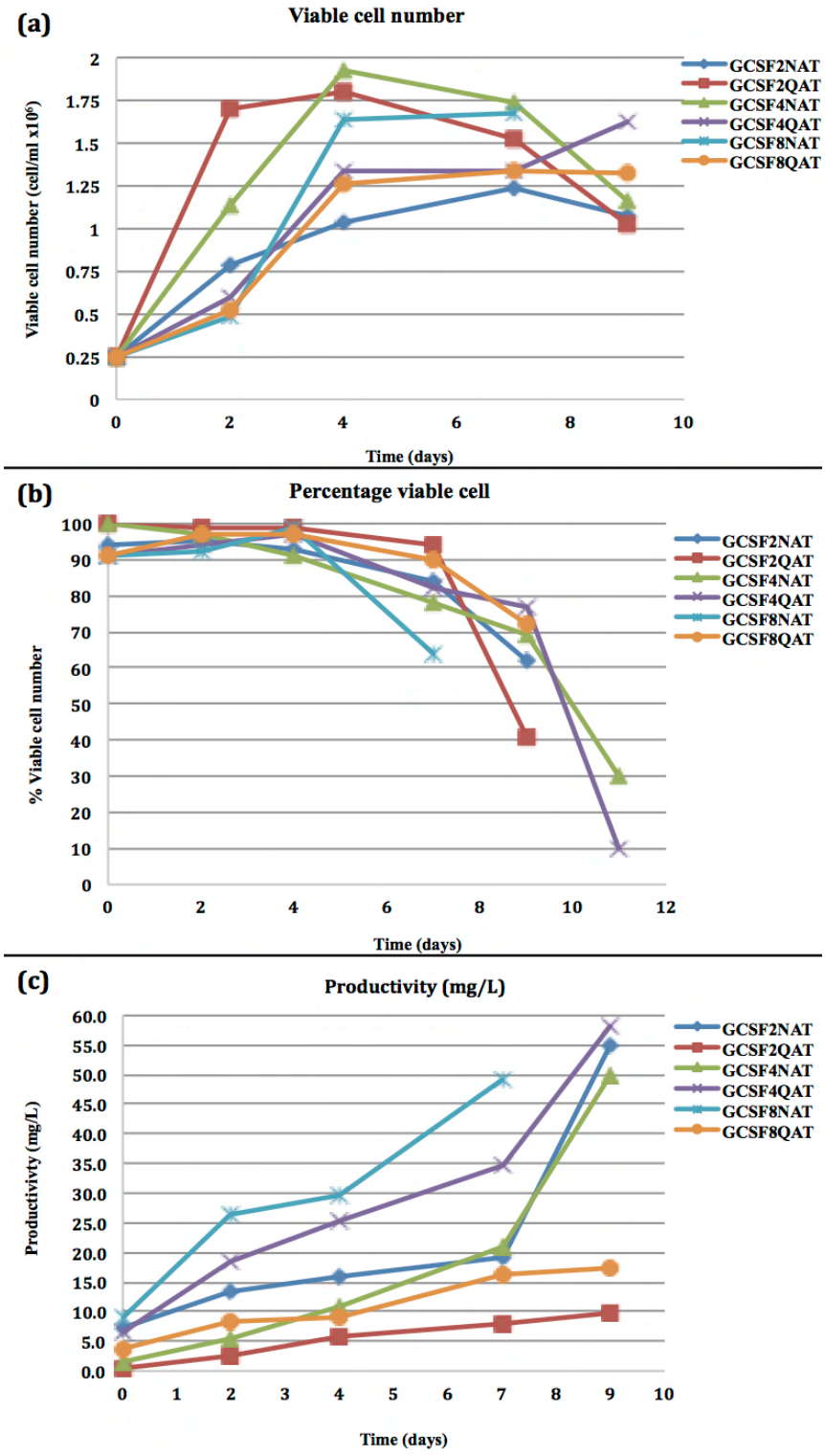


Figure 5-1: GCSF tandems expressing cells growth and productivity

The figure shows the total viable cells **(a)** and % viability of the tandem expressing cells **(b)** for all glycosylated tandem proteins (GCSF2NAT, GCSF4NAT & GCSF8NAT) together with their respective controls, non-glycosylated tandem proteins (GCSF2QAT, GCSF4QAT & GCSF8QAT), as well as the productivity of the cells within a 9-day growth period **(c).**

To confirm the integrity of the expressed GCSF tandem proteins, the protein samples from the GCSF tandems expressing cells were analysed by western blotting. The results revealed a consistent increase in the protein expression of all GCSF tandems with variable productivity in each tandem protein until day 9 (Figure 5.2). A significant loss of protein was equally observed for cells grown beyond day 9 (i.e. GCSF4NAT and GCSF4QAT), similar to what was earlier observed in the Elisa analyses. Consequently, culture media were routinely collected on day 9 (except 8NAT which was collected on day 7) for protein purification.

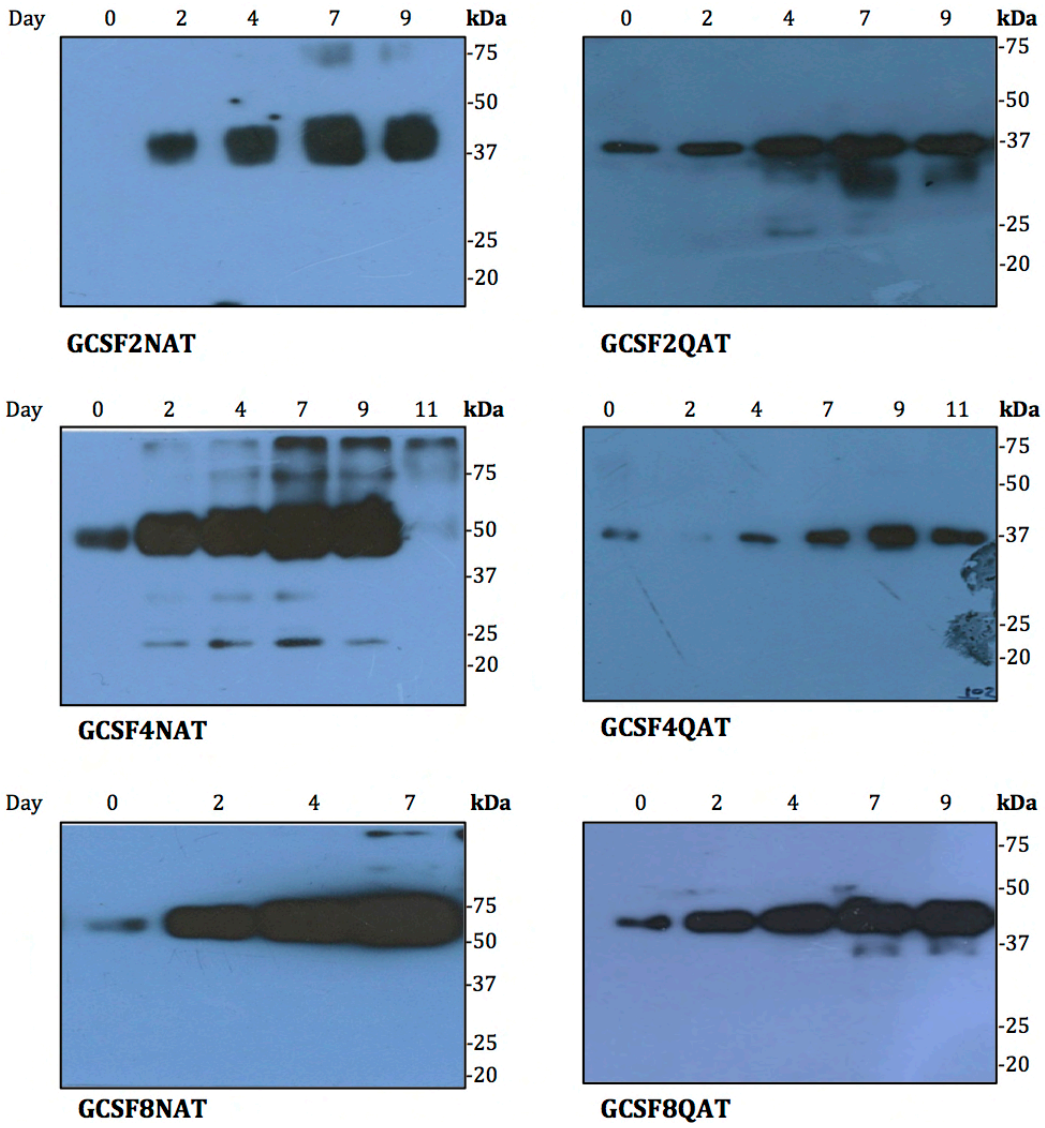


Figure 5-2: Western blot analysis of roller bottle media samples

Stable CHO Flp-In cells expressing the GCSF tandems were grown in roller bottle cultures and samples taken every 2 days up to a total of 11 days. 10µl of samples from culture medium for each tandem GCSF construct were analysed by western blotting.

5.3.2 Purification of GCSF Tandems Using IMAC

The culture media for each GCSF tandem was taken from roller bottle and spun down to pellet the cell debris by centrifugation (18000 g, JLA-16.250 Fixed-Angle Rotor for 30 minutes at 4°C). 10mM final concentration of Benzamidine HCl (serine protease inhibitor) was added to the 1L culture media to prevent protein degradation. Using a Vivaflow 200 concentrator, media samples were concentrated about 10-fold i.e. 100ml. The sample was diluted 1:1 with equilibration buffer prior to IMAC as previously described at section 3.7.1.

5.3.2.1 Purification of GCSF2NAT

Culture media collected from GCSF2NAT expressing cells was concentrated and purified on a Nickel IMAC column. Western blotting results showed that the majority of GCSF2NAT protein bound to the IMAC column with negligible amounts observed in the flow through (Figure 5.3.A). High purity protein (>90% pure) was eluted using imidazole containing buffers with concentrations ranging between 200mM and 350mM (shown in bold typeface, elutions 6,7,8 & 9 in figure 5.3.A). Elutions using imidazole concentrations below or above this range (100mM or 500mM respectively) were observed to be of low purity when analysed by SDS-PAGE and were thus excluded from the study. High purity elutions were collected and pooled together for dialysis. Dialysis was performed to remove any excess imidazole or other salts (Figure 5.3.B). Pre and post dialysed proteins were measured by Bradford assay (Table 5.1) and analysed by 10% SDS PAGE followed by western blot analysis to confirm the purification and size of GCSF2NAT. Coomassie stained 10% SDS -PAGE showed that the post-dialysis GCSF2NAT protein was stable with minimal protein degradation (visible as contaminated bands at 20-25kda). These lower molecular weight (LMW) bands may have resulted from handling or partial cleavage of the tandem GCSF2NAT linker (the observed molecule at this low molecular weight was predicted to be an equivalent of a GCSF ligand and non-glycosylated linker). Analysis of pellet samples post dialysis does show the

presence of a protein band at ~50kDa, however this was not detected by western blotting.

Western blotting also did not pick up the LMW bands of pre and post dialysed samples present post coomassie staining. It was presumed that these LMW bands are missing the antigenic site or are below detection limit of antibody. Also, the blots might not have been exposed long enough to see the contaminant bands. Nevertheless, the LMW proteins were of significantly lower intensity when compared to the intact molecule, suggesting an appreciably high purity yield during purification. A total of 1.96mg of GCSF2NAT was recovered of >90% purity from 1 litre a media.

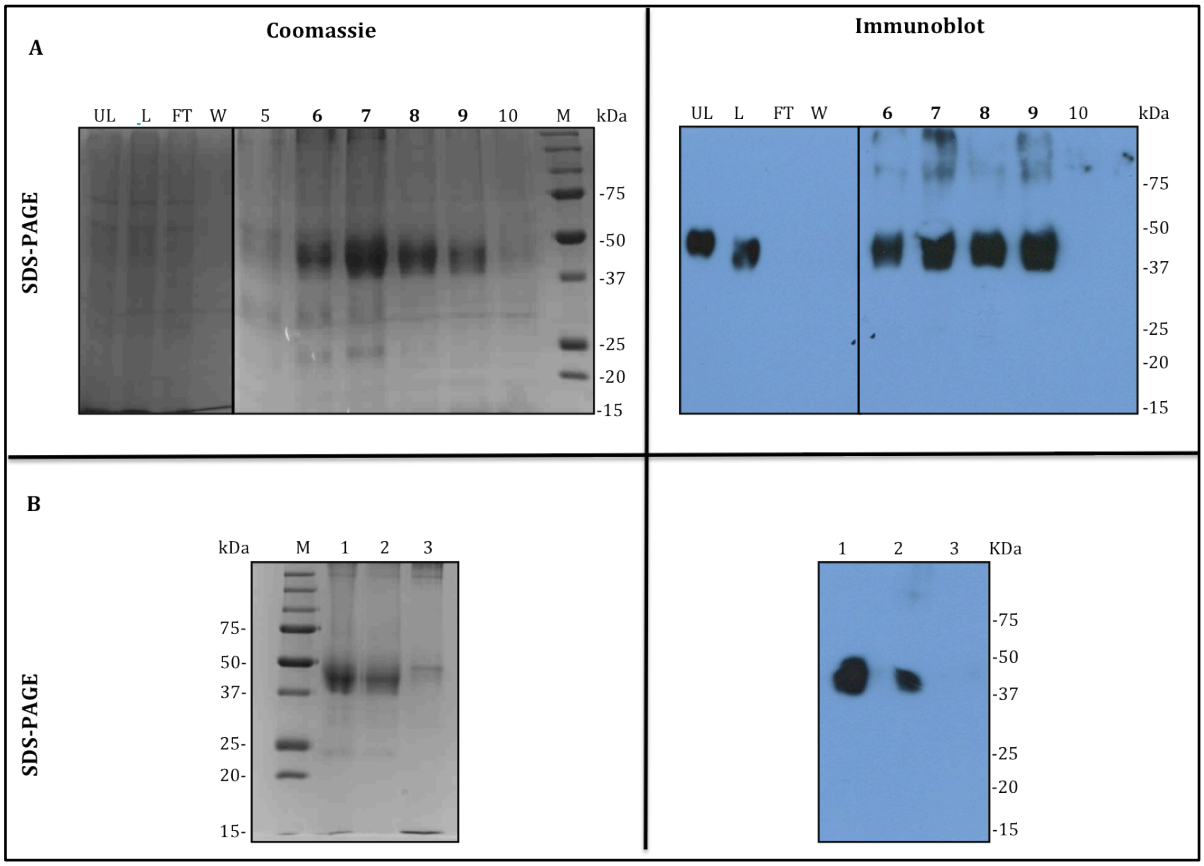


Figure 5-3: purification development of IMAC for GCSF2NAT

10µl of protein samples from each stage of the GCSF2NAT purification process were separated on a 10% SDS-PAGE and stained with coomassie blue for visualization (A, left figure), while 100ng of the same sample was analysed by western blotting (A, right figure) (UL=Unfiltered load, L=Load, FL=Flow through, W=Wash pH 7.4 & 6.0, M= Markers, 5-10= Elutions) Thereafter, active elution fractions were pooled, dialysed and analysed on a 10% SDS-PAGE gel by coomassie staining (B, left figure) and western blotting (B, right figure) (1= Pre-Dialysis, 2= Post-Dialysis and Post-Dialysis pellet).

Table 5-1: Concentrations of GCSF2NAT during the purification process

The table features the concentrations of GCSF2NAT at each stage of the purification process. The elutions in bold typeface **(6, 7, 8 & 9)** were pooled together giving 3.39mg of >90% pure GCSF2NAT tandem protein, which equated to a 1.95% recovery from total protein. However, ~ 42% of this was lost during dialyses giving a final amount of 1.96mg, which equated to 1.13% recovery from total protein.

Sample	Imidazole Concentration (nM)	Volume (ml)	Protein (µg/ml)	Total Protein (mg)	%Recovery
Unfiltered load	-	220	680.46	149.70	-
Load	-	220	790.48	173.90	100
Unbound	-	220	655.83	144.28	83
Wash pH 7.4	-	1	69.62	0.07	0.04
E5	200	1	839.74	0.84	0.48
E6	**200**	**1**	**539.24**	**0.54**	**0.3**
E7	**350**	**1**	**1031.86**	**1.03**	**0.59**
E8	**350**	**1**	**1333.99**	**1.33**	**0.76**
E9	**350**	**1**	**570.44**	**0.57**	**0.32**
E10	500	1	622.99	0.62	0.36
Pre-Dialysis	-	4	846.31	3.39	1.95
Post-Dialysis	-	4	491.07	1.96	1.13
Pellet	-	0.1	780.62	0.08	0.05

5.3.2.2 Purification of GCSF2QAT

Culture media collected from GCSF2QAT expressing cells was concentrated and purified on a Nickel IMAC column. Western blotting results showed that the majority of GCSF2QAT protein in the media was bound to the IMAC with negligible amounts observed in the flow through (Figure 5.4.A). High purity protein (>90% pure) was eluted using imidazole containing buffers with concentrations ranging between 200mM and 350mM (shown in bold typeface 7,8,9 & 10 in figure 5.4.A). Elutions using with imidazole concentrations below or above this range (100mM or 500mM respectively) were observed to be of low purity when analysed by SDS-PAGE and were thus excluded from the study. High purity elutions were collected from IMAC purification and pooled together for dialysis. Dialysis was performed to remove any excess imidazole or other salts (Figure 5.4.B). Pre and post dialysed protein were then measured by Bradford assay (Table 5.2) and analysed by 10% SDS PAGE followed by western blot analysis to confirm the purification and size of GCSF2QAT. Coomassie stained 10% SDS -PAGE post-dialysis protein analyses showed that the GCSF2QAT protein was highly stable with very minimal protein degradation (visible as contaminated bands at 20-25kda). These LMW bands may have resulted from handling or partial cleavage of the tandem GCSF2QAT linker. The observed molecule at this LMW was predicted to be equivalent to GCSF ligand and non-glycosylated linker since similar bands corresponding to these were observed on western blot at elution 8. Analysis of pellet samples post dialysis showed the presence of a protein band at ~40kDa of low intensity, however this was not detected by western blotting.

Western blotting also did not pick up the LMW bands of pre and post dialysed samples present post coomassie staining. However, it did pick up the LMW band in elution 8 and this might confirm that LMW bands are not missing the antigenic site or are below detection limit of antibody, but the blots might not have been exposed long enough to see the contaminant bands. Nevertheless, the LMW proteins were of significantly lower intensity when compared to the intact molecule, suggesting an appreciably high

purity yield during purification. A total of 1.81mg of GCSF2QAT was recovered of >90% purity from 1 litre a media.

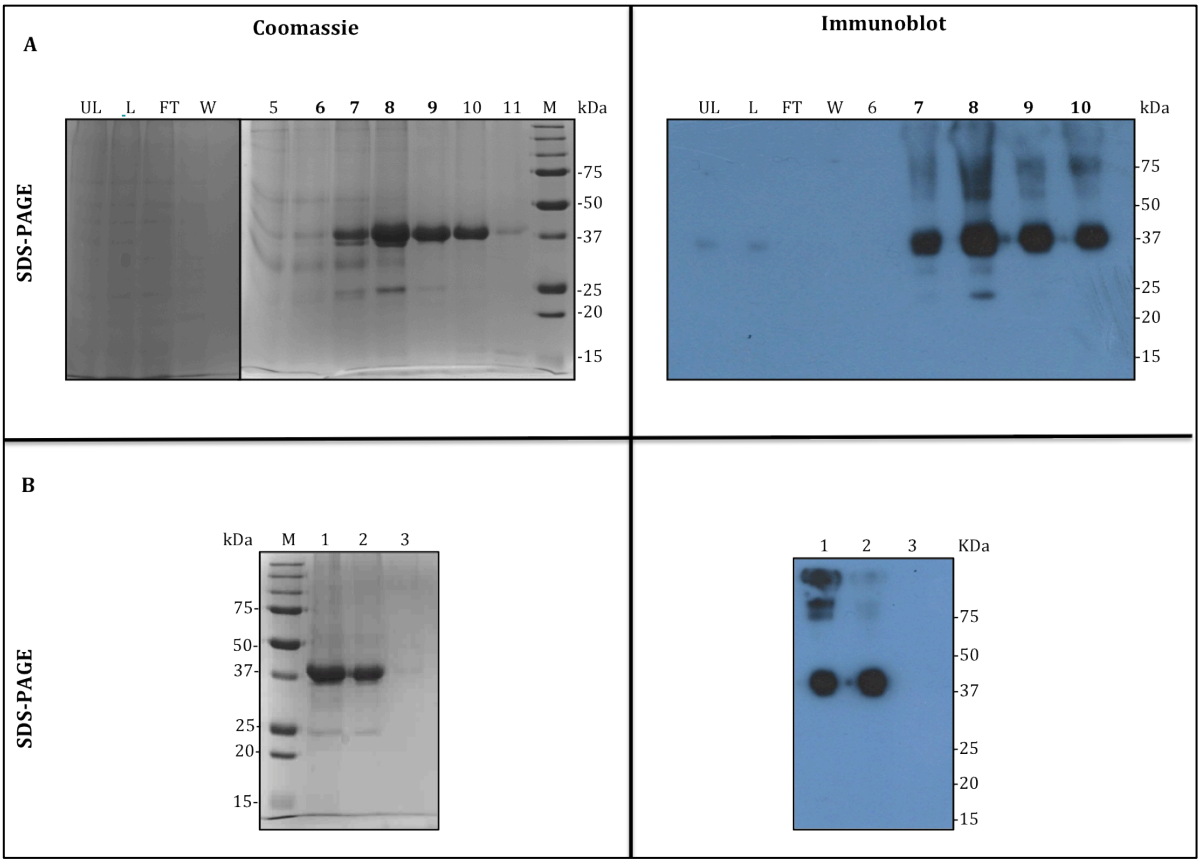


Figure 5-4: Purification development of IMAC for GCSF2QAT

10µl of protein samples from each stage of the GCSF2QAT purification process were separated on a 10% SDS-PAGE and stained with coomassie blue for visualization (A, left figure), while 100ng of same samples were analysed by Western blot. (A, right figure) (UL=Unfiltered load, L=Load, FL=Flow through, W=Wash pH7.4 & 6.0, M=Markers, 5-11= Elutions). Thereafter, active elution fractions were pooled, dialysed and analysed on a 10% SDS-PAGE gel by coomassie staining (B, left figure) and western blotting (B, right figure) (1= Pre-Dialysis, 2= Post-Dialysis and Post-Dialysis pellet).

Table 5-2: Protein concentrations of GCSF2QAT during the purification process

The table features the concentrations of GCSF2NAT at each stage of the purification process in the IMAC column. The elutions in bold typeface **(7, 8, 9 & 10)** were pooled together giving 2.47mg of >90% pure GCSF2QAT tandem protein, which equated to a 2.4% recovery from total protein, and a fraction (27%) of this was also lost during dialyses to give a final total amount of 1.81mg, which equated to 1.75% recovery from total protein.

Sample	Imidazole Concentration (nM)	Volume (ml)	Protein (µg/ml)	Total Protein (mg)	%Recovery
Unfiltered load	-	230	503.12	115.72	-
Load	-	230	448.93	103.25	100
Unbound	-	230	440.72	101.37	98
Wash pH 7.4	-	1	153.37	0.15	0.15
Wash pH 6.0		1	26.93	0.03	0.03
E1	50	1	150.08	0.15	0.15
E2	100	1	486.70	0.49	0.5
E3	100	1	260.10	0.26	0.25
E4	100	1	376.68	0.38	0.37
E5	200	1	567.16	0.57	0.55
E6	200	1	386.54	0.39	0.38
E7	**200**	**1**	**600.00**	**0.60**	**0.58**
E8	**350**	**1**	**987.52**	**0.99**	**0.96**
E9	**350**	**1**	**737.93**	**0.74**	**0.72**
E10	**350**	**1**	**476.85**	**0.48**	**0.46**
E11	500	1	148.44	0.15	0.15
Pre-Dialysis	-	6	616.42	2.47	2.4
Post-Dialysis	-	6	452.22	1.81	1.75
Pellet	-	0.1	56.49	0.01	0.01

5.3.2.3 Purification of GCSF4NAT

Culture media collected from GCSF4NAT expressing cells were concentrated and purified on a Nickel IMAC column. Western blot results showed that the majority of GCSF4NAT released from CHO cells were bound to the IMAC column with negligible amounts observed in the flow through (Figure 5.5.A). High purity protein (~95% pure) was eluted using imidazole elution buffers with concentrations ranging between 200mM and 350mM (shown in bold typeface 5,6,7,8 & 9 in figure 5.5.A). Elutions using with imidazole concentrations below or above this range (e.g. 100mM and 500mM) analysed were considered to be of low purity when analysed by SDS-PAGE and were thus excluded from the study. High purity elutions were collected from IMAC and pooled together for dialysis. Dialysis was performed to remove any excess imidazole or other salts (Figure 5.5.B). Pre and post dialysed proteins were measured by Bradford assay (Table 5.3) and analysed by 10% SDS PAGE followed by western blot analysis to confirm the integrity of purified GCSF4NAT protein and size.

In addition, A few other LMW bands observed by 10% SDS-PAGE (A, left figure) were confirmed to be contaminants since no bands corresponding to these were observed after dialysis (B, left figure). Western blotting also did not pick up the LMW bands of elutions samples present post coomassie staining. Nevertheless, the LMW proteins were of significantly lower intensity when compared to the intact molecule, suggesting an appreciably high purity yield during purification A total of 4.11mg of GCSF4NAT was recovered of >95% purity from 1 litre a media.

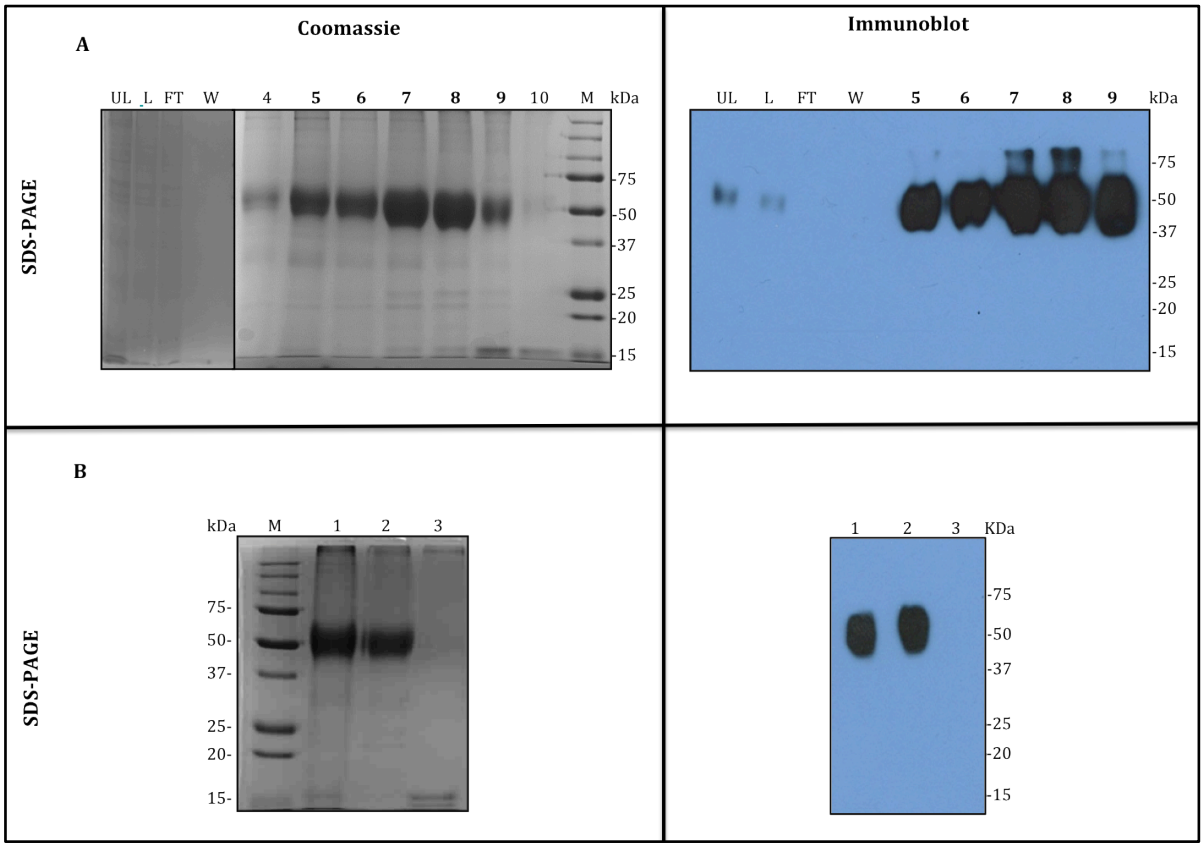


Figure 5-5: Purification analysis of IMAC for GCSF4NAT

10µl of protein samples from each stage of the GCSF2NAT purification process were separated on a 10% SDS-PAGE and stained with Coomassie blue for visualization (A, left figure), while 100ng of same samples were analysed by Western blot. (A, right figure) **(**UL=Unfiltered load, L=Load, FL=Flow through, W=Wash pH7.4 & 6.0, M=Markers, 4-10= Elutions). Thereafter, active elution fractions were pooled, dialysed and analysed on a 10% SDS-PAGE gel by coomassie staining (B, left figure) and western blotting (B, right figure) (1= Pre-Dialysis, 2= Post-Dialysis and Post-Dialysis pellet).

Table 5-3: Protein concentrations of GCSF4NAT during the purification process

The table features the concentrations of GCSF4NAT at each stage of the purification process in the IMAC column. The elutions in bold typeface **(5,6,7,8 & 9)** were pooled together giving 5.41mg of >95% pure GCSF4NAT tandem protein, which equated to a 5.2% recovery from total protein, and a fraction (24%) of this was also lost during dialyses to give a final total amount of 4.11mg, which equated to 4% recovery from total protein.

Sample	Imidazole Concentration (nM)	Volume (ml)	Protein (µg/ml)	Total Protein (mg)	Recovery
Unfiltered load	-	228	411.17	93.75	-
Load	-	228	452.22	103.11	100
Unbound	-	228	381.61	87.01	84
Wash pH 7.4	-	1	72.91	0.07	0.07
Wash pH 6.0		1	22.00	0.02	0.02
E1	100	1	146.80	0.15	0.15
E2	100	1	360.26	0.36	0.35
E3	100	1	205.91	0.21	0.2
E4	200	1	629.56	0.63	0.61
E5	**200**	**1**	**616.42**	**0.62**	**0.6**
E6	**200**	**1**	**856.16**	**0.86**	**0.83**
E7	**350**	**1**	**565.52**	**0.57**	**0.55**
E8	**350**	**1**	**1337.27**	**1.34**	**1.3**
E9	350	1	1141.87	1.14	1.1
E10	500	1	595.07	0.60	0.58
Pre-Dialysis	-	6	902.13	5.41	5.2
Post-Dialysis	-	6	685.46	4.11	4
Pellet	-	0.1	549.10	0.05	0.05

5.3.2.4 Purification of GCSF4QAT

Culture media collected from GCSF4QAT expressing cells were concentrated and purified on Nickel IMAC column. Western blotting results showed that the majority of GCSF4QAT protein bound to the IMAC column with negligible amounts observed in the flow through (Figure 5.6.A). High purity protein (>90% pure) was eluted using imidazole elution buffers with concentrations ranging between 200mM and 350mM (shown in bold typeface 9,10,11 & 12 in figure 5.6.A). Eluting using with imidazole concentrations below or above this range (100mM or 500mM respectively) were observed to be of low purity when analysed by SDS-PAGE and were thus excluded from the study. High purity elutions collected from IMAC purification were pooled together for dialysis. Dialysis was performed to remove any excess imidazole or other salts (Figure 5.6.B). Pre and post dialysed proteins were measured by Bradford assay (Table 5.4) and analysed by 10% SDS PAGE followed by western blot analysis to confirm the purification and size of GCSF4QAT.

Western blotting showed that the post-dialysis GCSF2NAT protein was stable with very minimal protein degradation (visible as contaminated bands at 20-25kda). These LMW bands may have resulted from handling or partial cleavage of the tandem GCSF4QAT linker (the observed molecule at this low molecular weight was predicted to be an equivalent of a GCSF ligand and non-glycosylated linker). Nevertheless, the LMW proteins were of significantly lower intensity when compared to the intact molecule, suggesting an appreciably high purity yield during purification. A total of 1.11mg of GCSF4QAT was recovered of >90% purity from 1 litre a media.

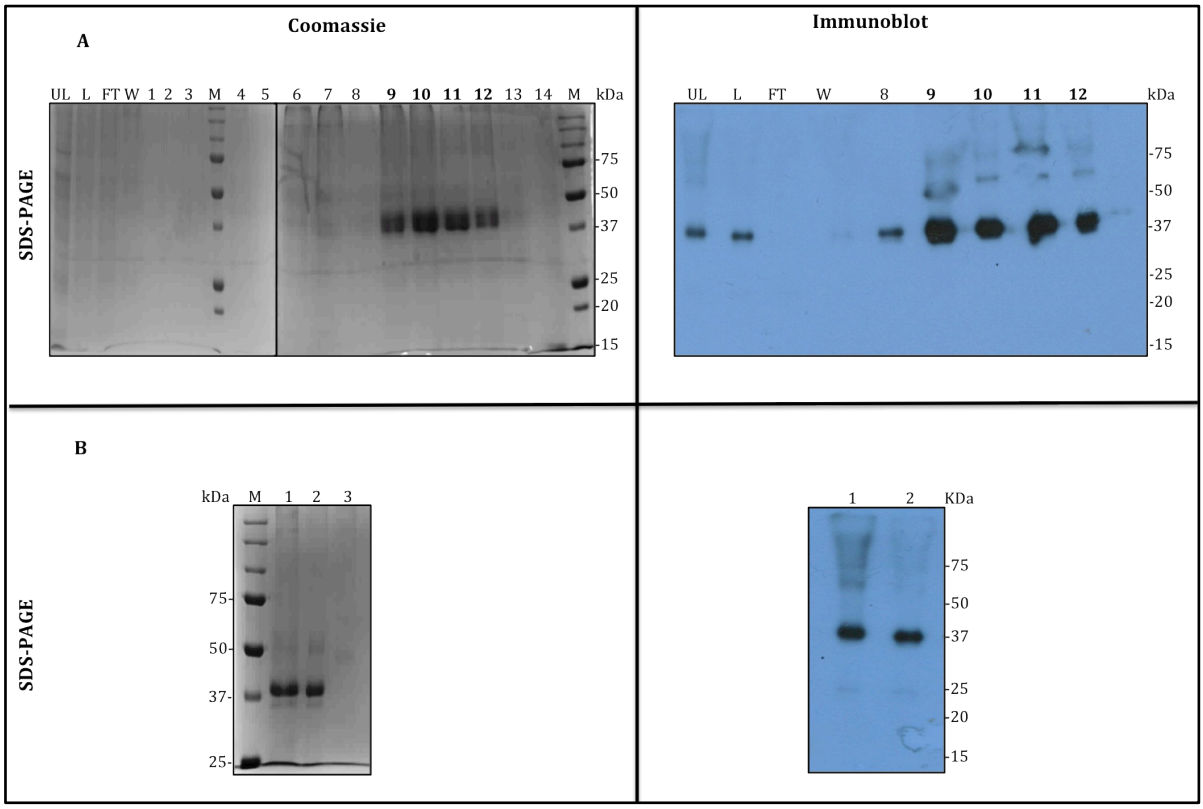


Figure 5-6: Purification analysis of IMAC for GCSF4QAT

10µl of protein samples from each stage of the GCSF4QAT purification process were separated on a 10% SDS-PAGE and stained with Coomassie blue for visualization (A, left figure), while 100ng of same samples were analysed by Western blot. (A, right figure) (UL=Unfiltered load, L=Load, FL=Flow through, W=Wash pH7.4, M=Markers, 1-14= Elutions). Thereafter, active elution fractions were pooled, dialysed and analysed on a 10% SDS-PAGE gel by coomassie staining (B, left figure) and western blotting (B, right figure) (1= Pre-Dialysis and 2= Post-Dialysis).

Table 5-4: Concentrations of GCSF4QAT during the purification process

The table features the concentrations of GCSF4QAT at each stage of the purification process in the IMAC column. The elutions in bold typeface **(9, 10, 11 &12)** were pooled together giving 1.83mg of >90% pure GCSF4QAT tandem protein, which equated to a 1.2% recovery from total protein, and a fraction (40%) of this was also lost during dialyses to give a final total amount of 1.11mg, which equated to 0.7% recovery from total protein.

Sample	Imidazole Concentration (nM)	Volume (ml)	Protein (µg/ml)	Total Protein (mg)	Recovery
Unfiltered load	-	217	1472.6	319.55	-
Load	-	217	700.82	152.08	100
Unbound	-	217	638.42	138.54	91
Wash pH 7.4	-	1	247.62	0.25	0.16
E1	50	1	0.00	0.00	0
E2	50	1	0.00	0.00	0
E3	50	1	0.00	0.00	0
E4	100	1	0.00	0.00	0
E5	100	1	0.00	0.00	0
E6	100	1	208.21	0.21	0.14
E7	200	1	188.51	0.19	0.13
E8	200	1	165.52	0.17	0.11
E9	**200**	**1**	**667.98**	**0.67**	**0.44**
E10	**350**	**1**	**671.26**	**0.67**	**0.44**
E11	**350**	**1**	**444.66**	**0.44**	**0.29**
E12	**350**	**1**	**214.78**	**0.21**	**0.14**
E13	500	1	0.00	0.00	0
Pre-Dialysis	-	5	366.83	1.83	1.2
Post-Dialysis	-	5	222.33	1.11	0.7
Pellet	-	0.1	41.71	0.00	0

5.3.2.5 Purification of GCSF8NAT

Analyses of GCSF8NAT protein pre- and post- purification by western blotting showed that the GCSF8NAT protein in the culture medium (secreted by CHO cells) was able to bind to the IMAC column. However, the results equally showed protein degradation for this particular tandem protein in both the crude culture media sample (Figure 5.7.A) and the purified elution samples (Figure 5.7.B). For the culture media samples, the western blot results (Figure 5.7.A) showed consistent increase in expressed GCSF8NAT protein by CHO cells in the culture media during the 9 days incubation period, running at the right molecular weight of ~70kDa. However, below the target protein bands, another band running at a LMW (~37kDa) was also observed. The degradation increased during purification (seen as triplet bands running at ~15-25kDa and another band at ~37kDa), suggesting that the tandem GCSF8NAT is not as stable as other tandems. Therefore, a number of measures were put in place to improve the stability of the protein.

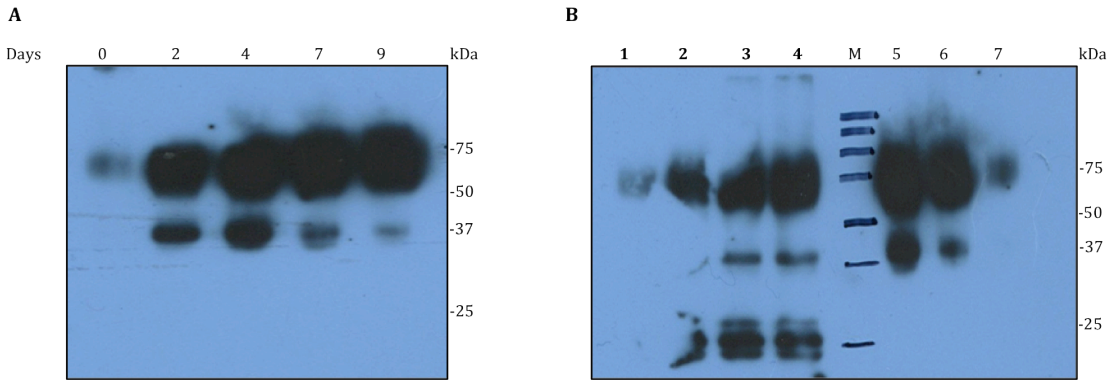


Figure 5-7: Western blot analysis of roller bottle media samples for GCSF8NAT

(A) Results of western blot shows increasing protein expression from day 0 to day 9 for GCSF8NAT. **(B)** Western blot analysis showing IMAC elutions and dialysed samples for GCSF8NAT (M=Markers, 1-4= Elutions, 5= Pre-Dialysis, 6= Post-Dialysis & 7= Post-Dialyses Pellet).

To improve the stability of GCSF8NAT, the incubation temperature for the stable cell growth and the culture volume were optimised. The CHO cells of GCSF8NAT were grown in 2 litres not 1 litre of Hyclone media at 31°C (as against previous 37°C). The cells were seeded at 0.25x10^6/ml, and allowed to grow for 7 days or until viability reached 70% (rather than 9 days or allowing the variability to decline just below 70%, respectively). Western blotting analyses of the protein samples in the culture media (Figure 5.8) showed a consistent increase in GCSF8NAT protein expression during the 7 days of cell growth (bands running at ~50-70kDa) with no other visible bands observed, which is an indication that this adjustments enhanced the stability of GCSF8NAT.

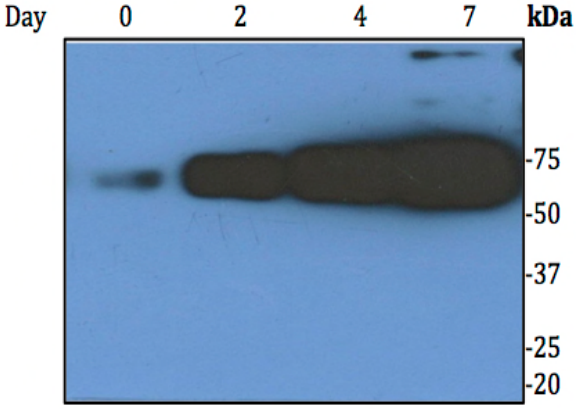


Figure 5-8: Western blot analysis of roller bottle media samples for GCSF8NAT

10µl samples taken at 2-3 days intervals from GCSF8NAT culture media grown for 7 days were separated by SDS-PAGE and analysed by western blotting. The western blot results shows a consistent increase in protein expression from day 0 to day 7 for GCSF8NAT.

Due to earlier challenges encountered with GCSF8NAT degradation and instability at 37°C, the GCSF8NAT protein collected from the culture under the optimised conditions was purified at 4°C to avoid potential protein degradation. Western blotting showed that the GCSF8NAT protein in the crude culture media expressed from CHO cells was able to bind the IMAC column and are stable when eluted (Figure 5.9.A). The bound protein was eluted with a gradient concentration of imidazole and majority of high purity protein was eluted between an imidazole concentration of 200nM and 350nM (shown in bold typeface 4, 5, 6 & 7 in Figure 5.9.A). Elutions using with imidazole concentrations below or above this range (100nM and 500nM respectively) were considered to be of low purity when analysed by SDS-PAGE and thus excluded from the study. High purity elutions were collected from IMAC and pooled together and dialysed in PBS so as to remove any excess imidazole or other salts (Figure 5.9.B). Pre and post dialysed protein were measured by Bradford assay (Table 5.5) and analysed by 10% SDS PAGE followed by coomassie staining and western blot analysis in order to confirm the purification and size of the GCSF8NAT. Coomassie stained 10% SDS -PAGE showed that the post-dialysis GCSF8NAT protein was stable with very minimal protein degradation (visible as contaminated bands at ~40kda). Analysis of pellet samples post dialysis does show the presence of a protein band at ~70kDa, however this was not detected by western blotting.

Western blotting also did not pick up the LMW bands of pre and post dialysed samples present post coomasie staining. It was presumed that these LMW bands are missing the antigenic site or are below detection limit of antibody. Also, the blots might not have been exposed long enough to see the contaminant bands. Nevertheless, the LMW proteins were of significantly lower intensity when compared to the intact molecule, suggesting an appreciably high purity yield during purification. A total of 2.91mg of GCSF8NAT was recovered of >95% purity from 2 litre a media.

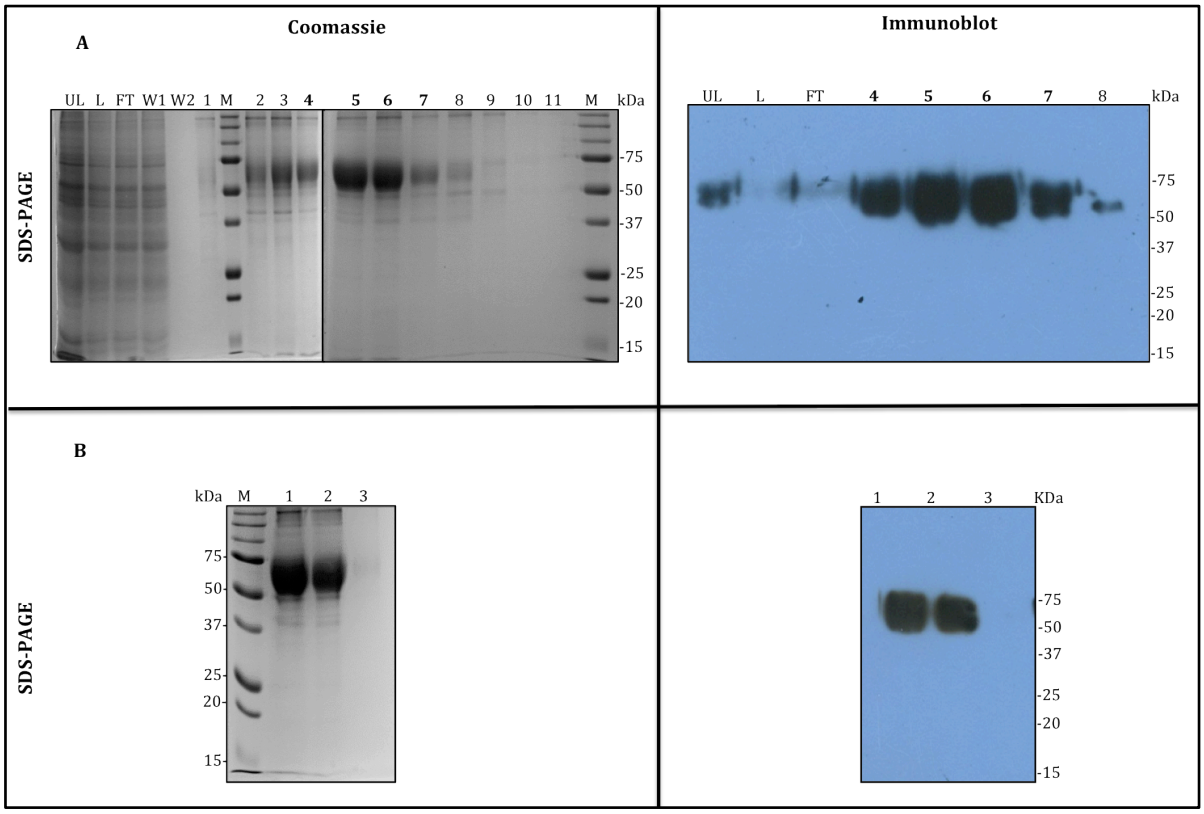


Figure 5-9: Purification analysis of IMAC samples for GCSF8NAT

10µl of protein samples from each stage of the GCSF8NAT purification process were separated on a 10% SDS-PAGE and stained with coomassie blue for visualization (A, left figure), while 100ng of same samples were analysed by western blot. (A, right figure) (UL=Unfiltered load, L=Load, FL=Flow through, W=Wash pH7.4, 6.0 M=Markers, 1-11= Elutions). Thereafter, active elution fractions were pooled, dialysed and analysed on a 10% SDS-PAGE gel by coomasie staining (B, left figure) and western blotting (B, right figure) (1= Pre-Dialysis and 2= Post-Dialysis & 3= Post-Dialysis pellet).

Table 5-5: Concentrations of GCSF8NAT during the purification process

The table features the concentrations of GCSF4NAT at each stage of the purification process in the IMAC column. The elutions in bold typeface **(4, 5, 6 & 7)** were pooled together giving 3.21mg of >95% pure GCSF8NAT tandem protein, which equated to a 0.67% recovery from total protein with a fraction (9%) of this was also lost during dialyses to give a final total amount of 2.91mg, which equated to 0.61% recovery from total protein.

Sample	Imidazole Concentration (nM)	Volume (ml)	Protein (µg/ml)	Total Protein (mg)	Recovery
Unfiltered load	-	220	3787.74	833.30	-
Load	-	220	2162.52	475.75	100
Unbound	-	220	2074.96	456.49	96
Wash pH 7.4	-	1	3952.36	3.95	0.83
Wash pH 6.0		1	22.42	0.02	0.004
E3	200	1	1129.25	1.13	0.24
E4	**200**	**1**	**814.01**	**0.81**	**0.17**
E5	**200**	**1**	**1560.07**	**1.56**	**0.33**
E6	**350**	**1**	**1528.55**	**1.53**	**0.32**
E7	**350**	**1**	**589.84**	**0.59**	**0.33**
E8	350	1	635.38	0.64	0.13
Pre-Dialysis	-	4	802.80	3.21	0.67
Post-Dialysis	-	4	727.50	2.91	0.61
Pellet	-	0.1	601.40	0.06	0.01

5.3.2.6 Purification of GCSF8QAT

Crude culture medium from stable CHO cells expressing tandem GCSF8QAT proteins were passed over an IMAC column. Western blotting results showed that the majority of the tandem protein was bound to the IMAC column and only very little was observed in the flow through (Figure 5.10.A Wrong). High purity protein was eluted between an imidazole concentration of 200nM and 350nM (shown in bold typeface 5, 6, 7 & 8 in figure 5.10.A). Other elutions below or above this range (100nM and 500nM respectively) were observed to be of low purity when analysed by SDS-PAGE and thus excluded from the study. High purity elutions were collected and pooled together for dialysis so as to remove any excess imidazole or other salts (Figure 5.10.B). Pre and post dialysed protein samples were measured by Bradford assay (Table 5.6) and analysed by 10% SDS PAGE followed by western blot analysis to confirm the purification and size of GCSF8QAT.

Coomassie stained 10% SDS –PAGE showed the post-dialysis GCSF8QAT protein was stable as no degradation or other visible contaminated bands were observed by western blot. However, a very minimal amount of contaminated bands (presumed to be degraded tandem GCSF8QAT protein) are observed on the coomassie stained gel, which are insignificant when compared to the amount of pure protein running at the predicted molecular weight. A total of 2.93mg of GCSF8QAT was recovered of >95% purity from 1 litre a media.

Also, dimer formation was observed pre and post dialysis which could be seen as bands running at twice the predicted molecular weight of GCSF8QAT in the western blotting image (Figure 5.10.B, right picture).

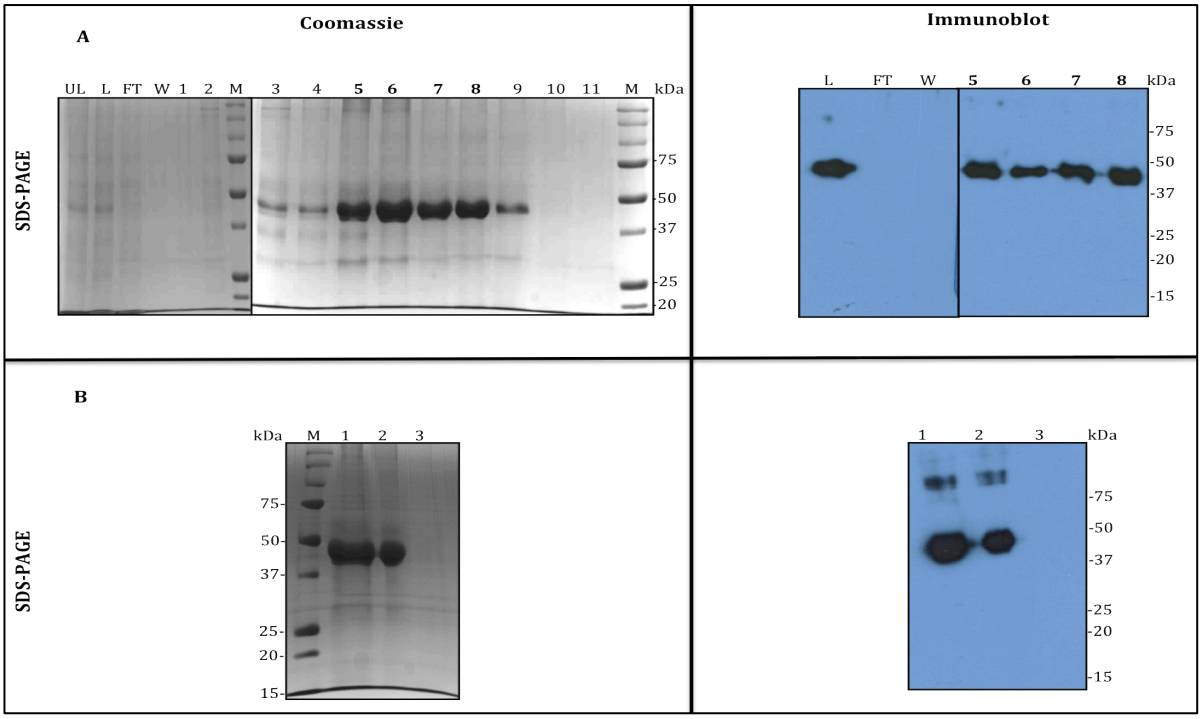


Figure 5-10: Purification development of IMAC for GCSF8QAT

10µl of protein samples from each stage of the GCSF8QAT purification process were separated on a 10% SDS-PAGE and stained with coomassie blue for visualization (A, left figure), while 100ng of same samples were analysed by Western blot. (A, right figure) (UL=Unfiltered load, L=Load, FL=Flow through, W=Wash pH7.4 & 6.0, M=Markers, 1-11= Elutions). Thereafter, active elution fractions were pooled, dialysed and analysed on a 10% SDSPAGE gel by coomassie staining (B, left figure) and western blotting (B, right figure) (1= Pre-Dialysis and 2= Post-Dialysis & 3= Post-D-Pellet).

Table 5-6: Protein concentrations of GCSF8QAT during the purification process

The table features the concentrations of GCSF4QAT at each stage of the purification process in the IMAC column. Elutions **5 to 8** were pooled together giving 3.96mg of >95% pure GCSF8QAT tandem protein, which equated to a 1.5% recovery from total protein pre-dialyses and a fraction (27%) of this was also lost during dialyses to give a final total amount of 2.93mg, which equated to 1.1% recovery from total protein.

Sample	Imidazole Concentration (nM)	Volume (ml)	Protein (µg/ml)	Total Protein (mg)	Recovery
Unfiltered load	-	230	1406.90	323.59	-
Load	-	230	1144.17	263.16	100
Unbound	-	230	779.64	179.32	68
Wash pH 7.4	-	1	204.93	0.20	0.08
E1	50	1	116.26	0.12	0.05
E2	50	1	418.39	0.42	0.16
E3	200	1	707.39	0.71	0.27
E4	200	1	592.45	0.59	0.22
E5	**200**	**1**	**1232.84**	**1.23**	**0.5**
E6	**350**	**1**	**1594.09**	**1.59**	**0.6**
E7	**350**	**1**	**1091.63**	**1.09**	**0.4**
E8	**350**	**1**	**1265.68**	**1.27**	**0.5**
E9	500	1	648.28	0.65	0.24
Pre-Dialysis	-	4	990.80	3.96	1.5
Post-Dialysis	-	4	732.75	2.93	1.1
Pellet	-	0.1	56.49	0.01	0.004

5.3.2.7 Summary of GCSF Protein Tandems Purification

It was possible to purify GCSF tandems linked by a flexible linker $(Gly_4Ser)_n$ from a mammalian cell line (CHO Flp-In cells). The glycosylated tandems showed an increase in molecular weight above that of their controls (non-glycosylated tandems) as assessed by SDS-PAGE and western blotting (Figure 5.11).

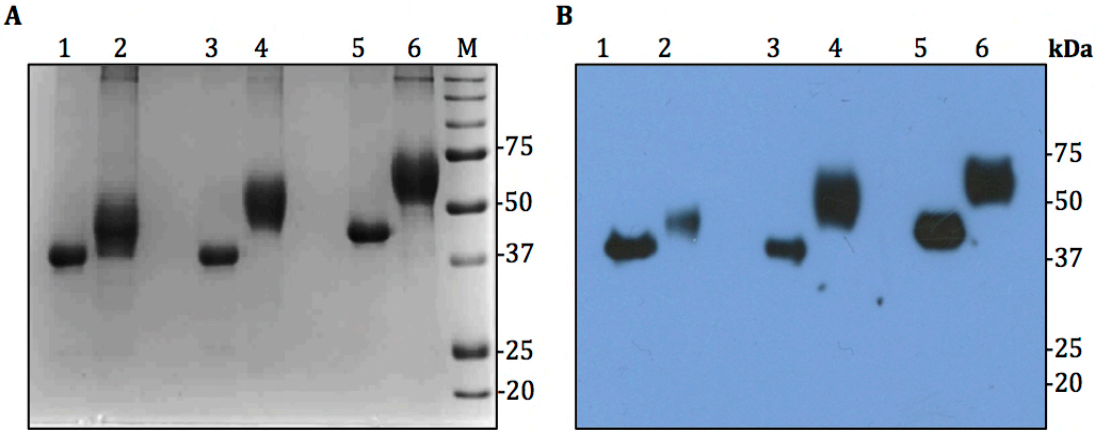


Figure 5-11: Purified GCSF tandems analysed by coomassie blue and western blot

(A) Purified GCSF tandem molecules analysed by coomassie blue. Lane 1; GCSF2QAT, Lane 2; GCSF2NAT, Lane 3; GCSF4QAT, Lane 4; GCSF4NAT, Lane 5; GCSF8QAT, Lane 6; GCSF8NAT, Lane M; 1kb marker. A total of 7.5µg protein was loaded per lane

(B) Purified GCSF tandem molecules analysed by western blot. Lane 1; GCSF 2QAT, Lane 2; GCSF2NAT, Lane 3; GCSF 4QAT, Lane 4; GCSF4NAT, Lane 5; GCSF 8QAT, Lane 6; GCSF8NAT. A total of 100ng protein was loaded per lane

In addition, both coomassie blue and western blotting analyses showed a large smeared band for each individual glycosylated GCSF tandem when compared to the corresponding non-glycosylated tandem control, which could be attributed to a large heterogeneous protein population. The reason for the observed increasing degree of population heterogeneity across the tandem GCSF proteins will be discussed later.

5.4 Discussion

The expression studies of the GCSF tandems containing a C-terminal histidine tag showed that the target genes could be stably expressed in a CHO cell line. However, glycosylated tandems showed high level of expression compared to non-glycosylated tandems (i.e. both 2QAT and 8QAT showed the lowest level of expression during culture growth). Since there is no difference between glycosylated and non-glycosylated tandems except for the linker, it seems that the presence of NAT motifs is enhancing DNA transcription more than QAT motifs, which in turn increased the productivity of protein.

The purification data indicated that IMAC was an appropriate method for purifying the GCSF tandems. All GCSF tandems were easily eluted from the IMAC by adding a high concentration of imidazole (an analogue of Histidine). IMAC permits a fast and easy way for purification due to the strong affinity of a nickel-complex Ni^{2+}-NTA to the histidine sequences of the GCSF tandems. Consequently, all constructs produced sufficient quantities of pure protein (between 1 to 4mg per litre as assessed by Bradford assay), which was sufficient for the PK/PD studies, and was considered to be 90-95% pure as assessed by SDS-PAGE followed by western blotting to confirm the molecular weight and integrity of each construct.

Some contaminating bands were observed for all GCSF tandems during the purification process as evidenced by coomassie staining. While these bands appeared to be degraded products, the amount of degradation observed was negligible when compared to the stable non-degraded proteins. Interestingly, most of the contaminated or degraded GCSF tandems were separated from the stable and pure proteins by dialysis through a semipermeable membrane with 10kDa molecular weight cut-off (e.g. a cellulose membrane with pores). Dialysis technique is an important step after purification for removing any excess imidazole, salts or other small molecules. Molecules that have sizes bigger than the cellulose membrane pores were retained inside the dialysis bag, while other smaller molecules or salts diffuse through the pores into the dialyses buffer Post-dialysis

analysis by coomassie staining or western blotting showed all GCSF tandems to be of the correct molecular weight with no degradation as there were no visible contaminants or degraded bands observed, suggesting that the dialysis was efficient for removing the contaminants for most of the purified samples.

However, for the majority of tandem molecules expression and purification was completed without difficulty, however for one tandem in particular (GCSF8NAT) we experienced difficulties with the stability during the expression and purification. It was difficult to determine what caused the degradation, as other tandems were stable when purified using the same purification method. To investigate the reason for GCSF8NAT instability, samples were taken during cell growth in roller bottles before, during and after the purification. The protein samples were analysed by western blotting, and the results revealed that the degradation occurred during both protein expression and purification processes. Consequently, the conditions for expressions and purifications were modified with a view to improving the stability of GCSF8NAT during the entire protein production processes.

Firstly the incubation time was reduced from 9 days to 7 days so that cells could be harvested at a higher viability (60% or higher) as it was earlier noticed that growing CHO cells for longer time points led to increased protein degradation, probably due to increased protease activity associated with the increased cell death.

The temperature at which growth and expression was also reduced to 31°C from 37°C. These adjustments made to the conditions of protein expression were necessary, as it has been shown that while longer linkers were preferable for the preservation of the independent folding and biological activities of two proteins, they could also be easily cleaved by proteases, because the structures and the adjacent regions of these linkers are more loosely connected (Liu et al., 2005).

A similar pattern was observed with the GCSF8NAT, which has a long linker similar to GCSF8QAT (non-glycosylated control). In the initial expression

studies under the same conditions, GCSF8NAT protein was degraded whereas GCSF8QAT was not. Another contributory factor to GCSF8NAT degradation could be associated with the presence of 8 N-linked glycosylation sites, which restricts intramolecular interaction between the two GCSF ligands and thus leaving the linker exposed to proteolytic attack. However, During protein expression, there is a probability that the GCSF molecules in GCSF8QAT (non-glycosylated control) for instance would bind to each other due to the formation of intra-molecular disulphide bond resulting from the interaction of free cysteine residues (Cys17) that were present on the individual GCSF molecule (which was observed to be the case as dimer formation was earlier observed for the tandem GCSF8QAT but not in GCSF8NAT) Consequently, the binding of the GCSF molecules in GCSF8QAT would protect the linker, to an extent, from proteolytic cleavage.

The beneficiary effect of decreasing culture temperature was seen on EPO production in CHO cell line. Consequently, a 2.5-fold increase in the maximum concentration of EPO was achieved by decreasing temperature from 37°C to 33°C (Yoon et al., 2003). Lowering the culture temperature increase protein expression by slowing down protein translation and thus facilitates correct folding of protein.

Lowering the temperature during culture of GCSF8NAT equally enhanced the protein production when compared to other tandem GCSF proteins. This is evidenced by a 49mg/L productivity with GCSF8NAT at day 7 during cell growth in roller bottles compared to 21mg/L, 19.3mg/L, 8.1mg/L, 34.5mg/L and 16.3mg/L yield with GCSF4NAT, GCSF2NAT, GCSF2QAT, GCSF4QAT and GCSF8QAT respectively.

Also, the adjustments made to the conditions of expression cells growth for the tandem GCSF8NAT was further justified by the abrogation of GCSF protein degradation that were earlier observed during initial protein expression and purification. To prevent potential protein degradation during the purification process, especially for GCSF8NAT, all the tandems GCSF purifications were purified at 4°C. The western blotting was used to

confirm the stability since degraded proteins were not easily visible in the gel with coomassie staining (~5-10µg purified protein), whereas western blot offers a better detection with ~0.1µg of protein load.

Tandem proteins analysed by SDS-PAGE and western blotting also revealed an obvious increase in the population heterogeneity in glycosylated tandems that was not apparent in the non-glycosylated tandem controls. A possible reason for the observed high heterogeneity of the population could be due to both macro- (glycosylation site occupancy) and micro- (structure of glycosylation relating length, composition and branching pattern), which is frequently present in the final population of recombinantly expressed glycosylated proteins (Sola and Griebenow, 2009 and Sinclair and Elliott, 2005).

In conclusion the glycosylated GCSF tandems and their respective non-glycosylated controls were successfully purified by IMAC and the purity was 90-95% as assessed by SDS-PAGE and western blotting. However, before being passed over for the future PK/PD *in vivo* study it was required to measure their biological activity. This will be discussed in the next chapter.

6. Results 3: *In vitro* Bioactivity Evaluation and Temperature Stability of GCSF Tandems

6.1 Summary

Purification of glycosylated and non-glycosylated GCSF tandem molecules was achievable, as evidenced by the previous chapter. The purified glycosylated tandem molecules also showed increased molecular weight above that of controls when analysed by SDS-PAGE and western blotting. However, before analysing the pharmacokinetics and pharmacodynamics of these tandem molecules, it is imperative to analyse their in vitro bioactivity and stability. In this study, the in vitro bioactivity of the GCSF tandems was tested using an AML-193 proliferation assay. These cells have been shown to proliferate in response to GCSF treatment. The short-term stability of GCSF tandems was investigated by two different ways. First by testing samples from the stability experiment in the AML-193 assay at 37°C for 3 days. Second by testing purified samples from stock at 3 different temperatures (4°C, room temperature (RmT) and -80°C freeze thaw (F/T) cycles) over an 8 day period. The results indicated significant increased bioactivity for GCSF tandems compared with rhGCSF (EC_{50} for tandems were about 3-fold lower than that for rhGCSF). All GCSF tandems showed good stability with no visible signs of degradation under all conditions studied (4°C, RmT and -80°C freeze thaw) over an 8 days period and also after incubation with AML-193 cells at 37°C for 3 days.

6.2 Introduction

Acute myeloid leukemia (AML) is a disorder of the myeloid lineage characterized by the quick growth of abnormal white blood cells (WBCs) that accumulate in the bone marrow and result of exhibiting a new morphological and immunophenotypic features to the myeloid lineage (Vardiman et al., 2009). AML-193 is one of 8 cell lines that were established from 50 patients with childhood acute leukemia (Lange et al., 1987, Valtieri et al., 1991). From those eight cell lines, AML-193 was derived from a 13-year-old female with acute myeloid leukemia, which was then the only cell line that required conditioned media to grow and proliferate. In vitro, cytokines such as GM-CSF and GCSF have been identified to support the growth and proliferation of hematopoietic stem cells. Interestingly, these cytokines have also been shown to support the growth and proliferation of AML-193 cells (Favreau and Sathyanarayana, 2012). Based on these findings, we hypothesized that AML-193 cell line would proliferate in response to GCSF tandem proteins in a manner comparable to the available rhGCSF. Therefore the proliferation bioassay was used to test the hypothesis.

The basic proliferation bioassay is a colorimetric method routinely used to determine the number of viable AML-193 cells. As earlier mentioned, AML-193 growth is stimulated in the presence of an active GCSF protein. A commercially available MTS reagent which contains a tetrazolium compound and an electron coupling reagent is used to estimate the proliferation of cells. Generally, metabolically active cells produce dehydrogenase enzymes that produce NADPH into the medium as a by-product. The produced NADPH reduces the MTS tetrazolium compound in the medium to give a yellow coloration (formazan product is soluble in tissue culture media), which can be read photometrically at 490nm (Berridge and Tan, 1993). The MTS compound is hydrolyzed by the NADPH producing cells into a yellow coloured product. The intensity of the colour produced is directly proportional to the concentration of the active enzymes present in the medium, which in turn is dependent on the number of cells

present (Cory et al., 1991). Therefore cell growth is directly related to the level of NADPH in the media and hence the activity of GCSF.

In the modified proliferation assay used in this study, the AML-193 cells were seeded in 96 well plates to the rhGCSF, GCSF tandems and their controls separately and grown for 3 days. MTS reagent was then added to estimate the proliferation of AML-193 cells that respond to the presence of GCSF in the medium. Readings were taken at 490nm every 40 minutes for 2 hrs using a 96 well plate reader.

6.2.1 Aim and Hypothesis

It is essential that the biological activity of a protein be retained following its purification. Maintaining the activity of the target protein is especially important for the future pharmacokinetic (PK) *in vivo* study. Also, proteins are best known to perform their biological functions when in the right conformation. Hence, it is essential to investigate the biological activities of the GCSF tandems, at different temperature conditions to determine the stability of the molecules. Therefore, the aim of this chapter is to measure the biological activity of the glycosylated GCSF tandems using AML-193 cell line, and to determine the short-term stability of these constructs under different temperatures. The stability was assessed at 3 different temperatures (4°C, RmT and -80°C freeze thaw (F/T) cycles) over 8 days period. Also, samples of GCSF tandems were taken after incubation with AML-193 cells at 37°C for 3 days and tested for stability. We hypothesized that our GCSF tandems will be biologically active and stable similar to the rhGCSF post-purification.

6.3 Results

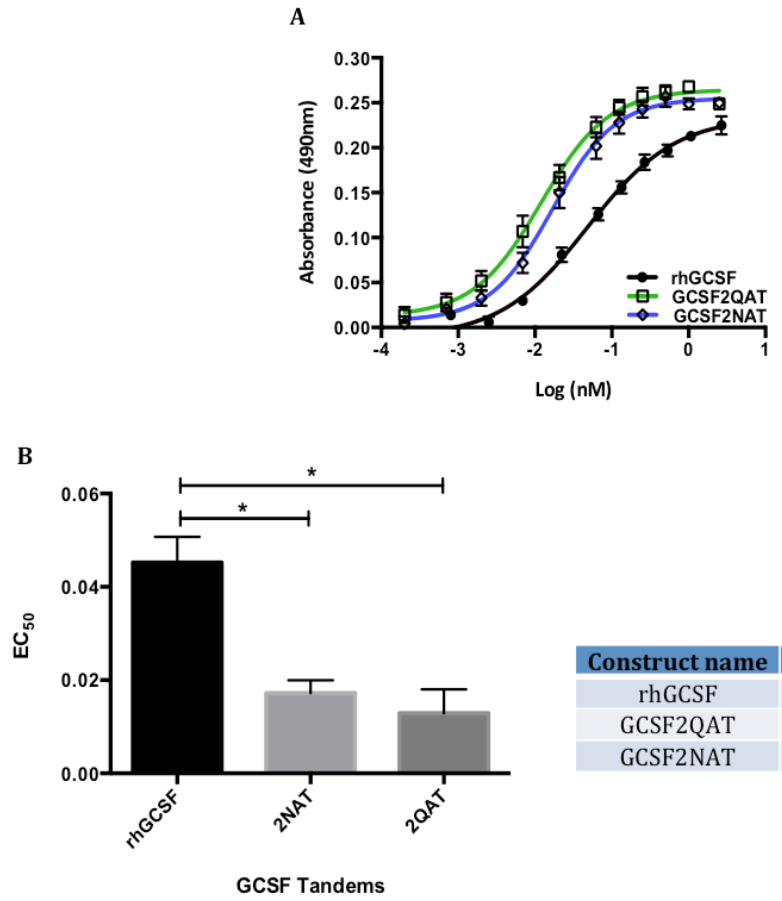
6.3.1 *In vitro* Bioactivity Evaluation

The biological activity of GCSF tandems was evaluated using an AML-193 cell-based proliferation assay. AML-193 cells were removed from liquid nitrogen storage and prepared as described in Section 3.8.5.1. 50µl of serial dilutions of rhGCSF and GCSF tandem proteins were added into the appropriate well of a 96-well microplate. 50µl of AML-193 cell suspension at a density of $5x10^5$ cells/ml was then added to the same plate and shaken gently to allow mixing of cell suspension with samples. Control wells contained assay medium and AML-193 cells (50µl + 50µl) and blank wells contained only assay medium (100µl) as previously described (Section 3.8.5.2).

Cells were exposed to different concentrations of test proteins in CO_2 incubator (5% CO_2, 37°C) and then 20µl of MTS (Celltiter 96 aqueous one solution cell proliferation assay from Promega) was added to each well to give a colour change that determines the number of viable cells in the assay. Readings were taken every 40 minutes for 2 hrs of incubation (5% CO_2, 37°C). The plates were finally read at 490nm using a BioTek FLX800 plate reader and Gen5 software as previously described (Section 3.8.5.2). The results from different concentrations of test protein were subtracted from control wells (AML-193 cells only) obtained at 490nm. Each experiment was repeated at least in triplicate.

6.3.1.1 *In vitro* Biological Activity of GCSF2NAT and Its Control

The results indicate that rhGCSF, purified GCSF2NAT and GCSF2QAT tandems can stimulate the proliferation of the AML-193 cell line. Both purified GCSF2NAT and GCSF2QAT tandems show a significant biological activity with both standard curves shifted to the left in comparison to rhGCSF. Besides this, both tandems show a greater maximal stimulation of AML193 cells compared to rhGCSF (Figure 6.1.A). Based on the findings of sigmoidal dose response curves (GraphPad Prism), the values of EC_{50} (the concentration which causes 50% of maximal response) for each protein were calculated (Figure 6.1.B).



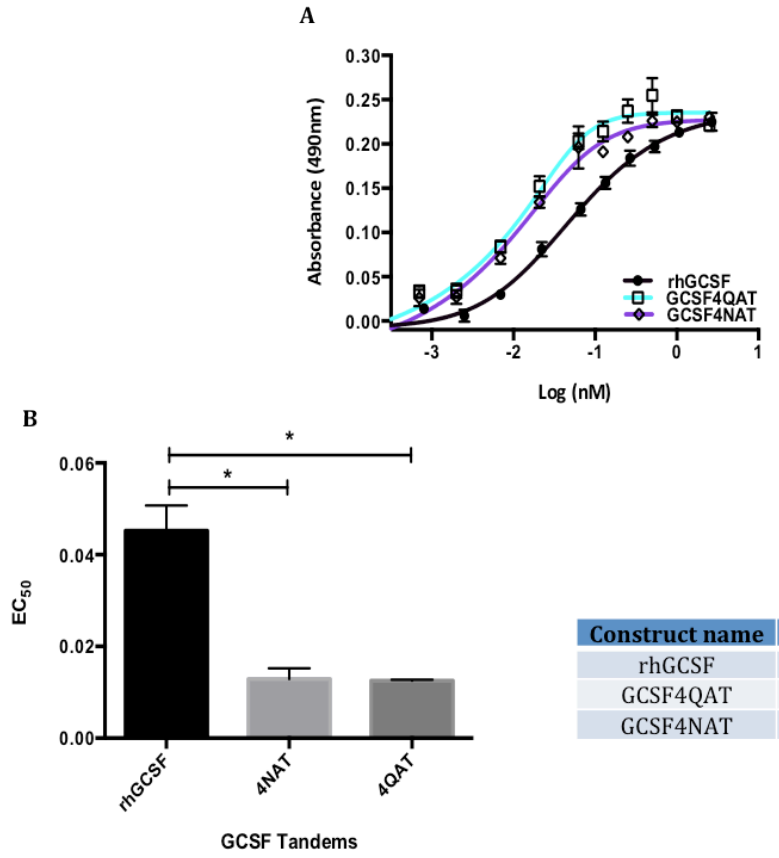
Construct name	EC50 (nM)	SEM
rhGCSF	0.045 (n=7)	0.0031
GCSF2QAT	0.0129* (n=3)	0.0062
GCSF2NAT	0.0172* (n=3)	0.0034

Figure 6-1: Proliferation of AML-193 cells in the presence of tandems and rhGCSF

(A) AML-193 cells were stimulated with GCSF2NAT, GCSF2QAT tandems and rhGCSF. The absorbance values, which are indicative of the number of live cells present, were plotted against the natural logarithm of the GCSF concentrations (rhGCSF and GCSF tandems). Sigmoidal dose-response fit (for variable slope) was used for the data analyses. **(B)** EC_{50} values calculated for GCSF2NAT, GCSF2QAT tandem & rhGCSF using GraphPad Prism software. Significance between EC_{50} values of GCSF tandems and rhGCSF was performed with Graphpad prism using Mann-Whitney test. Results are given as standard mean of error (SEM) for triplicate wells in graph A and triplicate EC_{50} in graph B. (* = p value of <0.05)

6.3.1.2 *In vitro* Biological Activity of GCSF4NAT and Its Control

The results below show that rhGCSF, GCSF4NAT and GCSF4QAT can stimulate the proliferation of the AML-193 cell line. Both tandems show significant increased bioactivity with both standard curves shifted to the left in comparison to rhGCSF (Figure 6.2.A). Based on the findings of sigmoidal dose response curves (GraphPad Prism), the values of EC_{50} were shown in the Figure 6.2.B.



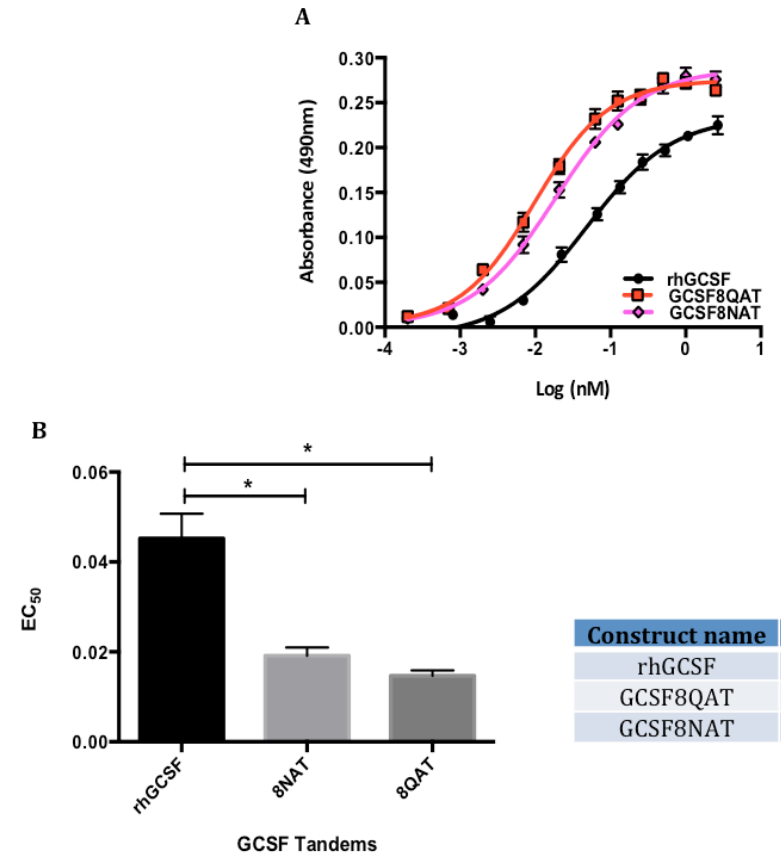
Construct name	EC50 (nM)	SEM
rhGCSF	0.045 (n=7)	0.0031
GCSF4QAT	0.0125* (n=3)	0.0002
GCSF4NAT	0.0129* (n=3)	0.0029

Figure 6-2: Proliferation of AML-193 cells in the presence of tandems and rhGCSF

(A) AML-193 cells were stimulated with GCSF4NAT, GCSF4QAT tandems and rhGCSF. **(B)** Shows EC_{50} values calculated for GCSF4NAT, GCSF4QAT tandem & rhGCSF Graph using Pad Prism software. Significance between EC_{50} values of GCSF tandems and rhGCSF was performed with GraphPad Prism using Mann-Whitney test. Results are given as SEM for triplicate wells in graph A and triplicate EC_{50} in graph B. (* = p value of <0.05)

6.3.1.3 *In vitro* Biological Activity of GCSF8NAT and Its Control

The bioactivity results for GCSF8NAT & GCSF8QAT show similar results to previous tandem proteins and together with rhGCSF can stimulate the proliferation of the AML-193 cell line. Both tandems show significant increased biological activity with both standard curves shifted to the left in comparison to rhGCSF. Both tandems also show a greater maximal level of activity than rhGCSF (Figure 6.3.A). Based on the findings of sigmoidal dose response curves (GraphPad Prism), the values of EC_{50} were calculated (Figure 6.3.B).



Construct name	EC50 (nM)	SEM
rhGCSF	0.045 (n=7)	0.0031
GCSF8QAT	0.0147*(n=3)	0.0015
GCSF8NAT	0.0191*(n=3)	0.0023

Figure 6-3: Proliferation of AML-193 cells in the presence of tandems and rhGCSF

(A) AML-193 cells were stimulated with GCSF8NAT, GCSF8QAT tandems and rhGCSF. **(B)** Shows EC_{50} values calculated for GCSF4NAT, GCSF4QAT tandem & rhGCSF using GraphPad Prism software. Significance between EC_{50} values of GCSF tandems and rhGCSF was performed with GraphPad Prism using Mann-Whitney test. Results are given as SEM for triplicate wells in graph A and triplicate EC_{50} in graph B. (* = p value of <0.05)

6.3.2 Short Term Stability of GCSF Tandem Molecules

The protein stability was assessed by two different ways. First by testing samples from the stability experiment in the AML-193 assay. Second by testing purified samples from stock at 3 different temperatures (4°C, room temperature (RmT) and -80°C freeze thaw (F/T) cycles) over 8 days period.

6.3.2.1 Protein Samples from the Stability Experiment in the AML-193 Assay

During the AML-193 stability experiment, 5µg of each tandem protein were incubated with AML-193 cells at 37°C for 3 days. Protein samples were taken and analysed by 10% SDS-PAGE gel and protein bands visualised by western blotting (Figure 6.4). This step is important to determine whether GCSF protein tandems are degraded or not during incubation**.**

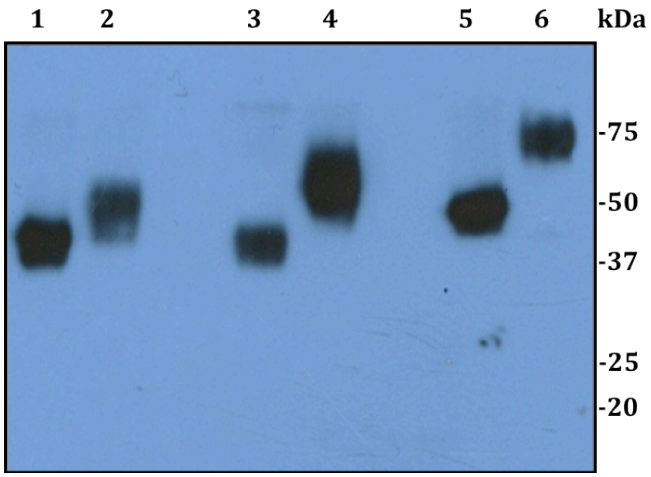


Figure 6-4: Stability of GCSF tandems after incubation with AML-193 cells

Western blot with anti GCSF antibodies show no visible degradation in all GCSF tandems (Lane 1; GCSF2QAT, Lane 2; GCSF2NAT, Lane 3; GCSF4QAT, Lane 4; GCSF4NAT, Lane 6; GCSF8QAT, Lane 7; GCSF8NAT). A total of 100ng protein was loaded per lane.

6.3.2.2 Temperature Stability of GCSF Tandem Molecules

Stock purified tandem proteins were taken from -80°C and diluted to 0.8mg/ml with filter sterile PBS and kept on ice. All manipulations were carried out under sterile conditions. Aliquots of protein at 0.8mg/ml were placed at room temperature (RmT), 4°C or -80°C in 1.5ml sterile eppendorf tubes. Samples were taken on days 0,1, 4 and 8 (samples taken on day zero, represent untreated sample controls) and immediately diluted with an equal volume of SDS-PAGE buffer (Laemmli buffer) and heated at 95°C for 5 minutes to denature. Samples were analysed by 10% SDS-PAGE gel and protein bands visualised by coomassie staining and western blotting. A total of 7.5µg protein was loaded per lane for SDS-PAGE gel and 100ng per lane for immunoblot. Samples were analysed by SDS-PAGE gels for are shown in Figures 6.5 to 6.10. The stability western blot analysis for each tandem protein is also shown adjacent to the respective SDS-PAGE gel.

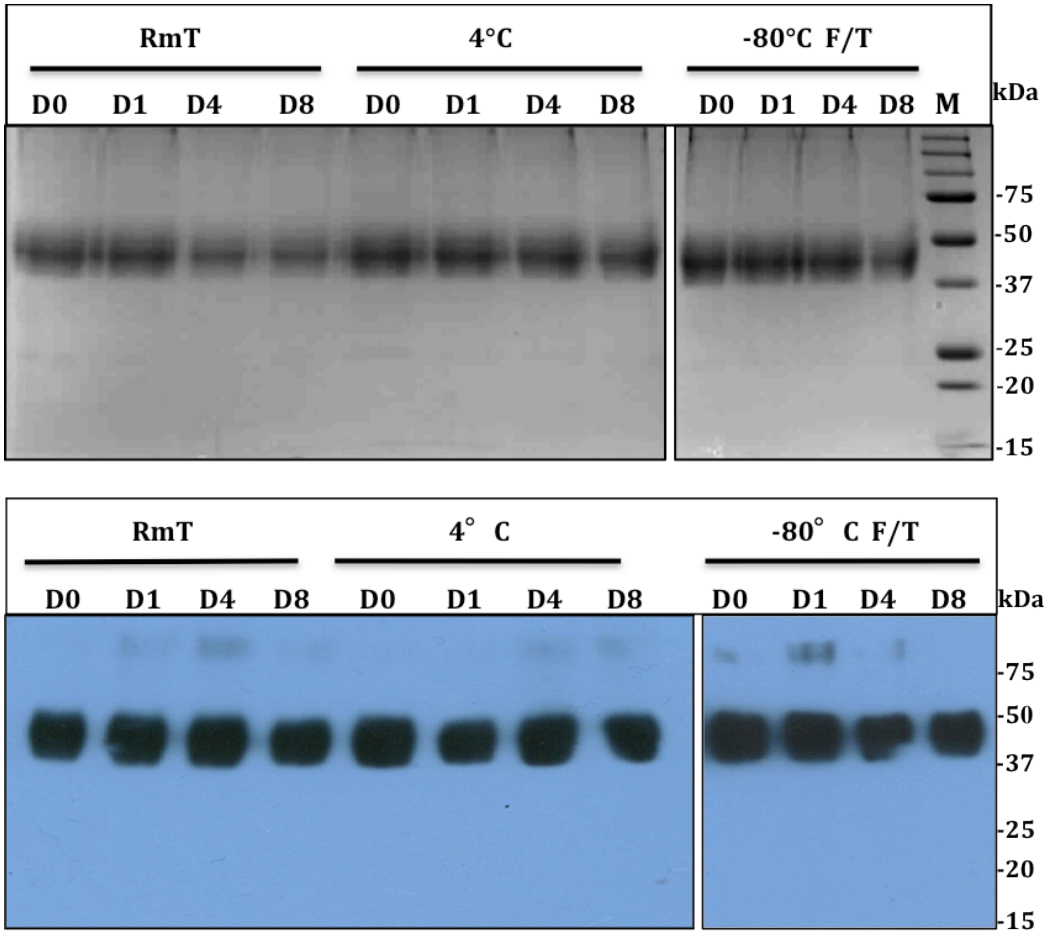


Figure 6-5: Temperature stability of GCSF2NAT

Coomassie blue and western blot with anti GCSF antibodies show GCSF2NAT under different temperatures stability conditions. In each gel the lanes contain samples taken at D0, D1, D4 and D8 stored at RmT, 4°C, and -80°C thawed/refrozen in each day. No visible degradation in protein tandems in any sample under all conditions studied (4°C, RmT and -80°C freeze thaw) over the 8 days period.

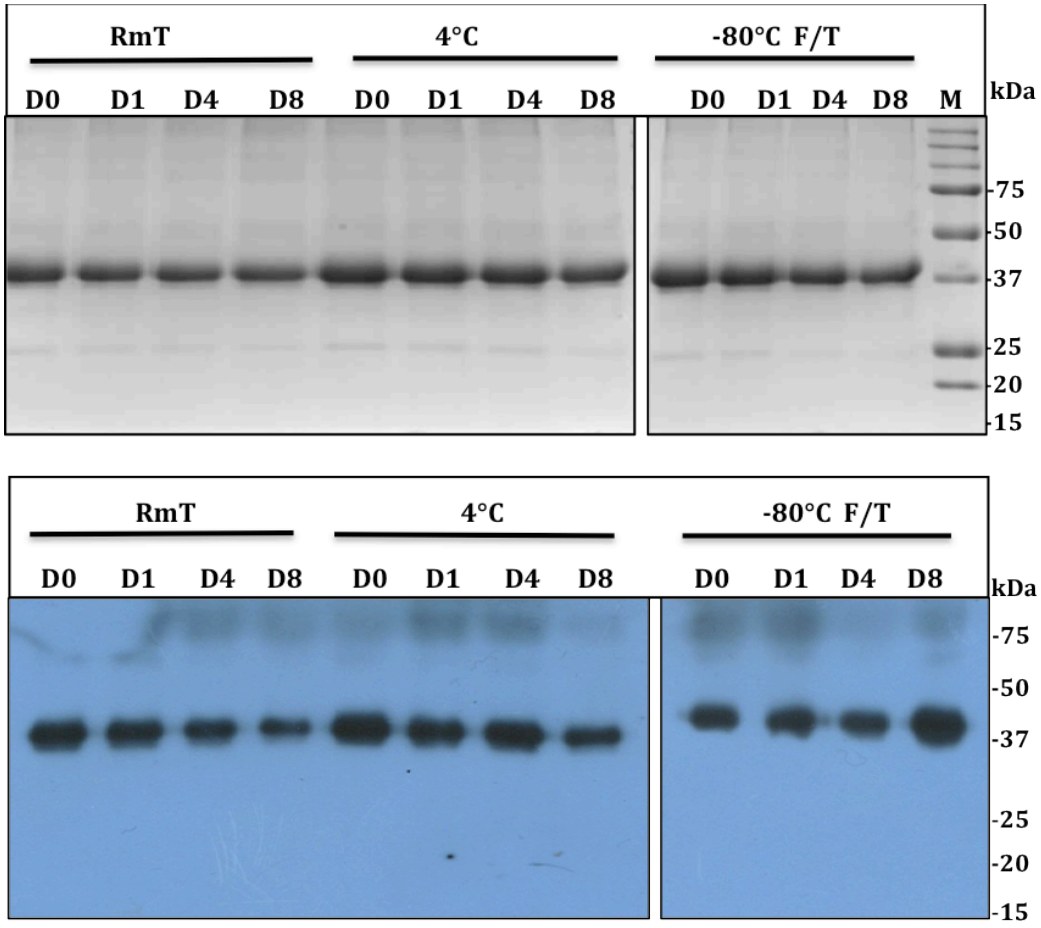


Figure 6-6: Temperature stability of GCSF2QAT

Coomassie blue and western blot with anti GCSF antibodies show GCSF2QAT under different temperatures stability conditions. In each gel the lanes contain samples taken at D0, D1, D4 and D8 stored at RmT, 4°C, and -80°C thawed/refrozen in each day. No visible degradation in protein tandems in any sample under all conditions studied (4°C, RmT and -80°C freeze thaw) over the 8 days period.

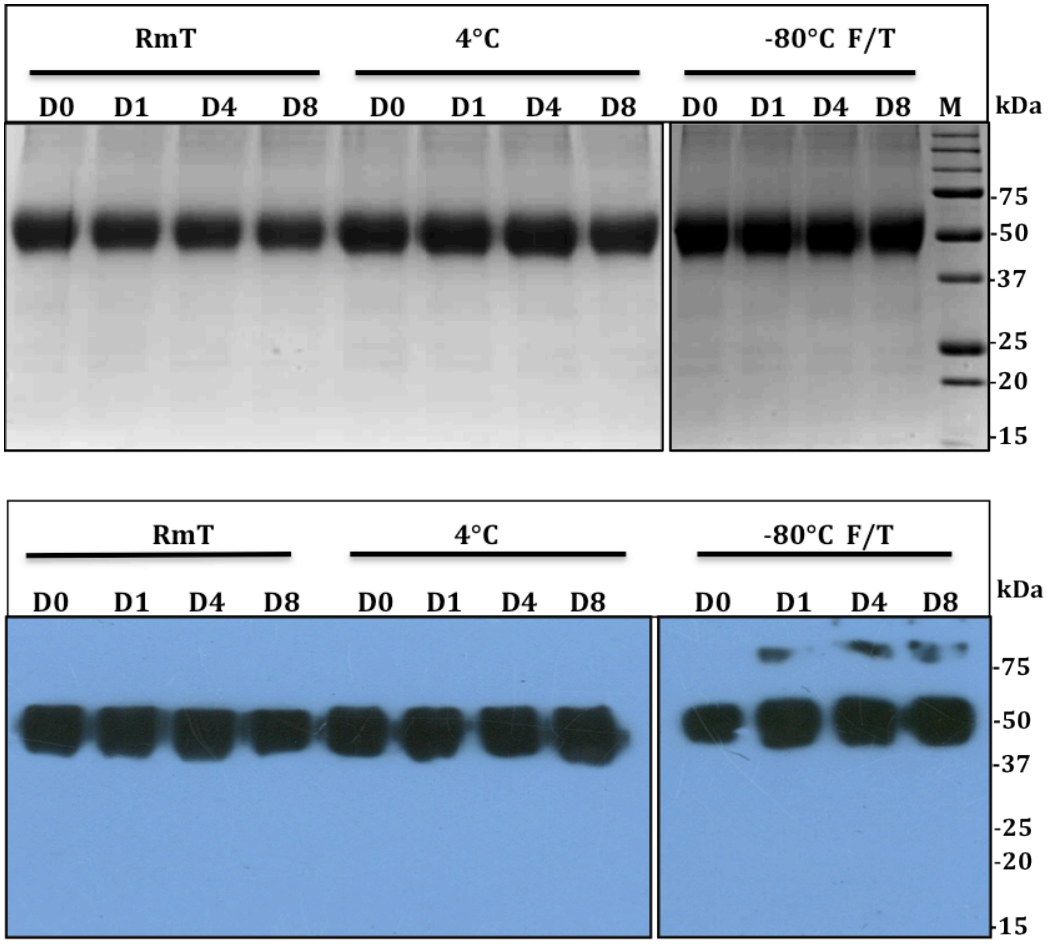


Figure 6-7: Temperature stability of GCSF4NAT

Coomassie blue and western blot with anti GCSF antibodies show GCSF4NAT under different temperatures stability conditions. In each gel the lanes contain samples taken at D0, D1, D4 and D8 stored at RmT, 4°C, and -80°C thawed/refrozen in each day. No visible degradation in protein tandems in any sample under all conditions studied (4°C, RmT and -80°C freeze thaw) over the 8 days period.

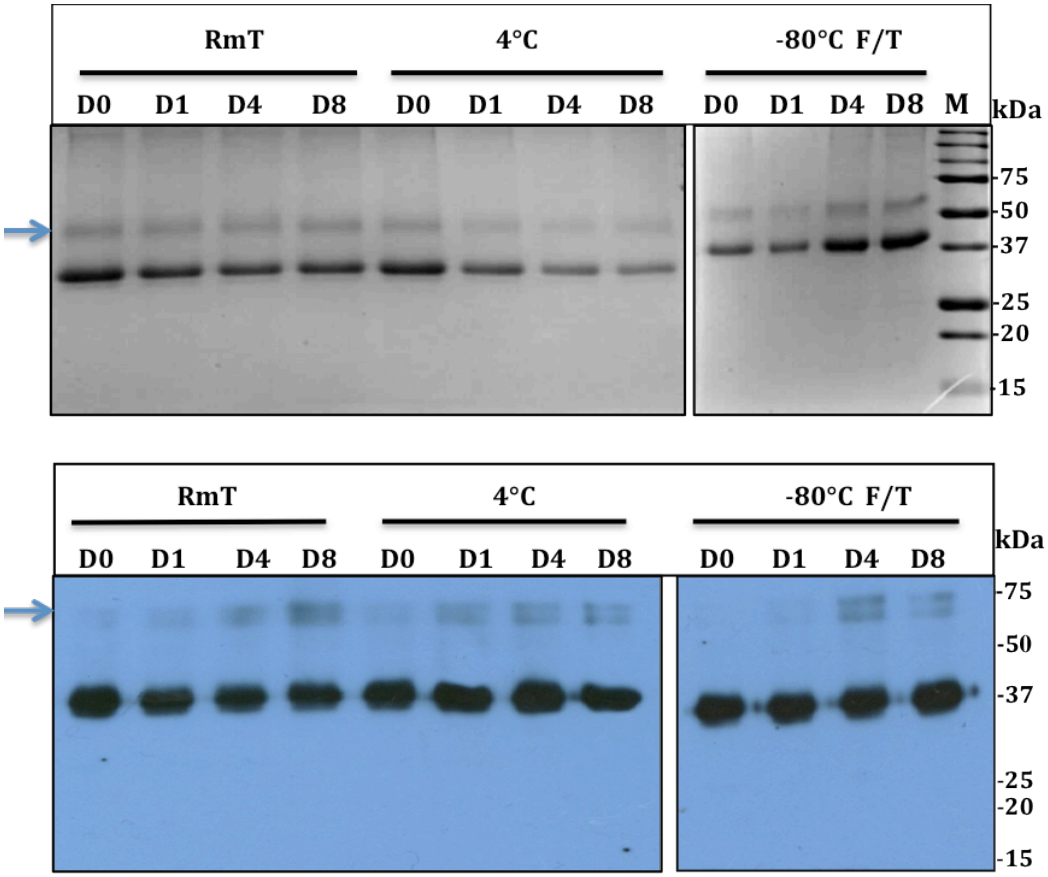


Figure 6-8: Temperature stability of GCSF4QAT

Coomassie blue and western blot with anti GCSF antibodies show GCSF4QAT under different temperatures stability conditions. In each gel the lanes contain samples taken at D0, D1, D4 and D8 stored at RmT, 4°C, and -80°C thawed/refrozen in each day. No visible degradation in protein tandems in any sample under all conditions studied (4°C, RmT and -80°C freeze thaw) over the 8 days period. Both arrows show around 20% potential dimer formation as judge by SDS-PAGE and western blot.

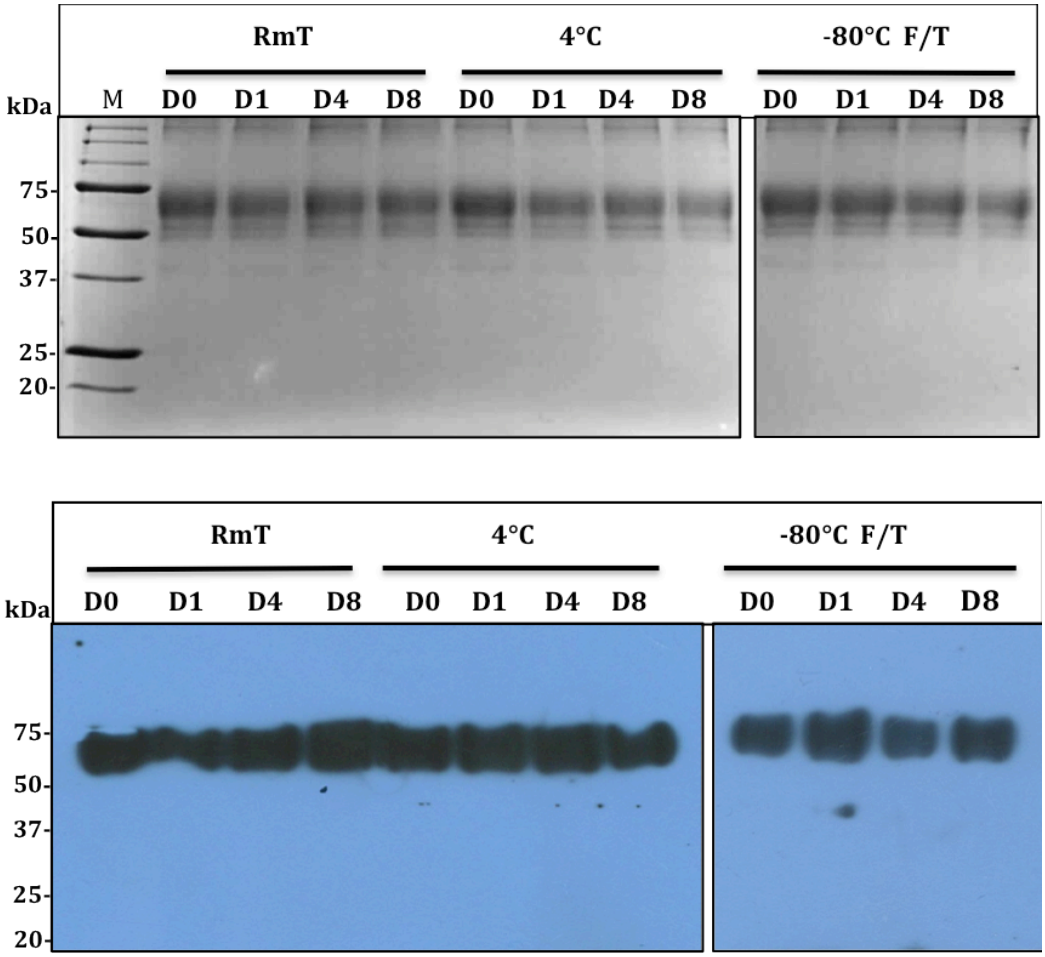


Figure 6-9: Temperature stability of GCSF8NAT

Coomassie blue and western blot with anti GCSF antibodies show GCSF8NAT under different temperatures stability conditions. In each gel the lanes contain samples taken at D0, D1, D4 and D8 stored at RmT, 4°C, and -80°C thawed/refrozen in each day. No visible degradation in protein tandems in any sample under all conditions studied (4°C, RmT and -80°C freeze thaw) over the 8 days period.

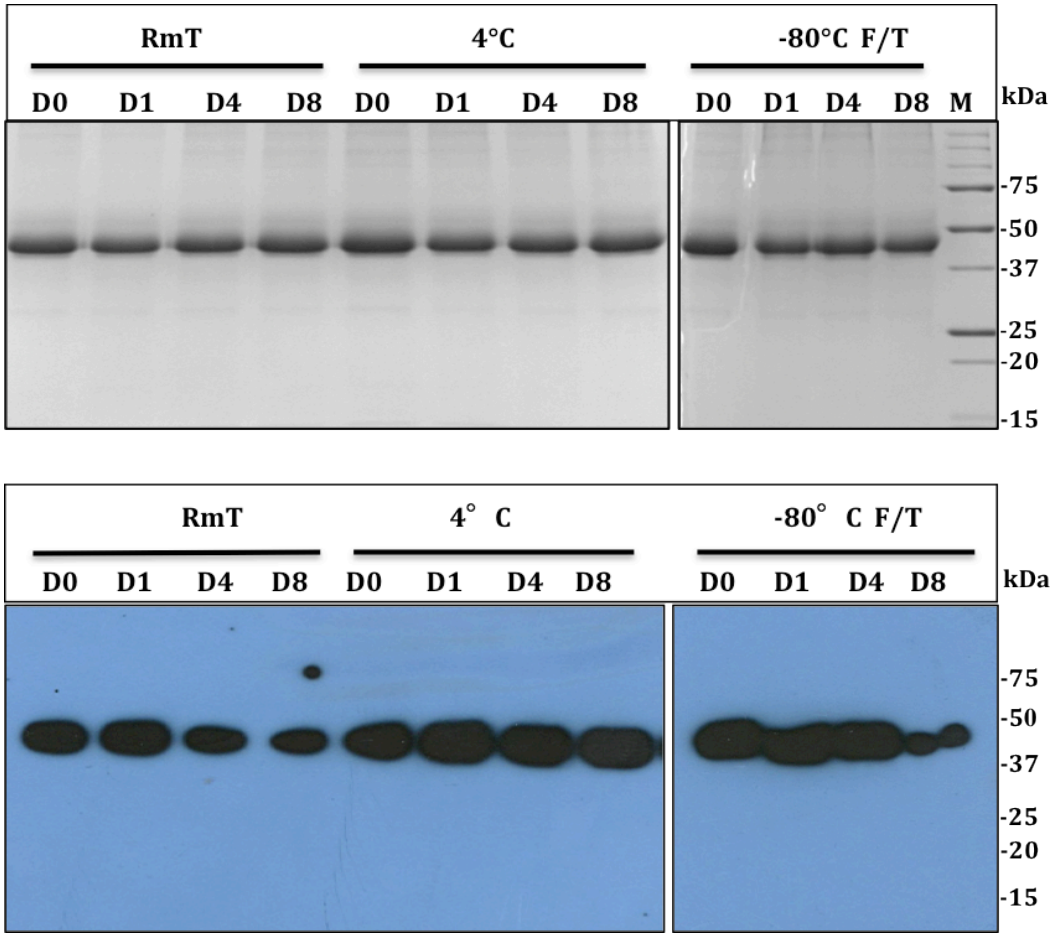


Figure 6-10: Temperature stability of GCSF8QAT

Coomassie blue and western blot with anti GCSF antibodies show GCSF8QAT under different temperatures stability conditions. In each gel the lanes contain samples taken at D0, D1, D4 and D8 stored at RmT, 4°C, and -80°C thawed/refrozen in each day. No visible degradation in protein tandems in any sample under all conditions studied (4°C, RmT and -80°C freeze thaw) over the 8 days period.

In summary, all samples of protein tandems show no visible degradation under all conditions studied (4°C, RmT and -80°C freeze thaw) over the 8 days period, which implies the high short term stability of these tandems. However, GCSF4QAT shows around 20% dimer formation upon analysing by SDS-PAGE and western blot. The probable reason for dimer formation in one of the constructs (GCSF4QAT) will be discussed in the final discussion.

6.4 Discussion

In this chapter, to determine the suitability of the purified glycosylated tandem proteins and their respective non-glycosylated controls for the PK *in vivo* studies; it was imperative to assess their biological activities. While active proteins are essential for use in the PK studies, it is equally important that the proteins be stable over the course of administration in an animal model and post-administration analyses of the protein samples. Therefore, AML-193 cell line was used to measure the bioactivity of these tandem proteins due to the ability of rhGCSF to proliferate this type of cell line *in vitro*. Furthermore, stability of these tandem proteins was also analysed using different temperatures (RmT, 4°C and -80°C F/T) and also after incubation with AML-193 cells at 37°C for 3 days.

The biological activity results show that AML-193 proliferation assay is a valid quantitative *in vitro* model for measuring the activity of glycosylated GCSF tandems and their respective non-glycosylated controls. Also, the results indicate that these tandem proteins exhibit agonistic action by stimulating the proliferation of AML-193 cells. Initial bioassay experiments showed that the best level of absorbance achievable, upon the induction of the cells to MTS reagent at 37°C for 2 hrs is around 0.25 at 490nm. Extending the time of induction beyond 2 hrs resulted in no significant increase in the colour change. It is apparent from the results that the tandem proteins exhibited significant increased biological activity compared to the rhGCSF (EC_{50} for tandems were about 3-fold lower than that for rhGCSF).

The observed increase in biological activity could be attributed to two factors; the design of the tandem molecule (two GCSF ligands linked with a flexible linker) and also the linker length. Normally, two GCSF ligands bind two GCSF receptors forming a homodimer (Larsen et al., 1990). The binding of the GCSF ligand to the receptor occurs at a 2:2 ligand:receptor subunit stoichiometry, forming a cross-over configuration between the receptor subunits. This structural configuration enables the trans-phosphorylation of the receptors and initiates downstream signal transduction. However, a lag in the availability or binding of a second GCSF ligand to the receptor would

result in delayed activation of the signaling pathway. This potential delay may be absent in the tandem molecule, as there are two covalently linked GCSF molecules in tandem that are readily available for binding to the receptors.

The observed potency of the tandem molecules is consistent with other published research studies where two ligands linked in tandem resulted in a novel molecule with higher bioactivity than the wild type. The binding of the EPO ligand to the receptor occurs at a 2:2 ligand:receptor subunit stoichiometry similar to that in GCSF. Sytkowski et al., (1999) generated a fusion of two hematopoietic growth factor erythropoietin (Epo) linked in tandem using a 17-amino acid flexible peptide linker. The specific biological activity of the Epo tandem at 1,007 IU/μg was found to be significantly greater than that of the native Epo protein at 352 IU/μg. Furthermore, the comparative studies of the pharmacokinetics of the Epo tandem and the native Epo proteins in mice showed the tandem molecule to be more potent, causing a significant increase in red blood cell production within 7 days whereas the conventional recombinant Epo control had no obvious stimulating effect (Sytkowski et al., 1999).

Furthermore, compared to the fusion constructs in the aforementioned research studies, in our GCSF tandem the linker length is a significant factor for consideration in the design of a fusion expression construct, as studies have shown that inappropriately short linkers could result in loss of biological activities or potency in tandem fusion molecules. Research by Qui et al., (1998), showed the bioactivity of a tandem Epo fusion construct with a shorter linker sequence (glycine (G) 3-7 peptide was compared with the Epo tandem with a longer linker sequence 17-amino acid (A-[G-G-G-G-S]$_3$-T) .The results revealed that, while both tandem molecules were biologically active the Epo tandem with the shorter linker showed no increase in bioactivity over the native Epo whereas the Epo tandem with the longer linker was significantly more biologically active (Qiu et al., 1998). This suggests that, tandem fusion ligands with linker length long enough would facilitate simultaneous binding of the ligands to the receptors, which in

effect induces the homodimerization and activation of the receptors, which in turn initiate the downstream signal transduction. Whereas, in a tandem fusion molecule with short linker sequence, the two ligands linker could potentially restrict the binding of one of the two ligands to the receptor, which inadvertently would exhibit similar activity to a 'free' ligand bound receptor. In the design of our three glycosylated GCSF tandems and their controls, linker lengths used were long enough and, all the tandem molecules had similar biologically activity. This suggests that the linker lengths between the tandem GCSF ligands in the six expression constructs investigated (49 amino acids for GCSF2NAT/2QAT, GCSF4NAT/4QAT and 94-amino acids for GCSF8NAT/8QAT) were long enough to allow both GCSF ligands in the tandem molecule to bind both receptors. This equally contributed to the higher biological activities observed in these constructs compared to the rhGCSF.

The data also indicate that the biological activity of glycosylated tandems is very similar to their non-glycosylated controls. For instance, the tandem molecule with 2 glycosylation sites showed similar results to the 4 and 8 glycosylation sites. It is apparent from the result that increasing the N-linked glycosylation sites has no noticeable negative impact on the *in vitro* biological activity of GCSF.

The GCSF tandems also showed a high degree of stability *in vitro* during the temperature stability test. The results (SDS-PAGE and western blot) showed no visible degradation in the analysed tandem proteins samples, under all conditions studied (4°C, RmT and -80°C freeze thaw), over the 8 days period and also after incubation with AML-193 cells at 37°C for 3 days. This suggests that the tandem molecules achieve the right conformation during protein synthesis in CHO cells.

The glycosylated tandem molecules are devoid of dimerisation, whereas, non-glycosylated tandems showed some dimer formation that was more obvious in GCSF4QAT (a control for GCSF4NAT tandem molecule). GCSF4QAT was estimated to contain approximately 20% dimer when

analysed by Coomassie blue and western blotting (Refer to Figure 6.8). However, all the controls showed similar biological activity in the proliferation bioassay. This indicated that the presence of a dimer does not affect the overall bioactivity of the GCSF4QAT control molecule.

Routinely, proteins are stored as frozen solutions at -80°C and this may preserve the stability of proteins for longer periods of time. However, it was observed that long time freezing at -80°C may be the cause of dimer formation as found in GCSF4QAT, which was stored at -80°C for about a year. Prior to freezing (i.e. during expression and purification) no dimerisation was observed in the GCSF4QAT just like the other test molecules.

Two factors may have contributed to the observed dimerisation of GCSF4QAT after longer storage at -80°C. First, the dimer formation could be attributed to the presence of free cysteine residue (Cys17) in the GCSF structure coupled with the absence of glycosylation in the molecule. Native GCSF contains five cysteine residues, two internal disulfide bonds at positions Cys36– Cys42 and Cys64– Cys74 leaving one free cysteine residue at position Cys17 with a free sulfhydryl group. It is possible that during longer storage in -80°C, the free cysteine (Cys17) of one GCSF tandem may form inter-molecular disulfide bonds with another GCSF tandem, which results in the formation of GCSF tandem dimer. This view is supported by a recent study where the Cys17 in the GCSF molecule was substituted with Ala17 (alanine instead of cysteine in wild-type GCSF) using direct mutagenesis and recombinant DNA technology. The resulting GCSF mutant exhibited enhanced *in vitro* stability and higher activity *in vivo* than the wild-type GCSF, perhaps through the elimination of dimerisation caused by the formation of intermolecular disulfide bonds (Jiang et al., 2011). While other GCSF tandems also contain GCSF units with free cysteine (Cys17), the glycosylation of the polypeptide chains may have induced unique conformational changes that prevent intermolecular disulphide bond formation with another tandem or with GCSF in the same tandem.

Another possible reason of dimerisation is the incorporation of glutamine (Q) into the linker of the non-glycosylated control GCSF tandem (i.e. GCSF4QAT) compared to asparagine (N) in the glycosylated GCSF tandems. It has been shown that incorporation of a poly-glutamine into a protein resulted in the formation of dimers. The glutamine repeats in proteins form hydrogen bonds as a polar zipper (two paired antiparallel β-sheet strands held together by hydrogen bonds between their amide groups) (Perutz, 1995, Hoffner and Djian, 2005). However, since the control tandem molecule has a single point mutation that allows for glutamine incorporation, the involvement of the incorporated glutamine in the dimer formation is rather ambiguous. Perhaps, analysing the control tandem protein on a reduced gel (i.e. containing DTT) would elucidate the type of bond that is present in the dimers, whether the dimers were covalently linked or not through disulphide bridges.

Moreover, the control tandem has no N-glycosylation sites at the linker and therefore would show less steric hindrance towards the formation of a dimer when compared to other glycosylated tandems. A number of published research studies have highlighted that glycosylated molecules have molecular characteristics that significantly affect the function of the molecule. For instance, it was recently reported that N-linked glycosylation modulates dimerisation of human protein disulfide isomerase (PDIA2) (important enzyme for the correct maturation and folding of proteins that reside or transit into the endoplasmic reticulum (ER)). This protein was shown to be glycosylated at the asparagine residues of three N-linked glycosylation sites (N127, N284 and N516). The finding was that mutation at N284 led to an increase in the dimer formation of PDIA2 (Walker et al., 2013). This suggests that, the site directed mutagenesis abrogated the glycosylation characteristics in the mutant molecule.

Glycosylation in the tandem molecules possibly creates steric hindrance. The presence of charge repulsion among molecules due to the negative charge and a huge relative volume occupied by sialic acid as described previously in the introduction chapter. However, the biological activity data

of non-glycosylated GCSF tandems were comparable to that observed for glycosylated GCSF tandems, indicating that dimerisation did not affect the *in vitro* biological activity.

In conclusion, The GCSF tandems demonstrated better biological activity compared to the rhGCSF and showed high stability at 4C, RmT, 37°C and multiple F/T cycles at -80°C.

7. Results 4: Pharmacokinetic & Pharmacodynamic Analysis of rhGCSF and GCSF Tandems in a Rat Model

7.1 Summary

Glycoproteins represent a major value for the next marketed and clinical generation of therapeutic proteins. A full understanding of the function and nature of the glycosylation and its influence on pharmacology properties is crucial in finding and developing efficient and safe glycoprotein biopharmaceuticals. In the previous chapter, glycosylated GCSF tandems together with non-glycosylated controls have shown better bioactivity compared to rhGCSF and high stability. *In vivo* pharmacokinetic and pharmacodynamics properties of these GCSF tandem proteins were measured in normal Sprague Dawley strain rats with full ethical approval. GCSF2NAT, GCSF4NAT and GCSF8NAT, containing 2, 4 & 8 glycosylation sites respectively, displayed a reduced rate of clearance compared to both rhGCSF and non-glycosylated GCSF tandem controls. Although the half-life of GCSF8NAT exhibited no further enhancement beyond that of GCSF4NAT, pharmacodynamics (PD) displayed a significant increase in the number of circulating neutrophils at 48 hrs in rats compared to 12 hrs reported for GCSF4NAT and GCSF2NAT, leading us to hypothesize that GCSF8NAT is a more efficient stimulator of neutrophils.

7.2 Introduction

Pharmacokinetics (PK) is described as the study that evaluates the time course of drug absorption, distribution, metabolism and excretion in living organisms such as rats. It is sometimes described as what the body does to a drug. In contrast, pharmacodynamics (PD) is described as what a drug does to the body, for example, in our case what GCSF does to neutrophil population in the organism. Protein drugs' PK/PD are typically affected by fast elimination in intravenous administration from the human body, via proteolytic, renal, hepatic, and receptor mediated clearance mechanisms (Tang et al., 2004, Mahmood and Green, 2005). Thus, PK/PD data that are produced from relevant species like mouse or rat support the prediction of PK/PD in humans and may help to generate safe and effective therapeutic applications for current therapies.

In humans, GCSF has a short serum *in vivo* half-life of 1.79 hrs (Tanaka et al., 1991). Consequently, individual patients with neutropenia require daily injections to increase neutrophils in the blood circulation, which leads to poor patient compliance. However, the current product on the market that is a long-acting form of GCSF called PEG-rhGCSF (with terminal half-life of 7.05 hrs) which is administered once per chemotherapy cycle to enhance the number of neutrophils (Tanaka et al., 1991). It has been reported that modifications of GCSF by covalently attaching a chemical, polyethylene glycol (PEG), can change the PK and PD properties of GCSF to significantly increase the time the modified native GCSF remains effective in the blood circulation (Delgado et al., 1992). Results from receptor binding, *in vitro* proliferation and neutrophil function studies show that PEG-rhGCSF and native GCSF have a similar mechanism of action in the circulation (Lord et al., 2001). In rat models, rhGCSF is mainly eliminated by the renal route; therefore, the presence of PEG moiety increases the molecular weight of GCSF and reduces its renal clearance by glomerular filtration (Jain and Jain, 2008).

Efforts have been made to improve the pharmacokinetic behaviour of therapeutic proteins by the addition of natural carbohydrates. In the

previous chapter, the addition of natural carbohydrates (via N-linked glycosylations) to the GCSF tandem has improved the bioactivity and stability. It has also been reported that hyperglycosylation can regulate activity and the pharmacokinetic profile of GCSF (using the mutant hGCSF (Phe140Asn)) leading to prolonged GCSF in the circulation and a more effective molecule than native GCSF in stimulating differentiation and proliferation of hematopoietic cells (Chung et al., 2011). Hyperglycosylation closely resembles PEGylation but has an added advantage over PEGyaltion. For instance, hyperglycosylation involves the biodegradable nature of carbohydrates, while in contrast the whole PEG molecule is often processed for excretion in the human body without undergoing an initial biodegradation which could be toxic to the body (Patel et al., 2014). This possible toxicity of PEGylation was supported by the detection of PEG only in bile (Caliceti and Veronese, 2003). Beside these advantages, this project used hyperglycosylation (N-linked glycosylated linker) to increase the molecular weight of GCSF and may also protect it from proteolysis due to the presence of terminal sialic acid that shields the underlying galactose peptides from protease recognition and cleavage within the circulation (Sola and Griebenow, 2009). This is supported by a study carried out by Raju and Scallon (2007) who demonstrated that removal of terminal sugars from Fc antibody fragments resulted in increased sensitivity to papain.

In this chapter, the construction of novel recombinant GCSF tandems results in molecules with reduced clearance while retaining bioactivity. It also alleviates potential problems with direct glycosylation of the ligand, which may reduce bioactivity and potentially introduce immunogenic sites.

7.3 Aim and Hypothesis

In vivo strength of therapeutic proteins is often strongly associated with residence time of blood circulating. This is a function of the drug's PK behaviour including, serum half-life, clearance rate and the minimum and maximum concentrations. In the previous chapter, glycosylated GCSF tandems and their respective controls have shown better bioactivity in comparison to rhGCSF and high stability. Thus, the aim of the current chapter is to determine the pharmacokinetic properties of rhGCSF & GCSF tandems in normal adult Sprague Dawley strain rats following intravenous injections and the effects of these GCSF tandems on white blood cell (WBCs) and neutrophil populations. We hypothesised that our glycosylated GCSF tandems will have longer circulating half-lives and more potent (i.e. mobilizing more neutrophils) compared to the rhGCSF.

7.4 Results

7.4.1 Preliminary Test for the Effect of Rat's Serum on Elisa Assay

Prior to measuring GCSF tandems in rat serum, it was necessary to measure the sensitivity of the Elisa assay by testing the effect of rat serum on rhGCSF and GCSF2NAT standard curves. This test was used to determine the specificity of the Elisa (i.e. testing if Elisa detects rat GCSF from serum) and to assess any interference of rat serum with the sensitivity of the Elisa.

Previous studies by Asterion on human growth hormone tandem molecule showed that 2% (v/v) of rat serum had no interference with the sensitivity of the growth hormone Elisa standard curve. Hence, this concentration of rat serum (2% v/v) was used in the initial trial Elisa experiments to determine the effect of rat serum on rhGCSF and GCSF2NAT Elisa's. The results of the experimental studies showed that 2% (v/v) of rat serum has no effect on the Elisa assay using rhGCSF and GCSF2NAT as standard curve controls. Both rhGCSF and GCSF2NAT standard curves generated in 2% (v/v) rat serum showed no OD change from LKC buffer only indicating that rat serum alone has no interference in the Elisa at this concentration (Figure 7.1).

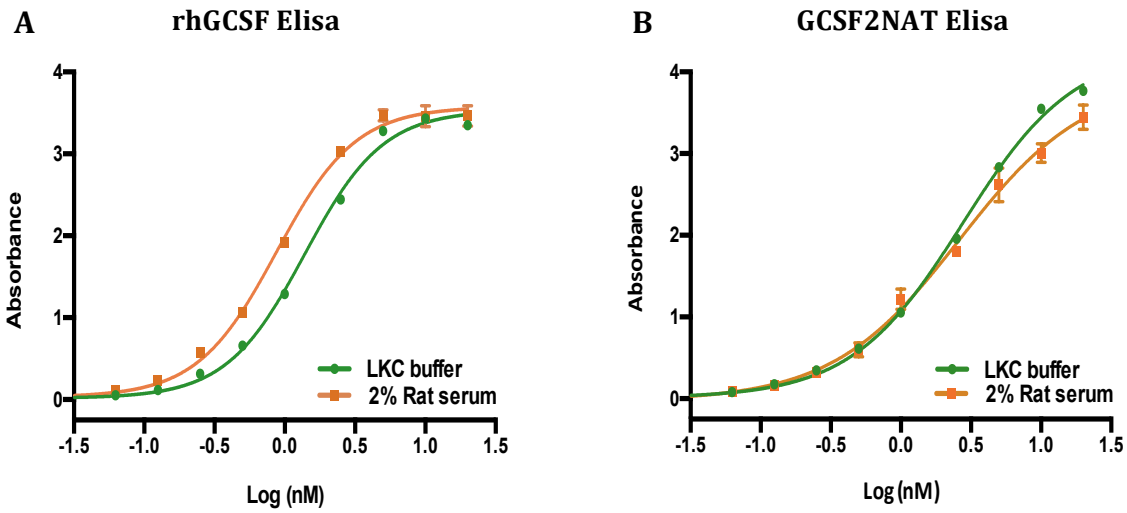


Figure 7-1: Effect of rat serum on the sensitivity of Elisa assay

(A) A comparison between rhGCSF diluted in 2% rat serum and rhGCSF diluted in LKC buffer only. (**B**) A comparison between GCSF2NAT diluted in 2% rat serum and GCSF2NAT diluted in LKC buffer only. Results are given as SEM for triplicate wells.

Sensitivity of the assay has not been impaired by the presence of rat serum: Therefore, if rat serum can be kept to 2% (v/v) final concentration in the assay then it will be acceptable as long as all samples receive the same concentration of serum.

7.4.2 Preliminary Pharmacokinetic Analysis in Normal Sprague Dawley Rats

7.4.2.1 Elisa Results

The pharmacokinetic performance of glycosylated GSCF tandem (GCSF2NAT, 4NAT, 8NAT & 8QAT) was evaluated in Sprague Dawley rats to assess the longevity of exposure in comparison to rhGCSF. 250µg/kg (75µg per rat) of rhGCSF, GCSF2NAT, GCSF4NAT, GCSF8NAT, GCSF8QAT and vehicle (PBS only) were given by intravenous injection (IV). 0.3 to 0.4ml blood sample were taken at specified time points (-24 pre-injection, 0.5, 1, 2, 4, 8, 12, 24 & 48, 72 hrs) and centrifuged for serum preparation as previously described in section 3.9.

GCSF Elisa was used to measure the concentration of proteins in each serum GCSF tandems samples and rhGCSF at -24, 0.5, 1, 2, 4, 8, 12, 24, 48, 72 hrs post-dose. The data analyses of these samples suggested that the pharmacokinetic profiles for GCSF tandems (GCSF2NAT, 4NAT, 8NAT & 8QAT) following intravenous dosing at 250µg/kg displayed a reduced rate of clearance compared to the published rhGCSF (Figures 7.2 – 7.7). However, the rhGCSF molecule in this study appeared to be mistakenly injected subcutaneously as against intravenous injection of other test GCSF tandems. This however was not confirmed by the contractor (more details provided in the sections below). For this reason, the rhGCSF samples were excluded from further data analyses, so as to avoid inconsistencies in data analyses.

7.4.2.1.1 Pharmacokinetics Analysis of rhGCSF in Normal Rats

Analysis of the pattern of clearance between the rats shows an apparent error during injection. Rat 2 and 5 both have high values at 0.5 hr, however, Rat 5 again show an unexpectedly high value at 2hrs. Rat 4 and 6 both have unexpectedly low values at all time points. Rat 1 and 3 both show peaks at 2hrs. This suggests that only Rat 2 was injected intravenously during the administration while other rats were injected subcutaneously by the contractor. Since the test GCSF tandems were injected intravenously these data were excluded in this research to avoid inconsistency in data analysis. A few of the serum samples were missing as the contractor who did the in vivo studies omitted these samples (shaded boxes in grey) (Figure 7.2).

A

	Rat					
Time (hr)	**1**	**2**	**3**	**4**	**5**	**6**
-24	0	3.1		0	0	2.6
0.5	1.9	2.3	21.7	2.5	0.6	23.8
1	4.6	3.6	4	25.4	8.1	7.1
2	23.8		4.1	45.8	5	24.1
4	0.7	5.2	1.8	11.3	5.1	8.3
8	9.2	2.6	5.9	4.5	3.1	
12	1.4			1.4	1.9	1.2
24	0	0	0	0		0
48	0	0	0	0	0	5
72	0	0	0	0	0	0

B

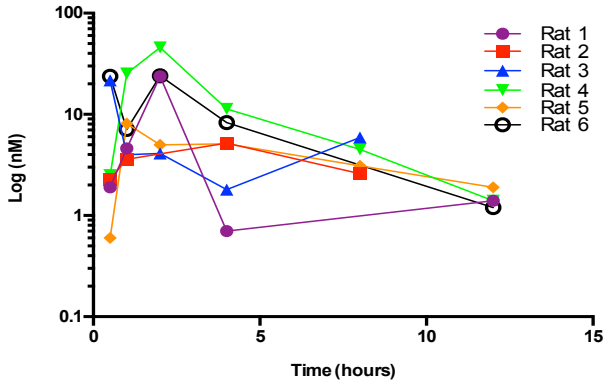


Figure 7-2: Elisa analysis of rhGCSF pharmacokinetics in normal rat

(A) Six rats were injected with 250µg/kg of native purified recombinant GCSF protein intravenously. Serum samples were taken 24 hrs before the injection and 0.5, 1, 2, 4, 8, 12, 24, 48, 72 hrs post injection. The samples were analysed by ELISA to determine the concentration (nM) of rhGCSF in the serum and clearance rate. The results for each rat at each individual sampling time points are tabulated. Missing samples shaded in grey. **(B)** The concentrations of rhGCSF were plotted against the time of sampling for each individual rat.

7.4.2.1.2 Pharmacokinetic Analysis of GCSF2NAT in Normal Rats

All rats in this group exhibited similar patterns of protein clearance with similar an early peak concentration of 0.5 hrs, which suggests that the rats were successfully injected intravenously. Rat 6 serum samples and a few of other serum samples were missing (shaded boxes in grey), as these were not received from the contractor who did the IV injection and sampling. Two samples were labelled with similar information (shaded boxes in red) and were therefore omitted from the study (Figure 7.3).

A

	Rat					
Time (hr)	**1**	**2**	**3**	**4**	**5**	**6**
-24	2	2.8	0	0.7	0	
0.5	90.7	42.1	20.5	49	40	
1	45.5	31.3	15.1		11.9	
4		12.4	10	5.8	9.2	
8	12.7	4.6	7.1	6.9	8.2	
12	8.5	6.8	6.2	5.7	2.4	
24	3.9	1.6	2	1.6	0	
48	8.4	0	0	1.9	0	
72	0	0	0	0	0	

B

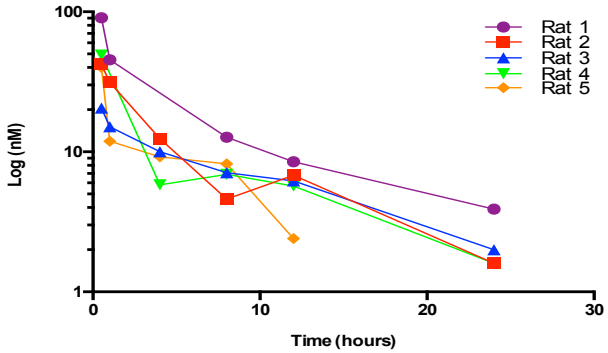


Figure 7-3: Elisa analysis of GCSF2NAT pharmacokinetics in normal rat

(A) Six rats were injected with 250µg/kg of purified GCSF2NAT intravenously. Serum samples were taken 24 hrs before the injection and 0.5, 1, 2, 4, 8, 12, 24, 48, 72 hrs post injection. The samples were analysed by an Elisa to determine the concentration (nM) of GCSF2NAT in the serum and clearance. The results for each rat at each individual sampling time point were tabulated. Missing samples shaded in grey. Omitted samples shaded in red. **(B)** The concentration of GCSF2NAT was plotted against the time of sampling, for each individual rat.

7.4.2.1.3 Pharmacokinetic Analysis of GCSF4NAT in Normal Rats

Analysis of the pattern of clearance between the rats shows that rats were successfully injected intravenously. Rat 6 serum samples and a few of the other samples were missing (shaded boxes in grey), as these were not received from the contractor who did the IV injection and sampling (Figure 7.4).

A

	Rat					
Time (hr)	**1**	**2**	**3**	**4**	**5**	**6**
-24		0	0	3.3	25.8	
0.5	81.6	63.1		4.7	61.6	
1	53	19.3	8.9	14.6	75.9	
2	36.7	14.6		21.2	40.5	
4	20.6	14.5	7.8	11.6	35.6	
8	12.3		11	14.5	45.7	
12		4.5	8.2	8.2	29.8	
24	5.8	1.7	1.9	12	13.3	
48	0	1.7	0	1.9	7.8	
72	1.4	1.8	0	0.7	22.8	

B

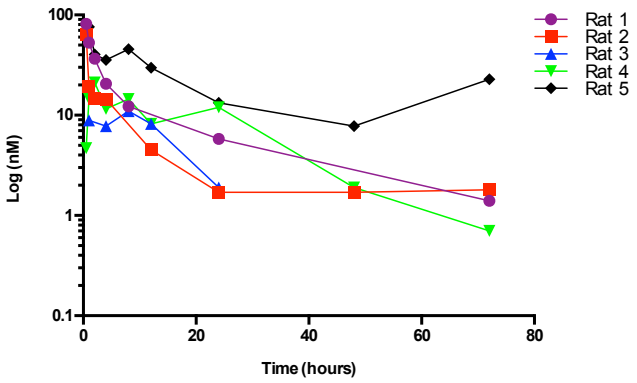


Figure 7-4: Elisa analysis of GCSF4NAT pharmacokinetics in normal rat

(A) Six rats were injected with 250µg/kg of purified GCSF4NAT intravenously. Serum samples were taken 24 hrs before the injection and 0.5, 1, 2, 4, 8, 12, 24, 48, 72 hrs post injection. The samples were analysed by Elisa to determine the concentration (nM) of GCSF4NAT in the serum and clearance. Missing samples shaded in grey **(B).** The results for each rat at each individual sampling time point were tabulated. The concentration of GCSF4NAT was plotted against the time of sampling, for each individual rat.

7.4.2.1.4 Pharmacokinetics Analysis of GCSF8NAT in Normal Rats

Analysis of the pattern of clearance between the rats shows that rats were successfully injected intravenously. Rat 6 serum samples and a few of other serum samples were missing (shaded boxes in grey), as these were not received from the contractor who did the IV injection and sampling. In addition, Rat 1 showed a very high protein concentration in all time points (shaded boxes in red) compared to other rats. Rat 1 was repeated but still give a very high value. Therefore, it was decided to omit from the current study after discussions with Phoenix Certara (half-life = 270.7 hrs as analysed by Winnonlin PK program provided by Phoenix Certara) (Figure 7.5).

A

	Rat						Rat repeated
Time (hr)	**1**	**2**	**3**	**4**	**5**	**6**	**1**
-24	22.6	0	2.1	5.7	0		8.4
0.5		86	17.7	51.3	31.4		
1	170.1	19.9	11.1	37.6	18.1		144.8
2		17.5	8.5	37.3	35.5		
4	16.4	14.8	7.7	27.6	18.6		9.9
8		13.1	11.3	19.2	12.1		
12	17.2	4.5	5.2	13	5.1		15.8
24	9.1	4.3	1.3	1.1	3.1		20.7
48		0.1	0	0	0		
72	21.8	1.3	0	0	0		16.8

B

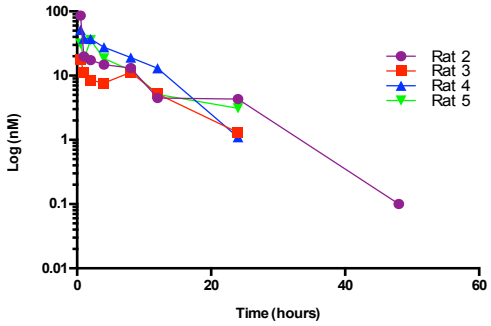


Figure 7-5: Elisa analysis of GCSF8NAT pharmacokinetics in normal rat

(A) Six rats were injected with 250µg/kg of purified GCSF8NAT intravenously. Serum samples were taken 24 hrs before the injection and 0.5, 1, 2, 4, 8, 12, 24, 48, 72 hrs post injection. The samples were analysed by Elisa to determine the concentration (nM) of GCSF8NAT in the serum and clearance. The results for each rat at each individual sampling time point were tabulated. Missing samples shaded in grey. Omitted samples shaded in red. **(B)** The concentration of GCSF8NAT was plotted against the time of sampling, for each individual rat.

7.4.2.1.5 Pharmacokinetic Analysis of GCSF8QAT in Normal Rats

Analysis of the pattern of clearance between the rats shows that rats were successfully injected intravenously. A few of the serum samples were missing (shaded boxes in grey), as these were not received from the contractor who did the IV injection and sampling. Additionally, Rat 5 and 6 showed a very low protein concentration in all time points compared to other rats. Both rats were repeated but they still give a very low protein concentration (shaded boxes in red), therefore, rat 5 & 6 were omitted from this study (Figure 7.6).

A

	Rat						Rat repeated	
Time (hr)	**1**	**2**	**3**	**4**	**5**	**6**	**5**	**6**
-24		0	0	0	0	-0.2	0	-0.2
0.5	66		74	9.4	2.3	0.6	4.1	4.7
1	48.1		46.2		0.8	0.1	0.8	0.1
2	16.6	9.6	11.9	10.5	2.7	1.3	0.2	4.1
4	7.7	9.5		11.1	7.7	-0.5	0.65	-0.5
8	5.6	7.4	10.4	8.3	3.3	-1.3	0.794	-1.3
12	0	2.7	3.9	2.5	2.7		0	
24	0.4	1.6	1.1	1	0.5	-1.9	0.5	-1.9
48	0	0	0	0	0	-2	0	-2
72	0	0	0	0.8	0	-1.2	0	-1.2

B

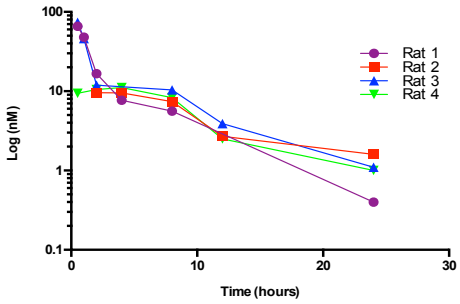


Figure 7-6: Elisa analysis of GCSF8QAT pharmacokinetics in normal rat.

(A) Six rats were injected with 250µg/kg of purified GCSF8QAT intravenously. Serum samples were taken 24 hrs before the injection and 0.5, 1, 2, 4, 8, 12, 24, 48, 72 hrs post injection. The samples were analysed by Elisa to determine the concentration (nM) of GCSF8QAT in the serum and clearance. The results for each rat at each individual sampling time point were tabulated. Missing samples shaded in grey. Omitted samples shaded in red. **(B)** The concentration of GCSF8QAT was plotted against the time of sampling, for each individual rat.

7.4.2.1.6 Pharmacokinetic Analysis of GCSF Tandems in Normal Rats

A

	Construct							
Time (hr)	**GCSF2NAT**	**SEM**	**GCSF4NAT**	**SEM**	**GCSF8NAT**	**SEM**	**GCSF8QAT**	**SEM**
0.5	48.5	12.9	52.8	19.2	46.6	17.1	49.8	24.9
1	26	9.0	34.3	14.4	21.7	6.5	47.2	1.3
2	---	---	28.3	7.1	24.7	8.1	12.2	1.8
4	9.4	1.6	18	5.4	17.2	4.8	9.4	1.0
8	7.9	1.5	20.9	9.6	13.9	2.1	7.9	1.2
12	5.9	1.1	12.7	6.7	7	2.3	2.3	0.9
24	1.8	0.7	6.9	2.7	2.5	0.9	1	0.3

B

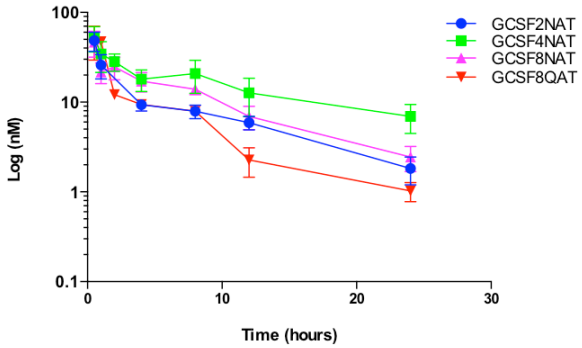


Figure 7-7: Elisa analysis of GCSF tandems in normal rat models

(A) 250µg/kg of GCSF2NAT, GCSF4NAT, GCSF8NAT, and GCSF8QAT proteins were given intravenously to Sprague Dawley rats. Serum samples were taken at specified time points (-24, 0.5, 1, 2, 4, 8, 12, 24 & 48, 72 hrs). The results of the mean concentrations for each tandem at each individual sampling time point in normal rats were tabulated. **(B)** The average protein concentrations of each GCSF tandem were plotted against the time of sampling, for each individual rat. Data are given as standard error of the mean (SEM) for at least four rats per group. Significance between Elisa results of GCSF tandems was performed with GraphPad Prism using Nonlin fit test.

From the Elisa data of GCSF tandems, it can be seen that for IV the maximum serum concentration was reached at the earliest sampling time-point of 0.5-hr for all samples. However, the rate of decline thereafter was much slower for GCSF tandems (GCSF2NAT, 4NAT, 8NAT & 8QAT). Interestingly, all the GCSF tandems displayed a reduced rate of clearance compared to the published rhGCSF when the half-life was calculated using Winnonlin 6.3 PK program developed by Phoenix Certara (more details about half-life in the next section).

7.4.2.2 Terminal Half-life Analyses of GCSF Tandems by the Non-Compartmental Method

Comparative analyses of the Elisa data for the glycosylated GCSF tandems (GCSF2NAT, 4NAT and 8NAT) and their non-glycosylated control (GCSF8QAT) suggested that all the GCSF tandem molecules have similar clearance with GCSF4NAT exhibiting a slightly higher serum concentration at 24 hrs of injection. However, the increase was not significantly different from other GCSF tandems. In order to determine in vivo terminal half-life of the GCSF tandems and control, the pharmacokinetic data were analysed using the non-compartmental method of data analysis. This involved the use of Winnonlin 6.3 PK program developed by Phoenix Certara to determine the terminal half-life of the tandem GCSF proteins in each individual rat used in this study. The results obtained from the analyses are shown below (Table 7.1 and Figure 7.8).

Table 7-1: Tandem GCSF proteins terminal half-life analyses by non-compartmental method for each rat

The terminal half-life of the tandem GCSF molecules were determined using a Winnonlin 6.3 PK program developed by Phoenix Certara. The results of the groups of rats used for each individual tandem molecule analyses (B1-B5: 2NAT, C1-C4: 4NAT, D2-D5: 8NAT, E1-E4: 8QAT). Cmax: concentration maximum (nM); Tmax: time maximum (hr); AUC0-24: Area Under the Curve during 24 hrs; Nº points: Terminal half-life points (at least 3 points); Lambda z: Slope of the drop; Half-life (hr); the time required for the protein concentration to fall to half its initial amount.

Construct	Rat	Cmax (nM)	Tmax (hr)	AUC0-24	No points	Lambda z	Half life (hr)
GCSF2NAT	B1	90.7	0.5	377.2	3	0.07	9.66
	B2	49	0.5	202.6	3	0.09	7.32
	B3	40	0.5	125	4	0.13	5.14
	B4	42.1	0.5	201.6	5	0.11	6.05
	B5	20.5	0.5	161.7	5	0.08	8.26
GCSF4NAT	C1	81.6	0.5	366.8	3	0.06	11.81
	C2	63.1	0.5	195.6	5	0.11	6.56
	C3	21.2	2	154.3	3	0.04	15.99
	C4	75.9	1	756.3	3	0.075	9.26
GCSF8NAT	D2	86	0.5	242.8	6	0.07	9.69
	D3	17.7	0.5	147.6	3	0.13	5.31
	D4	51.3	0.5	380	3	0.18	3.75
	D5	35.5	2	246.1	4	0.09	7.85
GCSF8QAT	E1	66	0.5	141.9	3	0.15	4.54
	E2	9.6	2	108.5	4	0.09	7.62
	E3	74	0.5	203.1	4	0.12	6.01
	E4	11.1	4	120.3	3	0.12	5.81

A

Construct	Rat Group	Cmax (hrs)	Tmax (hrs)	AUC0-24	No points	Lambda z	Half life (hrs)	SEM
Published GCSF	-	-	-	-	-	-	1.79	-
GCSF2NAT	B	48.5	0.5	192.3	3	0.09	7.38	0.89
GCSF4NAT	C	49.8	0.5	257.4	6	0.07	10.74	2.31
GCSF8NAT	D	46.6	0.5	254.8	5	0.1	6.74	1.52
GCSF8QAT	E	49.8	0.5	162.8	5	0.12	5.87	0.73

B

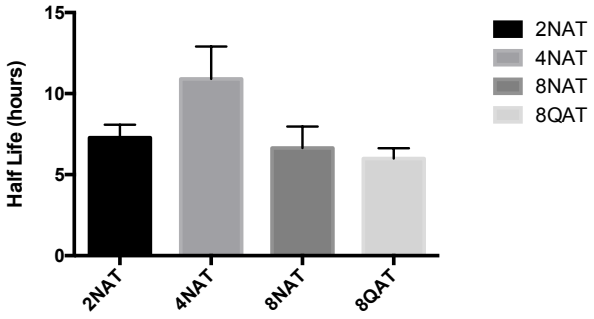


Figure 7-8: Tandem GCSF proteins terminal half-life analyses by non-compartmental method for each rats' group

(A) The terminal half-life of the tandem GCSF molecules were determined using a Winnonlin 6.3 PK program developed by Phoenix Certara. The results of the average for each treatment group together with published GCSF were tabulated. **(B)** The average terminal half-life for each molecule was graphically represented. Data are given as standard error of the mean (SEM) for at least four rats per group. Significance between terminal half-life data of GCSF tandems was performed with GraphPad Prism using One-Way ANOVA.

7.4.2.3 Pharmacodynamics of GCSF Tandems

Total cell blood count (CBC) was performed on selected whole blood samples (-24, 12, 24, 48 and 72 hrs). The counts of white blood cells (WBCs) were performed using a coulter counter instrument. Blood smears were fixed and stained using Giemsa stain by routine laboratory methods. The blood neutrophil count was performed on these stained slides using a light microscope to obtain the percentage of neutrophils.

Following injection with GCSF2NAT, 4NAT and 8NAT, 8QAT, rhGCSF and vehicle (PBS only), WBCs count showed no statistical difference in number of WBCs between any treatments above that of controls (Vehicle, rhGCSF & GCSF8QAT). However, WBC's peaked at 24 hrs post-injection for GCSF8NAT and its respective control GCSF8QAT before returning to baseline values at 72 hrs. This effect was not significantly different from the rhGCSF and vehicle control (Figure 7.9E).

In contrast, all glycosylated GCSF tandems (GCSF2NAT, 4NAT & 8NAT) showed increased percentage (%) of neutrophils at 12 hrs post injection, with this level of increase being higher for both 2NAT and 4NAT compared to controls (Vehicle, rhGCSF & GCSF8QAT). However, beyond 12 hrs post-injection, all three glycosylated GCSF tandems showed a decline in their neutrophil levels, which are more pronounced in 2NAT and 4NAT, up to 72 hrs. In contrast, GCSF8NAT exhibited a higher increase in neutrophil levels at 48 hrs post injection following a marginal decline at 24 hrs, compared to controls which may imply a longer duration of action (Figure 7.9A-D).

The obvious difference in the number of neutrophils mobilized between glycosylated tandems (at 12 hours for 2NAT/4NAT and 48 hours for 8NAT) is significant when the data was analysed by multiple T-test (Figure 7.9). However, as low number (~four) of rats were used in these studies, the neutrophils mobilization differences observed were not significant. Consequently, using more rats in the future studies is required to assess the statistical significance of the number of neutrophils mobilized by the glycosylated tandems following intravenous injection in the rats.

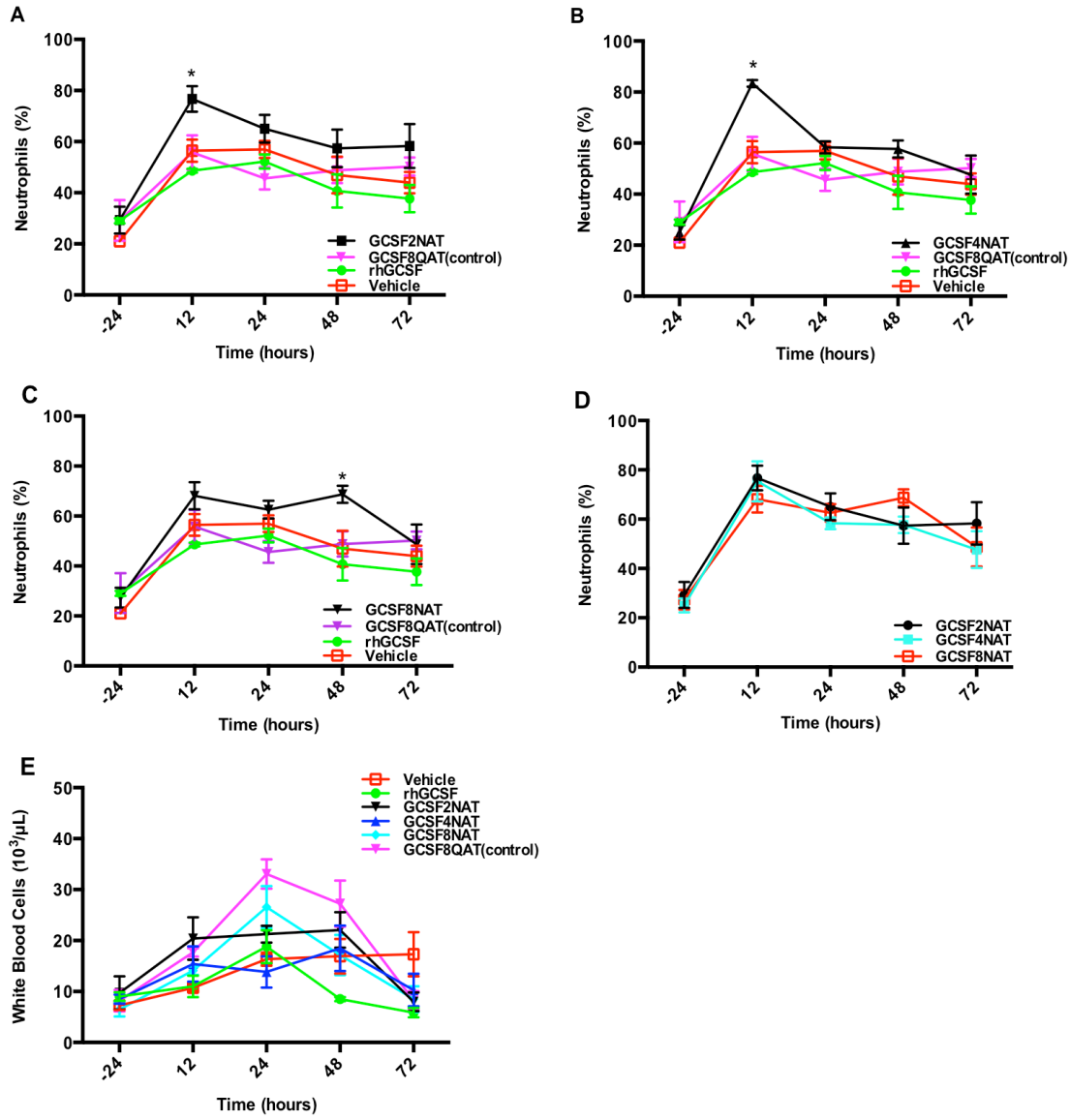


Figure 7-9: Percentage change in blood neutrophils following intravenous administration of GCSF tandems

(A) GCSF2NAT and controls (GCSF8QAT, rhGCSF & vehicle), **(B)** GCSF4NAT and controls (GCSF8QAT, rhGCSF & vehicle), **(C)** GCSF8NAT and controls (GCSF8QAT, rhGCSF & vehicle), **(D)** GCSF2NAT, GCSF4NAT and GCSF8NAT & **(E)** showing total WBC counts following intravenous administration of GCSF tandems (GCSF2NAT, 4NAT, 8NAT & 8QAT), rhGCSF and vehicle to normal rats. Data are given as standard error of the mean (SEM) for at least four rats per group. The pharmacodynamics studies of the GCSF tandems at these selected sampling time points (-24, 12, 24, 48 and 72 hrs) were analysed with GraphPad Prism using multiple T-test. Stars indicate neutrophil values that are significantly different between rats treated with GCSF tandems, rhGCSF and vehicle (* = p value of <0.05).

7.5 Discussion

In the previous chapter, the GCSF tandem molecules demonstrated better biological activity compared to the rhGCSF and showed high stability at 4°C, RmT, 37°C and multiple F/T cycles. Hence, these molecules were tested for the pharmacokinetics and pharmacodynamics (PK/PD) properties in an *in vivo* study. The pharmacokinetic performance of selected constructs (GCSF2NAT, 4NAT, 8NAT and its control GCSF8QAT) was evaluated in normal adult Sprague Dawley rats. Serum samples were quantified using Elisa technique to assess the delay in renal clearance (half-life) compared to the reported rhGCSF. This chapter equally evaluated the pharmacodynamics (PD) by measuring the effect of these GCSF tandems on the population of WBCs and neutrophils since GCSF is routinely used clinically to increase the number of neutrophils.

In the PK analysis the detection of maximum protein levels at the earliest sampling time-point of 0.5 hr, and decline thereafter for the rest of the time points, verified the successful IV injection of rats in all samples except rhGCSF. The results for rhGCSF were omitted from this study due to problems with the mode of injection. The rats should have been injected intravenously through the neck vein, but rats showed potential subcutaneous clearance for some of the time points when analysed by Elisa. The affected sample time points were at 1 and 2 hrs, which showed high protein levels compared to samples taken at time point 0.5 hr. It was presumed that the contractor might have missed the neck vein during the IV injection and injected the rats' muscle instead of the vein. Out of the 5 rats used for the in vivo PK/PD studies of rhGCSF, only 1 rat appeared to be injected intravenously. To annul any potential inconsistencies and misinterpretation of data, and erroneous data analyses, eliminating the rhGCSF control data from the in vivo study was considered appropriate. Consequently, the half-life of the GCSF tandems proteins were compared to the published rhGCSF due to time limitation of this project.

The pharmacokinetic profiles for all GCSF tandems following intravenous dosing at 250μg/kg showed an approximately 3 fold longer circulating half-

life compared to that reported for the rhGCSF (Tanaka et al., 1991). Additionally, the results also showed that GCSF4NAT had a slower rate of clearance (10.74 hrs) at 24 hrs post injection compared to other GCSF tandems (GCSF2NAT=7.38; GCSF8NAT=6.74; GCSF8QAT=5.87). However, this increase was only marginal, and not significantly different from other GCSF tandems. Interestingly, the GCSF8NAT with more glycosylation sites (8 sites) had a slightly lower clearance rate compared to the other two GCSF tandems with lower numbers of glycosylation sites (4 sites in GCSF4NAT and 2 sites in GCSF2NAT). This implies that there was a maximum glycosylation level in the GCSF tandems beyond which further glycosylation provided no additional benefit towards prolongation of the half-life as observed in GCSF8NAT.

The GCSF8QAT tandem has a similar clearance rate to all other glycosylated tandems, implying that the linker glycosylations are not required for the improved clearance of these molecules. It therefore appears that it is the increase in Mw of the tandem by virtue of containing two GCSF molecules that is responsible for the observed delayed clearance. Both GCSF molecules also have the potential to be O-link glycosylated which would also contribute to the increased Mw seen over that of native GCSF and therefore may contribute further to the delayed clearance. Study by Marinaro et al., (2000) showed that the removal of O-linked glycosylated from human IGFBP-6 (Insulin-like growth factor binding proteins-6 act as inhibitor of IGF actions) decrease the circulating half-life by 2.3 fold compared to glycosylated-IGFBP-6 (Marinaro et al., 2000).

This theory is supported by a study that was performed on recombinant human Follicle-Stimulating Hormone (rhFSH) containing different N-glycosylated linker inserts between the α- and β- chains in the same rhFSH ligand. In this study, the terminal half-life of three rhFSH molecules with increasing N-linked glycosylations: rhFSH-N1 (1 glycosylation site), rhFSH-N2 (2 glycosylation sites) and rhFSH-N4 (4 glycosylation sites) and one rhFSH with no N-linked glycosylation but only has the linker insert (rhFSH-N0), were assessed after intravenous injection into Sprague Dawley rats. In

the same study, two other molecules with O-linked glycosylation were equally tested, *in vivo*, to determine the half-life. These are rhFSH-CTP (comprising of a unique carboxy-terminal peptide (CTP) that contains 4 O-linked glycosylation sites inserts as a linker between the α- and β- chains in the same rhFSH ligand) and rhFSH-O1 (1 O-linked glycosylation site inserts as a linker between the α- and β- chains in the same rhFSH ligand). The half-lives of the O-linked glycosylated rhFSH molecules when compared to those of the N-linked showed that hyperglycosylation of either O- or N-linked of the rhFSH ligand increased the half-life of rhFSH. However, the increase was not linearly related to the carbohydrate sizes and numbers, as the rhFSH-2N, rhFSH-4N and rhFSH-CTP (4 O-linked glycosylation) had similar half-lives which were 2 fold longer compared to rhFSH, rhFSH-N0 and rhFSH-O1 (Klein et al., 2002, Weenen et al., 2004).

Furthermore, another independent study also described an increase in the half-life of rhFSH by introducing two N-linked glycosylation sites at the N terminus of the α-subunit (Perlman et al., 2003), which is different from the linker with glycosylation sites in the earlier mentioned studies on rhFSH. This suggests that the increase in the number of glycosylation sites in rhFSH, irrespective of the location or the type was responsible for the observed increase in terminal half-life of rhFSH in these studies.

The importance of hyperglycosylation to a longer terminal half-life is further highlighted in another study carried out to investigate the effects of introducing different number of N-linked glycosylation sites on the half-life of a small bispecific single-chain diabody (scDb CEACD3) (Used for the retargeting of cytotoxic T cells to CEA-expressing tumor cells). Stock et al. introduced 3, 6, or 9 N-glycosylation sites in the flanking linker of the scDb molecule and a C-terminal extension. Interestingly, the results showed a prolonged circulating half-life for all three scDb constructs compared to the unmodified scDb. However, the addition of 3 N-glycosylation sites is adequate to prolong circulation time and not significantly different from 6 or 9 N-linked glycosylation (Stork et al., 2008).

These studies revealed that the addition of a few N or O-glycans either in a linker or covalently bound to a protein improved the pharmacokinetic properties of the protein, thus producing a novel protein with moderately prolonged terminal half-life. However, there is a maximum benefit of glycosylation regarding the half-life, and additional glycosylation sites might not extend circulation time. This view is consistent with our findings in the PK studies, where all the proteins showed similar terminal half-life irrespective of the numbers of glycosylation sites in the linker.

The potency of each GCSF tandem was evaluated *in vivo* using white blood cell (WBC) and neutrophil counts. WBC counts showed no statistical difference in number following injection with either vehicle (PBS only), rhGCSF, GCSF2NAT and GCSF4NAT. However, WBC levels peaked at 24 hrs post-injection for GCSF8NAT and its respective control GCSF8QAT before returning to baseline values at 72 hrs. However, this effect was not significantly different from that observed in the rhGCSF and vehicle controls.

In contrast, all rats injected with GCSF tandems and rhGCSF showed an increase in the percentage of circulating neutrophils particularly at 12 hrs post injection for GCSF2NAT and GCSF4NAT compared to controls (Vehicle, rhGCSF & GCSF8QAT). This observed increase in the number of neutrophils at 12 hrs is consistent with that reported by Ulich et al. (1988). In their studies in mice, they observed that a single injection of GCSF induced a temporary neutropenia in the mice. However, the number of circulating neutrophils was found to increase significantly by 5-fold within 30 minutes and peaked at 12 hrs post-injection. While there was an increase in the number of circulating neutrophils, the number of mature neutrophils in the bone marrow became significantly reduced (Ulich et al., 1988).

The percentage of neutrophils in human blood circulation is around 60-65%, which is different from rats and mice. The percentage of neutrophils in normal rats is between 22-44.9% (Sharma, 2013) and that of normal mice is between 10-40% of total cell counts in the peripheral blood (www.ahc.umn.edu/rar/refvalues). In the *in vivo* studies, the percentage of

neutrophils at 12 hrs for GCSF2NAT, GCSF4NAT and GCSF8NAT was between 60-70% in normal rats compared to rhGCSF, GCSF8QAT and vehicle (~45%). The presence of high percentage of neutrophils, but very low percentage of other cells such as lymphocytes (normal range; 56-78%), monocytes, eosinophils or basophils, indicates that GCSF is selective for neutrophils. This observation is consistent with the report that the presence of mature and immature (or band) neutrophils in circulation, but not lymphocytes or eosinophils, was a result of GCSF induced stimulation of neutrophils from the bone marrow into the blood (Semerad et al., 2002).

In the current study, a higher increase in the percentage of circulating neutrophils at 12 hrs post injection was observed for both GCSF2NAT and GCSF4NAT and 48 hrs post injection for GCSF8NAT compared to controls (Vehicle, rhGCSF & GCSF8QAT). This confirms the positive role of glycosylation in improving the pharmacodynamics of our GCSF tandems. Although the half-life of GCSF8NAT exhibited no further enhancement beyond either the GCSF2NAT, GCSF4NAT or its control GCSF8QAT, the high percentage of neutrophils in GCSF8NAT compared to other tandems leads us to hypothesize that GCSF8NAT is a more efficient stimulator of neutrophils as it has a longer duration of action evidenced by high circulating neutrophil levels at 48 hrs. This could be due to a compound effect of being larger in size, more glycosylations (more terminal sialisation) or the presence of up to eight glycosylations sites could impede the interaction of these GCSF molecules (impede dimer formation) making it more active in the circulation.

The number of N-linked glycosylation sites within GCSF8NAT tandem linker increased its size to ~70kDa, which is bigger than both GCSF4NAT (~55kDa) and GCSF2NAT (~40kDa). We can hypothesis that GCSF4NAT and 2NAT are cleared more efficiently via filtration by the kidney, the increased size of GCSF8NAT delays its clearance by filtration, as kidney is reportedly known to filter molecules smaller than the human serum albumin (67kDa) (Dennis et al., 2002). Also terminal sialisation of GCSF8NAT tandem would also prevent clearance via filtration in the kidney due to over-expression of

a negatively charged barrier alongside the glomerular filtration barrier, preventing the passage of glycoproteins through charge repulsion (Varki, 2008).

Furthermore, the non-glycosylated GCSF tandem controls were observed to form dimers *in vitro*, which could be attributed to the formation of intermolecular or intramolecular disulphide bonds resulting from the interaction of the free cysteine residues (Cys17) present on individual GCSF molecules. Dimers are unable to induce signal transduction, which leads to less stimulation of neutrophils. Dimer formation is potentially easier in control tandems due to less steric hindrance from glycosylation whereas in 8NAT (and other glycosylated tandems) the presence of glycosylation within the linker could be inhibitory to the formation of dimers. For instance, 2NAT and 4NAT stimulated more neutrophils than 8QAT at 12 hrs but 8NAT stimulated more neutrophils than 8QAT at 48 hrs after intravenous injection. The presence of up to eight glycosylation sites could impede the interaction of these GCSF molecules better than Both GCSF2NAT and 4NAT (less glycosylation motifs). This was evidenced by stability results in the previous chapter, showed a slight dimer formation for both 2NAT and 4NAT, whereas, GCSF8NAT showed no dimer formation at all (please refer to Figure 6.5, 6.7 and 6.8). This may also be a contributing factor to the observed increase in neutrophil counts with these proteins and therefore increase clearance via the neutrophils-mediated pathway.

One of the main ways in which GCSF is cleared is via the neutrophil mediated pathway. It is this pathway that we hypothesise is the main route of clearance of our glycosylated GCSF tandems. This effect has been seen with Pegfilgratim, which was developed to improve PK and PD of GCSF molecules (i.e. decreased clearance via kidneys therefore increased clearance via neutrophils). For example, concentrations of Pegfilgrastim in patient serums following administration, remain high during neutropenia, but reduced when the numbers of neutrophils increase (Curran and Goa, 2002).

While these observations were different from our initial expectations regarding the terminal half-life of the GCSF tandems in circulation, as we hypothesized that increasing the glycosylation sites would increase the size and thus delay the clearance, our studies showed that the pharmacokinetic and pharmacodynamics properties of a therapeutic protein could be divergent. These observations are similar to the findings in an FSH study where the terminal half-life of different analogues of FSH with two (rhFSH-N2) and four (rhFSH-N4) glycosylation sites were found to have similar circulating half-lives but rhFSH-N4 was more potent in the stimulation of inhibin A and follicles than the rhFSH-N2 due to the addition of large, highly branched carbohydrates (Weenen et al., 2004).

Therefore, our studies of the tandem proteins revealed that all glycosylated tandem GCSF proteins are more potent than rhGCSF (containing intramolecular O-linked glycosylation) and GCSF8QAT (containing linker and O-linked glycosylation). However, despite a negligible difference in circulating half-lives of the tandem GCSF proteins, GCSF8NAT is the only tandem with a significant potency over rhGCSF and GCSF8QAT, being a more potent stimulator of neutrophils.

8. General Discussion

GCSF is a hormone produced by different tissues to stimulate the production of neutrophils from the bone marrow into the blood circulation. The rhGCSF has been shown to stimulate neutrophils for the treatment of neutropenic patients, and in stem cell mobilization in the circumstance of BM transplantation. The commercial products of rhGCSF fall under the category of either short acting half-life, such as, Filgrastim (NEUPOGEN®) and Lenograstim (Granocyte®) where in the pharmacokinetic properties and structural homology of the protein are similar to the human GCSF, or long acting half-life such as, Pegfilgrastim (Neulasta), where in chemical alteration with PEGylation has been applied to prolong the pharmacokinetics by reduced renal clearance. Since there is a need for less frequent (daily) dosing, the use of short-acting GCSF products have been limited and shifted towards the use of the long-acting products (e.g. Neulasta).

The fast growing market size will increase demand for generating similar long-acting rhGCSF. Therefore, it is essential to produce new GCSF compounds with improved properties over other products. Asterion is employing its proprietary protein fusion technology (ProFuse™) to produce a new form of long-acting GCSF molecule that retains the pharmacokinetic and efficiency of Neulasta, but with a more competitive manufacturing and cost-of-goods benefits over Neulasta.

A previous study by Asterion has shown that the use of glycosylated-linkers between two GH ligands to create protein-tandems resulted in their glycosylation and an increased MW whilst maintaining biological activity. This technology can be easily applied to other molecules such as, GCSF.

The data obtained in this study have shown that it is possible to clone, express and purify tandem GCSF molecules containing variably glycosylated linker regions from a CHO cell line. The purified GCSF tandems are stable with no degradation. The apparent stability could be attributed to the incorporated Gly4Ser flexible linker. This had been demonstrated in

recombinant single-chain Fv antibody production where the use of Gly4Ser linker offered the advantages of stability and lack of immunogenicity (Huston et al., 1993). Our data (western blots) suggest that these glycosylation consensus sequences (motifs) are glycosylated upon expression in CHO Flp-In cells (mammalian cell lines) as the MW of glycosylated tandem proteins are higher than their corresponding non-glycosylated controls. The MW of the non-glycosylated tandem GCSF controls 2QAT, 4QAT, and 8QAT were 45kDa, 45kDa and 49kDa respectively, whereas that of their corresponding glycosylated GCSF tandems (2NAT, 4NAT, and 8NAT) were approximately 52kDa, 60kDa and 70kDa. A study by Mann and Jensen (2003) highlighted that each unit of N-linked glycosylation contribute a minimum of approximately 800Da to the construct mass. This is consistent with our observation of increased MWs in the glycosylated tandems (~7kDa, 15kDa and ~21kDa) compared to the non-glycosylated controls. However, the increased MW observed was higher than what was expected using Mann and Jensen's prediction, which could be attributed to the unique design of our tandems to contain flexible linker. Each of our glycosylated GCSF tandem contains 2, 4 or 8 glycosylation motifs within the flexible linker contribute to the increase in MW. Increasing the linker length to accommodate more glycosylation sites (motifs) further increases the MW of the molecule. Also, glycosylation within the flexible linker region increases the apparent weight of the tandems not just by weight but also by hydrodynamic volume, which may have contributed to the difference in MWs between our study and Mann and Jensens'. In contrast, the non-glycosylated controls used in this work have shown an expected MW similar to those determined by DNASTAR laser gene (Table 3.2) and they were therefore suitable controls in this perspective of proof of concept testing.

All GCSF tandems studied showed similar *in vitro* biological activities in an AML-193 proliferation assay. This suggests that the use of hyperglycosylated linker technology is applicable for GCSF and other proteins since conjugation is not impeding or shielding residues within the

peptide that are necessary for activity. However *in vivo*, although all tandems showed similar terminal half-lives (range: 5.87 – 10.74 hrs) which were not significantly different, studies looking at neutrophil mobilization highlighted striking differences between these tandems. For instance, a higher increase in the percentage of circulating neutrophils was observed for both GCSF2NAT and GCSF4NAT at 12 hrs post injection and upto 48 hrs post injection for GCSF8NAT compared to controls (rhGCSF & GCSF8QAT). While this confirms the positive role of glycosylation in improving the pharmacodynamic of our GCSF tandems it equally showed GCSF8NAT as the most efficient stimulator of neutrophils as it has a longer duration of action evidenced by high circulating neutrophil levels at 48 hrs. This could be due to a compound effect of its larger in size, more glycosylations (more terminal sialylation) which limits its clearance by kidney filtration or the presence of 8NAT sites in the linker which impedes the interaction of the GCSF molecules (dimer formation) in the tandem making it more active in circulation.

Improving the circulating half-life and neutrophils mobilization has been the mainstay of the new generation of GCSF. GCSF is cleared from circulation via kidney filtration or the neutrophil mediated pathway. Enhancing the clearance via neutrophil mediated pathway increases the number of neutrophils in circulation, a characteristic that is beneficial for treating patients with neutropenia. The only commercially available longer acting GCSF, Pegfilgratim, was developed to improve the PK and PD of GCSF molecule by PEGylation (i.e. decreased clearance via kidneys therefore increased clearance via neutrophils). Similarly, our glycosylated GCSF tandems were designed to be cleared through this pathway. We hypothesized that increasing the glycosylation sites in our GCSF tandems would increase the size and thus delay the clearance. However, our tandems circulating half-lives (5.87 – 10.74 hrs) are not significantly different from that of Pegfilgratim (7.05 hrs) even though the molecular weights of our tandems (~45kDa-70kDa) are more than Pegfilgratim (38.8kDa).

Despite no apparent significant improvement in circulating half-life over Pegfilgrastim, our GCSF tandem molecules provide some physiological advantages. For instance, while PEGylation of GCSF in Pegfilgrastim has been shown to be non-biodegradable and toxic (Patel et al., 2014; Caliceti and Veronese, 2003), and induce vacuole formation in renal tubes thereby affecting the tissue distribution (Zhang et al., 2014). In contrast, our GCSF tandems are in a naturally occurring structure being hyperglycosylated, and therefore can undergo degradation within the human body with no potential toxicity of PEGylation was supported by the detection of PEG in bile (Caliceti and Veronese, 2003). Additionally, the cost of production of Pegfilgrastim (involves post-expression and post-purification chemical modification of the starting molecule rhGCSF) outweighs that of our GCSF tandems, which are easily expressed and purified using simple methods.

8.1 Future Work

To have a comprehensive overview of our tandems, a few more studies are required which were not covered during the course of this study. For instance, in this study, we observed high variations in the percentage of neutrophils in the normal rats used for the *in vivo* studies, as the normal range of neutrophils is 22-44%. This constituted a challenge in the analysis of the *in vivo* studies, as the variations in the neutrophil counts made it difficult to determine the exact percentage of neutrophils in the rats, especially since rat's neutrophils can increase even with stress during the injection or due to infection (Semerad et al., 2002). This observation was seen in rats injected with vehicle (PBS only) that exhibited high neutrophils similar to the test rats (please refer to Figure 7.9). Consequently, neutropenic rats are recommended for use in the future study, as we need a base line for the evaluation of the number or percentage of neutrophils.

Also, it was not conclusive if all the sites of glycosylation were occupied by glycan moieties during protein expression. Therefore, the level of glycosylation in the tandems could have been confirmed by Mass spectrometry (MS) analysis of the glycopeptides fragments following

treatment of glycosylated tandems with the PGNaseF digestive enzyme (Gervais et al., 2003). Also, the effect of O-linked glycosylation on clearance could have been determined, as all our GCSF tandems (N-linked glycosylated) and their respective non-glycosylated controls showed similar circulating half-life. Additionally, looking at the structure of the tandems may give insight as to whether proteins are monomer, dimer, and of any inter or intramolecular associations. A number of analytical methods could be used to assess these structural characteristics such as Analytical Ultracentrifugation (monomer/dimer formation), Size Exclusion Chromatography (monomer/Dimer), Dynamic Light Scattering, and Small X-Ray Scattering.

8.1.1 Future Work to Improve GCSF Tandems

Our study showed that tandem GCSF molecules with a variable N-linked glycosylated flexible linker have a therapeutic potential of stimulating neutrophils in neutropenic patients. However, improving the circulating half-life of our tandems over the commercially available long-acting GCSF (Pegfilgrastim) would provide novel longer acting GCSF products. This could be achieved by introducing two modifications: removing the free cysteine residue in GCSF molecule, and pH switching between cell-surface and endosome.

The native GCSF contains five cysteine residues, two internal disulfide bonds at positions Cys36– Cys42 and Cys64– Cys74 leaving one free cysteine residue at position Cys17 with a free sulfhydryl group. The free Cys17 of one GCSF tandem may form inter-molecular disulfide bonds with another GCSF tandem, which results in the formation of a GCSF tandem dimer, which we have seen in non-glycosylated GCSF tandems. However, it has been reported that the substitution of Cys17 with alanine (alanine instead of cysteine in wild-type GCSF) resulted in enhanced stability *in vitro* and higher bioavailability *in vivo* than wild-type GCSF, possibly through the elimination of dimerisation caused by the formation of inter-molecular disulfide bonds (Jiang et al., 2011). Therefore, substituting the Cys17 with alanine for both

GCSF ligands in our tandem could prevent dimer formation and thus increase their bioavailability.

GCSF binds its receptor with a high affinity and this results in its rapid depletion via receptor-mediated endocytosis by circulating neutrophils expressing GCSF-R, thus diminishing its therapeutic efficiency at a pharmacokinetic level. It has been demonstrated that substituting ligand residues at the binding site of GCSF with histidine switches protonation states between cell-surface and endosomal pH. Therefore, decreasing the physiological pH from ~7 at the cell surface to 5 or 6 in endosomes by selectively mutating amino acids at the GCSF binding site will deteriorate interactions at endosomal pH whilst maintaining the electrostatic interactions at extracellular pH, and consequently increase endocytic GCSF recycling (Sarkar et al., 2002). Six ligand residues located at the binding site of GCSF (Glu20, Gln21, Asp110, Asp113, Thr117, and Gln120) are potential targets for the site-directed mutagenesis with 6 histidine residues using two different single-histidine mutants (neutral and protonated histidine). The suggested mutation is based on the principle that neutral histidine might retain relatively tight binding on the cell surface while protonated histidine might lead to a weaker binding in endosomal partitions. Therefore, substituting the amino acids of each GCSF ligand binding site in our tandems with histidine would reduce its receptor binding affinity in intracellular endosomal partitions and resultantly leads to an increased recycling of GCSF ligand from the intracellular cell to the extracellular medium and thereby extend the circulating half-life of our GCSF tandems. This ultimately facilitates the endocytic GCSF ligand recycling and longer half-life in extracellular circulation.

9. Conclusion

The results obtained from this project exhibited that it is possible to clone, express and purify GCSF tandems. It also appeared that the use of glycosylated motifs (NAT) within a flexible linker between two GCSF ligands to generate protein-tandems results in molecules with increased molecular weight according to the number of glycosylation sites. Tandems of GCSF have increased *in vitro* bioactivity compared to monomeric GCSF but this was independent of glycosylation and glycosylation did not inhibit *in vitro* bioactivity. Tandems with and without glycosylation had three-fold greater half-lives than rhGCSF but this was not determined by the number of glycosylation sites. There was evidence that GCSF8NAT was biologically active *in vivo*. The results confirm the hypothesis that it is possible to predictably increase the molecular weight of GCSF tandems and retain biological activity but this was not associated with a predictable prolongation of the serum half-life.

Appendix A

Appendix A.1. Nucleotide Sequences of Primers

Primer	Nucleotide Sequence
GCSF Nhe1	5'-AAATTTGGATCCGCTAGCCACCATGGCTGGACC-3'
GCSF Xho1 Rev	5'-ATTCTCGAGGGGCTGGGCAAGGTGGCGTA-3'
GCSF BamH1	5'-AGGAGGGGATCCACCCCCCTGGG-3'
GCSF Age1 Rev	5'-AAGAAGACCGGTTCCACCGGTTCCACCTCCACCGGGCTGGGCAAGGTGGCG-3'
CMVFor	5'-TATTACCATGGTGATGCGGTTTTGG-3'
BGHRev	5'-TAGAAGGCACAGTCGAGG-3'
GHseq2Rev	5'-AAGGCCAGCTGGTGCAGACG-3'

Appendix A.2. Restriction Endonucleases Cut Sites

Enzyme	Restriction site
Nhe1	5' G/CTAGC 3' 3' CGATC/G 5'
Xho1	5' G/TCGAG 3' 3' GAGCT/G 5'
BamH1	5' G/GATCC 3' 3' CCTAA/G 5'
Age1	5' A/CCGGT 3' 3' TGGCC/A 5'

Appendix B

Appendix B.1. Nucleotide and Amino Acid Sequences of GH Tandem

Nucleotide sequence

GCTAGC**caccAtggctacaggctcccggacgtccctgctcctggcttttggcctgctctgcctgccctggct tcaagagggcagtgcc**TTCCCAACCATTCCCTTATCCAGGCTTTTTGACAACGCTATGCT CCGCGCCCATCGTCTGCACCAGCTGGCCTTTGACACCTACCAGGAGTTTGAAGAAGC CTATATCCCAAAGGAACAGAAGTATTCATTCCTGCAGAACCCCCAGACCTCCCTCTG TTTCTCAGAGTCTATTCCGACACCCTCCAACAGGGAGGAAACACAACAGAAATCCAA CCTAGAGCTGCTCCGCATCTCCCTGCTGCTCATCCAGTCGTGGCTGGAGCCCGTGCA GTTCCTCAGGAGTGTCTTCGCCAACAGCCTGGTGTACGGCGCCTCTGACAGCAACGT CTATGACCTCCTAAAGGACCTAGAGGAACGCATCCAAACGCTGATGGGGAGGCTGG AAGATGGCAGCCCCCGGACTGGGCAGATCTTCAAGCAGACCTACAGCAAGTTCGACA CAAACTCACACAACGATGACGCACTACTCAAGAACTACGGGCTGCTCTACTGCTTCA GGAAGGACATGGACAAGGTCGAGACATTCCTGCGCATCGTGCAGTGCCGCTCTGTGG AGGGCAGCTGTGGCTTCLinker:variableTTCCCAACCATTCCCTTATCCAGGCTTTT TGACAACGCTATGCTCCGCGCCCATCGTCTGCACCAGCTGGCCTTTGACACCTACCA GGAGTTTGAAGAAGCCTATATCCCAAAGGAACAGAAGTATTCATTCCTGCAGAACC CCCAGACCTCCCTCTGTTTCTCAGAGTCTATTCCGACACCCTCCAACAGGGAGGAAA CACAACAGAAATCCAACCTAGAGCTGCTCCGCATCTCCCTGCTGCTCATCCAGTCGT GGCTGGAGCCCGTGCAGTTCCTCAGGAGTGTCTTCGCCAACAGCCTGGTGTACGGCG CCTCTGACAGCAACGTCTATGACCTCCTAAAGGACCTAGAGGAAGGCATCCAAACGC TGATGGGGAGGCTGGAAGATGGCAGCCCCCGGACTGGGCAGATCTTCAAGCAGACCT ACAGCAAGTTCGACACAAACTCACACAACGATGACGCACTACTCAAGAACTACGGGC TGCTCTACTGCTTCAGGAAGGACATGGACAAGGTCGAGACATTCCTGCGCATCGTGC AGTGCCGCTCTGTGGAGGGCAGCTGTGGCTTC(ggtgga ggtgga) ACCGGT-catcatcaccatcaccat*

Amino acid Sequence:

matgsrtsIIIafgIIcIpwIqegsaFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIP KEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVF ANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDD ALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGFLinker:variableFPTIPLSRLF DNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQ QKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGR LEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRS VEGSCGFGGGGTGHHHHHH*

The GH tandem with a 4x glycine sequence highlighted in purple was ligated into the vector pGHSecTag between Nhe1 restriction site highlighted in blue and Age1 restriction site highlighted in pink. The GH tandem containing GH signal sequence is highlighted in bold lower case. The linker between two GH molecules is highlighted in red. A Hist tag is highlighted in green. This molecule will be used as the template to replace with GCSF.

Appendix B.2. Nucleotide and Amino Acid Sequences of GCSF Tandem

Nucleotide sequence

GCTAGCcacc**atggctggacctgccacccagagccccatgaagctgatggccctgcagctgctgctgtgg**
cacagtgcactctggacagtgcaggaagccACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAG
AGCTTCCTGCTCAAGTGCTTAGAGCAAGTGAGGAAGATCCAGGGCGATGGCGCAGC
GCTCCAGGAGAAGCTGTGTGCCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTGCT
GCTCGGACACTCTCTGGGCATCCCCTGGGCTCCCCTGAGCAGCTGCCCCAGCCAGGCC
CTGCAGCTGGCAGGCTGCTTGAGCCAACTCCATAGCGGCCTTTTCCTCTACCAGGGG
CTCCTGCAGGCCCTGGAAGGGATCTCCCCCGAGTTGGGTCCCACCTTGGACACACTG
CAGCTGGACGTCGCCGACTTTGCCACCACCATCTGGCAGCAGATGGAAGAACTGGGA
ATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTCGCCTCTGCTTTCC
AGCGCCGGGCAGGAGGGGTCCTGGTTGCCTCCCATCTGCAGAGCTTCCTGGAGGTGT
CGTACCGCGTTCTACGCCACCTTGCCCAGCCCLinker:variableACCCCCCTGGGCCCT
GCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAAGTGAGGAAGATC
CAGGGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAGCTGTGCCAC
CCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCCCTGAGC
AGCTGCCCCAGCCAGGCCCTGCAGCTGGCAGGCTGCTTGAGCCAACTCCATAGCGGC
CTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCCGAGTTGGGT
CCCACCTTGGACACACTGCAGCTGGACGTCGCCGACTTTGCCACCACCATCTGGCAG
CAGATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCG
GCCTTCGCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGTCCTGGTTGCCTCCCATCTGC
AGAGCTTCCTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCC(ggtgga
ggtgga) ACCGGT-catcatcaccatcaccat*

Amino acid Sequence:

MAGPATQSPMKLMALQLLLWHSALWTVQEATPLGPASSLPQSFLLKCLEQVRKIQ
GDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLAGCLSQLHSGLF
LYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPTQGAMPA
FASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQPlinker:variableTPLGPASSLPQSF
LLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQL
AGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAP
ALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQPGGGGTGHHHH
HH*

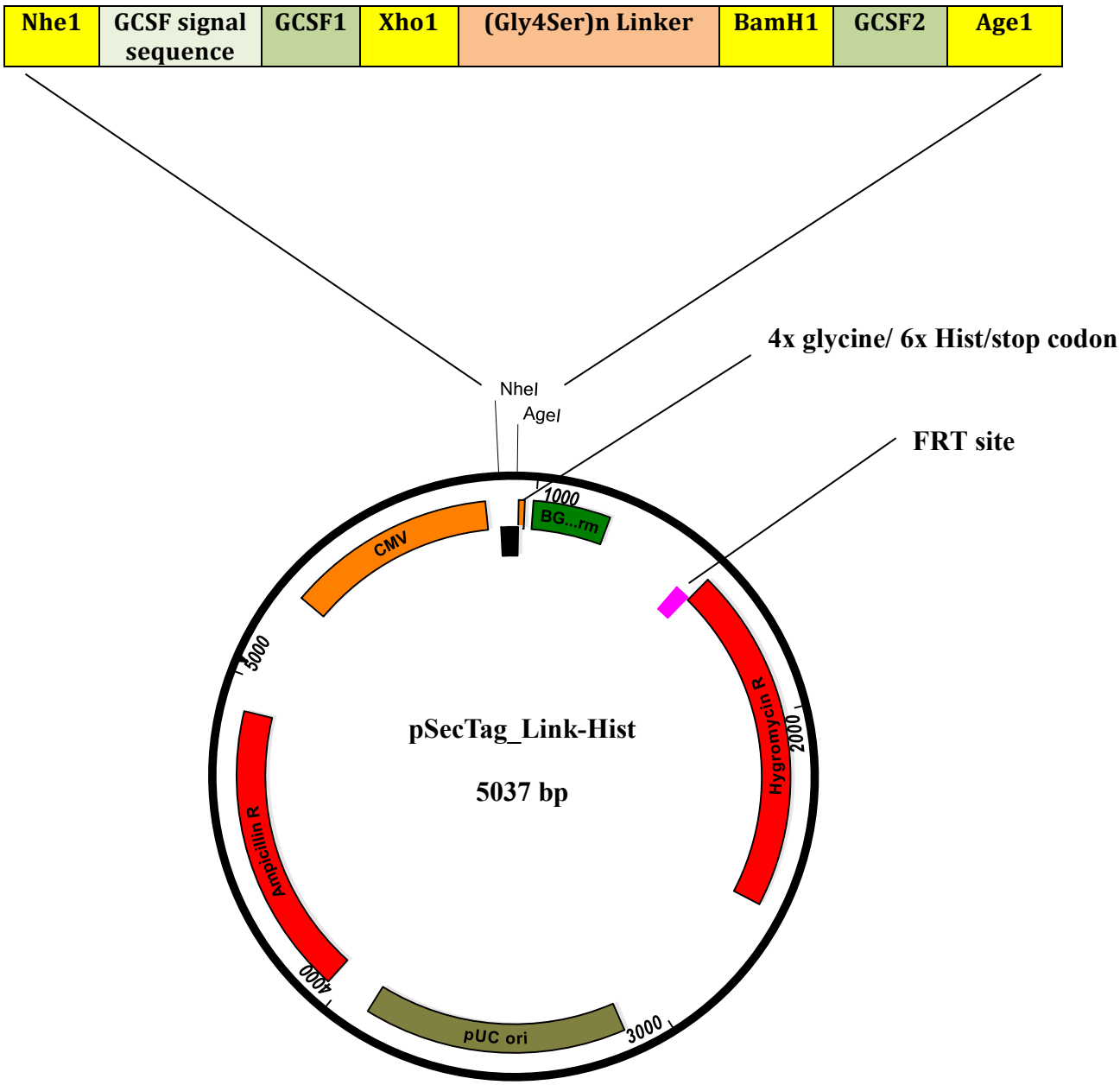
The GCSF tandem with a 4x glycine sequence highlighted in purple was ligated into the vector pGHSecTag between Nhe1 restriction site highlighted in blue and Age1 restriction site highlighted in pink. The GCSF tandem containing GCSF signal sequence is highlighted in bold lower case. The linker between two GCSF molecules is highlighted in red. A Hist tag is highlighted in green. This molecule will be used as the template to ligate in different linker constructions.

Appendix B.3. Nucleotide and Amino Acid Sequences of Linker Regions

Linker	Nucleotide sequence	Amino acid sequence
GCSF2NAT	CTCGAGGGTGGTGGAGGTAGTGGAGGAAACGCTA CAGGAGGTGGCGGGTCTGGTGGGGGGGGCTCTGG AGGTGGAGGGTCAGGCGGGGGAGGATCAGGGGGA GGCGGTTCCGGGGGCAACGCAACCGGGGGCGGAG GCTCCGGATCC	LEGGGGSGGNATGGGGS GGGGSGGGGSGGGGSGG GGSGGNATGGGGSGS
GCSF2QAT	CTCGAGGGTGGTGGAGGTAGTGGAGGACAGGCTA CAGGAGGTGGCGGGTCTGGTGGGGGGGGCTCTGG AGGTGGAGGGTCAGGCGGGGGAGGATCAGGGGGA GGCGGTTCCGGGGGCCAGGCAACCGGGGGCGGAG GCTCCGGATCC	LEGGGGSGGQATGGGGS GGGGSGGGGSGGGGSGG GGSGGQATGGGGSGS
GCSF4NAT	CTCGAGGGTGGAGGAGGTTCTGGAGGTAATGCTA CTGGAGGTGGTGGCAGCGGAGGCAACGCAACAGG GGGTGGCGGATCTGGAGGAAACGCAACCGGTGGA GGGGGATCTGGTGGGAACGCTACCGGCGGAGGGG GCTCTGGATCC	LEGGGGSGGNATGGGGS GGNATGGGGSGGNATG GGGSGGNATGGGGSGS
GCSF4QAT	CTCGAGGGCGGCGGTGGGTCCGGTGGCCAGGCTAC CGGAGGAGGCGGGAGTGGAGGCCAAGCCACAGGT GGCGGAGGGTCTGGCGGTCAGGCAACTGGCGGAG GAGGGTCAGGGGGGCAGGCCACGGGAGGTGGCGG GAGCGGATCC	LEGGGGSGGQATGGGGS GGQATGGGGSGGQATG GGGSGGQATGGGGSGS
GCSF8NAT	CTCGAGGGCGGCGGAGGGAGTGGCGGTAACGCTA CGGGAGGAGGAGGCTCTGGCGGCAATGCAACCGG CGGTGGCGGGAGTGGCGGGAATGCCACAGGTGGG GGGGGTTCAGGCGGGAATGCTACTGGCGGCGGCG GTTCCGGAGGCGGAGGGTCTGGTGGGAACGCAAC CGGTGGTGGTGGAAGCGGAGGGAATGCTACCGGT GGCGGAGGAAGCGGTGGTAACGCCACTGGAGGCG GCGGGTCCGGAGGCAACGCCACAGGGGGTGGAGG GTCAGGATCC	LEGGGGSGGNATGGGGS GGNATGGGGSGGNATG GGGSGGNATGGGGSGGG GSGGNATGGGGSGGNAT GGGGSGGNATGGGGSGG NATGGGGSGS
GCSF8QAT	CTCGAGGGCGGCGGAGGGAGTGGCGGTCAAGCTA CGGGAGGAGGAGGCTCTGGCGGCCAGGCAACCGGC GGTGGCGGGAGTGGCGGGCAAGCCACAGGTGGGG GGGGTTCAGGCGGGCAGGCTACTGGCGGCGGCGGT TCCGGAGGCGGAGGGTCTGGTGGGCAAGCAACAG GTGGTGGTGGAAGCGGAGGGCAGGCTACTGGTGG CGGAGGAAGCGGTGGTCAAGCCACTGGAGGCGGC GGGTCCGGAGGCCAGGCCACAGGGGGTGGAGGGT CAGGATCC	LEGGGGSGGQATGGGGS GGQATGGGGSGGQATG GGGSGGQATGGGGSGGG GSGGQATGGGGSGGQAT GGGGSGGQATGGGGSGG QATGGGGSGS

The table above represents the linker regions that were ligated between two GCSF molecules using Xho1 restriction site highlighted in pink and BamH1 highlighted in green. Glycosylation motifs highlighted in red NAT (AsnAlaThr) are amino acids in which N is recognized by cells for glycosylation. While, non-glycosylation motifs (Controls) highlighted in blue QAT (GlnAlaThr) are amino acids in which Q is not recognized by cells for glycosylation.

Appendix C. pSecTag_Link-Hist Modulating Vector:



Bibliography

AAPRO, M. S., BOHLIUS, J., CAMERON, D. A., DAL LAGO, L., DONNELLY, J. P., KEARNEY, N., LYMAN, G. H., PETTENGELL, R., TJAN-HEIJNEN, V. C., WALEWSKI, J., WEBER, D. C. & ZIELINSKI, C. 2011. 2010 update of EORTC guidelines for the use of granulocyte-colony stimulating factor to reduce the incidence of chemotherapy-induced febrile neutropenia in adult patients with lymphoproliferative disorders and solid tumours. *Eur J Cancer,* 47**,** 8-32.

ABUCHOWSKI, A., VAN ES, T., PALCZUK, N. C. & DAVIS, F. F. 1977. Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol. *J Biol Chem,* 252**,** 3578-81.

ANCLIFF, P. J., GALE, R. E., LIESNER, R., HANN, I. & LINCH, D. C. 2003. Long-term follow-up of granulocyte colony-stimulating factor receptor mutations in patients with severe congenital neutropenia: implications for leukaemogenesis and therapy. *Br J Haematol,* 120**,** 685-690.

ANDERSEN, J. T., DABA, M. B. & SANDLIE, I. 2010. FcRn binding properties of an abnormal truncated analbuminemic albumin variant. *Clin Biochem,* 43**,** 367-72.

ANDERSEN, J. T., DEE QIAN, J. & SANDLIE, I. 2006. The conserved histidine 166 residue of the human neonatal Fc receptor heavy chain is critical for the pH-dependent binding to albumin. *Eur J Immunol,* 36**,** 3044-51.

ANDERSEN, J. T. & SANDLIE, I. 2009. The versatile MHC class I-related FcRn protects IgG and albumin from degradation: implications for development of new diagnostics and therapeutics. *Drug Metab Pharmacokinet,* 24**,** 318-32.

ANDERSON, C. L., CHAUDHURY, C., KIM, J., BRONSON, C. L., WANI, M. A. & MOHANTY, S. 2006. Perspective-- FcRn transports albumin: relevance to immunology and medicine. *Trends Immunol,* 27**,** 343-8.

ARVEDSON, T. & GIFFIN, M. 2012. Structural Biology of G-CSF and Its Receptor. *In:* MOLINEUX, G., FOOTE, M. & ARVEDSON, T. (eds.) *Twenty Years of G-CSF.* Springer Basel.

ASANO, S., URABE, A., OKABE, T., SATO, N. & KONDO, Y. 1977. Demonstration of granulopoietic factor(s) in the plasma of nude mice transplanted with a human lung cancer and in the tumor tissue. *Blood,* 49**,** 845-52.

ATAERGIN, S., ARPACI, F., TURAN, M., SOLCHAGA, L., CETIN, T., OZTURK, M., OZET, A., KOMURCU, S. & OZTURK, B. 2008. Reduced dose of lenograstim is as efficacious as standard dose of filgrastim for peripheral blood stem cell mobilization and transplantation: a randomized study in patients undergoing autologous peripheral stem cell transplantation. *Am J Hematol,* 83**,** 644-8.

BAILON, P. & WON, C. Y. 2009. PEG-modified biopharmaceuticals. *Expert Opin Drug Deliv,* 6**,** 1-16.

BAKER, K., QIAO, S. W., KUO, T., KOBAYASHI, K., YOSHIDA, M., LENCER, W. I. & BLUMBERG, R. S. 2009. Immune and non-immune functions of the (not so) neonatal Fc receptor, FcRn. *Semin Immunopathol,* 31**,** 223-36.

BASU, S., HODGSON, G., KATZ, M. & DUNN, A. R. 2002. Evaluation of role of G-CSF in the production, survival, and release of neutrophils from bone marrow into circulation. *Blood,* 100**,** 854-61.

BEEKMAN, R. & TOUW, I. P. 2010. G-CSF and its receptor in myeloid malignancy. *Blood,* 115**,** 5131-6.

BEEKMAN, R., VALKHOF, M. G., SANDERS, M. A., VAN STRIEN, P. M., HAANSTRA, J. R., BROEDERS, L., GEERTSMA-KLEINEKOORT, W. M., VEERMAN, A. J., VALK, P. J., VERHAAK, R. G., LOWENBERG, B. & TOUW, I. P. 2012. Sequential gain of mutations in severe congenital neutropenia progressing to acute myeloid leukemia. *Blood,* 119**,** 5071-7.

BENSINGER, W., APPELBAUM, F., ROWLEY, S., STORB, R., SANDERS, J., LILLEBY, K., GOOLEY, T., DEMIRER, T., SCHIFFMAN, K., WEAVER, C. & ET AL. 1995. Factors that influence collection and engraftment of autologous peripheral-blood stem cells. *J Clin Oncol,* 13**,** 2547-55.

BERRIDGE, M. V. & TAN, A. S. 1993. Characterization of the cellular reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT): subcellular localization, substrate dependence, and involvement of mitochondrial electron transport in MTT reduction. *Arch Biochem Biophys,* 303**,** 474-82.

BLOCK, H., MAERTENS, B., SPRIESTERSBACH, A., BRINKER, N., KUBICEK, J., FABIS, R., LABAHN, J. & SCHAFER, F. 2009. Immobilized-metal affinity chromatography (IMAC): a review. *Methods Enzymol,* 463**,** 439-73.

BONEBERG, E. M., HARENG, L., GANTNER, F., WENDEL, A. & HARTUNG, T. 2000. Human monocytes express functional receptors for granulocyte colony-stimulating factor that mediate suppression of monokines and interferon-gamma. *Blood,* 95**,** 270-6.

BRADLEY, T. R. & METCALF, D. 1966. The growth of mouse bone marrow cells in vitro. *Aust J Exp Biol Med Sci,* 44**,** 287-99.

BROXMEYER, H. E., ORSCHELL, C. M., CLAPP, D. W., HANGOC, G., COOPER, S., PLETT, P. A., LILES, W. C., LI, X., GRAHAM-EVANS, B., CAMPBELL, T. B., CALANDRA, G., BRIDGER, G., DALE, D. C. & SROUR, E. F. 2005. Rapid mobilization of murine and human hematopoietic stem and progenitor cells with AMD3100, a CXCR4 antagonist. *J Exp Med,* 201**,** 1307-18.

BUSSOLINO, F., WANG, J. M., DEFILIPPI, P., TURRINI, F., SANAVIO, F., EDGELL, C. J., AGLIETTA, M., ARESE, P. & MANTOVANI, A. 1989. Granulocyte- and granulocyte-macrophage-colony stimulating factors induce human endothelial cells to migrate and proliferate. *Nature,* 337**,** 471-3.

BUTLER, M. & SPEARMAN, M. 2014. The choice of mammalian cell host and possibilities for glycosylation engineering. *Curr Opin Biotechnol,* 30**,** 107-12.

BYRNE, B., DONOHOE, G. G. & O'KENNEDY, R. 2007. Sialic acids: carbohydrate moieties that influence the biological and physical properties of biopharmaceutical proteins and living cells. *Drug Discov Today,* 12**,** 319-26.

CALICETI, P. & VERONESE, F. M. 2003. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates. *Adv Drug Deliv Rev,* 55**,** 1261-77.

CHAUDHURY, C., BROOKS, C. L., CARTER, D. C., ROBINSON, J. M. & ANDERSON, C. L. 2006. Albumin binding to FcRn: distinct from the FcRn-IgG interaction. *Biochemistry,* 45**,** 4983-90.

CHAUDHURY, C., MEHNAZ, S., ROBINSON, J. M., HAYTON, W. L., PEARL, D. K., ROOPENIAN, D. C. & ANDERSON, C. L. 2003. The major histocompatibility complex-related Fc receptor for IgG (FcRn) binds albumin and prolongs its lifespan. *J Exp Med,* 197**,** 315-22.

CHEERS, C., HAIGH, A. M., KELSO, A., METCALF, D., STANLEY, E. R. & YOUNG, A. M. 1988. Production of colony-stimulating factors (CSFs) during infection: separate determinations of macrophage-, granulocyte-, granulocyte-macrophage-, and multi-CSFs. *Infect Immun,* 56**,** 247-51.

CHUNG, H. K., KIM, S. W., BYUN, S. J., KO, E. M., CHUNG, H. J., WOO, J. S., YOO, J. G., LEE, H. C., YANG, B. C., KWON, M., PARK, S. B., PARK, J. K. & KIM, K. W. 2011. Enhanced biological effects of Phe140Asn, a novel human granulocyte colony-stimulating factor mutant, on HL60 cells. *BMB Rep,* 44**,** 686-91.

COOPER, K. L., MADAN, J., WHYTE, S., STEVENSON, M. D. & AKEHURST, R. L. 2011. Granulocyte colony-stimulating factors for febrile neutropenia prophylaxis following chemotherapy: systematic review and meta-analysis. *BMC Cancer,* 11**,** 404.

CORY, A. H., OWEN, T. C., BARLTROP, J. A. & CORY, J. G. 1991. Use of an aqueous soluble tetrazolium/formazan assay for cell growth assays in culture. *Cancer Commun,* 3**,** 207-12.

COX, G. N., CHLIPALA, E. A., SMITH, D. J., CARLSON, S. J., BELL, S. J. & DOHERTY, D. H. 2014. Hematopoietic properties of granulocyte colony-stimulating factor/immunoglobulin (G-CSF/IgG-Fc) fusion proteins in normal and neutropenic rodents. *PLoS One,* 9**,** e91990.

COX, G. N., SMITH, D. J., CARLSON, S. J., BENDELE, A. M., CHLIPALA, E. A. & DOHERTY, D. H. 2004. Enhanced circulating half-life and hematopoietic properties of a human granulocyte colony-stimulating factor/immunoglobulin fusion protein. *Exp Hematol,* 32**,** 441-9.

CURRAN, M. P. & GOA, K. L. 2002. Pegfilgrastim. *Drugs,* 62**,** 1207-1213.

D'SOUZA, A., JAIYESIMI, I., TRAINOR, L. & VENUTURUMILI, P. 2008. Granulocyte colony-stimulating factor administration: adverse events. *Transfus Med Rev,* 22**,** 280-90.

DAGDAS, S., OZET, G., ALANOGLU, G., AYLI, M., GOKMEN AKOZ, A. & EREKUL, S. 2006. Unusual extramedullary hematopoiesis in a patient receiving granulocyte colony-stimulating factor. *Acta Haematol,* 116**,** 198-202.

DALL'ACQUA, W. F., WOODS, R. M., WARD, E. S., PALASZYNSKI, S. R., PATEL, N. K., BREWAH, Y. A., WU, H., KIENER, P. A. & LANGERMANN, S. 2002. Increasing the affinity of a human IgG1 for the neonatal Fc receptor: biological consequences. *J Immunol,* 169**,** 5171-80.

DAMIANI, R., OLIVEIRA, J. E., VORAUER-UHL, K., PERONI, C. N., VIANNA, E. G., BARTOLINI, P. & RIBELA, M. T. 2009. Stable expression of a human-like sialylated recombinant thyrotropin in a Chinese hamster ovary

cell line expressing alpha2,6-sialyltransferase. *Protein Expr Purif,* 67**,** 7-14.

DELGADO, C., FRANCIS, G. E. & FISHER, D. 1992. The uses and properties of PEG-linked proteins. *Crit Rev Ther Drug Carrier Syst,* 9**,** 249-304.

DENNIS, M. S., ZHANG, M., MENG, Y. G., KADKHODAYAN, M., KIRCHHOFER, D., COMBS, D. & DAMICO, L. A. 2002. Albumin binding as a general strategy for improving the pharmacokinetics of proteins. *J Biol Chem,* 277**,** 35035-43.

DUMONT, J. A., LOW, S. C., PETERS, R. T. & BITONTI, A. J. 2006. Monomeric Fc fusions: impact on pharmacokinetic and biological activity of protein therapeutics. *BioDrugs,* 20**,** 151-60.

ELLIOTT, S., EGRIE, J., BROWNE, J., LORENZINI, T., BUSSE, L., ROGERS, N. & PONTING, I. 2004. Control of rHuEPO biological activity: the role of carbohydrate. *Exp Hematol,* 32**,** 1146-55.

ELLIOTT, S., LORENZINI, T., ASHER, S., AOKI, K., BRANKOW, D., BUCK, L., BUSSE, L., CHANG, D., FULLER, J., GRANT, J., HERNDAY, N., HOKUM, M., HU, S., KNUDTEN, A., LEVIN, N., KOMOROWSKI, R., MARTIN, F., NAVARRO, R., OSSLUND, T., ROGERS, G., ROGERS, N., TRAIL, G. & EGRIE, J. 2003. Enhancement of therapeutic protein in vivo activities through glycoengineering. *Nat Biotechnol,* 21**,** 414-21.

FAVREAU, A. J. & SATHYANARAYANA, P. 2012. miR-590-5p, miR-219-5p, miR-15b and miR-628-5p are commonly regulated by IL-3, GM-CSF and G-CSF in acute myeloid leukemia. *Leuk Res,* 36**,** 334-41.

FOSSIEZ, F., DJOSSOU, O., CHOMARAT, P., FLORES-ROMO, L., AIT-YAHIA, S., MAAT, C., PIN, J. J., GARRONE, P., GARCIA, E., SAELAND, S., BLANCHARD, D., GAILLARD, C., DAS MAHAPATRA, B., ROUVIER, E., GOLSTEIN, P., BANCHEREAU, J. & LEBECQUE, S. 1996. T cell interleukin-17 induces stromal cells to produce proinflammatory and hematopoietic cytokines. *J Exp Med,* 183**,** 2593-603.

FUKUNAGA, R., ISHIZAKA-IKEDA, E. & NAGATA, S. 1990a. Purification and characterization of the receptor for murine granulocyte colony-stimulating factor. *J Biol Chem,* 265**,** 14008-15.

FUKUNAGA, R., SETO, Y., MIZUSHIMA, S. & NAGATA, S. 1990b. 3 DIFFERENT MESSENGER-RNAS ENCODING HUMAN GRANULOCYTE COLONY-STIMULATING FACTOR RECEPTOR. *Proceedings of the National Academy of Sciences of the United States of America,* 87**,** 8702-8706.

FURMENTO, V. A., MARINO, J., BLANK, V. C. & ROGUIN, L. P. 2014. The granulocyte colony-stimulating factor (G-CSF) upregulates metalloproteinase-2 and VEGF through PI3K/Akt and Erk1/2 activation in human trophoblast Swan 71 cells. *Placenta,* 35**,** 937-46.

GABERC-POREKAR, V., ZORE, I., PODOBNIK, B. & MENART, V. 2008. Obstacles and pitfalls in the PEGylation of therapeutic proteins. *Curr Opin Drug Discov Devel,* 11**,** 242-50.

GAERTNER, H. F. & OFFORD, R. E. 1996. Site-specific attachment of functionalized poly(ethylene glycol) to the amino terminus of proteins. *Bioconjug Chem,* 7**,** 38-44.

GASCON, P. 2012. Presently available biosimilars in hematology-oncology: G-CSF. *Target Oncol,* 7 Suppl 1**,** S29-34.

GHETIE, V., HUBBARD, J. G., KIM, J. K., TSEN, M. F., LEE, Y. & WARD, E. S. 1996. Abnormally short serum half-lives of IgG in beta 2-microglobulin-deficient mice. *Eur J Immunol,* 26**,** 690-6.

HAMIDI, M., AZADI, A. & RAFIEI, P. 2006. Pharmacokinetic consequences of pegylation. *Drug Delivery,* 13**,** 399-409.

HAMILTON, J. A. 2008. Colony-stimulating factors in inflammation and autoimmunity. *Nat Rev Immunol,* 8**,** 533-44.

HARADA, M., QIN, Y., TAKANO, H., MINAMINO, T., ZOU, Y., TOKO, H., OHTSUKA, M., MATSUURA, K., SANO, M., NISHI, J., IWANAGA, K., AKAZAWA, H., KUNIEDA, T., ZHU, W., HASEGAWA, H., KUNISADA, K., NAGAI, T., NAKAYA, H., YAMAUCHI-TAKIHARA, K. & KOMURO, I. 2005. G-CSF prevents cardiac remodeling after myocardial infarction by activating the Jak-Stat pathway in cardiomyocytes. *Nat Med,* 11**,** 305-11.

HARTUNG, T., DOECKE, W. D., BUNDSCHUH, D., FOOTE, M. A., GANTNER, F., HERMANN, C., LENZ, A., MILWEE, S., RICH, B., SIMON, B., VOLK, H. D., VON AULOCK, S. & WENDEL, A. 1999. Effect of filgrastim treatment on inflammatory cytokines and lymphocyte functions. *Clin Pharmacol Ther,* 66**,** 415-24.

HASEGAWA, M. 1993. A thermodynamic model for denaturation of granulocyte colony-stimulating factor: O-linked sugar chain suppresses not the triggering deprotonation but the succeeding denaturation. *Biochim Biophys Acta,* 1203**,** 295-7.

HERMANS, M. H., VAN DE GEIJN, G. J., ANTONISSEN, C., GITS, J., VAN LEEUWEN, D., WARD, A. C. & TOUW, I. P. 2003. Signaling mechanisms coupled to tyrosines in the granulocyte colony-stimulating factor receptor orchestrate G-CSF-induced expansion of myeloid progenitor cells. *Blood,* 101**,** 2584-90.

HIROKAWA, M., LEE, M., MOTEGI, M. & MIURA, A. B. 1996. Reversible renal impairment during leukocytosis induced by G-CSF in non-Hodgkin's lymphoma. *Am J Hematol,* 51**,** 328-9.

HOFFNER, G. & DJIAN, P. 2005. Transglutaminase and diseases of the central nervous system. *Front Biosci,* 10**,** 3078-92.

HOGGATT, J. & PELUS, L. M. 2014. New G-CSF agonists for neutropenia therapy. *Expert Opin Investig Drugs,* 23**,** 21-35.

HÖLIG, K. 2013. G-CSF in Healthy Allogeneic Stem Cell Donors. *Transfusion Medicine and Hemotherapy,* 40**,** 225-235.

HUSTON, J. S., TAI, M. S., MCCARTNEY, J., KECK, P. & OPPERMANN, H. 1993. Antigen recognition and targeted delivery by the single-chain Fv. *Cell Biophys,* 22**,** 189-224.

ICHIKAWA, Y., PLUZNIK, D. H. & SACHS, L. 1966. In vitro control of the development of macrophage and granulocyte colonies. *Proc Natl Acad Sci U S A,* 56**,** 488-95.

INOKUCHI, R., MANABE, H., OHTA, F., NAKAMURA, K., NAKAJIMA, S. & YAHAGI, N. 2015. Granulocyte colony-stimulating factor-producing lung cancer and acute respiratory distress syndrome. *Clin Respir J,* 9**,** 250-2.

ISRAEL, E. J., WILSKER, D. F., HAYES, K. C., SCHOENFELD, D. & SIMISTER, N. E. 1996. Increased clearance of IgG in mice that lack beta 2-

microglobulin: possible protective role of FcRn. *Immunology,* 89**,** 573-8.

JAIN, A. & JAIN, S. K. 2008. PEGylation: an approach for drug delivery. A review. *Crit Rev Ther Drug Carrier Syst,* 25**,** 403-47.

JIANG, Y., JIANG, W., QIU, Y. & DAI, W. 2011. Effect of a structurally modified human granulocyte colony stimulating factor, G-CSFa, on leukopenia in mice and monkeys. *J Hematol Oncol,* 4**,** 28.

KASTURI, L., ESHLEMAN, J. R., WUNNER, W. H. & SHAKIN-ESHLEMAN, S. H. 1995. The hydroxy amino acid in an Asn-X-Ser/Thr sequon can influence N-linked core glycosylation efficiency and the level of expression of a cell surface glycoprotein. *J Biol Chem,* 270**,** 14756-61.

KAWAKAMI, M., TSUTSUMI, H., KUMAKAWA, T., ABE, H., HIRAI, M., KUROSAWA, S., MORI, M. & FUKUSHIMA, M. 1990. Levels of serum granulocyte colony-stimulating factor in patients with infections. *Blood,* 76**,** 1962-4.

KIM, P. J., LEE, D. Y. & JEONG, H. 2009. Centralized modularity of N-linked glycosylation pathways in mammalian cells. *PLoS One,* 4**,** e7317.

KINSTLER, O. B., BREMS, D. N., LAUREN, S. L., PAIGE, A. G., HAMBURGER, J. B. & TREUHEIT, M. J. 1996. Characterization and stability of N-terminally PEGylated rhG-CSF. *Pharm Res,* 13**,** 996-1002.

KLEIN, J., LOBEL, L., POLLAK, S., FERIN, M., XIAO, E., SAUER, M. & LUSTBADER, J. W. 2002. Pharmacokinetics and pharmacodynamics of single-chain recombinant human follicle-stimulating hormone containing the human chorionic gonadotropin carboxyterminal peptide in the rhesus monkey. *Fertil Steril,* 77**,** 1248-55.

KORNFELD, R. & KORNFELD, S. 1985. Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem,* 54**,** 631-64.

KREISIG, T., PRASSE, A. A., ZSCHARNACK, K., VOLKE, D. & ZUCHNER, T. 2014. His-tag protein monitoring by a fast mix-and-measure immunoassay. *Sci Rep,* 4**,** 5613.

LAMBERTINI, M., DEL MASTRO, L., BELLODI, A. & PRONZATO, P. 2014. The five "Ws" for bone pain due to the administration of granulocyte-colony stimulating factors (G-CSFs). *Crit Rev Oncol Hematol,* 89**,** 112-28.

LANGE, B., VALTIERI, M., SANTOLI, D., CARACCIOLO, D., MAVILIO, F., GEMPERLEIN, I., GRIFFIN, C., EMANUEL, B., FINAN, J., NOWELL, P. & ET AL. 1987. Growth factor requirements of childhood acute leukemia: establishment of GM-CSF-dependent cell lines. *Blood,* 70**,** 192-9.

LANGRISH, C. L., CHEN, Y., BLUMENSCHEIN, W. M., MATTSON, J., BASHAM, B., SEDGWICK, J. D., MCCLANAHAN, T., KASTELEIN, R. A. & CUA, D. J. 2005. IL-23 drives a pathogenic T cell population that induces autoimmune inflammation. *J Exp Med,* 201**,** 233-40.

LARSEN, A., DAVIS, T., CURTIS, B. M., GIMPEL, S., SIMS, J. E., COSMAN, D., PARK, L., SORENSEN, E., MARCH, C. J. & SMITH, C. A. 1990. Expression cloning of a human granulocyte colony-stimulating factor receptor: a structural mosaic of hematopoietin receptor, immunoglobulin, and fibronectin domains. *J Exp Med,* 172**,** 1559-70.

LI, H. & D'ANJOU, M. 2009. Pharmacological significance of glycosylation in therapeutic proteins. *Curr Opin Biotechnol,* 20**,** 678-84.

LIESCHKE, G. J., GRAIL, D., HODGSON, G., METCALF, D., STANLEY, E., CHEERS, C., FOWLER, K. J., BASU, S., ZHAN, Y. F. & DUNN, A. R. 1994. Mice lacking granulocyte colony-stimulating factor have chronic neutropenia, granulocyte and macrophage progenitor cell deficiency, and impaired neutrophil mobilization. *Blood,* 84**,** 1737-46.

LIONGUE, C. & WARD, A. C. 2014. Granulocyte colony-stimulating factor receptor mutations in myeloid malignancy. *Front Oncol,* 4**,** 93.

LITAM, P. P., FRIEDMAN, H. D. & LOUGHRAN, T. P., JR. 1993. Splenic extramedullary hematopoiesis in a patient receiving intermittently administered granulocyte colony-stimulating factor. *Ann Intern Med,* 118**,** 954-5.

LIU, F., WU, H. Y., WESSELSCHMIDT, R., KORNAGA, T. & LINK, D. C. 1996. Impaired production and increased apoptosis of neutrophils in granulocyte colony-stimulating factor receptor-deficient mice. *Immunity,* 5**,** 491-501.

LLAMAS-VELASCO, M., GARCIA-MARTIN, P., SANCHEZ-PEREZ, J., FRAGA, J. & GARCIA-DIEZ, A. 2013. Sweet's syndrome with subcutaneous involvement associated with pegfilgrastim treatment: first reported case. *J Cutan Pathol,* 40**,** 46-9.

LORD, B. I., MOLINEUX, G., POJDA, Z., SOUZA, L. M., MERMOD, J. J. & DEXTER, T. M. 1991. Myeloid cell kinetics in mice treated with recombinant interleukin-3, granulocyte colony-stimulating factor (CSF), or granulocyte-macrophage CSF in vivo. *Blood,* 77**,** 2154-9.

LORD, B. I., WOOLFORD, L. B. & MOLINEUX, G. 2001. Kinetics of neutrophil production in normal and neutropenic animals during the response to filgrastim (r-metHu G-CSF) or filgrastim SD/01 (PEG-r-metHu G-CSF). *Clin Cancer Res,* 7**,** 2085-90.

LÖWENBERG, B., VAN PUTTEN, W., THEOBALD, M., GMÜR, J., VERDONCK, L., SONNEVELD, P., FEY, M., SCHOUTEN, H., DE GREEF, G., FERRANT, A., KOVACSOVICS, T., GRATWOHL, A., DAENEN, S., HUIJGENS, P. & BOOGAERTS, M. 2003. Effect of Priming with Granulocyte Colony-Stimulating Factor on the Outcome of Chemotherapy for Acute Myeloid Leukemia. *New England Journal of Medicine,* 349**,** 743-752.

MAHMOOD, I. & GREEN, M. D. 2005. Pharmacokinetic and pharmacodynamic considerations in the development of therapeutic proteins. *Clin Pharmacokinet,* 44**,** 331-47.

MANZ, M. G., MIYAMOTO, T., AKASHI, K. & WEISSMAN, I. L. 2002. Prospective isolation of human clonogenic common myeloid progenitors. *Proc Natl Acad Sci U S A,* 99**,** 11872-7.

MARINARO, J. A., CASLEY, D. J. & BACH, L. A. 2000. O-glycosylation delays the clearance of human IGF-binding protein-6 from the circulation. *Eur J Endocrinol,* 142**,** 512-6.

MARTINO, M., LASZLO, D. & LANZA, F. 2014. Long-active granulocyte colony-stimulating factor for peripheral blood hematopoietic progenitor cell mobilization. *Expert Opin Biol Ther,* 14**,** 757-72.

MCCRACKEN, S., LAYTON, J. E., SHORTER, S. C., STARKEY, P. M., BARLOW, D. H. & MARDON, H. J. 1996. Expression of granulocyte-colony

stimulating factor and its receptor is regulated during the development of the human placenta. *J Endocrinol,* 149**,** 249-58.
METCALF, D. 2010. The colony-stimulating factors and cancer. *Nat Rev Cancer,* 10**,** 425-34.
MESHKIBAF, S. 2015. Granulocyte colony-stimulating factor: Its role in gut-homing macrophage generation and colitis, and production by probiotics. *Electronic Thesis and Dissertation Repository.* Paper 2867.
MILLA, P., DOSIO, F. & CATTEL, L. 2012. PEGylation of proteins and liposomes: a powerful and flexible strategy to improve the drug delivery. *Curr Drug Metab,* 13**,** 105-19.
MOLINEUX, G. 2004. The design and development of pegfilgrastim (PEG-rmetHuG-CSF, neulasta((R))). *Current Pharmaceutical Design,* 10**,** 1235-1244.
MOLINEUX G., FOOTE, M., ARVEDSON, T. 2012. Twenty years of G-CSF: clinical and nonclinical discoveries*. Springer Science & Business Media.*
MORIKAWA, K., MORIKAWA, S., NAKAMURA, M. & MIYAWAKI, T. 2002. Characterization of granulocyte colony-stimulating factor receptor expressed on human lymphocytes. *Br J Haematol,* 118**,** 296-304.
MORRIS, K. T., KHAN, H., AHMAD, A., WESTON, L. L., NOFCHISSEY, R. A., PINCHUK, I. V. & BESWICK, E. J. 2014. G-CSF and G-CSFR are highly expressed in human gastric and colon cancers and promote carcinoma cell proliferation and migration. *Br J Cancer,* 110**,** 1211-20.
NAGATA, S., TSUCHIYA, M., ASANO, S., KAZIRO, Y., YAMAZAKI, T., YAMAMOTO, O., HIRATA, Y., KUBOTA, N., OHEDA, M., NOMURA, H. & ET AL. 1986. Molecular cloning and expression of cDNA for human granulocyte colony-stimulating factor. *Nature,* 319**,** 415-8.
NATALELLO, A., AMI, D., COLLINI, M., D'ALFONSO, L., CHIRICO, G., TONON, G., SCARAMUZZA, S., SCHREPFER, R. & DOGLIA, S. M. 2012. Biophysical characterization of Met-G-CSF: effects of different site-specific mono-pegylations on protein stability and aggregation. *PLoS One,* 7**,** e42511.
NGUYEN-JACKSON, H., PANOPOULOS, A. D., ZHANG, H., LI, H. S. & WATOWICH, S. S. 2010. STAT3 controls the neutrophil migratory response to CXCR2 ligands by direct activation of G-CSF-induced CXCR2 expression and via modulation of CXCR2 signal transduction. *Blood,* 115**,** 3354-63.
NGUYEN-JACKSON, H. T., LI, H. S., ZHANG, H., OHASHI, E. & WATOWICH, S. S. 2012. G-CSF-activated STAT3 enhances production of the chemokine MIP-2 in bone marrow neutrophils. *J Leukoc Biol,* 92**,** 1215-25.
NICOLA, N. A. & METCALF, D. 1984. Binding of the differentiation-inducer, granulocyte-colony-stimulating factor, to responsive but not unresponsive leukemic cell lines. *Proc Natl Acad Sci U S A,* 81**,** 3765-9.
NICOLA, N. A., METCALF, D., MATSUMOTO, M. & JOHNSON, G. R. 1983. Purification of a factor inducing differentiation in murine myelomonocytic leukemia cells. Identification as granulocyte colony-stimulating factor. *J Biol Chem,* 258**,** 9017-23.
NIMUBONA, S., GRULOIS, I., BERNARD, M., DRENOU, B., GODARD, M., FAUCHET, R. & LAMY, T. 2002. Complete remission in hypoplastic acute myeloid leukemia induced by G-CSF without chemotherapy: report on three cases. *Leukemia,* 16**,** 1871-3.

NOMIYAMA, J., SHINOHARA, K. & INOUE, H. 1994. Erythema nodosum caused by the administration of granulocyte colony-stimulating factor in a patient with refractory anemia. *Am J Hematol,* 47**,** 333.

O'MALLEY, D. P., WHALEN, M. & BANKS, P. M. 2003. Spontaneous splenic rupture with fatal outcome following G-CSF administration for myelodysplastic syndrome. *Am J Hematol,* 73**,** 294-5.

OBER, R. J., MARTINEZ, C., LAI, X., ZHOU, J. & WARD, E. S. 2004. Exocytosis of IgG as mediated by the receptor, FcRn: an analysis at the single-molecule level. *Proc Natl Acad Sci U S A,* 101**,** 11076-81.

PANOPOULOS, A. D., BARTOS, D., ZHANG, L. & WATOWICH, S. S. 2002. Control of myeloid-specific integrin alpha Mbeta 2 (CD11b/CD18) expression by cytokines is regulated by Stat3-dependent activation of PU.1. *J Biol Chem,* 277**,** 19001-7.

PANOPOULOS, A. D. & WATOWICH, S. S. 2008. Granulocyte colony-stimulating factor: molecular mechanisms of action during steady state and 'emergency' hematopoiesis. *Cytokine,* 42**,** 277-88.

PATEL, A., CHOLKAR, K. & MITRA, A. K. 2014. Recent developments in protein and peptide parenteral delivery approaches. *Ther Deliv,* 5**,** 337-65.

PELUS, L. M., HOROWITZ, D., COOPER, S. C. & KING, A. G. 2002. Peripheral blood stem cell mobilization. A role for CXC chemokines. *Crit Rev Oncol Hematol,* 43**,** 257-75.

PERLMAN, S., VAN DEN HAZEL, B., CHRISTIANSEN, J., GRAM-NIELSEN, S., JEPPESEN, C. B., ANDERSEN, K. V., HALKIER, T., OKKELS, S. & SCHAMBYE, H. T. 2003. Glycosylation of an N-terminal extension prolongs the half-life and increases the in vivo activity of follicle stimulating hormone. *J Clin Endocrinol Metab,* 88**,** 3227-35.

PERUTZ, M. F. 1995. Glutamine repeats as polar zippers: their role in inherited neurodegenerative disease. *Mol Med,* 1**,** 718-21.

PICARD, V., ERSDAL-BADJU, E. & BOCK, S. C. 1995. Partial glycosylation of antithrombin III asparagine-135 is caused by the serine in the third position of its N-glycosylation consensus sequence and is responsible for production of the beta-antithrombin III isoform with enhanced heparin affinity. *Biochemistry,* 34**,** 8433-40.

PISAL, D. S., KOSLOSKI, M. P. & BALU-IYER, S. V. 2010. Delivery of therapeutic proteins. *Journal of Pharmaceutical Sciences,* 99**,** 2557-2575.

PRENDIVILLE, J., THIESSEN, P. & MALLORY, S. B. 2001. Neutrophilic dermatoses in two children with idiopathic neutropenia: association with granulocyte colony-stimulating factor (G-CSF) therapy. *Pediatr Dermatol,* 18**,** 417-21.

PUSIC, I. & DIPERSIO, J. F. 2008. The use of growth factors in hematopoietic stem cell transplantation. *Current Pharmaceutical Design,* 14**,** 1950-1961.

QIU, H., BELANGER, A., YOON, H. W. & BUNN, H. F. 1998. Homodimerization restores biological activity to an inactive erythropoietin mutant. *J Biol Chem,* 273**,** 11173-6.

ROBERTS, A. W. 2005. G-CSF: a key regulator of neutrophil production, but that's not all! *Growth Factors,* 23**,** 33-41.

SAINT-JORE-DUPAS, C., FAYE, L. & GOMORD, V. 2007. From planta to pharma with glycosylation in the toolbox. *Trends Biotechnol,* 25**,** 317-23.

SAMPSON, M., ZHU, Q. S. & COREY, S. J. 2007. Src kinases in G-CSF receptor signaling. *Front Biosci,* 12**,** 1463-74.

SCHNEIDER, A., KRUGER, C., STEIGLEDER, T., WEBER, D., PITZER, C., LAAGE, R., ARONOWSKI, J., MAURER, M. H., GASSLER, N., MIER, W., HASSELBLATT, M., KOLLMAR, R., SCHWAB, S., SOMMER, C., BACH, A., KUHN, H. G. & SCHABITZ, W. R. 2005. The hematopoietic factor G-CSF is a neuronal ligand that counteracts programmed cell death and drives neurogenesis. *J Clin Invest,* 115**,** 2083-98.

SEMERAD, C. L., LIU, F., GREGORY, A. D., STUMPF, K. & LINK, D. C. 2002. G-CSF is an essential regulator of neutrophil trafficking from the bone marrow to the blood. *Immunity,* 17**,** 413-23.

SHAH, J. & WELSH, S. J. 2014. The clinical use of granulocyte-colony stimulating factor. *Br J Hosp Med (Lond),* 75**,** C29-32.

SINCLAIR, A. M. & ELLIOTT, S. 2005. Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. *J Pharm Sci,* 94**,** 1626-35.

SIVAKUMAR, R., ATKINSON, M. A., MATHEWS, C. E. & MORE, L. 2015. G-CSF: A Friend or Foe? *Immunome Research,* 2015.

SOLA, R. J. & GRIEBENOW, K. 2009. Effects of glycosylation on the stability of protein pharmaceuticals. *J Pharm Sci,* 98**,** 1223-45.

SOLA, R. J. & GRIEBENOW, K. 2010. Glycosylation of therapeutic proteins: an effective strategy to optimize efficacy. *BioDrugs,* 24**,** 9-21.

SOLA, R. J., RODRIGUEZ-MARTINEZ, J. A. & GRIEBENOW, K. 2007. Modulation of protein biophysical properties by chemical glycosylation: biochemical insights and biomedical implications. *Cell Mol Life Sci,* 64**,** 2133-52.

SOUZA, L. M., BOONE, T. C., GABRILOVE, J., LAI, P. H., ZSEBO, K. M., MURDOCK, D. C., CHAZIN, V. R., BRUSZEWSKI, J., LU, H., CHEN, K. K. & ET AL. 1986. Recombinant human granulocyte colony-stimulating factor: effects on normal and leukemic myeloid cells. *Science,* 232**,** 61-5.

SPIRO, R. G. 2002. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology,* 12**,** 43R-56R.

STORK, R., ZETTLITZ, K. A., MULLER, D., RETHER, M., HANISCH, F. G. & KONTERMANN, R. E. 2008. N-glycosylation as novel strategy to improve pharmacokinetic properties of bispecific single-chain diabodies. *J Biol Chem,* 283**,** 7804-12.

SYTKOWSKI, A. J., LUNN, E. D., RISINGER, M. A. & DAVIS, K. L. 1999. An erythropoietin fusion protein comprised of identical repeating domains exhibits enhanced biological properties. *J Biol Chem,* 274**,** 24773-8.

TAMADA, T., HONJO, E., MAEDA, Y., OKAMOTO, T., ISHIBASHI, M., TOKUNAGA, M. & KUROKI, R. 2006. Homodimeric cross-over structure of the human granulocyte colony-stimulating factor (GCSF) receptor signaling complex. *Proc Natl Acad Sci U S A,* 103**,** 3135-40.

TANAKA, H., SATAKE-ISHIKAWA, R., ISHIKAWA, M., MATSUKI, S. & ASANO, K. 1991. Pharmacokinetics of recombinant human granulocyte colony-stimulating factor conjugated to polyethylene glycol in rats. *Cancer Res,* 51**,** 3710-4.

TANG, L., PERSKY, A. M., HOCHHAUS, G. & MEIBOHM, B. 2004. Pharmacokinetic aspects of biotechnology products. *J Pharm Sci,* 93**,** 2184-204.

TANHEHCO, Y. C., ADAMSKI, J., SELL, M., CUNNINGHAM, K., EISENMANN, C., MAGEE, D., STADTMAUER, E. A. & O'DOHERTY, U. 2010. Plerixafor mobilization leads to a lower ratio of CD34+ cells to total nucleated cells which results in greater storage costs. *Journal of Clinical Apheresis,* 25**,** 202-208.

TOUW, I. P. & VAN DE GEIJN, G. J. 2007. Granulocyte colony-stimulating factor and its receptor in normal myeloid cell development, leukemia and related blood cell disorders. *Front Biosci,* 12**,** 800-15.

TRAINER, P. J., DRAKE, W. M., KATZNELSON, L., FREDA, P. U., HERMAN-BONERT, V., VAN DER LELY, A. J., DIMARAKI, E. V., STEWART, P. M., FRIEND, K. E., VANCE, M. L., BESSER, G. M., THORNER, M. O., PARKINSON, C., KLIBANSKI, A., POWELL, J. S., BARKAN, A. L., SHEPPARD, M. C., MALDONADO, M., ROSE, D. R., CLEMMONS, D. R., JOHANNSSON, G., BENGT-ÅKE BENGTSSON, B.-Å., STAVROU, S., KLEINBERG, D. L., COOK, D. M., PHILLIPS, L. S., BIDLINGMAIER, M., STRASBURGER, C. J., HACKETT, S., ZIB, K., BENNETT, W. F., DAVIS, R. J. & SCARLETT, J. A. 2000. Treatment of Acromegaly with the Growth Hormone–Receptor Antagonist Pegvisomant. *New England Journal of Medicine,* 342**,** 1171-1177.

ULICH, T. R., DEL CASTILLO, J. & SOUZA, L. 1988. Kinetics and mechanisms of recombinant human granulocyte-colony stimulating factor-induced neutrophilia. *Am J Pathol,* 133**,** 630-8.

VALTIERI, M., BOCCOLI, G., TESTA, U., BARLETTA, C. & PESCHLE, C. 1991. Two-step differentiation of AML-193 leukemic line: terminal maturation is induced by positive interaction of retinoic acid with granulocyte colony-stimulating factor (CSF) and vitamin D3 with monocyte CSF. *Blood,* 77**,** 1804-12.

VARDIMAN, J. W., THIELE, J., ARBER, D. A., BRUNNING, R. D., BOROWITZ, M. J., PORWIT, A., HARRIS, N. L., LE BEAU, M. M., HELLSTROM-LINDBERG, E., TEFFERI, A. & BLOOMFIELD, C. D. 2009. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood,* 114**,** 937-51.

VARKI, A. 2008. Sialic acids in human health and disease. *Trends Mol Med,* 14**,** 351-60.

WADA, H., YOSHIDA, S., ISHIBASHI, F., MIZOBUCHI, T., MORIYA, Y., HOSHINO, H., OKAMOTO, T., SUZUKI, M., SHIBUYA, K. & YOSHINO, I. 2011. Pleomorphic carcinoma of the lung producing granulocyte colony-stimulating factor: report of a case. *Surg Today,* 41**,** 1161-5.

WALKER, A. K., SOO, K. Y., LEVINA, V., TALBO, G. H. & ATKIN, J. D. 2013. N-linked glycosylation modulates dimerization of protein disulfide isomerase family A member 2 (PDIA2). *FEBS J,* 280**,** 233-43.

WEENEN, C., PENA, J. E., POLLAK, S. V., KLEIN, J., LOBEL, L., TROUSDALE, R. K., PALMER, S., LUSTBADER, E. G., OGDEN, R. T. & LUSTBADER, J. W. 2004. Long-acting follicle-stimulating hormone analogs containing N-linked glycosylation exhibited increased bioactivity compared with o-linked analogs in female rats. *J Clin Endocrinol Metab,* 89**,** 5204-12.

WELTE, K., PLATZER, E., LU, L., GABRILOVE, J. L., LEVI, E., MERTELSMANN, R. & MOORE, M. A. 1985. Purification and biochemical characterization of human pluripotent hematopoietic colony-stimulating factor. *Proc Natl Acad Sci U S A,* 82**,** 1526-30.

WERNER, J. M., BREEZE, A. L., KARA, B., ROSENBROCK, G., BOYD, J., SOFFE, N. & CAMPBELL, I. D. 1994. Secondary structure and backbone dynamics of human granulocyte colony-stimulating factor in solution. *Biochemistry,* 33**,** 7184-92.

WHITE, S. M., ALARCON, M. H. & TWEARDY, D. J. 2000. Inhibition of granulocyte colony-stimulating factor-mediated myeloid maturation by low level expression of the differentiation-defective class IV granulocyte colony-stimulating factor receptor isoform. *Blood,* 95**,** 3335-40.

WHITE, S. M., BALL, E. D., EHMANN, W. C., RAO, A. S. & TWEARDY, D. J. 1998. Increased expression of the differentiation-defective granulocyte colony-stimulating factor receptor mRNA isoform in acute myelogenous leukemia. *Leukemia,* 12**,** 899-906.

WILKINSON, I. R., FERRANDIS, E., ARTYMIUK, P. J., TEILLOT, M., SOULARD, C., TOUVAY, C., PRADHANANGA, S. L., JUSTICE, S., WU, Z., LEUNG, K. C., STRASBURGER, C. J., SAYERS, J. R. & ROSS, R. J. 2007. A ligand-receptor fusion of growth hormone forms a dimer and is a potent long-acting agonist. *Nat Med,* 13**,** 1108-13.

WONGTRAKUL-KISH, K., KOLARICH, D., PASCOVICI, D., JOSS, J. L., DEANE, E. & PACKER, N. H. 2012. Characterization of N- and O-linked glycosylation changes in milk of the tammar wallaby (Macropus eugenii) over lactation. *Glycoconj J*.

WURM, F. M. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol,* 22**,** 1393-8.

YAMAGUCHI, T., MIYATA, Y., HAYAMIZU, K., HASHIZUME, J., MATSUMOTO, T., TASHIRO, H. & OHDAN, H. 2012. Preventive effect of G-CSF on acute lung injury via alveolar macrophage regulation. *J Surg Res,* 178**,** 378-84.

YE, P., RODRIGUEZ, F. H., KANALY, S., STOCKING, K. L., SCHURR, J., SCHWARZENBERGER, P., OLIVER, P., HUANG, W., ZHANG, P., ZHANG, J., SHELLITO, J. E., BAGBY, G. J., NELSON, S., CHARRIER, K., PESCHON, J. J. & KOLLS, J. K. 2001. Requirement of interleukin 17 receptor signaling for lung CXC chemokine and granulocyte colony-stimulating factor expression, neutrophil recruitment, and host defense. *J Exp Med,* 194**,** 519-27.

YOON, S. K., SONG, J. Y. & LEE, G. M. 2003. Effect of low culture temperature on specific productivity, transcription level, and heterogeneity of erythropoietin in Chinese hamster ovary cells. *Biotechnol Bioeng,* 82**,** 289-98.

YOWELL, S. L. & BLACKWELL, S. 2002. Novel effects with polyethylene glycol modified pharmaceuticals. *Cancer treatment reviews,* 28**,** 3-6.
ZHANG, F., LIU, M. R. & WAN, H. T. 2014. Discussion about several potential drawbacks of PEGylated therapeutic proteins. *Biol Pharm Bull,* 37**,** 335-9.
ZHAO, S., ZHANG, Y., TIAN, H., CHEN, X., CAI, D., YAO, W. & GAO, X. 2013. Extending the serum half-life of G-CSF via fusion with the domain III of human serum albumin. *Biomed Res Int,* 2013**,** 107238.