STRUCTURAL STUDIES OF CYTOKINES AND CYTOKINE RECEPTORS

by

Marc Christian Deller

Laboratory of Molecular Biophysics and The Wellcome Trust Centre for Human Genetics, Oxford

Corpus Christi College, Oxford

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at The University of Oxford

Michaelmas Term 1999

STRUCTURAL STUDIES OF CYTOKINES AND CYTOKINE RECEPTORS

The Laboratory of Molecular Biophysics, The
Wellcome Trust Centre for Human Genetics and
Corpus Christi College, Oxford

by Marc Christian Deller

ABSTRACT

Cytokines are a family of secreted polypeptide factors that function in controlling the survival, growth, differentiation and effector functions of tissue cells. They achieve this through receptor-mediated signal transduction in which extracellular receptor components of target cells bind circulating ligands, triggering a cascade of events that leads to the activation of various intracellular processes. Commonly, protein kinases are activated which modulate cellular proteins and transcriptional activity.

Two members of the gp130 cytokine family were studied; namely leukemia inhibitory factor receptor (LIFR) and oncostatin M (OSM). The use of a variety of eukaryotic expression systems was investigated for the production of soluble extracellular fragments of the ligand-binding domain of LIFR. Preliminary characterisation of LIFR was carried out, and its suitability for structural studies assessed.

A truncated variant of OSM was designed, expressed, characterised and crystallised, as wildtype material proved recalcitrant to crystallisation. Multiple wavelength anomalous dispersion techniques were applied and the atomic structure determined in the spacegroup $P2_12_12_1$ to a resolution of 2.2 Å.

The structure of OSM was compared with other helical cytokines of the growth hormone family and these comparisons, taken together with mutagenesis data, support a mode of receptor-ligand engagement displaying similarities to that of erythropoietin and granulocyte colony stimulating factor. This study presents information that allows further structural analysis of the epitopes involved in ligand binding by the gp130 and LIFR family of receptors.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Firstly, I would like to thank Professors E. Yvonne Jones and John K. Heath for helpful supervision and giving me the opportunity to work with them. I would also like to thank all members of the CRC Growth Factor Group, Birmingham, and members of the Laboratory of Molecular Biophysics and the Wellcome Trust Centre for Human Genetics, Oxford, for general assistance and making the laboratory a great place to work.

In Birmingham, I am particularly indebted to Keith Hudson, Katya Chobotova, Ann Vernalis, Dave Staunton and Mark Hall for introducing me to the wonderous techniques of molecular biology and for supplying all necessary cells and vectors. In Oxford, I am particularly grateful to Professor Louise Johnson for giving me the opportunity to work in such a stimulating laboratory environment. Thanks also extend to Shinji Ikemizu and Jeronimo Bravo for introducing me to the marvels of modern crystallography; without whom my exploration into crystallography would have never begun.

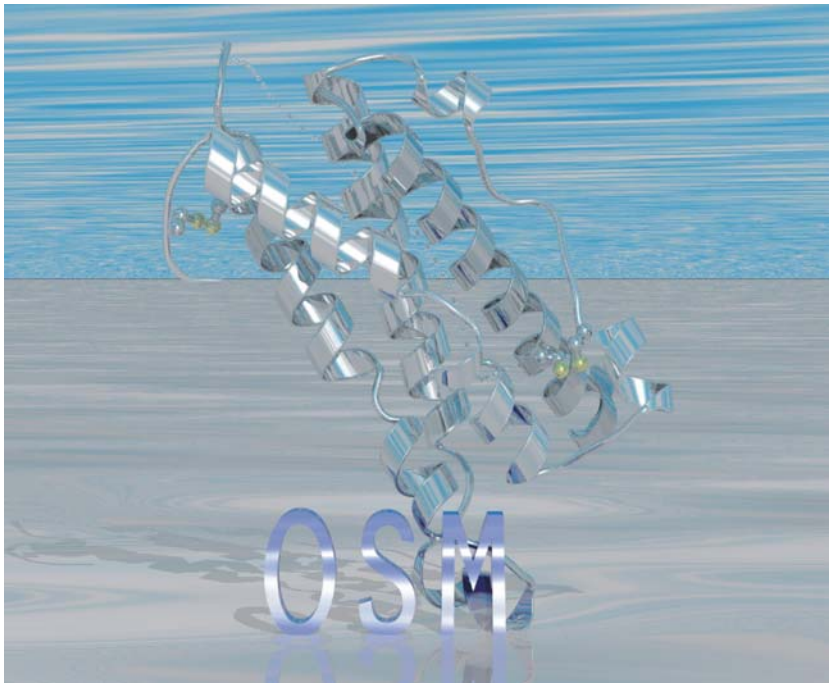
I would also like to thank Jon, Geoff, Thil, Malkit, Katsumi, Nathan, Wei Ching, Chris, Andy, Neil, Matt, Kurt, Amit, Ealish, Nick, Bob, Paul, Ed, Pete, Rao, Ren, Martin, Mary, José and anyone else who has shown interest in this project. I would like to thank Vivian Strojanoff (ESRF, France), Karl Harlos, Elspeth Garman, Dave Stuart and Dave Stammers for assistance with data collection and Richard Bryan, Kathryn Measures and Robert Esnouf for unsurpassed computing facilities. Many thanks also go to Robin Aplin, Dyson Perrins Laboratory, Oxford for persistence with the mass spec analysis. Financial income was supplied by The Medical Research Council and Louise Johnson, for which I am eternally grateful.

I would like to raise a glass to Leo Sayer, Geordie, Bo', Dr Deeps, Magoo, Taff, Nobby, Ballie, Porker, Mop, Sam, Buck, Dirk and 'Big' Kev for being great mates and helping me through the endless 'Engine Trouble'. A big shout also goes to the boy's of Nottingham Forest F.C., Stella Artois, John Player Specials and the Profanisaurus (unabridged version) for entertainment in the darker hours. Special mention must also go to teachers past, present and future, particularly Alison Parker and Gilbert Gimson for believing in me when all others had lost hope.

Finally, I would like to thank Mum, Dad and Andy for their help, support and encouragement over the years...cheers!

"...it was like climbing a mountain and falling in love at the same time."

Max Perutz.



For Mum and Dad

ABBREVIATIONS

293T:	Fibroblast Cell-Line 293T
A_{260}:	Absorbance at 260 nm
A_{280}:	Absorbance at 280 nm
AMP	Ampicillin
ARIA:	Acetylcholine Receptor-Inducing Activity
BDNF:	Brain-Derived Neurotrophic Factor
BMP:	Bone Morphogenic Protein
BAF:	B-cell Activating Factor
BP:	Base Pairs of DNA
BSA:	Bovine Serum Albumin
CHO:	Chinese Hamster Ovary Cells
CNTF/CNTFR:	Ciliary Neurotrophic Factor / Receptor
CSF/CSFR:	Colony-Stimulating Factor / Receptor
CNDF:	Cholinergic Neuronal Differentiation Factor
CCP4:	Collaborative Computational Project No 4
CCD:	Charged Coupled Device
DRF:	Differentiation Retarding Factor
DIF:	Differentiation Inhibiting Factor
D-Factor:	Differentiation Factor
DNA:	Deoxyribonucleic Acid
dNTP:	Deoxyribose Nucleoside Triphosphate
DTT:	Dithiothreitol
EC_{50}/IC_{50}	Effective/Inhibitory Concentration 50 %
EDTA:	Ethylene Diamine Tetra Acetate
EtBr:	Ethidium Bromide
ENA-78:	Epithelial cell-derived Neutrophil Activator
EGF/EGFR:	Epidermal Growth Factor / Receptor
Epo/EpoR:	Erythropoeitin / Receptor
ELISA:	Enzyme Linked ImmunoSorbant Assay
ESRF:	European Synchrotron Radiation Facility
Fc	Fragment Crystallisable
FCS:	Foetal Calf Serum
FFT:	Fast Fourier Transform
FGF:	Fibroblast Growth Factor
FPLC:	Fast Protein Liquid Chromatography
FN-III:	Fibronectin type-III
GRO:	Growth Related cytokine
GGF:	Glial Growth Factor
GS	Glutamine Synthetase
GST:	Glutathione S-Transferase
GCSF/GCSFR:	Granulocyte Colony Stimulating Factor/Receptor
GH/GH-R:	Growth Hormone/Receptor
GMCSF/GMCSFR:	Granulocyte Macrophage Colony Stimulating Factor/Receptor
GCP:	Granulocyte Chemotactic Peptide
GPI:	Glycosyl-Phosphatidylinositol
HCMV	Human Cytomegalovirus
HSF:	Hepatocyte Stimulating Factor
HILDA:	Human Interleukin for Da1 cells
HB-EGF:	Heparin-Binding EGF
HRG:	Heregulin
HGF:	Hepatocyte Growth Factor

HP:	Haemopoietin
HRP:	Horse Radish Peroxidase
Ig:	Immunoglobulin
IP:	Image Plate
IPTG:	Isopropyl β-D-galactopyranoside
IP-10:	Interferon-Inducible Protein-10
I309:	Human T-cell Activation gene-3
IFN:	Interferon
IGF:	Insulin-like Growth Factor
IL:	Interleukin
JAK	Janus Kinase
KIR	Killer Inhibitory Receptor
KGF:	Keratinocyte Growth Factor
LIF/LIFR:	Leukaemia Inhibitory Factor / Receptor
MAD:	Multiple Wavelength Anomalous Dispersion
MAPK	Mitogen Activated Protein Kinase
M-CSF:	Macrophage-CSF
Mϕ:	Macrophage
M-SCF:	Macrophage-Stem Cell Factor
MCP:	Monocyte Chemoattractant Protein
MES:	2-(N-morpholino) Ethane Sulphonic Acid
MIP:	Macrophage Inflammatory Protein
MIR:	Multiple Isomorphous Replacement
M6P:	Mannose-6-Phosphate
MBP:	Maltose Binding Protein
MLPLI:	Melanocyte-Derived Lipoprotein Lipase Inhibitor
MTT:	3-(4-5-dimethylthazol-2-yl)-2-5-diphenyl tetrazolium bromide
MSX	L-Methionine Sulphoximine
NAP:	Neutrophil Activating Peptide
NGF:	Nerve Growth Factor
NT-3/4-5:	Neurotrophin
NK:	Natural Killer
NDF:	Neu Differentiation Factor
OD:	Optical Density
OPD:	O-nitrophenyl Phosphate Disodium Hexahydrate
OSM:	Oncostatin-M
PAGE:	Polyacylamide Gel Electrophoresis
PBS:	Phosphate Buffered Saline
PCR:	Polymerase Chain Reaction
PDB:	Protein Database
PEG:	Polyethylene Glycol
PF:	Platelet Factor
PDGF:	Platelet-Derived Growth Factor
pGEX:	GST Fusion Protein Plasmid
Prl/PrlR:	Prolactin/Receptor
PMSF:	Phenylmethylsulphonyl Fluoride
PNGase F:	Peptide-N^4-(*N*-acetyl-β-D-Glucosaminyl)asparagine amidase
RANTES:	Regulated on Activation Normal T-cell Expressed and Secreted
S2:	Schneider's *Drosophila* Cell line 2
SCDGF:	Schwann Cell-Derived Growth Factor
SDS:	Sodium Dodecyl Sulphate
SIRAS:	Single Isomorphous Replacement With Anomalous Scattering
STAT	Signal Transducer and Activator of Transcription
S/T kinase:	Serine-Threonine Kinase

SCF:	Stem Cell Factor
TF	Tissue Factor
TFA:	Trifluoro Acetic Acid
TNF:	Tumour Necrosis Factor
TGF:	Transforming Growth Factor
TK:	Tyrosine Kinase
UV:	Ultra Violet
VEGF:	Vascular Endothelial Growth Factor
VVGF:	Vaccinia Virus Growth Factor
WWW:	World Wide Web

1. Chapter 1 – Introduction

This thesis concerns structural studies of key proteins of the haematopoetin cytokine system; namely oncostatin M (OSM) and leukemia inhibitory factor receptor (LIFR). The proteins studied form part of a regulatory network that controls the biological and physical state of cells in multicellular organisms. This control is primarily driven by secreted polypeptide factors called growth factors or cytokines that function in intercellular communication between cells. This signal initiates a cascade of events within the cell resulting in the regulation of gene expression or enzyme activity and ultimately leading to cell migration, programmed cell death (apoptosis), cell division or cell differentiation.

1.1 Medical Importance of Cytokines

The response of the target cell to the cytokine can be beneficial to the organism, for example the co-ordination of cell growth during development and control of the inflammatory response for fighting disease and infection. Alternatively, disruption of normal cytokine function can lead to aberrant cell growth as observed in the many cancers that plague society and developmental disorders such as those observed for the fibroblast growth factor family of cytokines (e.g. Aperts syndrome) (Wilkie *et al.*, 1995). As a result, several cytokines have been used as pharmacological agents, for example growth hormone (GH) in the treatment of dwarfism (Larons syndrome) (Laron *et al.*, 1966) and interferon-β (IFNβ) for the treatment of multiple sclerosis. Many more cytokines are being extensively studied as potential therapeutics, including leptin (obesity) (Scott, 1996), erythropoetin (Epo) (anaemia) (Foa, 1991), nerve growth factor (NGF) (Parkinson's disease), tumour necrosis factor (TNF) (cancer) and interleukin-11 (IL-11) (Crohn's disease and thrombocytopoenia) (Gordon, 1996). Modulation of cytokine activity is being widely researched, with a view to enhancement of therapeutic properties. Structural studies,

through the means of nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography, play a pivotal role in this research.

1.2 The Cytokine and Growth Factor Family of Proteins

Cytokines function through receptor-mediated signal transduction in which extracellular signaling proteins produced by one cell are recognised by specific receptors on neighbouring cells. This triggers mechanisms that function in controlling the survival, growth, differentiation and effector functions of tissue cells (Balkwill and Burke, 1989). Cytokines are regulators of haematopoesis and act as coordinators of both immune and inflammatory responses (Arai *et al.*, 1990). Included in the cytokine family of extracellular signaling proteins are colony-stimulating factors, interleukins, lymphokines, monokines, interferons, neurotrophic factors, differentiation factors and chemokines.

The action of this family of protein factors often appears to be paradoxical, as they interact with many different cell types producing pleiotropic effects dependent on the target cell type. However, one cell type often shows the same response to several cytokines and growth factors, so giving rise to a level of redundancy and overlap as observed through genetic analysis and biochemical reconstitution. Gene deletion experiments have revealed that few individual cytokines or growth factors are essential for life or cell survival. This emphasizes the fundamental importance of this protein class, as these vital cellular functions are backed up in a 'fail safe' type mechanism where by one cytokine or growth factor can compensate for the loss of another.

Many different families of cytokines exist and the focus of this study is the group utilising the glycoprotein 130 (gp130) receptor. The gp130 family of cytokines includes leukemia inhibitory factor (LIF), oncostatin M (OSM), interleukin-6 (IL-6), interleukin-11 (IL-11), ciliary neurotrophic factor (CNTF) and cardiotrophin-1 (CT-1). These are amongst the

most pleiotropic cytokines known and function in controlling haematopoesis, neuronal development and the immune system (Cosman, 1993; Nicola, 1994).

1.3 Cytokine or Growth Factor?

In common both cytokines and growth factors act on their target cells by binding cell surface receptors, which then transduce the stimulus to the inside of the cell, so eliciting a physiological response. The biological response is typically mediated through protein phosphorylation events commonly brought about by kinase activity. A cytokine is most commonly defined as a protein factor that binds to a cell surface receptor containing no intrinsic kinase activity. Conversely, the term growth factor is reserved for the subset of extracellular signaling proteins that bind to cell surface receptors containing intrinsic kinase activity. Cytokines can also be defined as a class of extracellular signaling proteins that bind structurally related receptors. These receptors are characterised by a conserved pattern of cysteine residues as well as a conserved WSXWSX motif. Cytokines generally act on target cells in contact with the secreting cell (juxtacrine) or cells in the immediate vicinity (paracrine) of the secreting cell. Several cytokines also function in direct regulation of the secreting cell (autocrine). In contrast, growth factors as exemplified by growth hormone and prolactin have a more classical long-range (endocrine) role.

However, the terms cytokine and growth factor are often used interchangeably; reflecting their similarities and for the purpose of this study no distinction will be made (see Table 1-1 below for comparison).

Property	Growth Factor	Cytokines
Site of Production	Few	Many
Cellular Target	Many	Few
Biological Role	Homeostasis	Fighting Infection Tissue Repair
Biological Redundancy	Low	High
Biological Pleiotropy	Low	High
In the Circulation?	Yes	Rarely
Sphere of Influence	Widespread	Autocrine Paracrine
Inducers	Physiological Variation	External Insults
Receptor Properties	Tyrosine Kinase Activity	No Enzymatic Activity

Table 1-1
Table showing a comparison of cytokines and growth factor properties. Adapted from Nicola, 1994.

1.4 Structural Classification of Cytokines and their Receptors

Despite the overall lack of sequence similarity in most members of the cytokine family, they can be sub-classified according to similarities in biological action, induction mechanisms and three-dimensional structure (Table 1-2).

1.4.1 Structural Classification of Cytokines

Cytokines adopt several distinct structural motifs and since the three-dimensional structure of a protein often reflects its evolutionary origins more accurately than sequence homology, this is a very useful means of classification. Cytokines adopt many diverse folds involving all alpha (α), all beta (β) or alternating mixtures of both alpha separated by beta (α/β) secondary structures (Richardson, 1981). As in other classes of proteins discrete domains of alpha and beta secondary structure (a+b) are uncommon (see Table 1-2 and Figure 1-1)

This distribution demonstrates that no particular fold is favoured for use in cytokine signaling. However, cytokines are generally relatively small, compact and adopt very stable structures; often with protective carbohydrates and intramolecular disulphide bridges. These features enhance solubility, stability and generally increase the half-life through protection from degradation and protease action. These protective features are evolutionarily beneficial considering the vigorous insults that cytokine signaling is often induced by.

Group	**Group 1: 4-α-Helical Bundle Cytokines**							
Structural Class	**Short-Chain**				**Long-Chain**			
Cytokines	IL-2 IL-4 IL-7 IL-9 IL-13	IL-3 GM-CSF IL-5	M-SCF CSF	IFNγ	IL-6 LIF OSM CNTF IL-11 CT-1	Epo G-CSF IL-12α	GH PRL	IL-10 IFNα IFNα IFNτ
Receptor Class	Haemopoietin Domain		Ig-TK II	IFN-R	Haemopoietin Domain			IFN-R
Biological Action	Proliferation/Differentiation of haemopoeitic Cells				Pleiotropic	Mono-Specific	Growth	Immunity
Cell Type	T-Cell B-Cell Mϕ	Neutrophil Eosinophil Mast Cell Mϕ	Stem Cell Mϕ	Many	Liver Fat Nerve Muscle Mϕ	Erythroid Neutrophil T-Cell	Many	Many

Group	**Group 2: β-Sheet Rich Cytokines**					
Structural Class	**Cysteine Knot**			**β-Jelly Roll**	**β-Trefoil**	
Cytokines	TGFβ Activin Inhibin BMP	PDGF-A PDGF-B VEGF PIGF	NGF BDNF NT-3 NT-4/5	TNFα TNFβ CD40L CD27L FASL	IL-1α IL-1β	FGF INT-2 KGF
Receptor Class	S/T Kinase I /II	Ig-TK III	Cys-TK V	TNF-R P75 P55	Ig-Like	Ig-TK IV Heparin
Biological Action	Differentiation Growth Inhibition Tissue Modeling	Growth	Survival Growth Differentiation	Immunity Cell Death	Immun ity Cell Death	Growth
Cell Type	Many	Epithelial Endothelial	Nerve Cells	Many	Many	Mesechymal Cells Neuroectoderm Cells

Group	**Group 3: α/β Cytokines**						
Structural Class	**β-Meander**	**S-S Rich**	**Chemokines**				
			CXC	**CC**	**C**	**CXXXC**	
Cytokines	EGF TGFα β-cellulin SCDGF VVGF Amphiregulin HB-EGF	Insulin IGF-I IGF-II Relaxin Bombyxin	IL-8 GRO PF-4 IP-10 MIP-2 NAP-2 GCP-2	MCP-1 MCP-2 MCP-3 RANTES MIP-1α MIP-1β I309	Lymphotactin	Fractalkine	
Receptor Class	Cys-Rich TK I	Cys-Rich TK II M6P-Receptor	Serpentine seven pass				
Biological Action	Proliferation Wound Healing	Metabolism Proliferation/Differentiation	Chemotaxis Immunity				
Cell Type	Epithelial Cells	Mesenchymal Cells	Neutrophil Macrophage Eosinophil				

Group	**Group 4: Mosaic Cytokines**		
Structural Class	**Ig-EGF-TM-Cyt**	**Kringle-Ser Protease**	**α-helix-haemopoietin**
Cytokines	HRGα/β NDF ARIA	HGF	IL-12
Receptor Class	Cys-TK I neu/erbB-2	TK VI Met	?
Biological Action	Growth Differentiation	Growth	Immuno-modulation
Cell Type	Nerve Cells Epithelial Cells	Hepatocytes Endothelial Cells Epithelial Cells	T-Cells NK-Cells

Table 1-2

Table showing the four structural classes of cytokines according to secondary structure (Richardson, 1981). Subgroupings within each group are shown according to structural motifs, receptor classes utilised and biological actions. Adapted from Nicola, 1994. See abbreviations for full description.

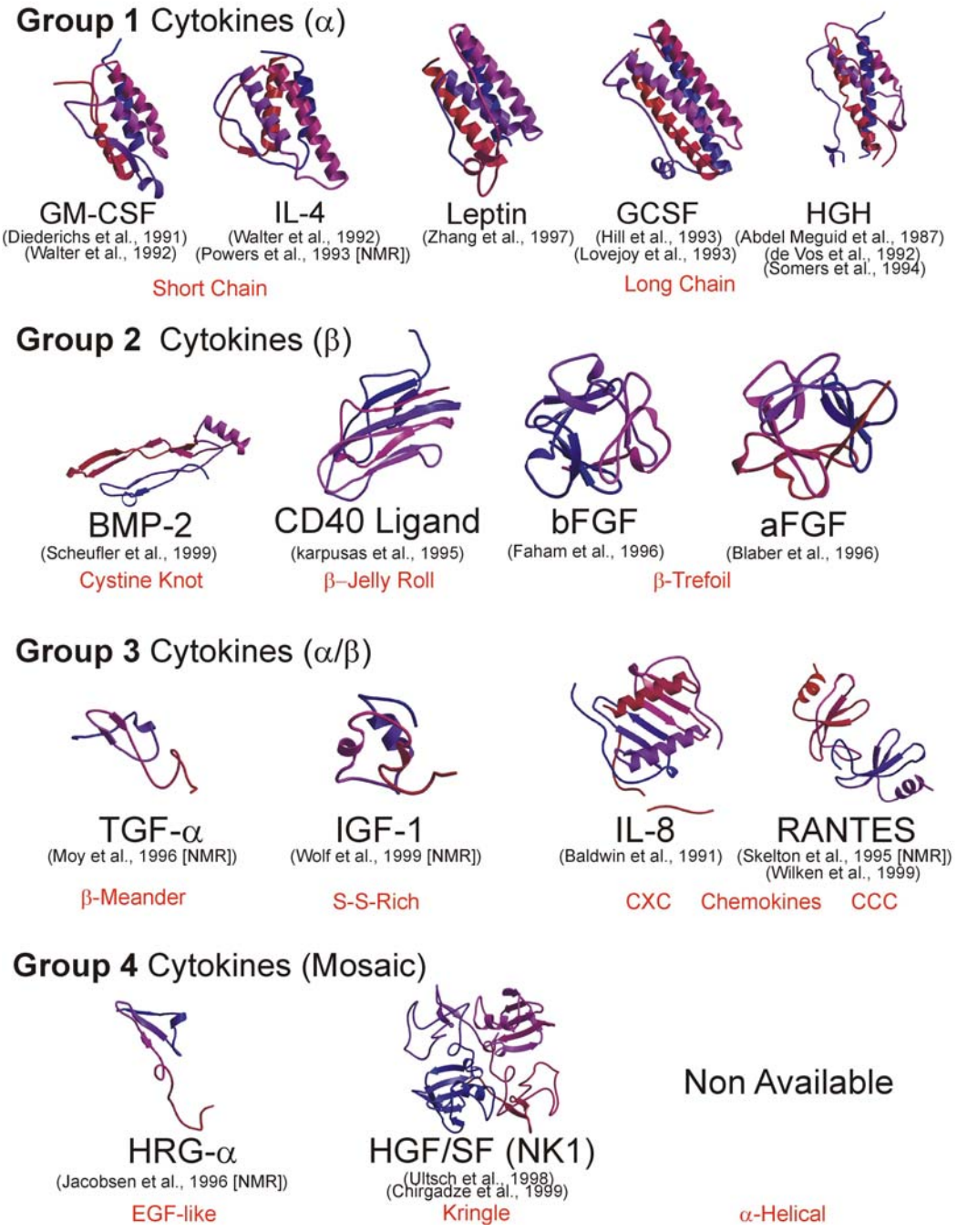


Figure 1-1

Gallery of cytokine folds arranged according to the four structural classes of secondary structure (Richardson, 1981). Subgroupings within each group are shown according to structural motifs as highlighted in red lettering. Adapted from Nicola, 1994. All structures solved by X-ray crystallography (unless otherwise stated). Figures produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994). The polypeptides are coloured from blue at the N-terminus, through red at the C-terminus. See abbreviations for full descriptions.

1.4.1.1 Group 1 Cytokines (α)

This is the largest family of cytokines and these structures adopt a 4-α-helical bundle as first reported for GH (Abdel Meguid *et al.*, 1987) (see Figure 1-2). Members of this family have attracted much attention due to their wide-ranging roles in the immune system. This fold has an unusual 'up-up-down-down' connectivity in which two long over-hand loops (connecting helices A and B and helices C and D) allow for helix direction reversal and for packing of the A:D helix pair against the B:C helix pair with a characteristic skew angle. This unusual topology has so far not been observed in any other family of proteins outside the cytokine realm. This group of cytokines can be further sub-divided into short-chain or long-chain helical cytokines, dependent on the length of the main helical segments (see Figure 1-2).

1.4.1.1.1 Short chain helical cytokines

Short-chain 4-α-helical cytokines include IL-2, IL-3, IL-4, IL-5, IL-7, IL-9, IL-13, GM-CSF, M-CSF, SCF and IFNγ (Figure 1-2). Short-chain cytokines have relatively short α-helices of ~15 amino acids in length and have a relatively large skew angle (~35^{o}) between the helical axis of the A:D pair and the B:C pair. Short chain cytokines often have two short twisted anti parallel β-strands in the A-B and C-D loops as exemplified by GM-CSF. Other short chain helical cytokines analysed structurally include IL-5 (Milburn *et al.*, 1993), M-CSF (Pandit *et al.*, 1992), IFN-γ (Ealick *et al.*, 1991; Samudzi *et al.*, 1991) and IL-3 (Feng *et al.*, 1996) which are dimeric, and GM-CSF (Diederichs *et al.*, 1991; Walter *et al.*, 1992a), IL-2 (Bazan, 1992) and IL-4 (Powers *et al.*, 1993; Smith *et al.*, 1994; Walter *et al.*, 1992b; Wlodaver *et al.*, 1992) which are monomeric.

1.4.1.1.2 Long chain helical cytokines

Long-chain cytokines include IL-6, IL-12, Epo, G-CSF, LIF, OSM, CT-1, CNTF, IL-11, leptin, GH, Prl, IFNα, IFNβ, IFNτ and IL-10 (see Table 1-2). This class of helical cytokines have relatively long α-helices of ~25 amino acids in length and a relatively small skew angle (~18^{o}) between the helical axis of the A:D pair and the B:C pair. Long chain helical cytokines resolved structurally include GH (Abdel Meguid *et al.*, 1987; de Vos *et al.*, 1992), Epo (Livnah *et al.*, 1998; Livnah *et al.*, 1996), Leptin (Zhang *et al.*, 1997a), GCSF (Hill *et al.*, 1993; Lovejoy *et al.*, 1993), IL-10 (Zdanov *et al.*, 1995), IFN-β (Senda *et al.*, 1995) and IFN-τ (Radhakrishnan *et al.*, 1999).

The long chain cytokine family is often further sub-divided into three further groups with GH, Epo, leptin and GCSF forming the classic four helix group and the IFN members adopting a topology based on a five helix bundle. LIF, CNTF, CT-1, IL-6, IL-11 and OSM form another group through structural and sequence homology and the shared use of the gp130 receptor (see Figure 1-2). Long chain cytokines differ from the short chain variants in having between 50 and 100 extra amino acids, so producing a more elongated cylindrical shape as compared to the ellipsoidal shape adopted by the short chain forms.

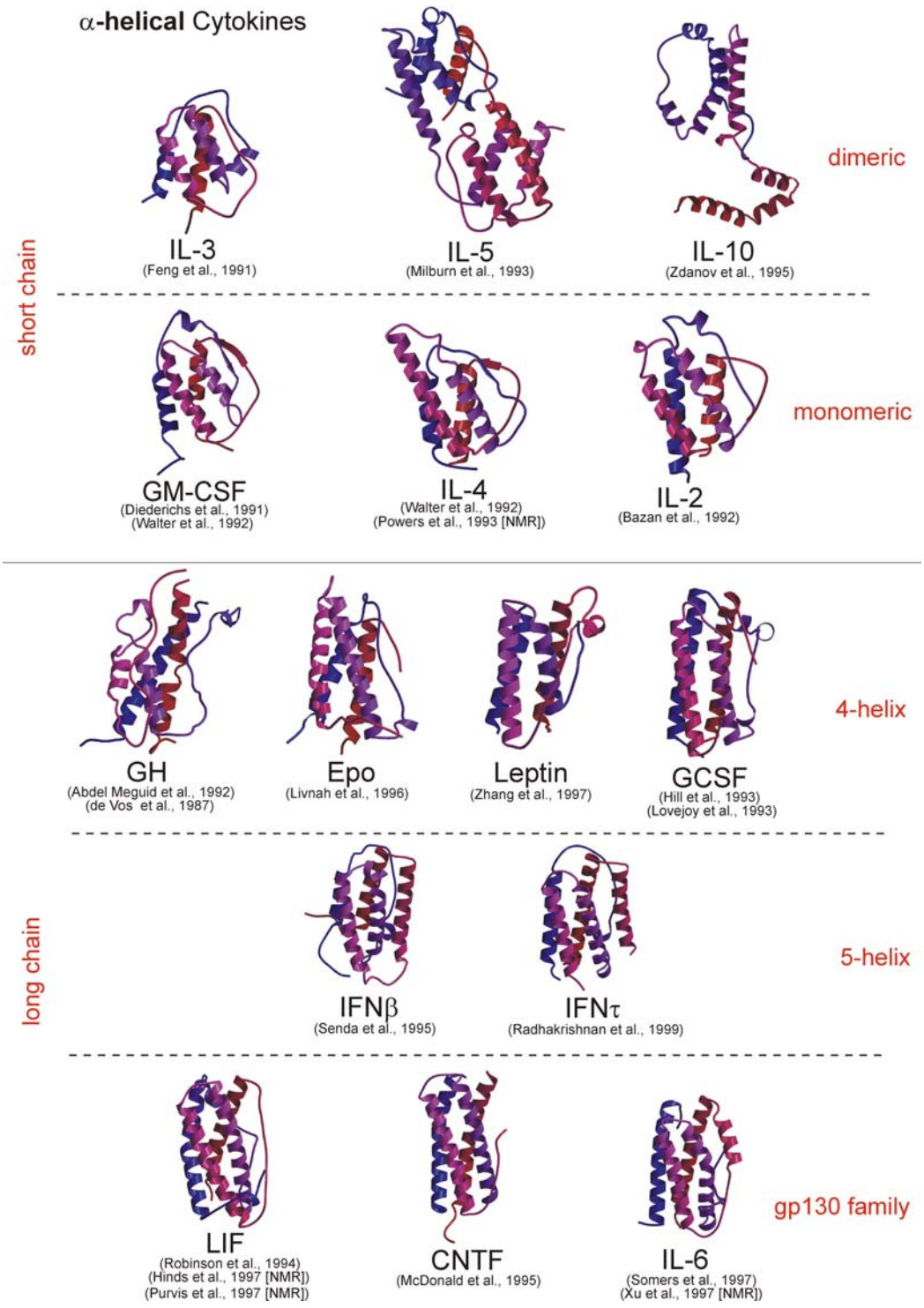


Figure 1-2

Gallery of group 1 α-helical cytokines arranged in their sub-groupings of long and short chain (Nicola, 1994). Further sub-divisions according to dimerisation state, number of main helices and receptor usage are shown in red lettering. All structures solved by X-ray crystallography (unless otherwise stated). Figures produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994). The polypeptides are coloured from blue at the N-terminus through, red at the C-terminus. See abbreviations for descriptions.

1.4.1.2 Group 2 Cytokines (β)

Members of this family include TNF, IL-1, FGF, PDGF, NGF, VEGF, CD40 ligand, and TGFβ (see Figure 1-2). All members of this family contain long chain β-sheet structures and can be further subdivided according to β-jelly roll, β-trefoil or cystine knot motifs.

1.4.1.2.1 β-jelly roll motif

The TNF sub-family of group 2 cytokines includes TNFα, TNFβ, OX-40 antigen, CD40, CD27 and Fas ligands (see Table 1-2). The TNF family are cell surface associated ligands and form symmetric homotrimers based on the β-jelly roll motif as observed in many viral coat proteins (e.g. satellite tobacco necrosis virus) and cell surface proteins. The β-jelly roll motif is best considered as a β-sandwich structure formed by two anti-parallel β-pleated sheets as observed in the crystal structures of the CD40 ligand (Karpusas *et al.*, 1995) and TNF (Baeyens *et al.*, 1999; Eck and Sprang, 1989; Jones *et al.*, 1989; Reed *et al.*, 1997) (see Figure 1-1).

1.4.1.2.2 β-trefoil motif

The IL-1 (Le and Vilcek, 1987) and FGF (Faham *et al.*, 1996) sub-family of group 2 cytokines contain a β-trefoil motif (see Table 1-2), as first identified in Soya bean trypsin inhibitor (Blow *et al.*, 1974). This motif consists of three sets of 4-β-strands each of which adopts a trefoil shape (Y-shape). The 12 strands of β–sheet form six hairpins, with three of the hairpins forming a β–barrel structure and the other three forming a triangular structure. Members of this family that have been determined structurally include basic FGF (Faham *et al.*, 1996), acidic FGF (Blaber *et al.*, 1996) and IL-1β (Priestle *et al.*, 1988) (see Figure 1-1).

1.4.1.2.3 Cystine knot motif

The cystine knot sub-family of group 2 cytokines is particularly divergent and includes TGFβ, BMP, PDGF, VEGF, NGF and BDNF (see Table 1-2). Only the overall fold and

position of the three disulphide bonds unites the group. The cystine knot motif consists of an elongated asymmetric monomer produced by an interlocking set of three disulphide bonds and at least four strands of twisted anti-parallel β–sheet. Members of the group delineated structurally include BMP-2 (Scheufler *et al.*, 1999), NGF (McDonald *et al.*, 1991), VEGF (Muller *et al.*, 1997), PDGF (Oefner *et al.*, 1992), TGF-β2 (Daopin *et al.*, 1992) and chorionic gonadotrophin (Wu *et al.*, 1994) (see Figure 1-1).

1.4.1.3 Group 3 Cytokines (α/β)

Members of this family include EGF, TGF-α, β-cellulin, insulin, IGF-1 and the chemokine sub-family of cytokines including IL-8, GRO, MIP-2, RANTES and MIP-1α (see Table 1-2). All members of this family contain short chain α/β structures. Structurally the group 3 cytokines can be further sub-divided according to those containing many disulphides, β-meander motifs or a conserved pattern of cysteines (CXXXC, CXC, CC or C chemokines).

1.4.1.3.1 β-meander motif

The β-meander sub-family of group 3 cytokines includes EGF, TGFα, β-cellulin, amphiregulin and vaccinia virus growth factor (see Table 1-2. These cytokines are small proteins of ~50 residues in length, that contain characteristic EGF-like domains (Carpenter and Cohen, 1990). EGF-like domains consist of at least two anti-parallel β-strands connected to the intervening loops by three disulphide bonds. The EGF domain is released from the transmembrane precursor by proteolysis and has been shown to bind calcium. EGF-like domains are found in several non-cytokine and growth factor proteins, for example the adhesion molecule L-selectin and the blood coagulation protein factor X, which has also been shown to bind calcium. Members of this group solved structurally include TGF-α (Moy *et al.*, 1993) and EGF (Campbell *et al.*, 1989)(see Figure 1-1).

1.4.1.3.2 Chemokines

The chemokine sub-families are of particular importance due to their roles in inflammation (Baggiolini and Dahinden, 1994) and chemotaxis (Miller and Krangel, 1992). Recently it has been shown that entry of human immunodeficiency virus-1 (HIV-1) into cells depends on the chemokine receptors CCR5 and CXCR4 to which RANTES and SDF-1α bind, respectively, so antagonizing infectivity. Members of this family adopt a β-sandwich motif with a C-terminal α-helical segment. This fold is exemplified by IL-8, which forms a symmetrical dimer consisting of a six-stranded β-sheet with two α-helices straddling the sheet. This produces a structure reminiscent of the peptide binding groove observed in the major histocompatibility complex (MHC) class I molecules. The chemokine family can be further sub-grouped according to the amino acid sequence around a conserved set of cysteine residues; C, CC, CXC or CXXXC (see Table 1-2). Members of the CXC chemokine family analysed structurally include IL-8 (Baldwin *et al.*, 1991; Clore and Gronenborn, 1991), NAP-2 (Malkowski *et al.*, 1995), SDF-1a (Dealwis *et al.*, 1998). Members of the CC chemokine family which have been analysed structurally include MCP-1 (Lubkowski *et al.*, 1997), MIP1-α and RANTES (Skelton *et al.*, 1995; Wilken *et al.*, 1999).

1.4.1.3.3 Disulphide rich

Members of this sub-family include insulin, IGF-I, IGF-II, relaxin and bombyxin (see Table 1-2). The polypeptide hormone insulin can be included in this sub-family along with the structurally related IGF-1 (see Figure 1-1) and IGF-2. Insulin consists of a α/β structure derived through folding and cleavage of proinsulin.

1.4.1.4 Group 4 Cytokines (mosaic)

Members of this family combine several structural features and as such may also belong to more than one structural group. Members include EGF, HGF and IL-12. This is a very

diverse group of cytokines and is often further sub-divided according to the domain types involved (i.e. EGF-like, kringle and α-helical)(see Table 1-2). For example HGF is a heterodimer consisting of an α- and β-domain. The α-subunit contains four Kringle domains and the β-subunit contains a serine protease domain. Kringle domains consist of a five-stranded antiparallel beta sheet structure with a conserved pattern of three internal disulphide bridges, as observed in the NK1 fragment of HGF (Chirgadze *et al.*, 1999; Ultsch *et al.*, 1998). IL-12 is also a heterodimer in which the α-subunit contains features of 4-α-helical bundle cytokines and the β-subunit contains features of a haemopoietin receptor domain.

1.4.2 Structural Classification of Cytokine Receptors

Crystallographic and NMR studies of cytokine receptors are severely hampered by their physical and chemical properties. Firstly, cytokine receptors are relatively large; often being composed of multiple immunoglobulin related domains, so rendering them problematic for expression in simple prokaryotic systems. The large size also limits the use of NMR techniques. The multiple extended domain structure can also be problematic for crystallographic studies as the innate flexibility in the domain linker regions can adversely effect crystallisation. Secondly, cytokine receptors generally have transmembrane regions, thus rendering them insoluble in aqueous solutions and liable to aggregation. Thirdly, cytokine receptors are often heavily and heterogeneously glycosylated, which can again be problematic for crystal packing and X-ray structure determination. Methods used to overcome these problems include the use of eukaryotic expression systems that allow expression of larger and more homogeneous proteins and the formation of ligand receptor complexes which imposes greater rigidity to the receptor domains.

There is a close correlation between the structural class of the cytokine and the receptor type to which it binds. For example, almost all of the 4-α-helical bundle cytokines bind to receptors that contain a 200 amino acid ligand-binding domain known as the haemopoietin domain (see Figure 1-3). Exceptions to this rule include the interferons and IL-10, which bind to interferon receptor domains (Ho *et al.*, 1993). However, the haemopoietin domain is structurally related to the interferon receptor domain and these domains are often referred to as haemopoietin class II. These domains are more distantly related to the fibronectin type III and immunoglobulin domains. Seven distinct cytokine receptor superfamilies can be defined on the basis of biochemical function, structural homology and sequence similarity. These are the haemopoietin receptor superfamily, the interferon/IL-10 receptor superfamily, the TNF/NGF receptor superfamily, the TGFβ/activine receptor superfamily, the immunoglobulin receptor superfamily, the tyrosine kinase receptor superfamily and the guanine nucleotide-binding protein-coupled receptor superfamily (G-protein) (see Figure 1-3).

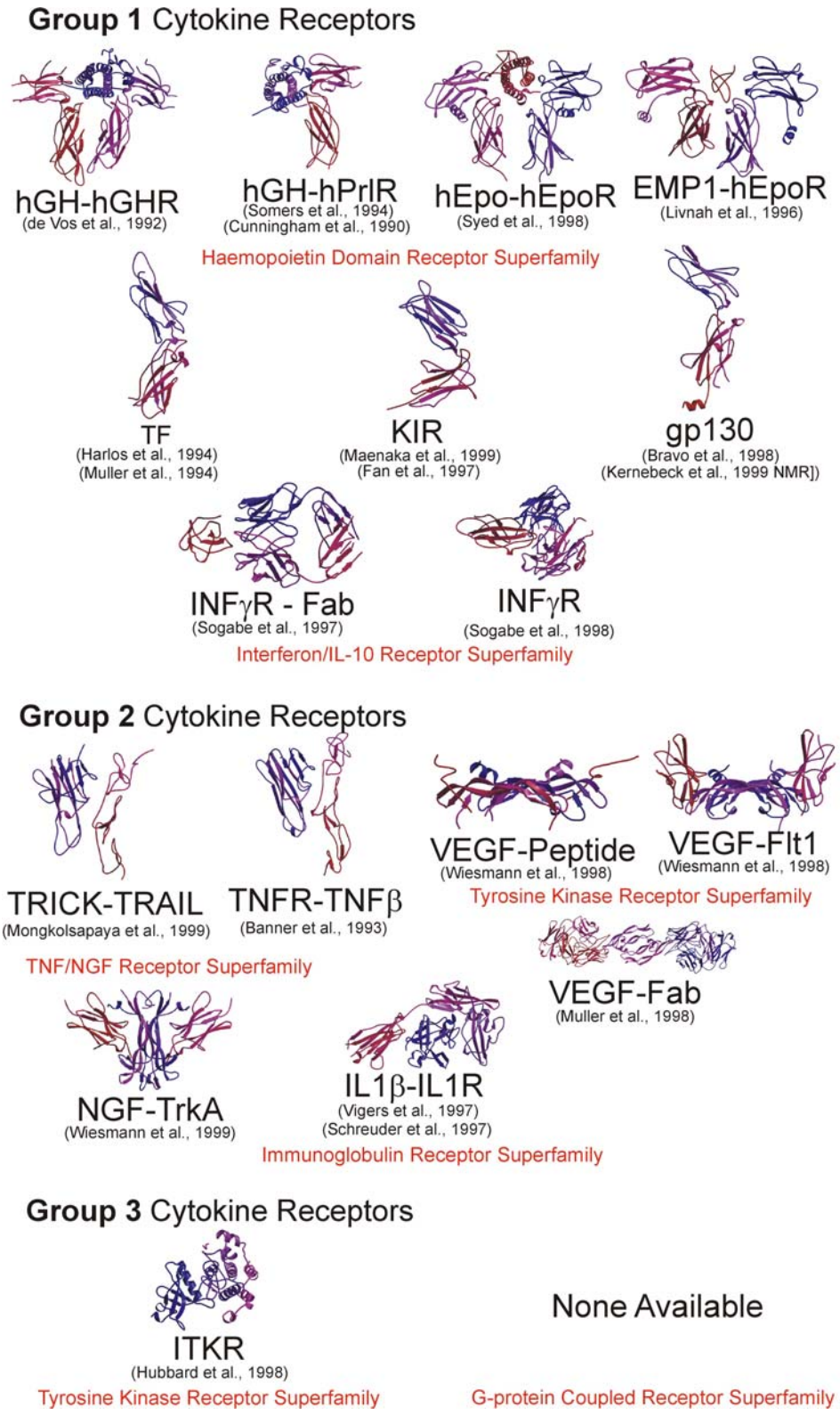


Figure 1-3

Gallery of cytokine receptor folds arranged according to the cytokine ligands that they bind. All structures solved by X-ray crystallography (unless otherwise stated). Figures produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994). The polypeptides are coloured from blue at the N-terminus through to red at the C-terminus. Red lettering denotes the main superfamilies of receptors.

1.4.2.1 Haemopoietin Receptor Superfamily

This superfamily includes receptors for all of the group 1 helical cytokines, except the interferons and IL-10 family of ligands (Cosman *et al.*, 1990; Cosman, 1993) (see Figure 1-4). This superfamily is characterised by the conserved nature of the ligand-binding domain, which is 200 amino acids in size and referred to as a haemopoietin domain. Members of the haemopoietin receptor superfamily are type I membrane glycoproteins, orientated with the N-terminus in an extracellular position through the use of an N-terminal signal sequence, which is post-translationally cleaved. The haemopoietin domain is composed of two fibronectin type III-like modules. These modules consist of 100 amino acids, which form seven β-strands that are folded in an antiparallel fashion to form a β-barrel-like structure. These receptors share a conserved pattern of cysteines in the N-terminal fibronectin type III-like domain which is required for correct folding of the haemopoietin domain and functions in disulphide bond formation between the β-sheet (Rozakis Adcock and Kelly, 1991). The C-terminal fibronectin type III-like domain contains a conserved WSXWS motif, which also performs an important structural function. Substitution of these amino acids in IL-6 results in low levels of expression suggesting incorrect protein folding and poor protein stability (Yawata *et al.*, 1993).

The haemopoietin domain is often combined with other structurally distinct domains such as the immunoglobulin domain, as observed in IL-3Rα, GM-CSFRα, IL-5Rα, IL-6Rα, gp130 and GCSFR. Another structural domain also utilised is the fibronectin type III module as observed in IL-6Rα, gp130, LIFR and G-CSFR. The haemopoietin domain can also be duplicated as observed in IL-3R, IL-5R and GM-CSFR (see Figure 1-4).

.

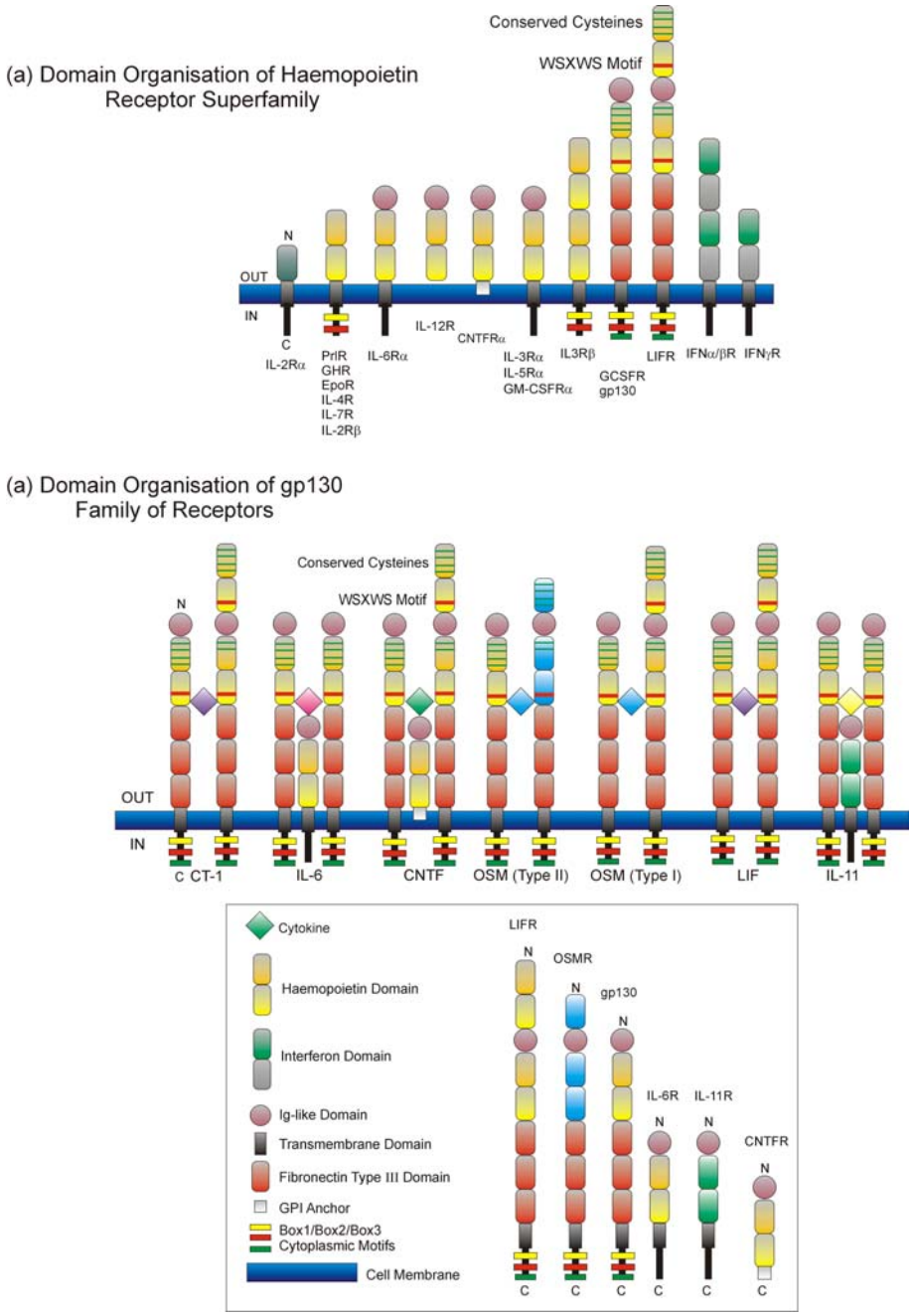


Figure 1-4

Diagrammatic representation of domain organisation observed in the (a) haemopoietin and (b) gp130 family of cytokine receptors. Adapted from (Nicola, 1994). See abbreviations for descriptions.

1.4.2.1.1 Structure-Function Relationship of Growth Hormone

This was the first haemopoietin receptor-ligand complex to be solved and revealed that one growth hormone molecule binds two receptor sub-units (de Vos *et al.*, 1992) so forming a cytokine plus receptor homodimer complex (see Figure 1-3). The structure of each receptor unit consists of two β-barrel motifs each composed of seven β-strands (conventionally labelled strands A, B, C, C', E, F, G in common with the immunoglobulin subset C2). The β-barrels are produced by folding of two β-sheets produced by E, B, A and G, F, C, C' so forming a buried hydrophobic interior and a hydrophilic surface. Three disulphide bridges are present in the N-terminal domain which function in linking strands A to B, E to C' and F to G. Also the WSXWS motif as observed in other members of the haemopoietin receptor family is replaced with the motif YGEFS in the C-terminal domain on the loop connecting strands F and G (see Figure 1-6).

In common with many other proteins, particularly the complementarity determining regions (CDRs) of antibodies, the β-barrel structures provide a stable framework on which the functionally important loops can reside. The hinge region connecting the two β-barrels adopts a 3_{10} helical conformation, which functions in fixing the interdomain angle at approximately 90^0. The hinge region connecting the two domains functions as the ligand interaction site.

High concentrations of hGH are capable of antagonising receptor dimerisation. This led to the proposal of a two site binding model, whereby hGH binds to a molecule of hGHR first at site 1 and following that forms a weaker interaction through site 2 with a second hGHR molecule (Cunningham *et al.*, 1991b). Site directed mutagenesis (Cunningham *et al.*, 1989; Cunningham and Wells, 1989) has revealed that the functional epitopes map onto the centre of these interaction sites. The receptor binding sites on hGH are located on opposite sides of the four helix bundle. Site 1 has a slightly concave character and is

formed by residues of helix A, helix D and the AB loop. The total surface buried at this hormone-receptor interface is approximately 1230 Å^2. In contrast, the second binding site on hGH is relatively flat and is made up of exposed side chains from helix A and helix C and the N-terminal loop region. Also, the surface buried in the site 2 interface is much smaller at approximately 900 Å^2, and it has been proposed that this allows for weaker site 2 interactions, so giving rise to a sequential binding mechanism. Once formed, the complex is then further stabilised by burial of approximately 1000 Å^2 of surface in the membrane proximal halves of each receptor.

Around 30 side chains from each protein make contact and mutagenesis of contact residues in hGHR, have revealed that a central hydrophobic region dominated by two tryptophan residues, accounts for more than three-quarters of the binding free energy (Clackson and Wells, 1995). It has been shown that both site 1 and site 2 of the hormone interact with the same part of the hGHR molecule (see Figure 1-5). This region consists of around 11 residues with Trp 104 and Trp 169 (one from each fibronectin type III-like domain) contributing most of the binding energy through hydrophobic burial. Trp 104 has been shown to fit into a pocket of the ligand produced by Gly 120 and mutation of this residue on the ligand sterically blocks the formation of a 1:2 complex (Clackson *et al.*, 1998). The same two residues plus Asn 128 of the receptor are important in site 2 interactions. These functional epitopes lie on loops extending out of the β-sheet, with Trp 104 residing on the EF loop of the N-terminal domain and Trp 169 residing on the BC loop of the C-terminal domain (see section 1.4.2.1.2 and Figure 1-6). Asn 128 is positioned in the helical linker region between the two domains. These loops have been shown to have similar main-chain conformations in both hGH-hGHR and hPrl-hPrlR complexes (Somers *et al.*, 1994) (see Figure 1-6).

Analysis of the binding determinants in the hGH-hGHR complex has revealed a pattern in which energetically important residues form dominant hydrophobic patches and the less important hydrophilic residues surround this region. It has been proposed that the hydrophobic core of these ligand-binding surfaces drives the interaction entropically, through Van der Waals contact and the associated exclusion of water from the hydrophobic surfaces. In contrast, the purpose of the peripheral hydrophilic residues is to provide the fidelity for the interaction through specific hydrogen bonding and salt bridge formation. In contrast, a reverse of this pattern has been reported from analysis of the crystal structure of the IL4-IL4Rα complex (Hage *et al.*, 1999), in which Asp 72 of the receptor forms a core hydrophilic element on the receptor which is surrounded by a shell of hydrophobic residues. This residue forms a salt bridge with Arg 88 of the ligand, which itself is surrounded by a plethora of hydrophobic residues. The inner core of hydrophilic residues is thought to provide specificity to the binding and the positioning of Asp 72 in the hydrophobic environment enhances both hydrogen and ionic bonding potentials. The hydrophobic shell again provides the entropic driving force for the interaction. Asp 72 of the IL4Rα is the major determinant of binding and as such can be compared with Trp 104 of hGHR.

1.4.2.1.2 Structure-Function Relationship of Prolactin

The hPrlR binds both hGH and hPrl. The crystal structure of hPrlR bound to hGH, in 1:1 stoichiometry, has been solved and this structure shows homology to the hGH receptor-ligand complex (Cunningham *et al.*, 1990; Somers *et al.*, 1994) (see Figure 1-3). As in hGHR the WSXWS box is located in the FG loop of the C-terminal domain. The hGH binds to the PrlR at site 1 in a similar fashion to that observed in the hGH-hGHR system, with residues from the AB loop and the D helix participating in the interface. Similarly to hGHR the hPrlR also displays two conserved tryptophan residues; Trp 72 in the EF loop (analogous to Trp 104 of hGHR) and Trp 139 in the BC loop (analogous to Trp 169 of

hGHR). Although hGH binds to both the hGHR and hPrlR, hPrl does not bind to the hGHR and mutagenesis studies have shown that the binding sites for both receptors on hGH overlap, with Arg 167 and Glu 174 being key specificity determinants (Somers *et al.*, 1994). It has also been shown that Glu 174 and His 18 of hGH function in chelating Zn^{2+} ions along with Asp 217 and His 218 of the receptor. This contributes a considerable enthalpy to the binding reaction, increasing the affinity of hGH for hPrlR by upto 8000-fold (Somers *et al.*, 1994).

1.4.2.1.3 Structure-Function Relationship of Erythropoetin

Several structures of EpoR have been determined including Epo bound (Syed *et al.*, 1998), peptide agonist bound (Livnah *et al.*, 1996), peptide antagonist bound (Livnah *et al.*, 1998) and ligand free receptor (Livnah *et al.*, 1999). These structures show general homology to the hGH-hGHR complex (see Figure 1-3). The structure of EpoR has been determined in complex with a dimeric 20 residue cyclic peptide unrelated in sequence to Epo, termed EMP1 (Epo mimetic peptide 1; Livnah *et al.*, 1996; Wrighton *et al.*, 1996). EMP1 is capable of inducing an erythropoetic response and the signaling pathways activated appear to be identical to those activated by native Epo (Wrighton *et al.*, 1996). Signaling by the mimetic shows ~5% the efficiency of the natural ligand (Johnson *et al.*, 1998). This mimetic dimerises EpoR using a smaller set of residues, as compared to the hGH-hGHR system, again with key aromatic residues (Phe 93 and Phe 205) forming the functional epitope (Middleton *et al.*, 1999). Comparison of the EMP1-EpoR complex with the hGH-hGHR and Epo-EpoR complexes has suggested that the efficiency of signaling through these receptors is dependent on the receptor orientation (Syed *et al.*, 1998). In the EMP1-EpoR structure the membrane proximal haemopoietin domains are twisted away from each other by ~45^{o}, whereas these two domains in all other haemopoietin receptor-ligand complexes lie in the same plane. It has been proposed that this shift is responsible for the reduced

signaling efficiency of EMP1, as the intracellular domains are not oriented optimally to allow signaling through the JAK pathway (see 1.4.3)(Syed *et al.*, 1998).

This hypothesis has gained further support from *in vivo* protein fragment complementation assays, in which EpoR-JAK2 fusions have been made constitutively active through the inclusion of linker regions between the receptor and kinase domians (Remy *et al.*, 1999). This allows optimal association of signaling components in the absence of ligand induced receptor dimerisation. Further support is also provided by the crystal structure of the ligand free EpoR, in which crystal packing contacts imply a propensity for formation of an inactive dimer structure in which the intracellular domains would be too far apart to permit phosphorylation of JAK2 (Livnah *et al.*, 1999). It is proposed that these preformed inactive dimers could exist on the cell surface and that ligand binding induces a reaarangement that activates the receptor.

1.4.2.1.4 Structure-Function Relationship of gp130

The gp130 receptor structure has been determined crystallographically in a ligand free state (Bravo *et al.*, 1998) (see Figure 1-3 and Figure 1-6). This receptor shows homology to other haemopoietin receptors as discussed above (see 1.4.2.1.1) (see Figure 1-6). However, one notable feature unique to gp130 is the packing observed in the WSXWSX motif. This motif forms two-successive β-bulges resulting from the intercalation of Trp 288, Trp 291 and Trp 247 between the aliphatic residues Met 280, Arg 278, Arg 276 and Arg 240 which are located on strand F, so forming an extended π-cation stacking system (see Figure 1-6). This is more extended than other members of the family due to inclusion of Trp 247 and Arg 240, which are not classic members of the motif. The second major structural difference is the division of strand G of the N-terminal domain into two equal portions, again resulting from a second β-bulge structure similar to that observed in the WSXWS motif itself. In contrast to hGHR, the interdomain 3_{10} helical region is followed by a

region of short β-strand. This functions in hydrogen bonding to residues of both fibronectin type III-like domains, so resulting in increased interdomain rigidity.

Sites responsible for ligand binding have been mapped to residues including Phe 169 of the N-terminal domain EF loop (analogous to Trp 104 of hGHR) and Val 230 in the BC loop (analogous to Trp 169 of hGHR) of the C-terminal domain. Val 230 is relatively buried in the crystal structure although mutagenesis does affect binding to IL-6-IL-6R (Horsten *et al.*, 1997), so suggesting this residue has an indirect effect on binding resulting from perturbed local structure. In GHR site 2 interactions are mediated through Asn 128 and this residue is structurally equivalent to Asp 283 of gp130 (see Figure 1-6). However, no data are yet available concerning the function of this residue in gp130.

The gp130 family of receptors includes IL-6R (Ward *et al.*, 1995), LIFR (Gearing *et al.*, 1991) (Hilton and Nicola, 1992), OSMR (Linsley *et al.*, 1989; Mosley *et al.*, 1996), CNTFR (Davis *et al.*, 1993), IL-11R (Hilton *et al.*, 1994), CT-1R (Robledo *et al.*, 1997) and gp130 (Taga and Kishimoto, 1995). These receptors are structurally related and show many overlapping biological activities. Genetic analysis (Gough, 1992) and biochemical reconstitution (Baumann *et al.*, 1993) has revealed a common receptor complex structure in which the ligand specific α-subunit is combined with one or two receptor components that are shared between the various cytokine family members. These common components are gp130 and LIFR, with the gp130 acting as a signal-transducing component in all cases. IL-6 and IL-11 have been shown to bind to ligand-specific receptors (i.e. IL-6R and IL-11R, respectively) so inducing homodimerisation of gp130 and initiation of signaling (see Figure 1-4).

Other members of the family such as LIF and OSM induce heterodimerisation of gp130 with LIFR. OSM can also induce heterodimerisation of gp130 with a specific OSMRβ

subunit (Mosley *et al.*, 1996), so forming a so called type II assembly consisting of OSM, gp130 and OSMRβ as shown in Figure 1-5 below. In a similar fashion CNTF interacts with CNTFR and then forms a higher order complex also containing gp130 and LIFR. Analogous trimeric receptor assemblies are also proposed to exist for CT-1, which is thought to bind a ligand-specific receptor followed by gp130 and LIFR. CNTFR differs from other members of the family as it is attached to the membrane via a glycero-phosphatidyl inositol (GPI) anchor, therefore lacking both transmembrane and cytoplasmic domains.

1.4.2.2 Interferon/IL-10 Receptor Superfamily

INFγR and IL-10R belong to this superfamily which is structurally related to the haemopoietin receptor superfamily and as such also functions in binding α-helical bundle cytokines (see Table 1-2). This superfamily is often referred to as haemopoietin class II. The fibronectin type III-like modules of this family contain a conserved pattern of cysteine residues at the C- and N-termini, but the WSXWS motif observed in the haemopoietin class I family is not present (Taga and Kishimoto, 1992). Crystal structures solved include tissue factor (TF) (Harlos *et al.*, 1994; Muller *et al.*, 1996), killer cell inhibitory receptor (KIR) (Fan *et al.*, 1997; Maenaka *et al.*, 1999) and IFNγR (Sogabe *et al.*, 1997). These structures have revealed variability in the interdomain angle ranging from ~50^{o} in KIR to ~120^{o} in TF and IFNγR. This compares with the ~90^{o} angle observed in EpoR, GHR, PrlR and gp130 (see Figure 1-3 for comparison).

1.4.2.3 TNF/NGF Receptor Superfamily

Members of this superfamily include TNFR, NGFR, FasR, CD27R, CD30R, CD40R, 4-1BBR and OX-40R (see Table 1-2). This receptor superfamily is structurally distinct and is characterised by long chain β-sheet structures, with a characteristic set of cysteine-rich repeats. Activation of TNFR requires receptor trimerisation, which is brought about by

aggregation of receptor subunits on binding a TNF trimer (Banner *et al.*, 1993). Other structures determined include the apoptosis inducing ligand TRAIL in complex with its receptor DR5 (Mongkolsapaya *et al.*, 1999) (see Figure 1-3).

1.4.2.4 TGF/Activine Receptor Superfamily

This is a large superfamily of transmembrane serine/threonine kinases, which function in binding long chain β-sheet cytokines such as TGF and activine (see Table 1-2). Receptors of this superfamily contain short cysteine-rich extracellular domains and intracellular kinase domains. The structure of the signaling complex is not fully understood, although it is thought that oligomerisation between TGF receptors type I, type II and type III occurs (Nicola, 1994).

1.4.2.5 Immunoglobulin Receptor Superfamily

The IL-1 group of cytokines (IL-1α, IL-1β and IL-1γ) bind to this receptor superfamily (Sims *et al.*, 1988). These receptors contain an immunoglobulin-like domain, which consists of seven anti-parallel β-strands, folded into two sheets, which are stabilised by a conserved disulphide bond. Structures determined include IL1R in complex with both IL1β (Vigers *et al.*, 1997) and its naturally occurring antagonist (Schreuder *et al.*, 1997) (see Figure 1-3).

1.4.2.6 Tyrosine Kinase Receptor Superfamily

Members of this superfamily include CSFR, c-kit, EGFR, IGF-1R, FGFR and PDGFR. These receptors have a cytoplasmic domain with tyrosine kinase activity. All members of this family undergo ligand induced homodimerisation, which brings the cytoplasmic domains together, so activating the kinase domain and initiating signal transduction (Ullrich and Schlessinger, 1990). Members of this superfamily studied structurally include the Fms-like tyrosine kinase receptor (Flt-1) (Wiesmann *et al.*, 1997) (see Figure 1-3). Intracellular portions of these receptors have also been solved, for example the activated

insulin receptor tyrosine kinase in complex with peptide substrate and ATP analogue (Hubbard, 1997) (see Figure 1-3).

1.4.2.7 Guanine Nucleotide-Binding Protein-Coupled Receptor Superfamily

Chemokines such as IL-8 bind to this superfamily of structurally distinct receptors (see Figure 1-1 and Table 1-2) Receptors of this superfamily consist of a "serpentine seven pass" motif containing seven hydrophobic membrane-spanning domains and relatively short N- and C-terminal regions. Activation of the receptor complex leads to the formation of GTP bound α-subunits of a G-protein complex (G_α-GTP) which is subsequently released from the receptor and functions as an intracellular second messenger (Gilman, 1987; Neer and Smith, 1996). Structural studies of this superfamily are limited due to the hydrophobic nature of such membrane spanning proteins, although NMR studies have been carried out on IL-8 in complex with N-terminal peptide fragments of the CXCR1 receptor (Skelton *et al.*, 1999).

1.4.3 Receptor Activation and Signaling Through Haemopoietin Receptors

Signal transduction through cytokine receptors generally requires ligand induced receptor homodimerisation or heterodimerisation (Heldin, 1995). For example, homodimerisation is required for activation of M-CSFR, SCFR, EpoR, G-CSFR, GHR and PrlR (Watowich *et al.*, 1992). However, as discussed above the active signaling complex of many other cytokines, for example the gp130 family, requires heteroligomerisation of different receptor subunits. In general ligands first bind to specific receptor α-subunits so forming a low-affinity receptor-ligand complex. Further receptor components (receptor β-subunits and receptor γ-subunits) may bind and this oligomerisation leads to the formation of high-affinity activated receptor-ligand complexes.

The cytoplasmic regions of most cytokine receptors do not possess any intrinsic enzymatic activity and with the exception of IL-2R and EpoR sequence similarity in these regions is

low. There are, however, conserved regions in the intracellular domain termed box 1, box 2 and box 3 (Murakami *et al.*, 1991) and the recently identified suppresser of cytokine signalling (SOCS) box (Zhang *et al.*, 1999). These regions interact with cytosolic signaling molecules and so regulate signal transduction (Fukunaga *et al.*, 1993; Murakami *et al.*, 1993).

The membrane proximal motif, box 1, is characteristically rich in proline residues and is essential for association with the Janus Kinase (JAK) family of proteins (Ihle and Kerr, 1995; Ihle *et al.*, 1994; Taga and Kishimoto, 1995). JAKs responsible for gp130 signaling include JAK1, JAK2 and Tyk2, whereas the PrlR and GHR systems utilise only JAK2. JAK molecules are brought together through receptor dimerisation and this leads to phosphorylation of the conserved KEYY motif of the kinase activation loop. The associated increase in kinase activity leads to tyrosine phosphorylation of the receptor.

Once phosphorylated, this region becomes a binding site for Src homology 2 domains (SH2) of various signaling proteins such as Shc and Grb of the mitogen-activated protein (MAP) kinase pathway. Alternatively, signal transducers and activators of transcription (STAT) molecules can also associate with the phosphorylated receptors through SH2 domains. Phosphorylation of the STAT molecules induces formation of homo- or heterodimers, which then migrate to the nucleus where they bind interferon-γ activation sites (GAS) of various promoter elements. This functions in activating transcription (Darnell *et al.*, 1994). STAT1 and STAT3 are activated in response to gp130 stimulation, whereas, PrlR and EpoR utilize STAT 5 and GHR activates STAT1 (Ihle and Kerr, 1995).

This system is also modulated by the SOCS family of proteins, which act as intracellular inhibitors of several cytokine signal transduction pathways. Their expression is induced by cytokine activation of the JAK/STAT pathway and they act as a negative feedback loop by subsequently inhibiting the JAK/STAT pathway either by direct interaction with activated

JAKs or with the receptors. These interactions are modulated, in part, by a cytoplasmic element termed the SOCS box which is thought to target activated cell signalling proteins to the protein degradation pathway (Zhang *et al.*, 1999).

In a similar fashion, box 2 is important for association with JAKs and also mediates other biological effects. Box 1 is conserved in all members of the haemopoietin receptor superfamily, whereas box 2 has so far only been observed in G-CSFR, IL-2Rβ, EpoR, GM-CSFR, IL-3R and the gp130 related receptors. The role of the box 3 motif is not clear; however, all members containing the motif such as gp130, LIFR and OSMRβ activate the STAT3 transcription factor. It has been suggested that JAK1 is the major JAK involved in gp130 signaling and this possibly functions in phosphorylating the box 3 motif of the receptor, so allowing subsequent recruitment and activation of STAT3.

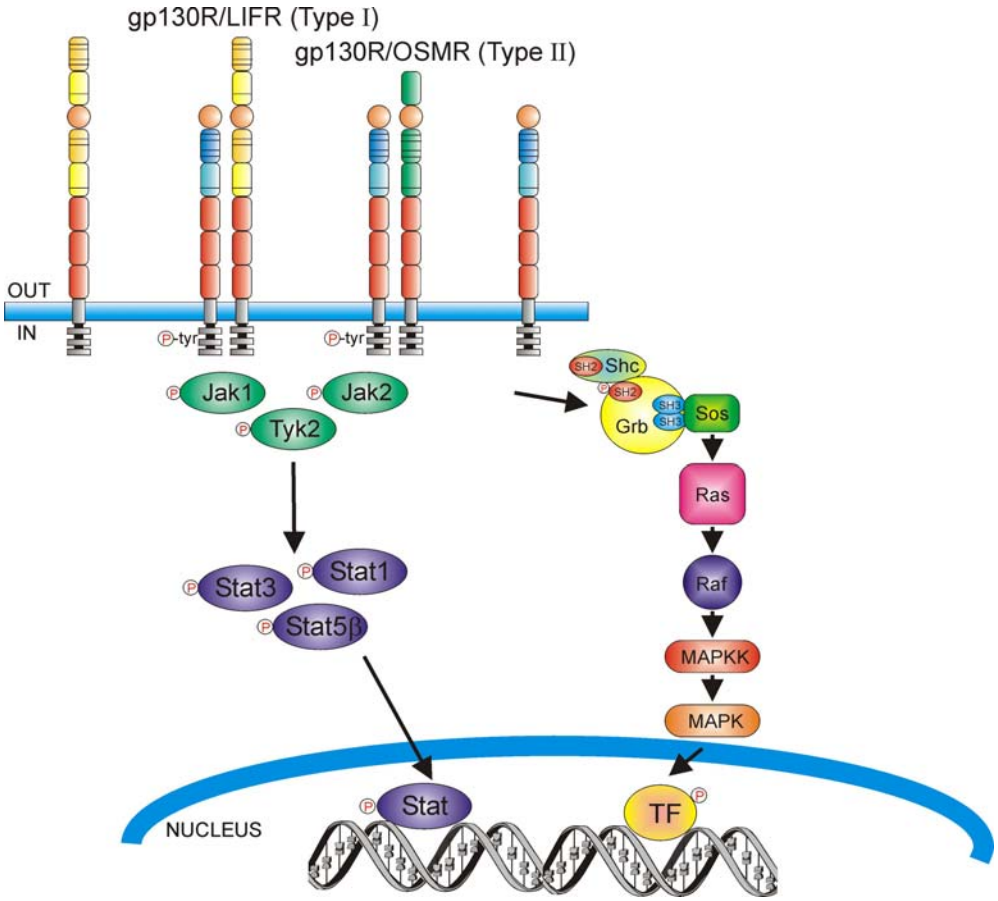


Figure 1-5

Diagrammatic representation of the JAK-STAT signaling pathway utilised by the gp130 family of cytokines as exemplified by OSM. Type I (gp130-LIFR-OSM) and Type II (OSMRβ-LIFR-OSM) signaling assemblies are shown on the cell surface and the associated JAK-STAT and MAP kinase pathways. Adapted from (Ihle *et al.*, 1994).

1.5 gp130 Family of Cytokines

As discussed previously LIF, IL-6, IL-11, CNTF, OSM and CT-1 are structurally and functionally related sharing the same signal transducer and protein fold. However, there is no strict functional compensation between the gp130 subfamily of haemopoietin cytokines and each member has unique biological actions (Piquet Pellorce *et al.*, 1994), thus necessitating detailed structure-function studies of each.

1.5.1 Leukemia Inhibitory Factor and its Receptor

1.5.1.1 Biological Action of Leukemia Inhibitory Factor

LIF is a polyfunctional cytokine that plays a role in many systems of the organism, eliciting a diversity of biological effects dependent on the cell type (Hilton, 1992). *In vitro*, LIF acts as a growth regulator in adipocytes, glial cells (Patterson, 1992), motor neurons (Li *et al.*, 1995), hepatocytes, myoblasts, embyonic stem cells and certain myeloid leukaemic cell-lines (Moreau *et al.*, 1988; Hilton *et al.*, 1988). The precise role of LIF *in vivo* is less defined, with the over expression of LIF resulting in rapid weight loss, behavioural disorders, ectopic calcification and bone abnormalities. In contrast, deletion of the LIF gene in mice results in little change towards known cellular targets of LIF. However, female reproduction is inhibited due to the prevention of embryonic implantation (Stewart *et al.*, 1992). This suggests that LIF is one of the few cytokines absolutely essential for life (Heath, 1992). Targeted disruption of the LIFR gene has also been shown to cause placental, skeletal, neural and metabolic defects resulting in perinatal death (Li *et al.*, 1995).

1.5.1.2 Structure-Function Relationship of Leukemia Inhibitory Factor and its Receptor

As discussed above the biological effects of LIF are mediated by interaction with a specific LIFR subunit. Formation of a high-affinity ($K_D \sim 10^{-9}$ M) signal-transducing complex requires association of the LIF-LIFR complex with the signal transducing component gp130. The hGH-hGHR complex has provided a model for possible modes of gp130-

LIF-LIFR engagement (see section 1.4.2.1.1, Figure 1-6). In common with the other haemopoietin cytokines it has been shown that several discrete sites of the molecule are involved in receptor engagement; sites 1 and 2 in similar locations to that observed for hGH. However, a third site, termed site 3, has been identified which allows for extra interactions not observed in the hGH-hGHR system. Site 1 and site 3 are the contact points for LIFR and site 2 is the contact point for gp130 (see Figure 1-6)

Tests of biological activity in human Ba/F3 and murine Da-1a assays have indicated that murine and human LIF, although exhibiting closely related three-dimensional structures, show distinct differences in biological specificity (Hudson *et al.*, 1996; Layton *et al.*, 1994a; Layton *et al.*, 1994b; Owczarek *et al.*, 1993). For example, both murine and human LIF interact with human LIFR, whereas murine LIF binds only murine LIFR. Data obtained from human-murine LIF chimeras suggests that the species-specific regions lie in the C-terminal region between residues 99 and 180. In particular, residues Ser 107, Val 155, Lys 58, His 112 and Ser 113 have been shown to be important for human LIFR binding. This region corresponds to site 1 in the growth hormone model and lies on helix D and the AB loop (see Figure 1-6). Using alanine scanning mutagenesis critical residues in site 1 have been localised to Val 175 with contributions from Asp 57, Lys 58, Lys 170 and Ala 174 (Hudson *et al.*, 1996).

Similarly, mutagenesis of human LIF has revealed a third site consisting of residues Lys 159 and Phe 156, located at the N-terminus of helix D, which provides the majority of free energy for binding to LIFR (Hudson *et al.*, 1996) (see Figure 1-6). Mutation of these residues individually reduces binding by ~100-fold and biological activity by ~1000-fold. Double mutations virtually eliminate receptor binding and biological activity. These residues are solvent exposed in close proximity to each other and are surrounded by a group of several residues that also affect binding. This arrangement of several pivotal

residues shielded by ones of less importance has also been observed in the hGH-hGHR complex (Clackson and Wells, 1995). Other cytokines with sequence similarity to human LIF (murine LIF, OSM, CNTF and CT-1), that utilise the LIFR in their signaling complexes, all have the human LIF residues Phe 156 and Lys 159 conserved (Robinson *et al.*, 1994). In contrast, IL-6 and IL-11, which do not bind the LIFR, lack these residues. Therefore, Phe 156 and Lys 159 (or their equivalents) may function as a LIFR binding epitope. This epitope is referred to as site 3 and forms high affinity contacts with LIFR whilst site 1 functions as a secondary component of LIFR binding. Regions of LIFR responsible for the site 3 interaction have recently been localised to the so-called LIGHT motif of the immunoglobulin-like domain (Chobotova and Heath, 2000).

Residues of human LIF involved in binding gp130 have also been identified through the use of multiple simultaneous substitution mutagenesis (Hudson *et al.*, 1996). This has revealed several residues in helix-A (Gln 25, Ser 28 and Gln 23) and helix-C (Ile 121, Gly 124 and Ser 127) that participate in gp130 binding (see Figure 1-6). None of these residues are entirely conserved through the gp130-binding family of cytokines, so suggesting that a variable subset of these residues is involved in binding gp130. This motif is therefore thought to function in regulating the affinity of gp130 binding. This region is analogous to site 2 in the hGH-hGHR model and non-conservative mutations in this region result in loss of biological activity (see Figure 1-6).

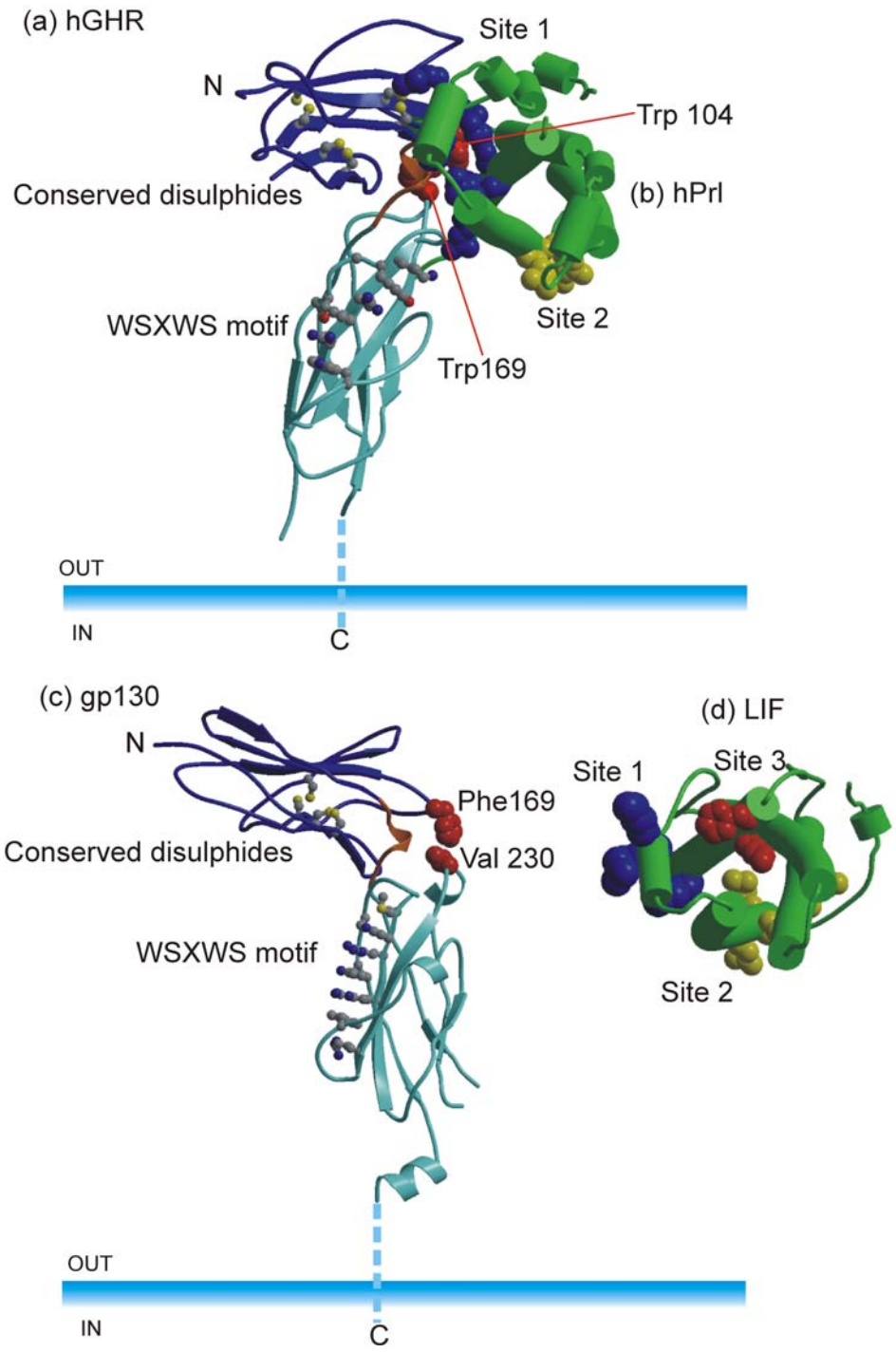


Figure 1-6

Figure showing the structural and functional motifs observed in the haemopoietin cytokine family. The C-terminal (light-blue) and N-terminal (dark-blue) fibronectin type III-like domains of (a) hGHR (Somers *et al.*, 1994) and (c) gp130 (Bravo *et al.*, 1998) are shown. The Linker region between domains is shown in orange along with key residues of the receptor in red. The ligands (b) hPrl and (d) LIF (Robinson *et al.*, 1994) are shown in green. Site 1 (blue), site 2 (yellow) and site 3 (red) residues for each ligand are shown. Figures produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

1.5.1.3 Leukemia Inhibitory Factor Receptor

LIFR is 1097 amino acids long with a 44 amino acid leader sequence. The single transmembrane domain is 26 amino acids in size and consists of a stretch of hydrophobic residues. The extracellular domain is 789 residues long and consists of two haemopoietin domains separated by a putative immunoglobulin-like domain. This ligand-binding region is separated from the cell surface by three fibronectin type III domains. The cytoplasmic domain is 238 residues in size and is rich in serines, threonines, prolines and acidic residues. The predicted molecular weight of LIFR is 111 kDa, although the observed molecular weight on SDS-PAGE is 200 kDa due to extensive glycosylation at several of the 20 potential N-glycosylation sites. Murine and human LIFR show 80% sequence identity with both showing extensive homology to the entire sequence of gp130.

1.5.2 Oncostatin M and the gp130 Receptor

1.5.2.1 Biological Action of Oncostatin M

OSM and LIF function in a highly pleiotropic and yet redundant fashion. This is reflected through their syntenic linkage on chromosome 22 (Giovannini *et al.*, 1993; Rose *et al.*, 1993), and the shared use of gp130 (Taga and Kishimoto, 1997). Among the shared biological activities of OSM and LIF are the ability to induce differentiation of murine myeloid leukemic cells into macrophage-like cells (Bruce *et al.*, 1992) and the inhibition of differentiation of embryonic stem cells (Rose *et al.*, 1994). Both LIF and OSM also play a role in the inflammatory response, for example induction of acute phase protein synthesis in hepatocytes (Richards *et al.*, 1992) and cartilage destruction (Hui *et al.*, 1996).

However, several important biological activities specific to OSM have also been identified including growth inhibition of a variety of solid tissue tumours (Horn *et al.*, 1990) and human melanoma cells (Brown *et al.*, 1987; Malik *et al.*, 1989; Zarling *et al.*, 1986). OSM also displays mitogenic activity in rabbit aorta smooth muscle cells (Grove *et al.*, 1993). OSM, and not LIF or IL-6, has also been shown to induce spindle cell morphology in

acquired immunodeficiency syndrome (AIDS)-associated Kaposi's sarcoma (AIDS-KS cells) (Cai *et al.*, 1994; Miles *et al.*, 1992; Nair *et al.*, 1992). The lack of biological response of AIDS-KS cells to IL-6 and LIF can be explained by the absence of IL-6R and LIFR (Murakami Mori *et al.*, 1995).

1.5.2.2 Structure-Function Relationship of Oncostatin M and gp130

OSM shares approximately 25% sequence identity with LIF, so allowing structural comparison with the crystal structures of both human (Robinson *et al.*, 1994) and murine LIF (Robinson *et al.*, 1995) and more distantly to the hGH-hGHR system (see section 1.5.1.2 and 1.4.2.1.1).

The most notable difference to the LIF system is the use of multiple OSM receptor assemblies; namely type I (Linsley *et al.*, 1989) and type II (Mosley *et al.*, 1996). Type I assemblies consist of a low affinity ($K_D \sim 10^{-9}$ M) association with LIFR (Gearing *et al.*, 1991) followed by high affinity ($K_D \sim 10^{-11}$ M) binding to gp130, so forming the heterodimeric signal transducing complex. OSM is the only gp130 cytokine that can bind, albeit with low affinity, to gp130 alone (Sporeno *et al.*, 1994). The low affinity OSM-gp130 complex is incapable of inducing proliferative responses in various cell lines and LIFR is required to form a signal-transducing complex. Conversely, type II complexes have been identified, in a number of cell lines that utilise a high affinity OSM-specific receptor β subunit (OSMRβ) that do not bind LIF (Gearing and Bruce, 1992; Mosley *et al.*, 1996; Thoma *et al.*, 1994). This dual OSM receptor system provides an insight into the mechanisms of OSM signal transduction, which shares many similarities with LIF, whilst at the same time exhibiting many OSM–specific biological activities.

A three site binding model has been suggested with similarity to that of the LIF-LIFR-gp130 system (see 1.5.1.2, Figure 1-6). Site 3 involves the conserved residues Phe 160 and

Lys 163 which are predicted to reside at the N-terminal end of helix D. Mutagenesis of both these residues has been shown to negate LIFR and OSMRβ binding and biological activity. In Ba/F3 cell based assays mutagenesis of Phe 160 has been shown to have approximately 30-fold lower activity than equivalently mutated Lys 163 on binding to LIFR-gp130, so suggesting that Phe 160 predominates in LIFR binding. Conversely, in Ba/F3 cell based assays Phe 160 has been shown to display approximately 7-fold lower activity on binding to OSMRβ-gp130 over Lys 163, so suggesting that Lys 163 plays a greater contribution towards OSMRβ interactions. Site 3 interactions are presumably mediated through interactions with the LIGHT motif of the putative immunoglobulin-like domain of LIFR (Chobotova and Heath, 2000).

Residues of OSM important for gp130 binding have been mapped to a solvent exposed 'hot spot' centred around residues Gly 120, Asn 124 of helix C and residues Gln 16 and Gln 20 of helix A. These residues are structurally equivalent to those of LIF site 2 (i.e. Gln 25, Ser 28, Gln 23, Ile 121, Gly 124 and Ser 127), all presumably lying within 2-3 Å, and forming a hydrophobic pocket between helices A and C. The three site binding model is further complicated by the finding that OSM can also bind the N-terminal immunoglobulin-like domain of gp130 through site 3 interactions (Staunton *et al.*, 1998). This is thought to allow formation of a pseudosymmetrical complex containing one LIFR (or OSMRβ), one gp130 and two oncostatin M molecules.

Two intramolecular disulphide bridges are present spanning Cys 6-Cys 127 and Cys 49-Cys 167. It has been shown that mutation of Cys 49 and Cys 167 reduces activity ~10-fold (Kallestad *et al.*, 1991) whereas mutations in the first disulphide bridge are not deleterious. Deletion mutants and neutralising antibody studies have revealed that residues either side of the second disulphide are required for activity. This is possibly because the second disulphide functions in anchoring the N-terminal region of helix D on which the site 3

residues Phe 160 and Lys 163 reside, presumably disruptions in this area would adversely affect LIFR and gp130 binding.

1.5.2.3 Oncostatin M

OSM is a monomeric glycoprotein with a molecular weight of ~28 kDa first isolated from supernatants of U-937 histiocytic lymphoma cells treated with phorbol ester (Zarling *et al.*, 1986). OSM is initially expressed as a 252 amino acid precursor polypeptide containing a 25 residue N-terminal hydrophobic signal sequence (Malik *et al.*, 1989). Post-translational processing results in removal of the signal sequence. A further 31 residues are removed from the C-terminus at a pair of basic residues (196 and 195) which act as a proteolytic site, so resulting in the predominant 196 residue form (Linsley *et al.*, 1990). The protein contains two potential N-linked glycosylation sites at residues 100 and 227, which are glycosylated in samples produced in eukaryotic expression systems (Malik *et al.*, 1992).

1.6 Aims of Thesis

This thesis aims to delineate structural features of two key members of the haemopoietin cytokine family; namely OSM and LIFR. This thesis reports the X-ray structure of a rationally designed OSM mutant along with investigations of several expression systems used for production and characterisation of LIFR. The structure of OSM, reported here for the first time, represents a structural framework for the understanding of mutagenesis data. This data will allow functional dissection of the modes of receptor engagement used by the gp130 family of cytokines. The established protocols for production and expression of LIFR will also allow structural investigation of site 3 interactions, which have yet to be analysed structurally. This information will provide further insights into the molecular basis of signal transduction by the haemopoietin cytokines and the epitopes involved.

2. Chapter 2 - Engineering of Truncated Oncostatin M

The aims of this chapter include the design and production of OSM truncation mutants suitable for use in crystallographic studies. Two OSM C-terminal truncation mutants were engineered as GST fusions in the pGEX expression vector. Subsequent expression in *E. coli* and characterisation of OSM is reported in Chapter 3.

Appendix I provides a summary of the reagents and buffers used in the engineering of truncated OSM. The methods involved are discussed in greater detail in texts such as (Maniatis *et al.*, 1989). An overview of the processes involved is shown in Figure 2-1 below.

2.1 Outline of Methods Employed

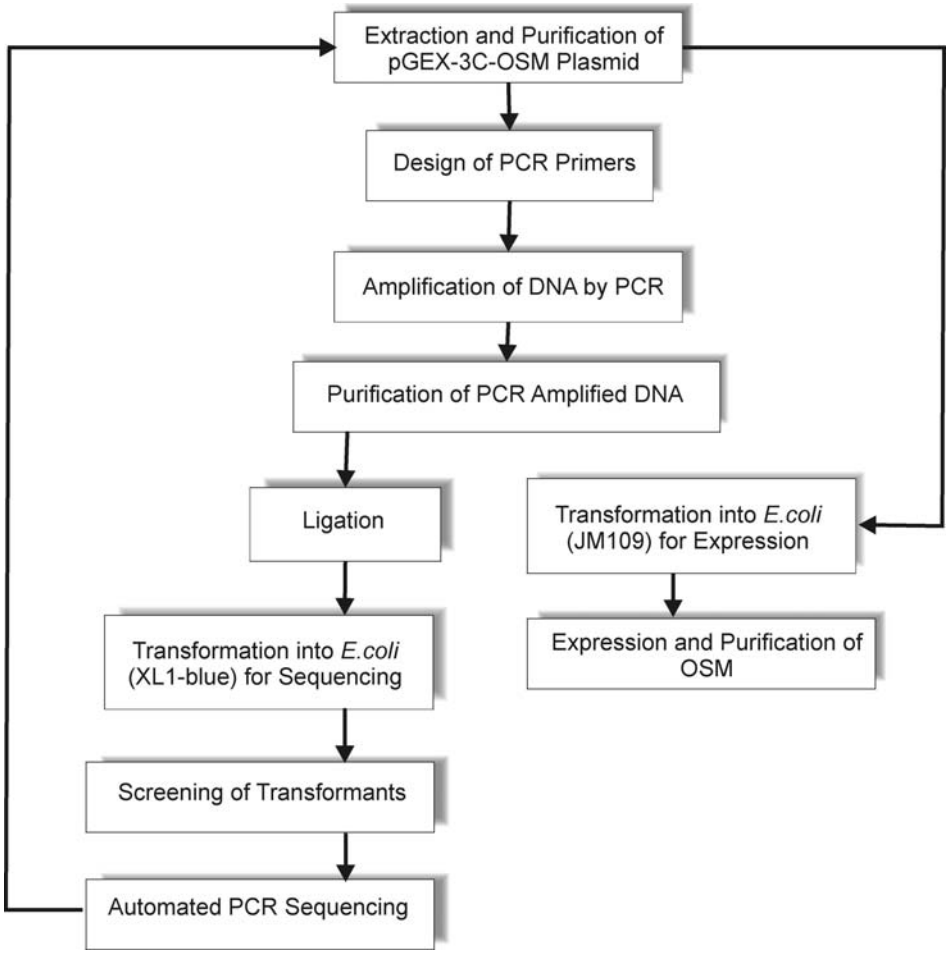


Figure 2-1
Flowchart detailing methods and techniques employed in the course of chapter 2.

2.2 Methods of Engineering

2.2.1 Rational Design of Truncated Oncostatin M

OSM was previously cloned from U-937 cells induced with phorbal myristate acetate (PMA) (Staunton *et al.*, 1998; Zarling *et al.*, 1986). All OSM sequences referred to correspond to that previously published (GenBank accession number M27286; (Malik *et al.*, 1989). A template for the OSM molecule as a glutahione S-transferase (GST) fusion protein was kindly provided in the pGEX vector (Pharmacia)(Keith Hudson and John K. Heath, CRC Growth Factor Group, Birmingham) (see Figure 2-2).

The pGEX-3C-OSM-196 vector consisted of the predominant form of full length mature OSM (OSM-196; 196 residues). A GST moiety (247 residues) was included for use as a purification tag (Smith and Johnson, 1988). An intervening rhinovirus type 14 protease 3C site (protease 3C) was included for cleavage of the GST fusion protein. This protease site consisted of the 11 residue sequence GSEVLFQ⇓GPGS, where ⇓ represents the position of cleavage and underline represents core residues required for efficient cleavage (Leong *et al.*, 1992; Walker *et al.*, 1994). The pTac promoter was utilised for high levels of expression and the Lac gene included for inducible expression (see Figure 2-2).

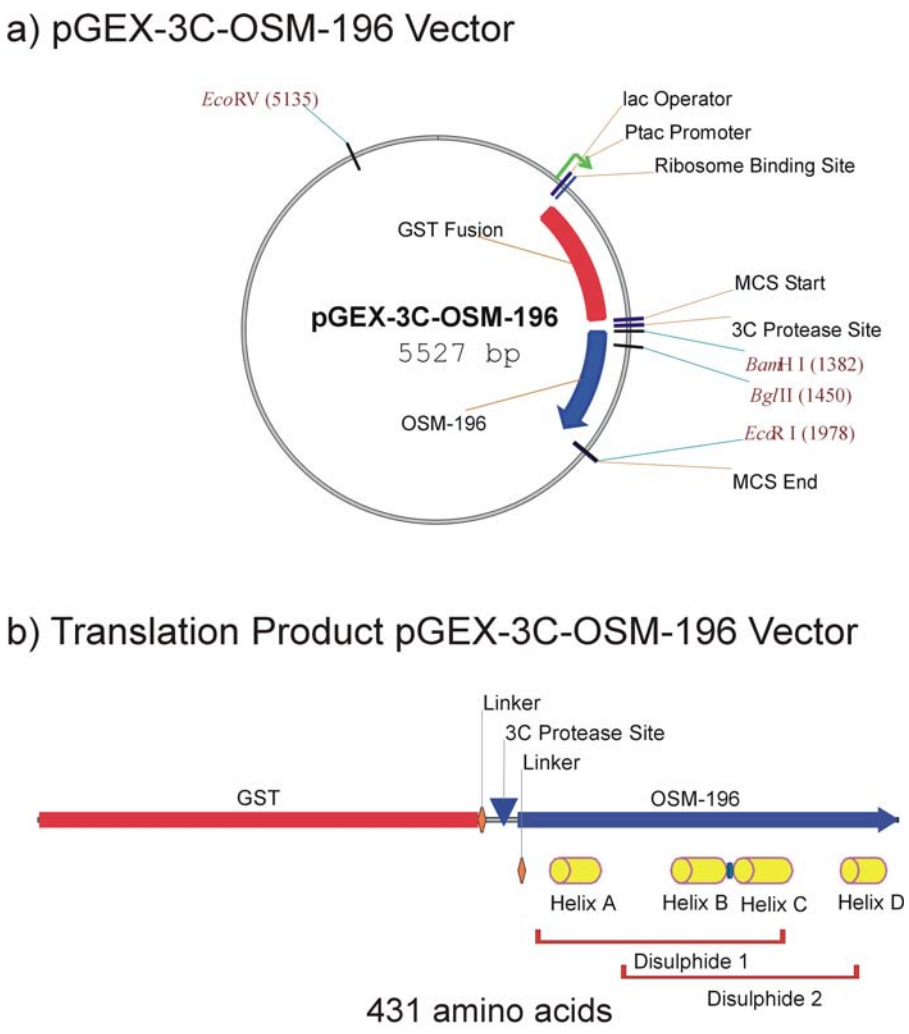


Figure 2-2

Diagrammatic representation of (a) pGEX-3C-OSM-196 vector used for expression of soluble (b) OSM-3C-GST fusions. Multiple cloning site (MCS) is shown along with unique restriction sites in red lettering. Translation product (b) shows predicted helices (yellow cylinders) and disulphide bridges (red).

Previously, it has been shown that OSM-196 is prone to aggregation, low solubilty and dimer formation (Staunton *et al.*, 1998). Crystallisation attempts with this protein proved unsuccessful, due to low solubility and excessive precipitation. In order to produce material suitable for crystallographic studies mutant OSM proteins were engineered. These mutants were based on homology models designed through structural comparison with LIF (Robinson *et al.*, 1994).

Structural alignments of murine LIF and human OSM-196 were carried out using the program THREADER (Miller *et al.*, 1996). Default values were used to produce an alignment resulting in the lowest pairwise energy sum between aligned residues. The resulting alignment was used for modelling as implemented by the package MODELLER (Sanchez and Sali, 1997). Default values were used and a model was generated from murine LIF coordinates (PDB code 1LKI)(Robinson *et al.*, 1994) and the OSM sequence using molecular dynamics for satisfaction of spatial restraints, so relieving any regions of poor stereochemistry.

2.2.2 Design of Oligonucleotides to Produce Truncation Mutants

Oligonucleotide primers were designed to enable amplification of the N-terminal 184 residues (Δ185_OSM) or 187 residues (Δ188_OSM) of OSM-196 using the polymerase chain reaction (PCR) (see Table 2-1). The two C-terminal truncation primers contain a cleavage site for the restriction enzyme *Eco*RI (shown underlined), so allowing cleavage of the resulting PCR amplified DNA fragment. The two C-terminal truncation primers also contain the triplet sequence ATC, which codes for the stop codon UAG (Amber), immediately following the C-terminal residue. These primers all bind directly to the pGEX-3C-OSM-196 plasmid sequence.

Primer Description	Oligonucleotide Sequence
5' pGEX	5'-GCA-TGG-CCT-TTG-CAG-GG-3'
3' Δ185_OSM	5'-GGA-ATT-CTA-GAA-GAC-CCG-CCC-CAC-TGA-GTG-3' Stop Phe Val Arg Gly Val Ser His
3' Δ188_OSM	5'-GGA-ATT-CTA-CCA-CTT-GCT-GAA-GAC-CCG-CCC-3' Stop Trp Lys Ser Phe Val Arg Gly

Table 2-1

Table showing primers used for amplification of Δ185_OSM and Δ188_OSM. EcoRI restriction enzyme site is denoted by underlined bases and corresponding amino acids are written underneath each codon. Stop codon UAG (Amber) is denoted by 'Stop'.

2.2.3 Preparation of Oligonucleotides

Oligonucleotides were prepared by John Fox (AltaBiosciences, Birmingham), using a 308B Applied Biosystems DNA synthesiser. Primers were prepared on a 0.2 μM scale and purified on Sephadex G25. Primers were dissolved in sterile water to 1 mg/ml and stored at –20°C. For use in PCR all primers were further diluted to 200 ng/μl in water. Measurements were based on an absorbance reading of 1 at 260 nm corresponding to an oligonucleotide concentration of 20 μg/ml.

2.2.4 Preparation of Plasmid DNA

E. coli, strain JM109, containing pGEX-3C-OSM-196 were streaked onto LB-Amp plates (Lund-Bertani-ampicillin) (12g agar, 8g tryptone, 4 g yeast extract, 8 g sodium chloride and

50 mg of ampicillin upto 1000 cm^3) (see Appendix I). Plates were grown for 16 hours at 37°C. A single colony was used to inoculate a 5 ml starter culture of LB-Amp media (10 g typtone, 5 g yeast, 10 g sodium chloride and 50 mg of ampicillin upto 1000 cm^3) (see Appendix I). The culture was then grown for 8 hours at 37°C with shaking. This culture was diluted to 500 ml in LB-Amp media and grown for a further 16 hours at 37°C with vigorous shaking. The cells were harvested by centrifugation at 6000 x g for 15 minutes at 4°C and plasmid extracted using a Maxi Prep™ Kit (Qiagen). The purified pGEX-3C-OSM-196 was used as a template for PCR (see 2.2.5) or was digested with *Bam*HI and *Eco*RI and purified (see 2.2.6) for use in ligation with insert (see 2.2.7).

2.2.5 Amplification of DNA by PCR

PCR was carried out in thin walled PCR tubes (Sigma) to aid heat transfer thus reducing the number of cycles required. Several polymerases were tested including Amplitaq™ (Perkin Elmer) and Vent Polymerase™ (New England Biolabs). Amplitaq™ has no 3'-5' exonuclease 'proof reading' activity, so low nucleotide concentrations were used to minimise error levels. Alternatively, Vent Polymerase™ (New England Biolabs) which does possess 3'-5' exonuclease activity was used as the enzyme of choice. This enzyme is known to work on a smaller subset of template and primer combinations. The reaction mixtures used are shown in Table 2-2 below. PCR buffer (10x) contained 100 ng/µl bovine serum albumin (BSA), 1 % Triton X-100, 50 mM ammonium sulphate, 100 mM potassium chloride and 200 mM Tris-HCl, pH 8.2.

Deoxynucleotide triphosphates (dNTP) were supplied as 100 mM solutions in water. A cocktail containing each nucleotide was produced with a final concentration for each nucleotide of 10 mM. A control reaction, containing no template DNA, was run in parallel with the amplification reaction.

	Amplitaq™	Vent Polymerase™
Component	Volume (μl)	Volume (μl)
10x PCR Buffer	20.0	20.0
25 mM magnesium chloride	12.0	_
10 mM dNTP's	1.0	4.0
5' pGEX Primer (200 ng/μl)	1.0	4.0
C-Terminal Truncation PCR Primer (200 ng/μl)	1.0	4.0
pGEX-3C-OSM-196 Template DNA (100 ng)	2.0	2.0
Polymerase	1.0	1.0
Water	162.0	165.0
Total	200.0	200.0

Table 2-2
Table showing composition of PCR reaction mixtures used.

Each reaction mixture was split into two equal aliquots and one drop of mineral oil (Sigma) added to each to prevent condensation. The PCR reaction cycle was carried out automatically using a PTC-200 DNA Engine (MJ Research Inc.). A 94°C 'hot start' was used to prevent 'primer dimer' formation. Preparation of the reaction mixture was carried out on ice and polymerase added last. The reaction mixture was placed immediately into the PCR machine equilibrated to 94°C.

Initial denaturation of the double stranded DNA was carried out by incubation at 96°C for 1 minute, subsequent denaturation cycles were carried out for 30 seconds. Annealing of primers to their complementary sequences was carried out by rapid cooling to 45°C followed by 45 seconds incubation. Primer extension was carried out by incubation of the reaction mixture for 45 seconds at 72°C, with the final extension being increased to 5 minutes. Twenty five cycles of denaturation-annealing-extension were carried in order to strike a balance between complete amplification and maintenance of fidelity. The reaction cycles are summarised below in Table 2-3.

Reaction Step						
1) Denaturation		2) Annealing		3) Extension		Cycles
Temp	Time (min)	Temp	Time (min)	Temp	Time (min)	
96°C	—	—	—	—	—	Start to 1
96°C	1.00	45°C	0.75	72°C	0.75	1 to 2
96°C	0.50	45°C	0.75	72°C	0.75	2 to 23
—	—	—	—	72°C	5.00	23 to End
4°C	∝	—	—	—	—	End

Table 2-3

Table showing temperatures and incubation times used for automated PCR amplification.

The mineral oil was carefully removed by pipette and 5 µl samples of each reaction mixture were added to 1 µl of ethidium bromide (10 mg/ml stock) and 1 µl of DNA loading buffer containing 0.25 % bromophenol blue, 0.25 % xylene cyanol FF and 30 % glycerol (see Appendix I). The samples were run on a 1 % agarose gel in TBE buffer (54 g Tris-base, 27.5 g boric acid and 20 ml 0.5 M ethylenediaminetetraacetic acid (EDTA), pH 8 upto 500 ml for 5x stock) (see Appendix I). The gel was run for approximately 30 minutes at 60 volts and the bands observed using a UV transilluminator (Ultra-Products Inc.). The amount of PCR product was estimated by comparison with the brightness of a known quantity of equimolar λ *Hind*III digestion fragments (Gibco).

2.2.6 Purification of PCR Amplified DNA

The PCR amplified DNA was doubly digested with *Eco*RI and *Bam*HI to remove template DNA. The reaction mixture was setup as below in Table 2-4. Restriction buffer B contained 1 mM β-mercaptoethanol, 100 mM sodium chloride, 5 mM magnesium chloride and 10 mM Tris-HCl, pH 8.0.

Volume (µl)	Component
40.0	PCR Product
8.0	10x Restriction Buffer B
1.0	*Eco*RI (40 units/ml)
6.0	*Bam*HI (10 units/ml)
25.0	Sterile Distilled Water
80.0	Total

Table 2-4
Table showing composition of restriction enzyme reaction mixture used to remove template DNA from amplified PCR products.

Digestion was carried out by incubation in a water bath at 37°C for 1.5 hours. DNA loading buffer and ethidium bromide were added as previous and the samples run in two large lanes on a 1 % agarose gel in TAE buffer (242 g Tris-base, 57.1 ml glacial acetic acid and 100 ml 0.5 M EDTA, pH 8 upto 500 ml for 50x stock) (see Appendix I). The lower band was excised from the gel using a sterile razor blade and then trimmed tightly to remove unwanted agarose. The band was placed in a 1.5 ml eppendorf tube and extraction of the digested PCR product was carried out using a Qia-Quick PCR Purification Kit™ (Qiagen).

2.2.7 Ligation

Prior to ligation the concentration of the final purified PCR product was determined from a 1 % agarose-TAE gel. Ligation was carried out at a molar ratio of 10:1 (insert:vector). Reaction mixtures containing the enzyme T4 ligase (Boehringer Mannheim) were used as in Table 2-5 below. Ligation buffer contained 10 mM ATP, 50 mM magnesium chloride, 10 mM DTT and 660 mM Tris-HCl, pH 7.5. The vector was previously digested with *Bam*HI and *Eco*RI and purified as in 2.2.6. Acetylated BSA was included to stabilise and protect the T4 ligase enzyme. A control reaction containing no PCR product was set-up along side and both reactions incubated at room temperature for 4 hours to allow ligation to proceed.

Volume (µl)	Component
0.5	Purified PCR Product
1.0	pGEX-3C Vector (*Bam*HI–*Eco*RI cut) (2 ng/ml)
1.0	10x Ligation Buffer
1.0	Acetylated BSA (100 ng/µl)
1.0	T4 Ligase (5 units/ml)
6.0	Sterile Distilled Water
10.0	Total

Table 2-5

Table showing composition of ligation reaction mixture used to ligate PCR product to vector DNA.

2.2.8 Preparation of Competent Cells

E. coli strain XL1-Blue (Invitrogen), were streaked onto LB plates and grown for 16 hours at 37°C. A single colony was picked and used to inoculate 5 ml of LB media. The culture was grown, with shaking, overnight at 37°C. This starter culture was diluted to 200 ml of LB and then further grown to an OD_{600} ~0.5. The cells were harvested at 3000 x g and the pellet resuspended gently in 80 ml of sterile ice-cold 50 mM calcium chloride. The cells were left on ice for 20 minutes and then harvested at 3000 x g. Finally, the cells were gently resuspended in 20 ml of sterile ice cold 100 mM calcium chloride containing 15 % glycerol. The cells were aliquoted in 0.5 ml portions and stored at -70°C.

2.2.9 Transformation

An aliquot of competent *E. coli*, strain XL1-Blue (Invitrogen), was thawed slowly at room temperature. The cells were gently mixed and three transformation reactions set up; experimental, negative control and positive control. For the experimental ligation 100 µl of competent cells were gently added to 5 µl of ligation reaction mixture as in Table 2-5. The cells were incubated on ice for 20 minutes and then 'heat shocked' at 42°C for 90 seconds. The cells were returned to ice for 1 minute. After addition of 700 µl of LB media containing 20 mM sucrose, 10 mM magnesium sulphate and 10 mM magnesium chloride the culture was grown for 1 hour at 37°C. The cells were harvested at 5000 x g for 2 minutes and 700 µl of the supernatant was removed. The cells were gently mixed and then plated out onto an LB-Amp plate. The plates were grown for 16 hours at 37°C and the number of transformant colonies on each plate recorded. Positive and negative controls using 5 µl of pUC DNA (200 pg/µl) or 5 µl of water instead of ligation reaction mixture were prepared in a similar fashion.

2.2.10 Screening of Transformants

Colonies grown on LB-Amp plates were isolated and used to inoculate 5 ml of LB-Amp media. The cultures were grown for 4 hours in 50 ml centrifuge tubes at 37°C with shaking. For each culture 3 ml of cell suspension was removed and the cells harvested by centrifugation at 15000 x g into eppendorf tubes. The plasmid DNA was extracted using a Mini-Prep Plasmid Extraction Kit™ (Qiagen).

Extracted plasmid samples were doubly digested with *Eco*RV and *Bgl*II to check for correct ligation. The vector contains a unique *Eco*RV site and the insert contains a unique *Bgl*II site. Restriction enzyme digests were set up as in Table 2-6 and the reaction incubated at 37°C for 2 hours. The digested samples were run on a 1 % agarose-TAE gel.

Volume (µl)	Component
4.00	Extracted Plasmid DNA
1.00	10x Restriction Buffer B
0.50	*Bgl*II (40 units/µl)
0.25	*Eco*RV (40 units/µl)
4.25	Sterile Distilled Water
10.00	Total

Table 2-6

Table showing composition of restriction enzyme reaction mixture used to analyse pGEX-3C-OSM_Δ185 and pGEX-3C-OSM_Δ188 plasmids.

2.2.11 Automated PCR Sequencing

Automated PCR sequencing was carried out using cycle sequencing with dye-labelled terminators. Sequencing reactions were setup using the ABI PRISM BigDye™ Terminator Cycle Sequencing Kit (PE Applied Biosystems) according to the recipe in Table 2-7. Three primers (5', 3' and internal to the insert) were used to ensure accurate reading through all areas of the insert (see Table 2-8). The BigDye™ reaction mixture contains all necessary components including BigDye terminators™, dNTP's, AmpliTaq FS™ polymerase and magnesium chloride.

5' - Sequencing Reaction		Internal - Sequencing		3' - Sequencing Reaction	
Volume (µl)	Component	Volume	Component	Volume (µl)	Component
6.0	Plasmid	6.0	Plasmid DNA	6.0	Plasmid DNA
8.0	BigDye™	8.0	BigDye™	8.0	BigDye™
1.0	Primer	1.0	Primer	1.0	Primer
5.0	Sterile water	5.0	Sterile water	5.0	Sterile water
20.0	Total	20.0	Total	20.0	Total

Table 2-7
Table showing composition of reaction mixtures used for automated PCR sequencing.

Primer Description	Oligonucleotide Sequence
5' pGEX Primer	5'-GCA-TGG-CCT-TTG-CAG-GG-3'
Internal OSM Primer	5'-AGT-GAG-GAG-ACC-CTG-AGG-3'
3' pGEX Primer	5'-TCC-GCT-TAC-AGA-CAA-GCT-G-3'

Table 2-8
Table showing oligonucleotides used for automated PCR sequencing.

Amplification of DNA for sequencing was carried out using automated PCR cycling on a DNA Thermal Cycler 480™ (Perkin Elmer Applied Biosystems). Denaturation of the double stranded DNA was carried out by incubation at 96°C for 10 seconds. Annealing of primers to their complementary sequences was carried out by incubation for 5 seconds at 50°C. Primer extension was carried out by incubation of the reaction mixture for 4

minutes at 60°C. Twenty five cycles of denaturation-annealing-extension were carried out. The reaction cycles are summarised below in Table 2-9.

Reaction Step						
1) Denaturation		2) Annealing		3) Extension		Cycles
Temp	Time (sec)	Temp	Time (sec)	Temp	Time (sec)	
96°C	10	50°C	5	60°C	240	25

Table 2-9
Table showing temperatures and incubation times used for automated PCR sequencing.

Following PCR, 20 µl of reaction mixture was pipeted into thin walled PCR tubes and 2 µl of 3 M sodium acetate, pH 5.2 and 50 µl of ethanol were added. The contents were mixed and incubated on ice for 10 minutes. The sample was centrifuged at 15000 x g for 30 minutes at room temperature to pellet the DNA. The pellet was aspirated and washed with 250 µl of 70 % ethanol. The DNA was dried under vacuum ensuring that all ethanol was removed. Samples were resuspended in DNA loading buffer and loaded onto a 36 cm 5% Long Ranger™ gel (PE Applied Biosystems) and run on an ABI PRISM 377 DNA Sequencer (AltaBiosciences, Birmingham).

The output sequence was aligned with the target pGEX-3C-OSM-196 sequence using the BESTFIT command within the GCG Package (Wisconsin Package Version 9.0, Genetics Computer Group, Madison, Wisconsin, USA). This produces an optimal alignment using the local homology algorithm of Smith and Waterman (Waterman *et al.*, 1984) so enabling easy analysis of mutations.

Clones containing correct OSM sequences were selected and plated out on LB-Amp plates and grown overnight at 37°C. Individual colonies were selected and grown in 5 ml of LB-

Amp media overnight at 37°C. Plasmid was extracted as in 2.2.4 and used to transform *E. coli*, strain JM109, as detailed in 2.2.9.

2.3 Results of Engineering

2.3.1 Rational Engineering of Truncated Oncostatin M

Structural/sequence alignments of human OSM-196 and murine LIF reveal an extended C-terminal region in human OSM-196 (see panel (a) Figure 2-3). Homology models based on this alignment, using secondary structure patterns observed in murine LIF, also suggest that OSM-196 may have an extended main chain conformation for residues beyond the equivalent end points of helix D in murine LIF (see panel (b) Figure 2-3). Truncations were rationally engineered to remove the C-terminal region of human OSM-196, on the premise that these additional residues, while surplus to structural stability and function could actually be hindering crystallisation.

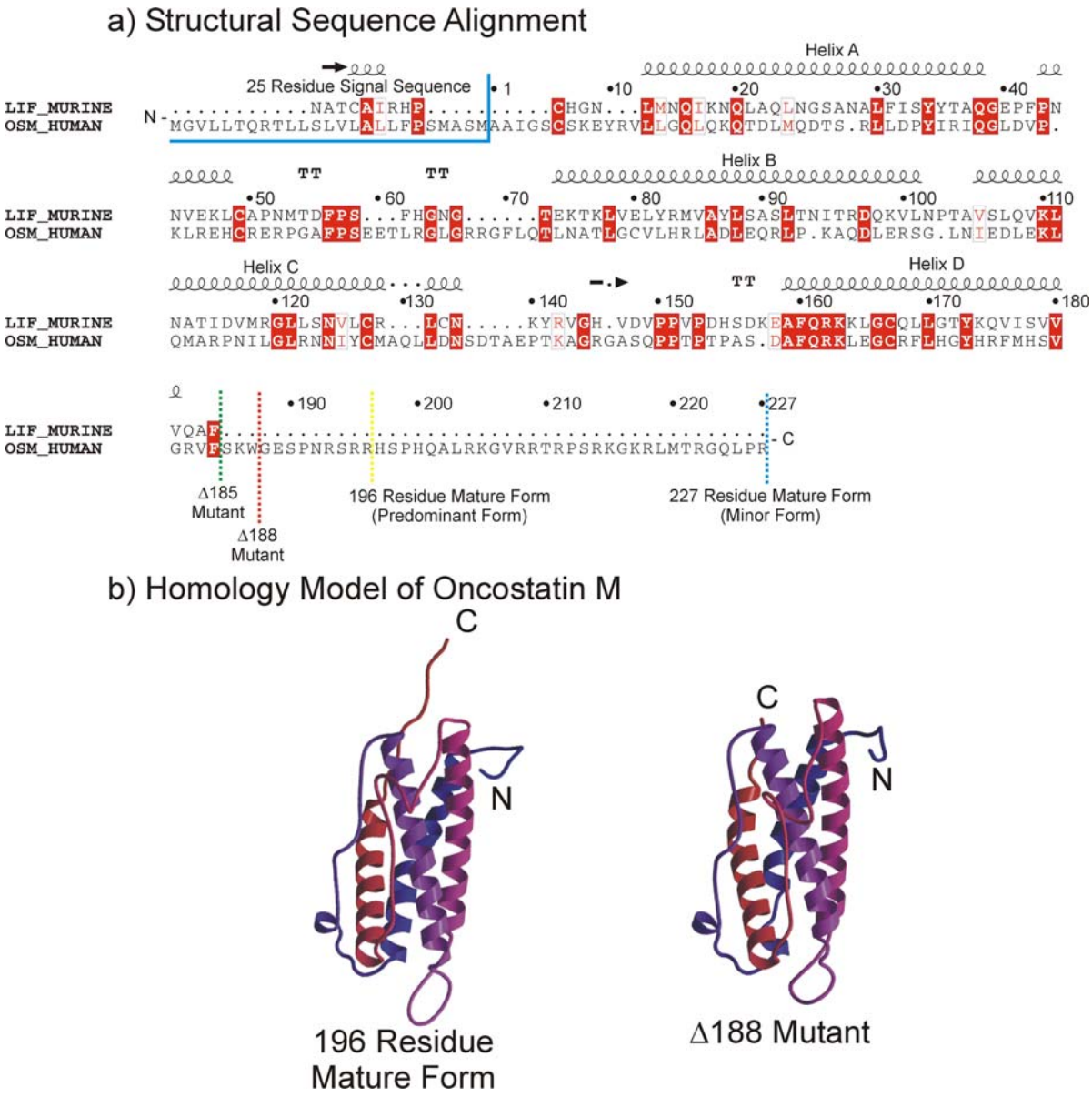


Figure 2-3

Figure showing (a) structural/sequence alignment of human OSM and murine LIF and (b) homology models of human OSM based on this alignment. The 25 residue signal sequence of human OSM is denoted by the blue line and numbering is according to residues of mature human OSM. The various truncation points of human OSM are denoted by dashed lines. Secondary structure of murine LIF (Robinson *et al.*, 1994) is shown as arrows (β-strand), coils (α-helix) or turn (TT). Fully conserved residues are denoted as white on red and conservative substitutions are shown in red lettering. Figure produced using DSSP (Kabsch and Sander, 1983), THREADER (Miller *et al.*, 1996) and Espript (Gouet *et al.*, 1999). Polypeptide is shown coloured from blue at the N-terminus through to red at the N-terminus. Models shown in (b) produced using MODELLER (Sanchez and Sali, 1997), BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

2.3.2 Amplification of DNA by PCR

Two truncation mutants, Δ185_OSM and Δ188_OSM, were produced by amplification using PCR. Figure 2-4 below shows a 1 % agarose gel of the PCR products of Δ185_OSM amplification. The DNA products were visualised under UV illumination as a result of ethidium bromide binding. Lane 4 shows the five main marker bands of the λ *Hind*III digest. Starting at the top these bands correspond to 9146, 6882, 2322, 2027 and 564 (not visible in photograph) base pairs. Lanes 2 and 3 contain duplicate samples of the Δ185_OSM reaction mixture after PCR amplification. These lanes show a single band running at ~700 base pairs corresponding to the amplified PCR product of Δ188_OSM. Lane 1 shows the control reaction containing no template DNA in which no bands can be observed. Analysis of Δ188_OSM samples produced a similar result (data not shown).

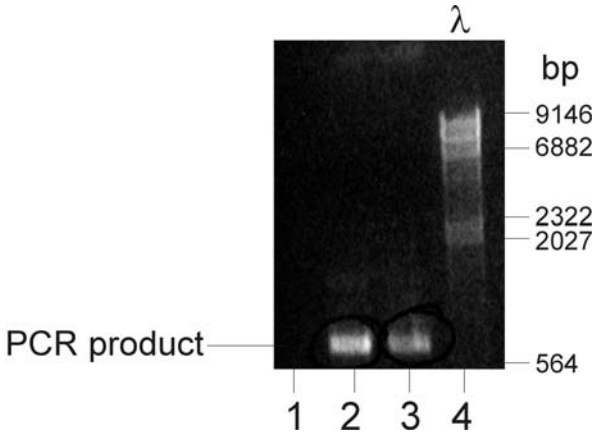


Figure 2-4

Figure showing 1 % agarose gel of Δ185_OSM amplified PCR products. Lanes 4 show the five main marker bands of the λ *Hind*III digest, running at 9146, 6882, 2322, 2027 and 564 (not visible in photograph) base pairs. Lanes 2 and 3 show the amplified PCR product of Δ188_OSM at ~700 base pairs. Lane 1 shows a control PCR reaction, containing no template DNA.

2.3.3 Purification of PCR Amplified DNA

Figure 2-5 below shows a 1 % agarose gel of purified PCR product which has been digested with *Bam*HI and *Eco*RI. Lane 5 shows the five main marker bands of the λ *Hind*III digest. Lanes 3 and 4 contain duplicate *Bam*HI and *Eco*RI digested Δ185_OSM and each shows two distinct bands corresponding to PCR product at ~700 base pairs and *Bam*HI-*Eco*RI digestion fragment at ~560 base pairs. Lane 1 shows a control digestion reaction containing no DNA and no bands are observed. Lane 2 shows a control digestion reaction containing no restriction enzyme and one band is observed corresponding to PCR product. Analysis of Δ188_OSM samples produced a similar result (data not shown).

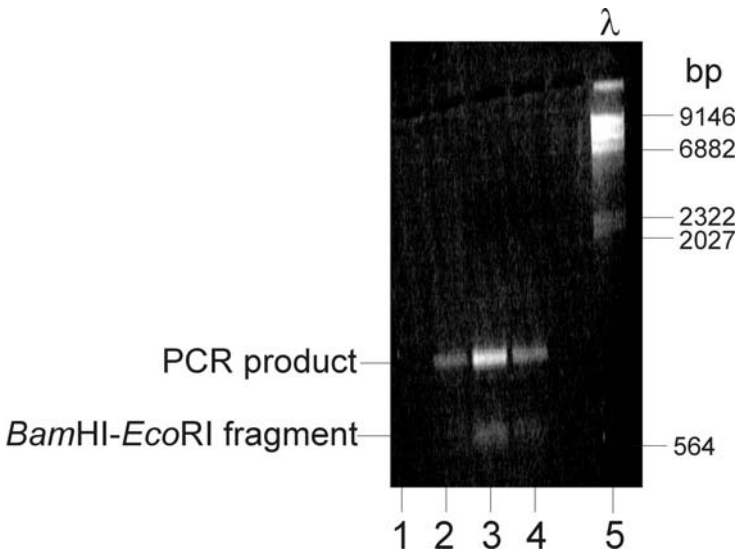


Figure 2-5

Figure showing 1 % agarose gel of *Bam*HI and *Eco*RI digestion of purified Δ185_OSM PCR product. Lane 5 shows the five main marker bands of the λ *Hind*III digest, running at 9146, 6882, 2322, 2027 and 564 (not visible in photograph) base pairs. Lanes 3 and 4 each show two distinct bands corresponding to PCR product at ~700 base pairs and *Bam*HI-*Eco*RI digestion fragment at ~560 base pairs. Lane 1 shows a control digestion reaction containing no DNA and no bands are observed. Lane 2 shows a control digestion reaction containing no restriction enzyme and one band is observed.

2.3.4 Transformation

Table 2-10 below shows the observed transformation frequencies of pGEX-3C-Δ185_OSM and pGEX-3C-Δ188_OSM into *E. coli* JM109. Typically, 3 and 12 transformants were produced for pGEX-3C-Δ185_OSM and pGEX-3C-Δ188_OSM, respectively. Over 2000 colonies were observed for cells transformed with pUC DNA. No colonies were observed in control cells transformed with water.

Plate	Number of Colonies Observed
pUC transformed cells (positive control)	>2000
Non-transformed cells (negative control)	0
pGEX-3C-Δ185_OSM transformed cells	3
pGEX-3C-Δ188_OSM transformed cells	12

Table 2-10

Table showing the number of colonies obtained after transformation of E.coli JM109 with pGEX-3C-Δ185_OSM and pGEX-3C-Δ188_OSM. Negative and positive controls are also shown.

2.3.5 Screening of Transformants

Figure 2-6 below shows a 1 % agarose gel of pGEX-3C-Δ185_OSM which has been digested with *Eco*RV and *Bgl*II. Lane 4 shows the five main marker bands of the λ *Hind*III digest. Lanes 1, 2 and 3 each show two distinct bands at 3685 and 1815 base pairs.

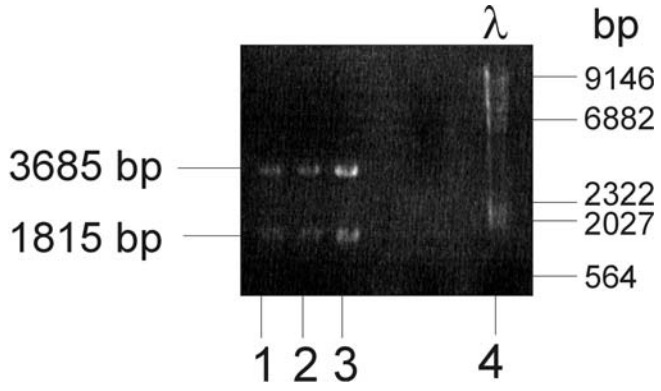


Figure 2-6

Figure showing 1 % agarose gel of *Eco*RV and *Bgl*II digestion of pGEX-3C-Δ185_OSM. Lane 4 shows the five main marker bands of the λ *Hind*III digest, running at 9146, 6882, 2322, 2027 and 564 (not visible in photograph) base pairs. Lanes 1, 2 and 3 contain triplicate digested samples and each shows two distinct bands at 3685 and 1815 base pairs.

2.3.6 Automated PCR Sequencing

pGEX-3C-Δ185_OSM and pGEX-3C-Δ188_OSM sequences were verified against the published OSM sequence (GenBank accession number M27286; (Malik *et al.*, 1989). Figure 2-7 below shows the position of insertion into the pGEX-3C-GST plasmid and the four primers used for sequencing are marked. All sequences were correct with any discrepancies in the alignment explained by misreading or ambiguity in interpretation of the fluorescence signal.

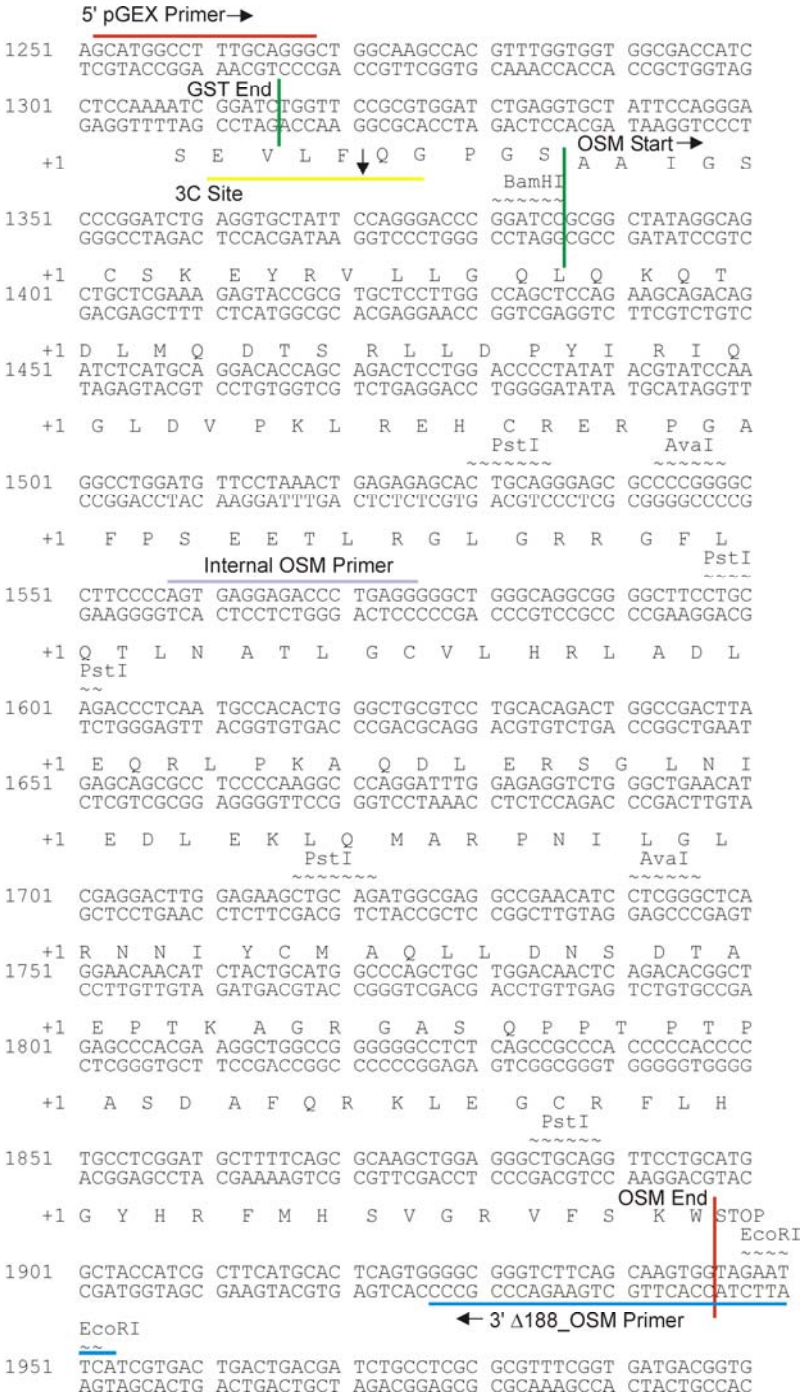


Figure 2-7

Figure showing sequence of pGEX-3C-OSM_Δ188 construct in region around OSM insert. 'OSM start' and 'OSM end' denote OSM insert. Primers used are shown along with their reading direction. OSM sequence corresponds to that previously published (GenBank accession number M27286; (Malik *et al.*, 1989).

2.4 Discussion of Engineering

Truncation and amplification of both Δ185_OSM and Δ188_OSM only proved successful using Amplitaq™ (Perkin Elmer) polymerase. No amplification was observed (data not shown) using Vent Polymerase™ (New England Biolabs). Vent Polymerase™ contains 3'-5' exonuclease activity for proof reading of amplified DNA and for normal applications is used as the enzyme of choice. Vent Polymerase™ is known to work on a smaller subset of template and primer combinations, so suggesting a possible reason for the lack of amplification in this study. In an attempt to maintain the fidelity of amplification low dNTP and magnesium concentrations were used for all PCR reactions, as Ampltaq™ contains no 3'-5' exonuclease 'proof reading' activity. Lanes 2 and 3 of Figure 2-4 show amplified DNA running at ~700 base pairs corresponding to Δ185_OSM PCR products. Comparison with lane 1 of Figure 2-4 shows that this band is not present in reactions containing no template DNA.

The amplified DNA was successfully digested with the restriction enzymes *Bam*HI and *Eco*RI to remove the template DNA. Lane 3 and 4 of Figure 2-5 each show two distinct bands at ~560 base pairs and ~700 base pairs, corresponding to cut and uncut PCR products of Δ185_OSM, respectively.

The purified insert was successfully ligated into the pGEX-3C-GST vector. The observed ligation efficiencies were very low (see Table 2-10) with typically only 3 and 12 transformant colonies for Δ185_OSM and Δ188_OSM respectively. Control plates transformed with pUC DNA contained over 2000 colonies, so confirming the competence of the cells used. Presumably, Δ185_OSM constructs are toxic to cells and basal levels of expression result in the death of transformed cells. Possible explanations for this observation include aberrant protein folding resulting from incorrect prediction of the

length of helix D. This is supported by the observation that Δ188_OSM constructs appear less toxic, so suggesting a more suitable truncation point has been found in this mutant.

All colonies were screened for correct ligation using digestion with *Eco*RV and *Bgl*II. Lanes 1, 2 and 3 of Figure 2-6 shows two distinct bands at 3685 and 1815 base pairs, so confirming the presence of correctly ligated insert. All transformants showed correct ligation and subsequent PCR sequencing, using three primers (internal to insert, 3' to insert and 5' to insert), confirmed that both constructs showed agreement with published pGEX-3C-OSM-196 sequences.

3. Chapter 3 - Expression, Purification and Characterisation of Oncostatin M

The aims covered this chapter include the development of protocols to allow production of OSM in sufficient quantities for crystallographic studies. This chapter also describes work towards characterising OSM and assessment of its suitability for such studies. Expression and purification of a biologically functional OSM mutant is reported. Subsequent crystallisation and X-ray structure determination is reported in Chapter 4 and Chapter 5.

Appendix II is included as a summary of the reagents and buffers used in the expression, purification and characterisation of OSM. The methods involved are discussed in greater detail in texts such as (Maniatis *et al.*, 1989), (Coligan *et al.*, 1995) and (Deutscher, 1990). An overview of the processes involved is shown in Figure 3-1 below.

3.1 Outline of Methods Employed

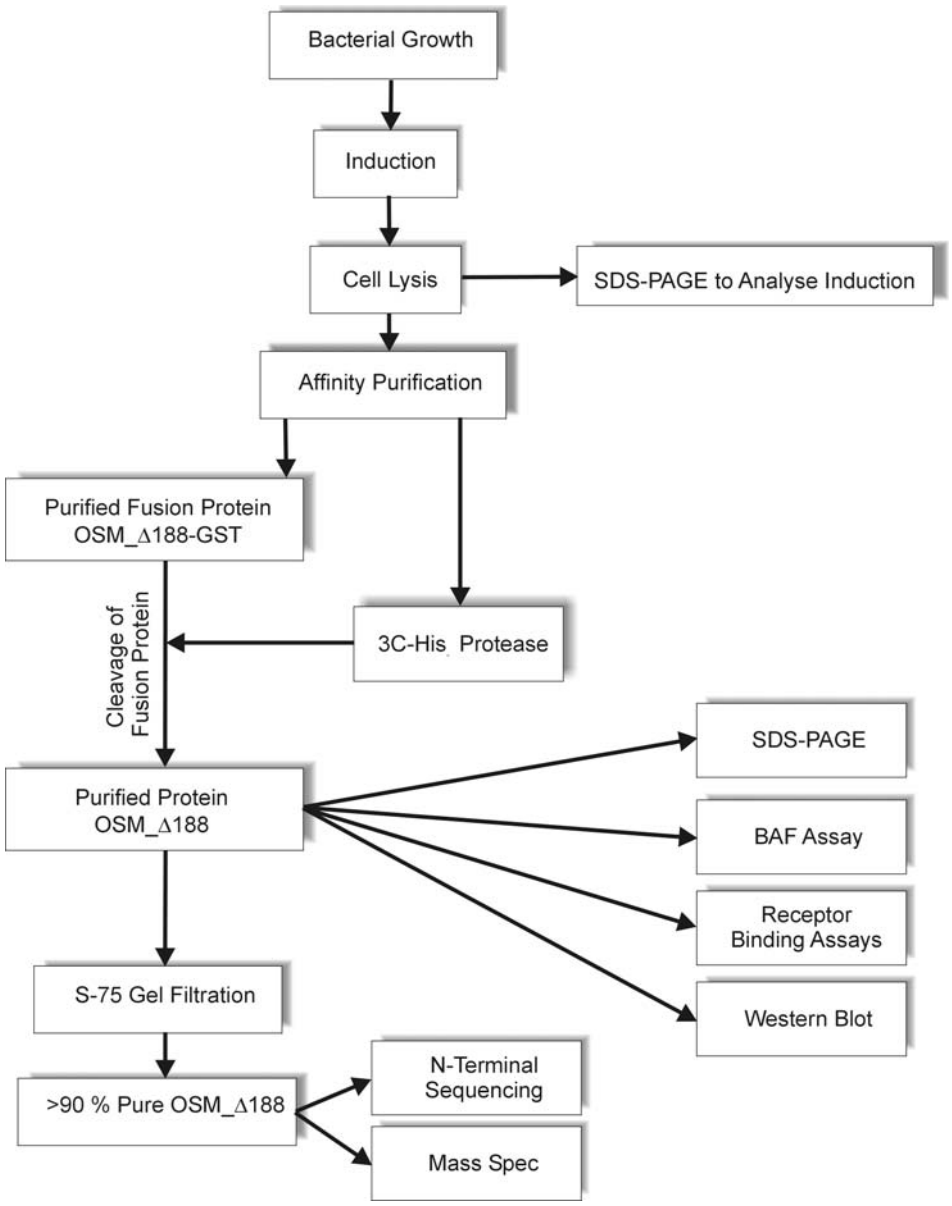


Figure 3-1
Flowchart detailing methods and techniques employed in the course of chapter 3.

3.2 Methods of Expression, Purification and Characterisation of Oncostatin M

3.2.1 Bacterial Growth and Lysis

Typically 500 ml aliquots of LB-Amp media were inoculated with a colony of *E.coli* strain JM109 transformed with pGEX-3C-OSM. The cultures were incubated at 37°C, with shaking, overnight. Samples of culture were diluted 100-fold and the culture further grown at 37°C to an OD_{600} ~0.6. The culture was transferred to 26°C and after 20 minutes isopropyl β-D-galactopyranoside (IPTG) was added to 0.1 mM and the culture then grown for a further 24 hours at 20°C. The cells were harvested by centrifugation at 6000 x g for 10 minutes and the pellet resuspended in 5 ml of ice-cold 150 mM sodium chloride, 16 mM disodium hydrogen phosphate and 4 mM sodium dihydrogen phosphate (MTPBS) (see Appendix II). MTPBS buffer was supplemented with Complete™ protease inhibitor tablets (Boehringer Mannheim), 0.1 % (v/v) glycerol and 0.1 % (v/v) Tween-20. The cells were lysed by sonication with four 30 second pulses at full power and Triton X-100 added to 1 % (v/v). The lysate was left on ice for 5 minutes and then centrifuged at 16000 x g for 20 minutes at 4°C. The supernatant was purified by affinity chromatography using glutathione sepharose resin as below (see 3.2.2). Samples were retained for analysis on sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (see 3.2.4).

3.2.2 Affinity Purification using Glutathione Sepharose

A chromatography support (Bio-Rad Poly-Prep) was packed with 2 ml of glutathione sepharose (Pharmacia) and washed with 10 ml of MTPBS (see Appendix II) followed by 10 ml of MTPBS containing 1 % (v/v) Triton X-100. The supernatant was passed through the column under gravity at 4°C and the column washed with 10 ml of MTPBS containing 1 % (v/v) Triton X-100, followed by 10 ml of MTPBS. Finally, the column was washed with 150 mM sodium chloride, 10 mM EDTA, 1 mM dithiothreitol (DTT) and 50 mM Tris-HCl, pH 8.0 (TNED)(see Appendix II). GST-fusions were eluted with 0.5 ml aliquots

of 10 mM reduced glutathione in 20 mM Tris-HCl, pH 8.0 (see Appendix II). Alternatively, cleavage of fusion proteins was carried out with 3C-His_6 protease (see 3.2.3). Fractions were assayed using Coomassie Protein Plus Assay (Pierce) (see 3.2.5) and SDS-PAGE (see 3.2.4).

3.2.3 Cleavage of Fusion Protein

Glutathione sepharose resin was equilibrated with 2 volumes of TNED buffer (see Appendix II). 3C-His_6 protease was added to the resin so forming a 50 % (v/v) slurry. Protease was used at 100 μg for every 1 mg of GST-fusion protein. The resin was incubated at room temperature for 24 hours, with occasional mixing. Protein was eluted using 0.5 ml aliquots of MTPBS and the fractions analysed using Coomassie Protein Plus Assay (Pierce) (see 3.2.5), western blot and SDS-PAGE (see 3.2.4).

3.2.3.1 Production of 3C Protease

pTrc-3C-His_6 (Invitrogen) (kindly provided by M. A. Hall and Dr. K. Hudson, Birmingham) were grown and processed as above (see 3.2.1) except for the omission of protease inhibitors and the addition of 50 mM imidazole to all buffers used in the purification. The supernatant was purified by affinity chromatography using nickel agarose resin as below (see 3.2.3.2).

3.2.3.2 Affinity Purification of 3C Protease

A chromatography support (Bio-Rad Poly-Prep) was packed with 1 ml of nickel agarose resin (Qiagen) and washed with 10 ml of MTPBS (see Appendix II) containing 50 mM imidazole. The supernatant was passed through the column under gravity at 4°C. The column was washed with 10 ml of MTPBS containing 50 mM imidazole followed by 10 ml of TNED buffer containing 50 mM imidazole. Finally, the protein fractions were eluted with 0.5 ml aliquots of TNED buffer supplemented 250 mM imidazole. The fractions

were assayed for protein using Coomassie Protein Plus Assay (Pierce) (see 3.2.5) and SDS-PAGE (see 3.2.4).

3.2.3.3 Analysis of Fusion Protein Cleavage Using Protease 3C

Analysis of cleavage was carried out using the eight reactions as in Table 3-1 below. The reactions were incubated at room temperature for 16 hours. Aliquots of each reaction were removed and run on 12 % SDS polyacrylamide gels (see 3.2.4):

Tube	OSM-196-3C-GST (μg)	3C-GST (μg)	DTT (mM)
1	10	1	_
2	10	5	_
3	10	20	_
4	10	40	_
5	10	1	0.9
6	10	5	0.8
7	10	20	0.6
8	10	40	0.4

Table 3-1
Table showing reactions used for analysis of fusion protein cleavage by 3C protease.

3.2.4 SDS-PAGE Analysis

Typically, 1ml samples of culture, before and after induction with IPTG, were centrifuged at 14000 x g and resuspended in 100 μl of protein loading buffer consisting of 10 % (v/v) glycerol, 5 % (v/v) β-mercaptoethanol, 2 % (w/v) SDS, 0.001 % (w/v) bromophenol blue and 62.5 mM Tris-HCl, pH 6.8 (see Appendix II). Magnesium chloride was added to 1 mM, to precipitate genomic DNA, and the sample vortexed. 20 μl of samples from all other steps of the preparation were retained and added to 60 μl of protein loading buffer. Samples were heated to 95^{o}C for 5 minutes, to denature any proteases, and then centrifuged at 14000 x g for 10 minutes. Typically, 20 μl samples were loaded onto 12 % SDS-polyacrylamide gels (Laemmli, 1970) and electrophoresed at 150 V for approximately

1.5 hours. Broad range rainbow molecular weight markers ranging from 210 kDa to 8 kDa (Bio-Rad) were run along side the samples (see Appendix II). Gels were stained using 0.2 % Coomassie Blue R-250 in 80 % (v/v) methanol and 20 % (v/v) acetic acid (see Appendix II) for ~0.5 hours and destained in 80 % (v/v) methanol and 20 % (v/v) acetic acid overnight.

3.2.5 Coomassie Protein Assay

Protein concentrations were determined using a Coomassie Protein Plus™ Assay Kit (Pierce). Standard curves were constructed using bovine serum albumin (BSA) concentrations in the range 0 to 1.6 mg/ml. Typically, 50 µl of sample was added to 1.5 ml of Coomassie Protein Plus™ Reagent and the absorbance measured at a wavelength of 595 nm using a Beckman DU-640 spectrophotometer. Protein concentrations were determined by comparison with the absorbance measurements of BSA standards. The accuracy of this assay was confirmed by comparison with direct absorbance measurements at 280 nm.

3.2.6 Western Blot Analysis

Proteins were separated on a 4-15 % gradient SDS polyacrylamide gel (Bio-Rad). The gel was equilibrated in western blot transfer buffer consisting of 0.3 % (w/v) Tris-base, 1.4 % (w/v) glycine and 10 % (v/v) methanol (see Appendix II) for 1 hour. Electrophorectic transfer of proteins was routinely performed using a Mini-Protean II wet blotter (Bio-Rad). The gel was placed on a pre-wetted section of 0.45 µM nitrocellulose (Schleicher and Schuell) trimmed to the dimensions of the gel. Three layers of pre-wetted filter paper (Whatmann 3MM™) were placed on each side and the sandwich blotted overnight at 30 V or for 1 hour at 100 V at 4 ^{o}C.

After transfer, the nitrocellulose was removed and blocked overnight with 5 % (w/v) milk (Marvel™) in phosphate buffered saline (PBS) (see Appendix II). Primary polyclonal antibody (sheep α-oncostatin m; Binding Site, U.K., 8.7 mg/ml) was added at a dilution of 1: 10000 in PBS containing 5 % (w/v) milk and 0.1 % (v/v) Tween-20. The blot was incubated at room temperature for 1 hour, with agitation, and then extensively washed in PBS containing 0.1 % (v/v) Tween-20. Secondary polyclonal antibody (donkey α-sheep IgG1-peroxidase conjugated, Binding Site, U.K) was applied at a dilution of 1:5000 in PBS as above. The blot was incubated at room temperature for 1 hour, with agitation, and then extensively washed in PBS as above. Finally, the blot was washed with PBS and luminescent detection of the horseradish peroxidase-labelled secondary antibody carried out using ECL™ SuperSignal Substrate (Amersham). Blots were wrapped in ClingFilm™ and then covered with HyperFilm™ (Amersham) and exposed for 10–60 seconds. Films were developed automatically (X-Omat).

3.2.7 Seleno-methionine Labelled Oncostatin M

Selenium labelled OSM (^{78}Se-Met) was prepared according to the method of Hendrickson (Hendrickson *et al.*, 1990). pGEX-3C-Δ188_OSM constructs were transformed as previously (see Chapter 2 for method) into methionine auxotrophic strains of *E.coli* in an attempt to maximise ^{78}Se-met incorporation. Strains tested included B834 and DL41, which were both obtained from the European culture collection (ECACC).

Transformed cells were grown in defined LeMasters medium (LeMaster and Richards, 1982) containing glucose, seventeen amino acids, nucleotide bases, various salts and sulphates. The media is produced in two portions; a non-autoclavable portion and a autoclavable portion (see Table 3-2). Autoclavable portions were made upto 4 L with water and the supersaturated solution was autoclaved immediately. The non-autoclavable portion was made upto 400 ml with water and then filter sterilised through a 0.22 μM filter (Whatman). Both the autoclavable and non-autoclavable fractions were stored at room temperature, to prevent precipitation. Once cool the non-autoclavable fraction was added to the autoclavable fraction and the medium used immediately.

Cell growth was carried out as previously (see 3.2.1) except only 2 ml of starter culture was used in an attempt to reduce the incorporation of non labelled methionine. Affinity purification and cleavage of the fusion proteins was also carried out as previously (see 3.2.2 and 3.2.3) except for the addition of 1 mM DTT during purification in an attempt to minimise oxidation.

Non-autoclavable portion (400 ml)		
Chemical class	Component	Quantity
Nutrient source	Glucose	40.00 g
	$MgSO_4.7H_2O$	1.00 g
	$FeSO_4$	16.7 mg
pH adjustment	conc. H_2SO_4	33.2 µl
Vitamin	Thiamine-HCl	20 mg
Label	Seleno-Methionine	200 mg

Autoclavable portion		
Chemical Class	Component	Mass (g)
Amino acid	Alanine	2.00
	Asparagine	2.32
	Aspartic acid	1.60
	Cystine	0.12
	Glutamic acid	2.68
	Glutamine	1.32
	Glycine	2.16
	Histidine	0.24
	Isoleucine	0.92
	Leucine	0.92
	Lysine-HCl	1.68
	Phenyalanine	0.52
	Proline	0.40
	Serine	8.32
	Threonine	0.92
	Tyrosine	0.68
	Valine	0.92
Nucleotide bases	Adenine	2.00
	Guanosine	2.68
	Thymine	0.68
	Uracil	2.00
Salts/Buffers	Sodium acetate	6.00
	Succinic acid	6.00
	Ammonium	3.00
	Sodium hydroxide	4.32
	Dibasic potassium hydroxide	42.0

Table 3-2

Table showing chemical composition of the non-autoclavable and autoclavable portions of LeMasters defined medium.

3.2.8 Receptor Competition and Titration Assay

Assays were carried out as represented in Figure 3-2 below. MaxiSorp™ 96 well plates (Nunc) were coated with 2-5 μg/well of Protein-A (Pharmacia). The plates were covered and left for 16 hours at room temperature. The Protein-A was discarded and the plate then blocked with PBS containing 1 % (w/v) BSA, for 2 hours, followed by two washes with PBS containing 0.05 % (w/v) Tween-20. Supernatants containing the receptor of interest, as Fc-fusions, were diluted 1:4 in PBS and 200 μl added to each well. The plate was left for 2 hours and after discarding the plate contents, a further two washes with PBS containing 0.05 % (w/v) Tween-20 were carried out.

Binding reactions were setup using a serial 4-fold dilutions of stock OSM_Δ188, starting at a saturating concentration of 10000 ng/ml. OSM-196 and murine LIF were included as controls. Dilutions were carried out using a stock solution of PBS containing biotinylated ligand (kindly provided by K. Chobotova, CRC Growth Factors Group. Birmingham.) as the competing agent. Biotinylated human OSM-196 and biotinylated human LIF were routinely used for competition. The plate was left at room temperature for 1 hour to equilibrate. The plate contents were discarded and the plate was then washed twice with PBS containing 0.05 % Tween-20. Table 3-3 below shows the receptor-ligand combinations tested.

Receptor	Ligand	Competing Ligand
	[Human OSM-196] (ng/ml)	[Human Biotinylated OSM-196] (ng/ml)
Human gp130	10000 – 0.6	50
Murine LIFR	10000 – 0.6	400

Table 3-3

Table showing receptor and ligand concentrations used for titration and competition assays.

The plate was washed with PBS and streptavidin-horseradish peroxidase conjugate (Amersham™), diluted 1:1000 in PBS, was added to each well and the plate incubated for 1 hour at room temperature. After washing each well with PBS, substrate development was carried out using 1,2-phenylenediamine (OPD) (Sigma) as chromogen. 8 mg of OPD was dissolved in 12 ml of citric acid phosphate buffer consisting of 0.035 M citric acid, 0.067 M $Na_2HPO_4.12H_2O$, pH 5.0 (see Appendix II). 5 µl of hydrogen peroxide was added. Chromogen solution was added to each well and the colour developed, in the dark, for 10-15 minutes. The reaction was terminated by the addition of 75 µl of 1 M sulphuric acid to each well. The absorbance was read at 490 nm.

Receptor titration assays were carried out as for the competition assays above, except for the omission of competing biotinylated ligand. Binding reactions were setup using serial 4-fold dilutions of stock biotinylated human OSM-196, starting at a saturating concentration of 10000 ng/ml. All receptor supernatants were used at a dilution of 1:4 in PBS and murine LIFR, murine OSMR and human FGFR2 were included as controls.

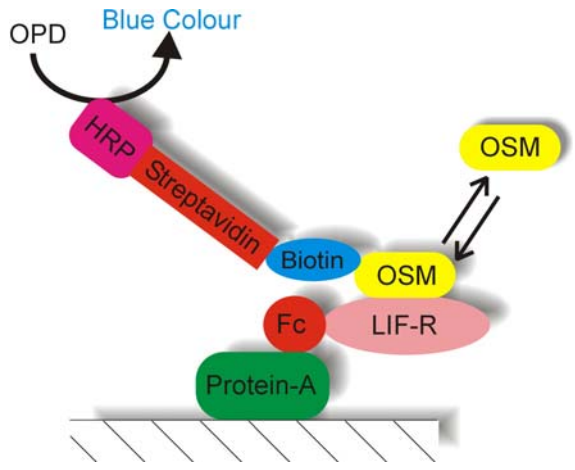


Figure 3-2
Figure showing schematic diagram of format used for receptor competition assays. Receptor titration assays were assembled in a similar fashion with the omission of competing biotinylated-OSM.

3.2.9 Ba/F3 Cell Survival Assay

Ba/F3 cells were cultured in RPMI medium (Gibco) supplemented with 10% dialysed foetal calf serum (Gibco), 50 IU/ml penicillin and 50 μg/ml streptomycin, glutamine and 2 μg/ml murine IL-3. Cultures were grown at 37°C in a humidified atmosphere of 5 % CO_2 in air. Cells were maintained at a density of between 5 x 10^4 and 1 x 10^6 cells/ml. All operations involving cell-lines were carried out aseptically, in a class II microbiological flow cabinet.

Stock samples of test cytokine were produced by dilution in minimal RPMI medium (as above with the exception of murine IL-3). OSM_Δ188 was diluted to 500 ng/ml. Control samples of human OSM-196 and murine LIF were prepared in a similar fashion at concentrations of 500 ng/ml and 50 ng/ml, respectively. All samples were centrifuged at 15000 x g for 15 minutes at 4°C. Serial 4-fold dilutions of each cytokine were prepared, by dilution in minimal RPMI medium as above, and 50 μl of each was transferred aseptically to the wells of a 96 well culture plate. Duplicate wells containing minimal RPMI with and without IL-3 were included as positive and negative controls, respectively.

Cells were harvested by centrifugation at 1000 x g for 5 minutes and carefully resuspended in 50 ml of minimal RPMI medium, to remove all traces of IL-3. This wash step was repeated 4 times and the cells then resuspended in 5 ml of minimal RPMI medium supplemented with 10% FCS. Viable cells were counted in the presence of trypan blue dye using a haemacytometer. The culture was diluted to a density of 2 x 10^5 cells/ml and 50 μl of culture pipeted into each well. The plate was covered with clingfilm to limit evaporation and then incubated at 37°C in a humidified atmosphere of 5 % CO_2 in air, for 72 hours.

Cell survival and proliferation was determined colorimetrically according to the modified method of Mossman (Chapdelaine, 1989). This assay is dependent on the reduction of the tetrazolium salt 3-(4-5-dimethylthazol-2-yl)-2-5-diphenyl tetrazolium bromide (MTT) by the mitochondrial dehydrogenase of viable cells to form a blue formazan product (see 3.4 for more detailed discussion). Typically, 10 µl of 5 mg/ml MTT was added to each well and the plate incubated for a further 4 hours. The formazan product was solubilised by the addition of 10% (w/v) SDS in 0.01 M HCl to each well. The plate was incubated at 37°C for 16 hours and the absorbance measured at 570 nm.

3.2.10 Ion-exchange and Gel Filtration Chromatography

Samples of affinity purified OSM_Δ188 were buffered to various pH's in the range of 3-10 and then applied to either Resource-Q, Resource-S, Mono-Q or Mono-S columns (Pharmacia). Samples were typically run at a flow rate of 5 ml/min using a linear gradient of 0 to 2.0 M sodium chloride on a Biosys 510 FPLC or Gold System HPLC (Beckman).

Samples of affinity purified OSM_Δ188 were applied to a Superdex-75 HiLoad column (bed volume 320 ml; 260 x 600 mm; separation range 3-70 kDa) (Pharmacia). Columns were typically run at a flow rate of 1 ml/min in 150 mM sodium chloride, 50 mM Tris-HCl, pH 7.5 on a Biosys 510 FPLC (Beckman). Columns were calibrated by running a cocktail of protein markers consisting of 0.5 mg each of cytochrome C (12.4 kDa), carbonic anhydrase (29.0 kDa), albumin bovine serum (66.0 kDa), alcohol dehydrogenase (150.0 kDa) and blue dextran (2000.0 kDa).

3.2.11 N-terminal Sequencing

N-terminal sequencing was carried out for 10 cycles using an Applied Biosystems 494A/473A Sequencer. (Tony Willis, MRC Immunochemistry Unit Protein Sequencing Facility, Oxford). Typically, 230-480 pMol samples of HPLC purified material were submitted and blotted onto PVDF membrane after separation on a 10% NuPAGE/MOPS or 4-12% NuPAGE/MES gel.

3.2.12 Mass Spectrometry

Typically, ~400 pMol samples of HPLC purified OSM_Δ188 were submitted for mass spectrometry. Analysis was carried out using electrospray ionization on a Micromass BioQ II triple, quadrapole, atmospheric pressure mass spectrophotometer equipped with an electrospray interface operating in the positive ion mode (Dr. Robin Aplin, Oxford Centre for Molecular Sciences, Oxford). Samples were also analysed in Matrix Assisted Laser Deabsorption and Ionisation (MALDI) mode, using a Kratos Kompact mass spectrophotometer (Dr. Lisa Wallington, CRC Growth Factor Group, Birmingham).

3.3 Results of Expression, Purification and Characterisation of Oncostatin M

3.3.1 SDS-PAGE of Protease 3C-His_6 Expression and Affinity Purification

Figure 3-3 below shows a 12 % SDS-polyacrylamide gel of the expression and affinity purification of 3C-His_6 protease. Lane 1 shows the total cellular protein before induction with IPTG and lane 2 shows the total cellular protein after ~20 hours induction with IPTG. By comparison of the first two lanes a relative increase in the density of bands at ~27 kDa and ~50 kDa can be observed on induction, so suggesting production of 3C-His_6. Lane 3 shows the cleared cell lysate, which shows major bands at ~27 kDa and 50 kDa. Lane 4 shows the first eluted fraction of 3C-His_6 running at ~27 kDa. Lane 5 shows the nickel agarose resin after passage of the cleared lysate and elution of the fusion protein and no bands are observed. Lane 6 shows the six components of the broad range rainbow molecular weight markers running at 210, 135, 80, 42, 31 and 18 kDa (Bio-Rad).

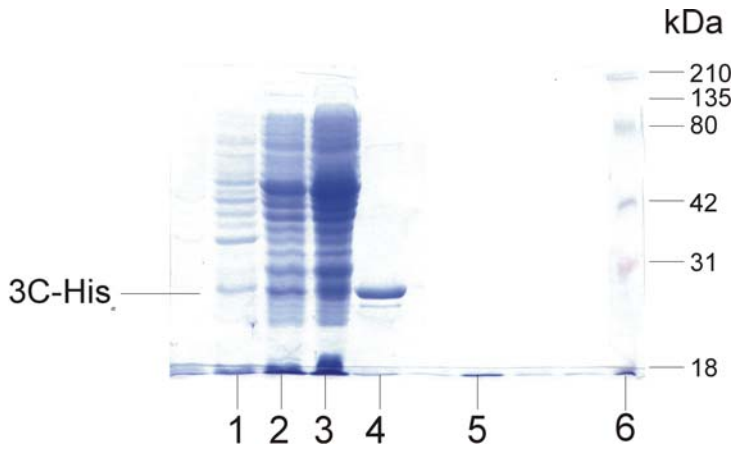


Figure 3-3

Figure showing 12 % SDS-PAGE gel of the expression and affinity purification of protease 3C-His_6. Lane 1 and 2 show the total cellular protein before and after induction, respectively. Lane 3 shows the cleared cell lysate and lane 4 shows the eluted protein fractions after cleavage with 3C protease. Lane 5 shows the nickel agarose resin after elution and lane 6 show broad range molecular weight markers running at 210, 135, 80, 42, 31 and 18 kDa (Bio-Rad).

3.3.2 Analysis of Fusion Protein Cleavage Using Protease 3C

Figure 3-4 below shows a 12 % SDS polyacrylamide gel used to determine optimal cleavage conditions. Lanes 1 to 8 correspond to reaction mixtures containing OSM-196-3C-GST and various concentrations of 3C-GST protease (as in Table 3-1). Lanes 1 to 4 contain increasing concentrations of 3C-GST protease and all contain no DTT. Lanes 1 to 4 show a main band at ~47 kDa, which corresponds to 3C-GST, and a band at ~50 kDa, which corresponds to OSM-196-GST. Minimal cleavage products are observed in lanes 1 to 4. Lanes 5 to 8 contain increasing concentrations of 3C-GST protease and all contain ~0.5 mM DTT. Lanes 5 to 8 show a main band at ~47 kDa, which corresponds to 3C protease, and a band at ~25 kDa, which corresponds to cleaved OSM-196. Lanes 5 to 8 demonstrate a loss of the OSM-196-GST band at ~50 kDa. Lane 9 contains a 3C-GST control running at ~47 kDa and lane 10 contains OSM-196-GST running at ~50 kDa. Lane 11 contains broad range rainbow molecular weight markers (Bio-Rad).

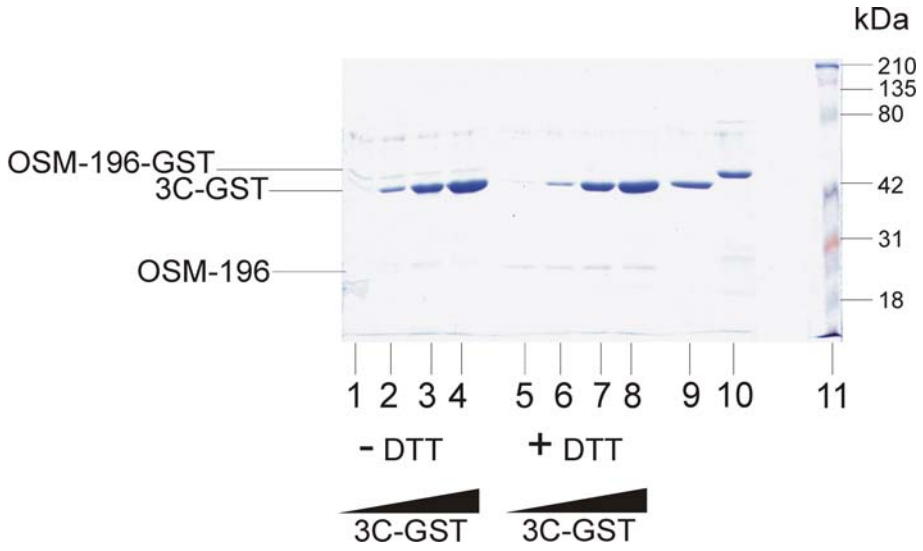


Figure 3-4

Figure showing 12 % SDS polyacrylamide gel used to determine optimal cleavage conditions. Lanes 1 to 8 correspond to reaction mixtures containing OSM-196-3C-GST and various concentrations of 3C-GST. Reactions in lanes 1 to 4 were carried out in the absence of DTT whereas lanes 5 to 8 contained ~0.5 mM DTT. Lane 9 contains 3C-GST and lane 10 contains OSM-196-GST as control. Lane 11 contains broad range molecular weight markers running at 210, 135, 80, 42, 31 and 18 kDa (Bio-Rad).

3.3.3 SDS-PAGE of OSM_Δ188 and OSM_Δ185 Expression.

SDS-PAGE analysis showed that no expression of OSM_Δ185-3C-GST was observed (data not shown). After induction with IPTG no band for OSM_Δ185 (predicted M_r ~50 kDa) was observed and after affinity purification no protein was detected using the Coomassie Protein Plus™ Assay Kit (Pierce). Soluble and insoluble fractions were analysed and no bands of the correct size were observed.

Figure 3-5 below shows a 12 % SDS-polyacrylamide gel of the expression and affinity purification of (a) OSM_Δ188-3C-GST and (b) OSM-196. In both cases, lane 1 shows the total cellular protein before induction with IPTG and lane 2 shows the total cellular protein after ~20 hours induction with IPTG. In both gels a relative increase in the density of a band at ~50 kDa can be observed on induction, so suggesting production of fusion protein. Lane 3 and 4 show duplicate samples of cleared cell lysate, which contains a major band at ~50 kDa, again suggesting production of soluble fusion protein. Lane 5 shows samples of glutathione sepharose resin after passage of the cleared lysate. This shows that two major species are bound to the affinity column at ~50 kDa and ~25 kDa, presumably corresponding to fusion protein and GST respectively. Lane 6 contains a wash fraction and displays no major bands. Lane 7 shows glutathione sepharose resin after overnight cleavage with 3C protease and only one major band at ~32 kDa can be seen, which presumably corresponds to sequestered GST purification tag. Lanes 8 and 9 show the first eluted fractions from the affinity column, after cleavage with protease 3C. One major band corresponding to OSM-196 or OSM_Δ188 can be clearly seen at ~25 kDa, although contaminating bands at ~60 kDa and ~90 kDa can also be seen. The later bands are thought to correspond to heat shock protein 60 and 90 (HSP60 and HSP90), which are known to co-purify with the pGEX system. Lane M shows the seven components of the high range molecular weight markers (Amersham).

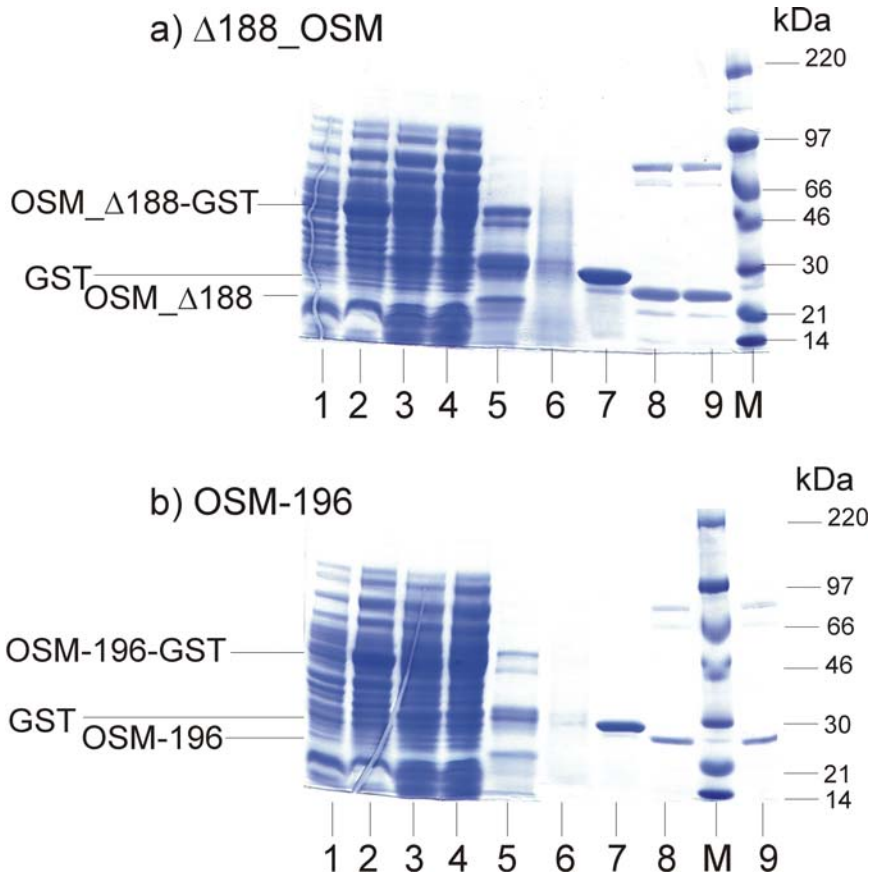


Figure 3-5

Figure showing 12 % SDS-PAGE gel of the expression and affinity purification of (a) OSM_Δ188 and (b) OSM-196. Lanes 1 and 2 show total cellular protein pre- and post-induction. Lanes 3 and 4 shows the cleared cell lysate. Lane 5 shows material bound to the affinity column. Lane 6 shows a column wash and lane 7 shows material remaining on the column after cleavage with 3C protease. Lanes 8 and 9 show cleaved fractions eluted from the affinity column. Lane M shows the high range molecular weight markers running at 220, 97, 66, 46, 30, 21 and 14 kDa (Amersham).

3.3.4 Western Blot Analysis

Figure 3-6 below shows western blot analysis of affinity purified OSM_Δ188 separated on a 12-15 % gradient SDS polyacrylamide gel. Lane 1 contains ~8 μg of OSM_Δ188 and shows approximately 7 main bands. Around 30 % of the density runs at ~25 kDa. Lane 2 shows broad range rainbow molecular weight markers (Bio-Rad). Lane 3 shows a control sample containing LIFR, which shows no bands and no cross reactivity to the anti-OSM antibodies.

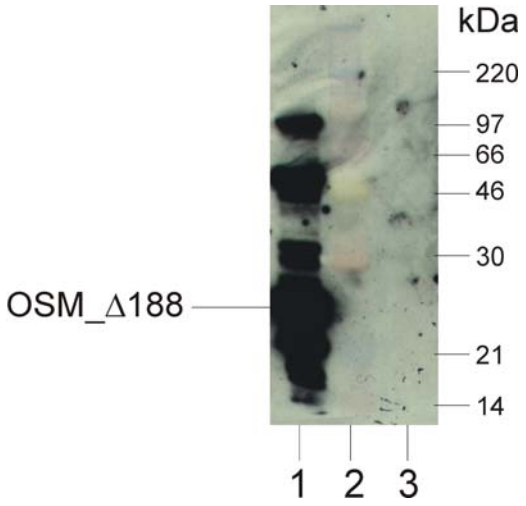


Figure 3-6

Figure showing western blot of affinity purified OSM_Δ188. Lane 1 contains OSM_Δ188. Lane 2 contains broad range molecular weight rainbow markers running at 210, 135, 80, 42, 31 and 18 kDa (Bio-Rad). Lane 3 contains a control sample of LIFR.

3.3.5 Receptor Competition and Titration Assay

Figure 3-7 below shows graphs of (a) receptor titration and (b) receptor competition assays. Graph (a) demonstrates that the concentration range of ligand used leads to saturation of the gp130 receptor under the conditions employed. This also shows that specific binding is occurring due to the sigmoidal increase in absorbance with increasing concentration of human biotinylated OSM-196. Binding is not observed to human OSMR and murine LIFR, as both show background responses similar to the control FGFR2 samples. Binding of human biotinylated OSM-196 shows an effective concentration 50% (EC_{50}) value of 348.9 ng/ml. From graph (b) it can be seen that increasing the concentration of both OSM-196 and OSM_Δ188, causes a sigmoidal decrease in the absorbance reading due to competition with biotinylated OSM-196. The OSM-196 sample agrees closely with the OSM_Δ188 sample and no fall in absorbance is demonstrated by the control murine LIF sample. Human OSM-196 displays an EC_{50} value of 209.2 ng/ml and human OSM_Δ188 displays an EC_{50} value of 182.0 ng/ml.

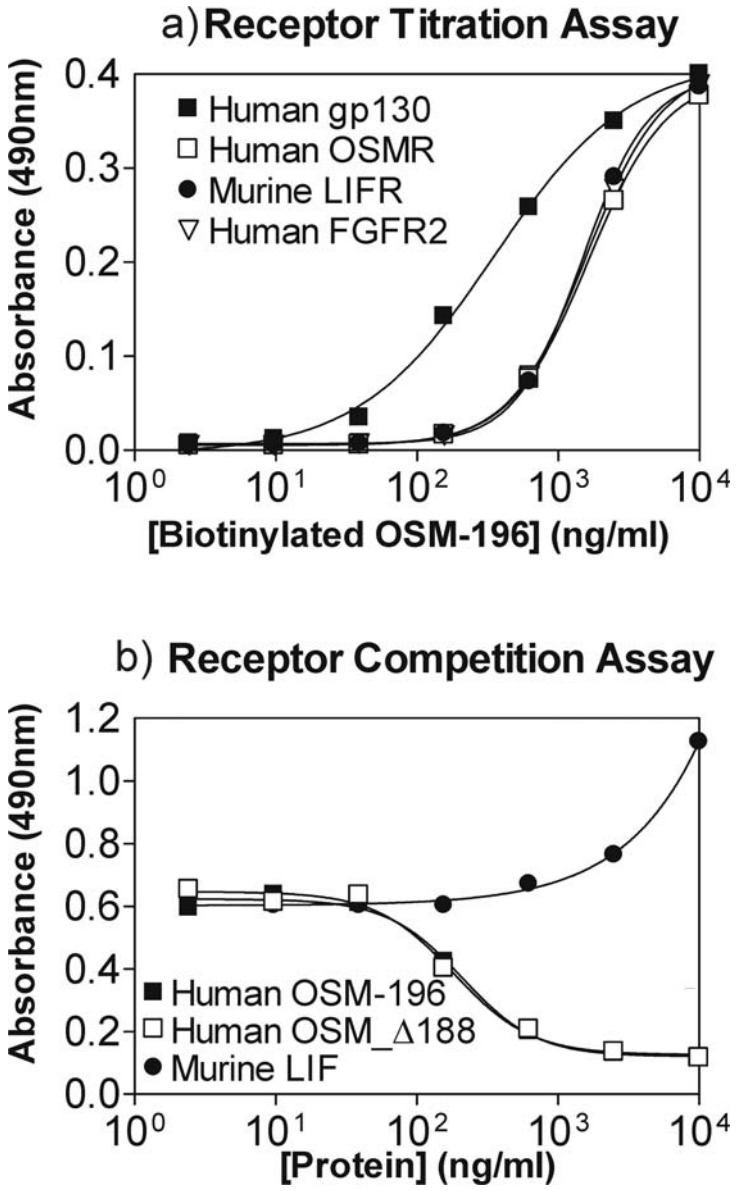


Figure 3-7

Figure showing graph of (a) receptor titration and (b) receptor competition assays. Panel (a) shows binding of human biotinylated OSM-196 to human gp130 (■) (EC_{50} 348.9 ng/ml), human OSMR (□), murine LIFR (●) and as control human FGFR2 (▽). Panel (b) shows competitive binding to human gp130 in the presence of human biotinylated OSM-196. Human OSM-196 (■) (EC_{50} 209.7 ng/ml), human OSM_Δ188 (□) (EC_{50} 182.0 ng/ml) and as control murine LIF (●) are shown.

3.3.6 Ba/F3 Cell Survival Assay

Figure 3-8 below shows a graph of a Ba/F3 cell survival assay for various cytokines. Analysis of the OSM_Δ188 and OSM-196 curves shows that a sigmoidal dose response is observed. In both cases increasing protein concentration results in an increase in absorbance reading. Murine LIF shows no dose response in the concentration range tested. Human OSM-196 shows an EC_{50} value of 0.39 ng/ml and human OSM_Δ188 displays an EC_{50} value of 0.26 ng/ml.

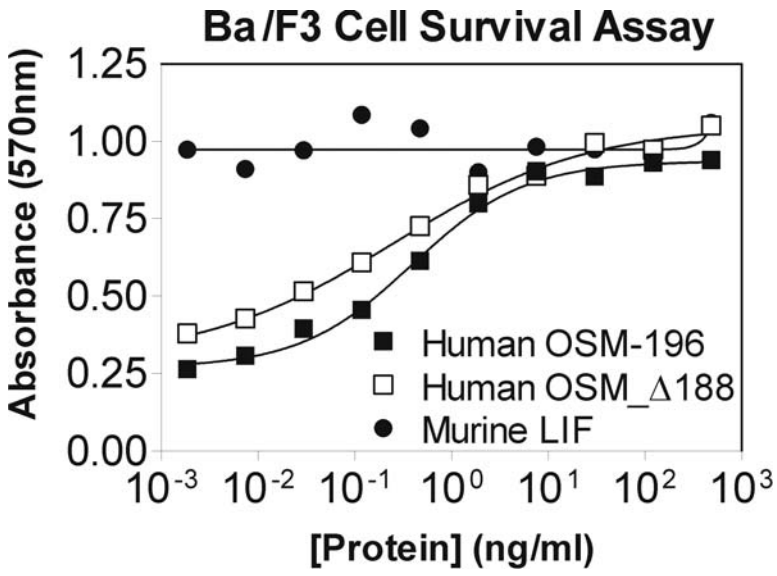


Figure 3-8

Graph showing Ba/F3 cell survival assay curves for various cytokines. Human OSM-196 (■) (EC_{50} 0.39 ng/ml), human OSM_Δ188 (□) (EC_{50} 0.26 ng/ml) and murine LIF (●) were tested at concentrations ranging from 500-0.001 ng/ml.

3.3.7 Ion-exchange and Gel Filtration Chromatography

Ion-exchange chromatography proved unsuccessful using both mono-Q and mono-S columns in the pH range 4-10. No specific binding of protein was observed and only aggregates were recovered from the column at high salt concentrations.

Figure 3-9 below shows the elution profiles of gel filtration using a Sephadex-75 column. Panel (a) shows the elution profile of molecular weight standards, with five main peaks at 2000, 150, 66, 29 and 12 kDa. Panel (b) shows the elution profile of affinity purified OSM_Δ188 in the absence of Triton X-100 which shows two major peaks eluted from the Sephadex-75 column at 170 minutes (25.1 kDa) and 155 minutes (34 kDa). Lane 8 of panel (d) shows that the 155 minute peak runs at ~30 kDa on SDS-PAGE. Lane 9 of panel (d) shows that the 170 minute peaks runs at ~25 kDa on SDS-PAGE. Panel (c) shows the elution profile of affinity purified OSM_Δ188 in the presence of Triton X-100 and this shows two major peaks eluted from the Sephadex-75 column at 170 minute (25.1 kDa) and 140 minute (50 kDa). Lane 7 of panel (d) shows that the 140 minute peak runs at ~25 kDa on SDS-PAGE.

Figure 3-10 below shows elution profiles of a Sephadex-75 column with OSM-Se-Met expressed in *E. coli* strain (a) B834 and (b) DL41. These chromatograms show two major peaks eluted at 170 minutes (25.1 kDa) and 140 minutes (50 kDa). Lane 9 of panel (c) shows that the peak eluted at 170 minutes runs at ~25 kDa on SDS-PAGE.

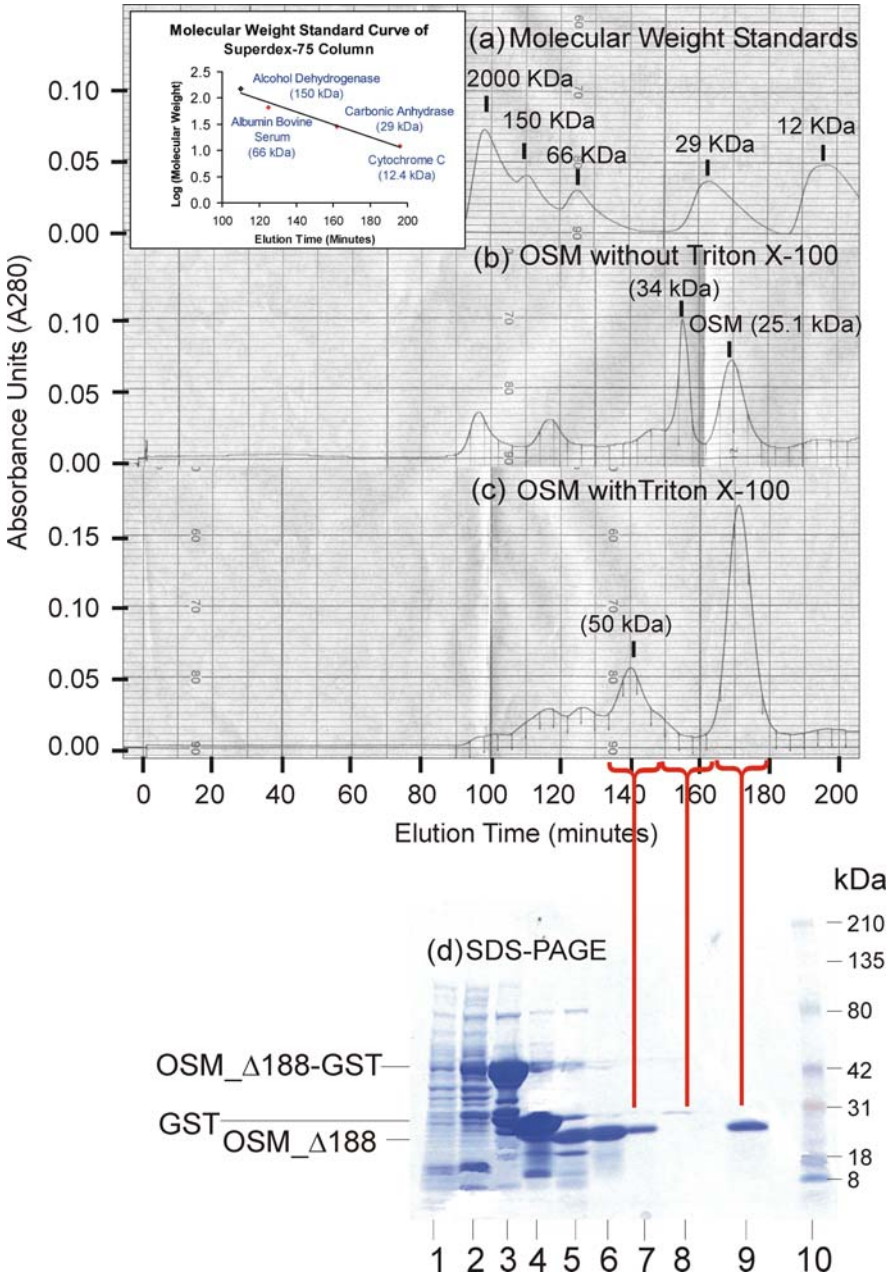


Figure 3-9

Chromatogram showing FPLC elution profile of Superdex-75 prep grade gel filtration column with (a) molecular weight standards, b) affinity purified OSM_Δ188 sample without Triton X-100 and (b) affinity purified OSM_Δ188 sample with Triton X-100. 4ml samples were loaded at a flow rate of 1 ml/min and the absorbance monitored at 280 nm. In the presence of Triton X-100 two peaks are eluted at 170 minutes (25.1 kDa) and 170 minutes (50 kDa). In the absence of Triton X-100 a second peak is eluted at 155 minutes (34 kDa). SDS-PAGE (d) of preparation, with lane 1 and 2 showing total cellular protein pre- and post-induction. Lane 3 contains material bound to affinity column. Lanes 4 shows material bound to affinity column after cleavage with 3C protease and lane 5 shows OSM_Δ188 eluted from affinity column. Lane 6 shows material loaded onto gel filtration. Lanes 7 shows the 140 min peak and lane 8 shows the 155 min peak. Lane 9 shows concentrated peak of purified OSM_Δ188 eluted at ~170 minutes. Lane 10 contains broad range molecular weight markers (Bio-Rad).

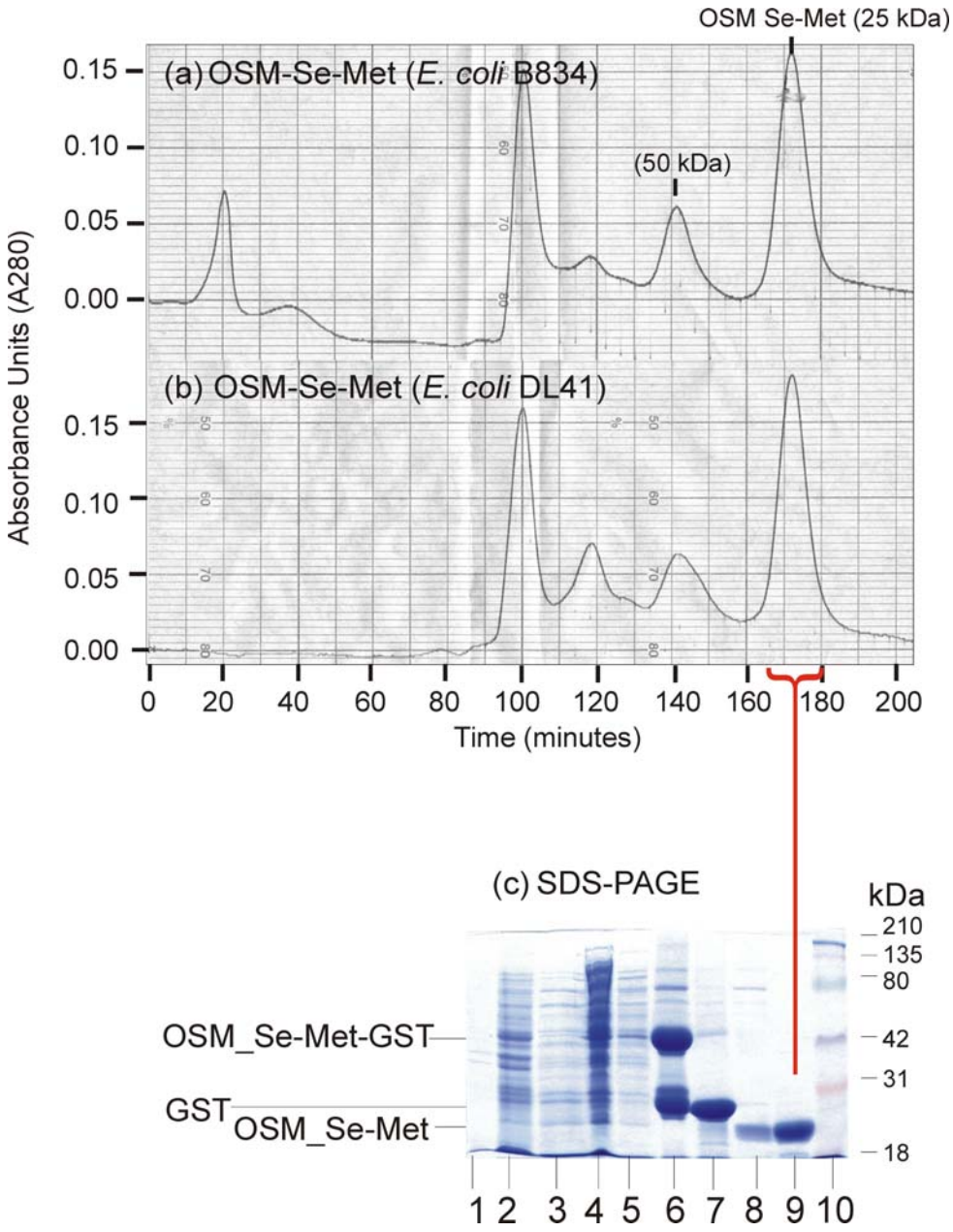


Figure 3-10

Chromatogram showing FPLC elution profile of Superdex-75 prep grade gel filtration column with a) affinity purified OSM_Se-Met from *E. coli* strain (a) B834 and (b) DL41. 4ml samples were loaded at a flow rate of 1 ml/min and the absorbance monitored at 280 nm. Two main peaks are eluted at 170 minutes (25.1 kDa) and 140 minutes (50 kDa). SDS-PAGE (c) of preparation with lane 1 and 2 showing total cellular protein pre- and post-induction. Lane 3 shows the material that did not bind to the affinity column. Lanes 4 contains the cleared lysate and lane 5 shows the column wash. Lane 6 shows material bound to affinity column and lane 7 shows material remaining on column after cleavage with 3C protease. Lanes 8 contains material eluted from affinity column after cleavage with 3C protease and lane 9 contains the concentrated peak of purified OSM-Se-Met eluted at ~170 minutes. Lane 10 contains broad range molecular weight markers.

3.3.8 N-Terminal Sequencing

The results of N-terminal sequencing are shown in Table 3-4 below. This data shows that *E. coli* strain JM109, produces only one species of native OSM_Δ188 with the addition of a four residues 3C linker site (GPGS) at the N-terminus. *E. coli* strain JM109 produced two minor species when grown on minimal media supplemented with Se^{78}-Met. The *E. coli* methionine auxotrophic strain B834 produced only one species.

E. coli strain	Protein	N-terminal Sequence	Yield (%)
JM109	Native (S^{32}-Met)	GPGSAAIGS-SKEYR	38.3
JM109	Derivative (Se^{78}-Met)	GPGSAAIGS-SKEYRVLLG	3.8
JM109	Derivative (Se^{78}-Met)	AAIGS-SKEYRVLLG	4.2
B834	Derivative (Se^{78}-Met)	GPGSAAIGS-	38.1

Table 3-4
Table showing results of N-terminal sequencing.

3.3.9 Mass Spectrometry

Figure 3-11 (a) below shows mass spectra of OSM_Δ188 samples run in electrospray mode. This shows one major peak at 21579.98±9.05 kDa. The theoretical mass of OSM_Δ188 is 21409.50 kDa as calculated using the ExPASy molecular biology server (The University of Geneva; www.expasy.ch).

Figure 3-11 (b) shows OSM-Se-Met material run in electrospray mode and this shows five major peaks, corresponding to sequential substitution of S^{32}-Met with Se^{78}-Met. The five species are observed at 21593.83±1.62 kDa (no substitution), 21641.27±10.75 kDa (one selenium substitution), 21683.85±1.39 kDa (two selenium substitutions), 21745.63±1.43 kDa (three selenium substitutions) and 21765.79±1.18 kDa (four selenium substitutions; i.e. susbstitution of all methionines).

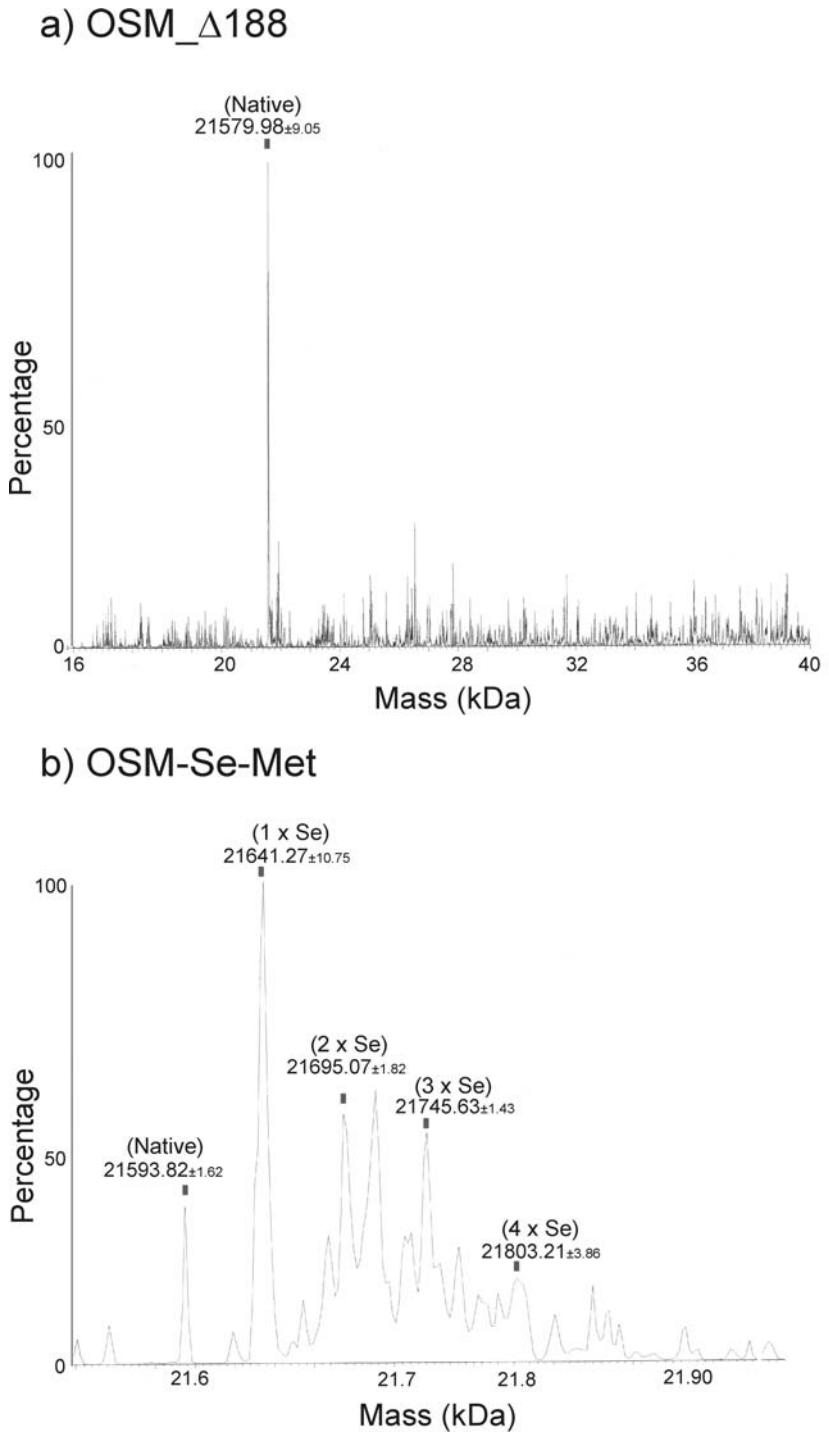


Figure 3-11

Figure showing mass spectra of (a) OSM_Δ188 and (b) OSM-Se-Met samples. Typically, ~400 pMol samples of HPLC purified protein in water were analysed by electrospray ionization on a Micromass BioQ II spectrophotometer. OSM_Δ188 samples display one major peak at 21579.98±9.05. OSM-Se-Met samples display five major peaks at 21593.82±1.62 (native), 21641.27±10.75 (1 x Se), 21695.07±1.82 (2 x Se), 21745.63±1.43 (3 x Se) and 21803.21±3.86 (4 x Se).

3.4 Discussion of Expression, Purification and Characterisation of Oncostatin M

Production of 3C protease in sufficient quantities proved successful. Figure 3-3 shows that ~90 % pure preparations of 3C-His_6 protease could be produced. Figure 3-4 demonstrates that DTT is essential for efficient cleavage of fusion proteins with 3C protease. Cleavage was achieved using very low concentrations of 3C protease as either GST or His_6 fusions. Typically, 3C-His_6 protease was used at a ratio of 1:1000 of protease to fusion protein and removal of 3C-His_6 was carried out by passage over nickel affinity resin.

Expression of OSM_Δ185 proved unsuccessful. No induction with IPTG was observed and no protein was observed following affinity purification with glutathione sepharose. It is possible that OSM_Δ185 was not folded properly and as such would be present in the insoluble fraction. Possible reasons for mis-folding include incorrect prediction of the C-terminal end of helix D, as determined through homology modelling with LIF. This proposal is further supported by the low transformation frequencies observed for Δ185_OSM (as reported in Chapter 2) which presumably result from the toxicity associated with this mutant protein.

Figure 3-5 shows that expression of OSM_Δ188-GST was successful. OSM_Δ188 was separated to greater than 90 % purity, as assessed by SDS-PAGE, after one affinity step. Yields of ~1 mg/L were achieved. Western blot analysis as shown in Figure 3-6 revealed extensive fragmentation of OSM_Δ188, so suggesting further purification. Purification using ion-exchange proved unsuccessful due to aggregation of OSM_Δ188 in all pH's tested. Figure 3-9 shows that OSM_Δ188 was resolved as a distinct peak using gel filtration. Elution profiles show two main peaks with the predominant peak corresponding to OSM_Δ188 monomers (25 kDa) and a minor peak corresponding to OSM_Δ188 dimers (50 kDa). Dimer formation was favoured in the presence of Triton X-100 and in

the absence of detergent a minor contaminating band at 34 kDa was observed. The nature of this band is unknown.

N-terminal sequencing was used to verify the correct nature of the purified OSM_Δ188 fraction. This revealed the inclusion of the four residue sequence GPGS at the N-terminus of the mature OSM sequence. These residues result from 3C protease and *Bam*HI cleavage of the engineered constructs. Mass spectrometry as shown in Figure 3-11 revealed one major species at a mass of 21579.98±9.05 kDa. This value disagrees with the theoretical mass of OSM_Δ188 of 21259 kDa, the difference presumably due to bound ions, detergent, glycerol or other buffer components.

Receptor binding assays as shown in Figure 3-7 suggest Δ188_OSM possesses similar properties to the wildtype protein *in vitro*. Comparison of the EC_{50} values for binding to the gp130 receptor in titration and competition assays suggests OSM_Δ188 has a similar affinity to OSM-196. Sigmoidal dose responses are observed in receptor competition assays with EC_{50} values of 209.7 ng/ml and 182 ng/ml for wildtype and mutant, respectively.

Ba/F3 cell survival assays demonstrate that OSM-196 and Δ188-OSM also have similar properties *in vivo*. Ba/F3 cells are an IL-3 dependent murine cell-line that grows in suspension (Palacios and Steinmetz, 1985). This cell-line contains few endogenous cytokine receptors, so can be used as a model system through the introduction of exogenous receptors. Ba/F3 cells were previously stably transfected with human gp130 and human LIFR (Vicky Barton, CRC Growth Factors Group, Birmingham) so providing a cell-line responsive to both human LIF and human OSM. OSM_Δ188 and OSM-196 show similar biological responses in Ba/F3 cell assays as shown in Figure 3-8. EC_{50} values

for OSM_Δ188 and OSM-196, show close aggreement with values of 0.26 ng/ml and 0.39 ng/ml, respectively. These data suggest that Δ188 C-terminal truncations do not affect the receptor binding and biological activity of OSM.

Figure 3-10 shows that expression of OSM-Se-Met was only successful when carried out in the methionine auxotrophic strains of *E. coli* DL41 or B834. Mass spectrometry as shown in Figure 3-11 shows that Se^{78}-Met derivatisation of OSM_Δ188 was not complete. Five main protein species are observed corresponding to no selenium substitutions, one selenium substitution, two selenium substitutions, three selenium substitutions and four selenium substitutions. Given the relatively good yield of protein obtained this observation is rather puzzling and suggests that some form of suppression of auxotrophism is occurring; possible sources of S^{32}-Met include the starter culture or contaminants in the LeMasters medium. The poor substitution observed could also be due to preferential use S^{32}-Met, which given the buried nature of the methionine residues in the refined structure would be more favourable for correct protein expression and folding. N-terminal sequence analysis of OSM-Se-Met material, produced in strain B834, has the correct sequence with the inclusion of four residues from the 3C protease site.

4. Chapter 4 – Expression, Purification and Characterisation of LIFR

The aims of this chapter include the production of leukemia inhibitory factor receptor (LIFR) of sufficient quality and quantity for crystallographic studies. This chapter also aims to characterise the LIFR-LIF receptor-ligand complex and its suitability for crystallographic studies. The use of several expression systems is reported, along with the production, purification, deglycosylation and preliminary characterisation of the LIFR-LIF receptor-ligand complex.

Appendix III is included as a summary of the reagents and buffers used in the expression, purification and characterisation of LIFR. The methods involved are discussed in greater detail in texts such as (Maniatis *et al.*, 1989), (Coligan *et al.*, 1995) and (Deutscher, 1990). A detailed analysis of cell and tissue culture is given by (Paul, 1980) and an overview of the processes involved in the course of this chapter is shown in Figure 4-1 below.

4.1 Outline of Methods Employed

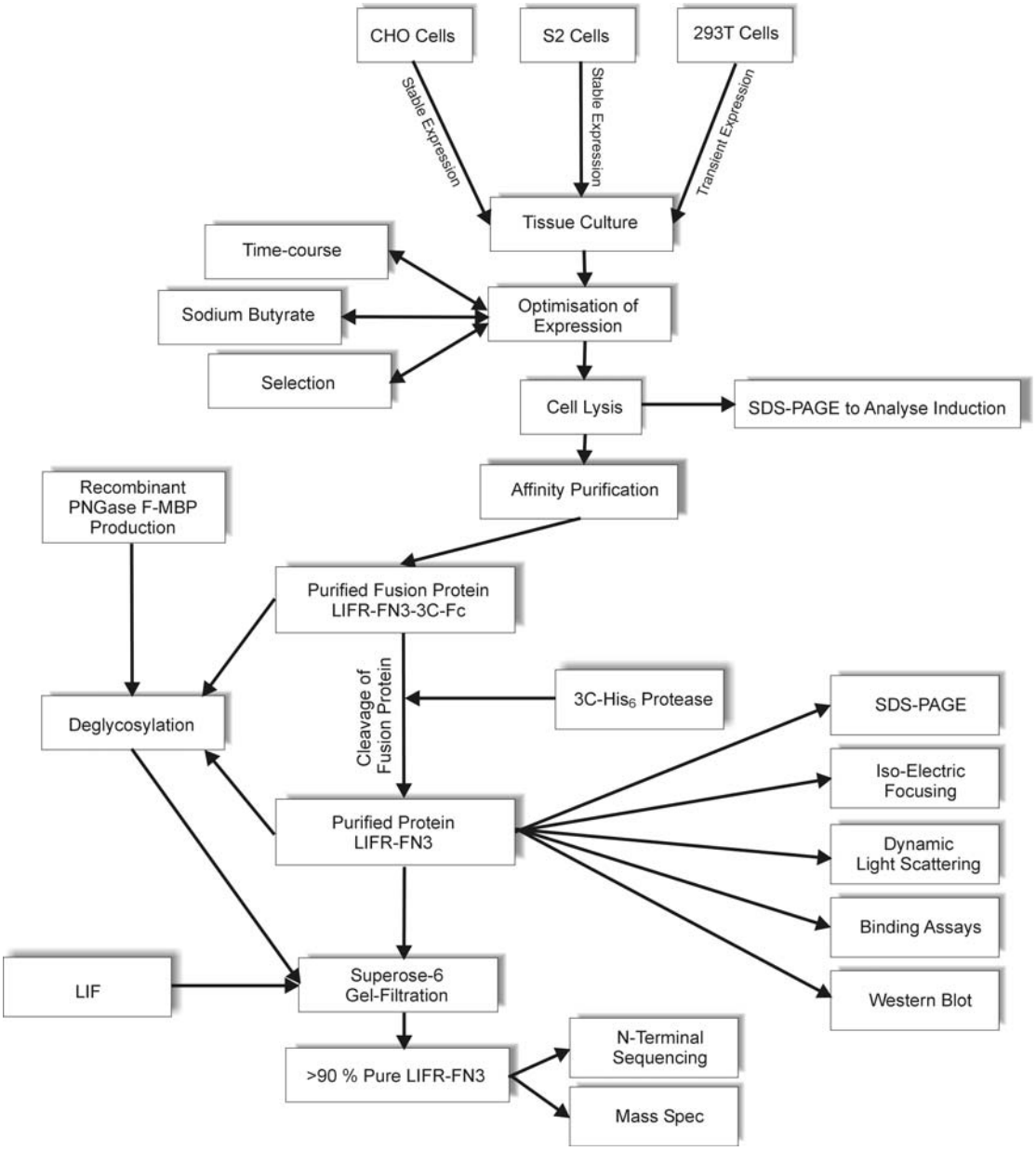


Figure 4-1
Flowchart detailing methods and techniques employed in the course of chapter 4.

4.2 Methods of Expression, Purification and Characterisation of LIFR

Previous attempts to produce LIFR protein in *E. coli* expression systems proved unsuccessful due to the large size of constructs involved (typically >1000 amino acids) and the inability to correctly determine domain boundaries (D. Staunton, J. K. Heath, CRC Growth Factors Group, Birmingham, personal communication).

In an attempt to produce large quantities of LIFR protein for crystallographic studies various mammalian expression systems were investigated. Full length constructs of the extracellular portion of LIFR consisting of the two N-terminal haemopoietin domains an intervening Ig-like domain and the three C-terminal fibronectin type III (FN) motifs were kindly provided in pEE14 (Bebbington, 1991), pRmHa-3 (Bunch *et al.*, 1988) and pIG-1 (Simmons, 1992) vectors as soluble Fc fusions (LIFR-FN3-3C-Fc; K. Chobotova, K. R. Hudson and J. K. Heath, CRC Growth Factor Group, Birmingham; see Figure 4-2 below). In all cases an Fc purification tag was fused to the C-terminus of the fibronectin type III motif to aid in affinity purification using sepharose immobilised Protein-A (Pharmacia). The Fc tag consisted of the heavy chain constant domains 2 and 3 and the hinge region (hinge-C_H2-C_H3) of human IgG1. In all cases the intervening linker region between the Fc and the LIFR contained 11 amino acids of the rhinovirus type 14 protease 3C recognition site to facilitate removal of the Fc purification tag (see Chapter 2).

The mature LIFR-FN3 moiety consists of 829 residues with a theoretical non-glycosylated molecular weight of 93.8 kDa and a theoretical pI of 7.2. All cells were obtained from the European Collection of Cell Cultures (ECACC), unless otherwise stated.

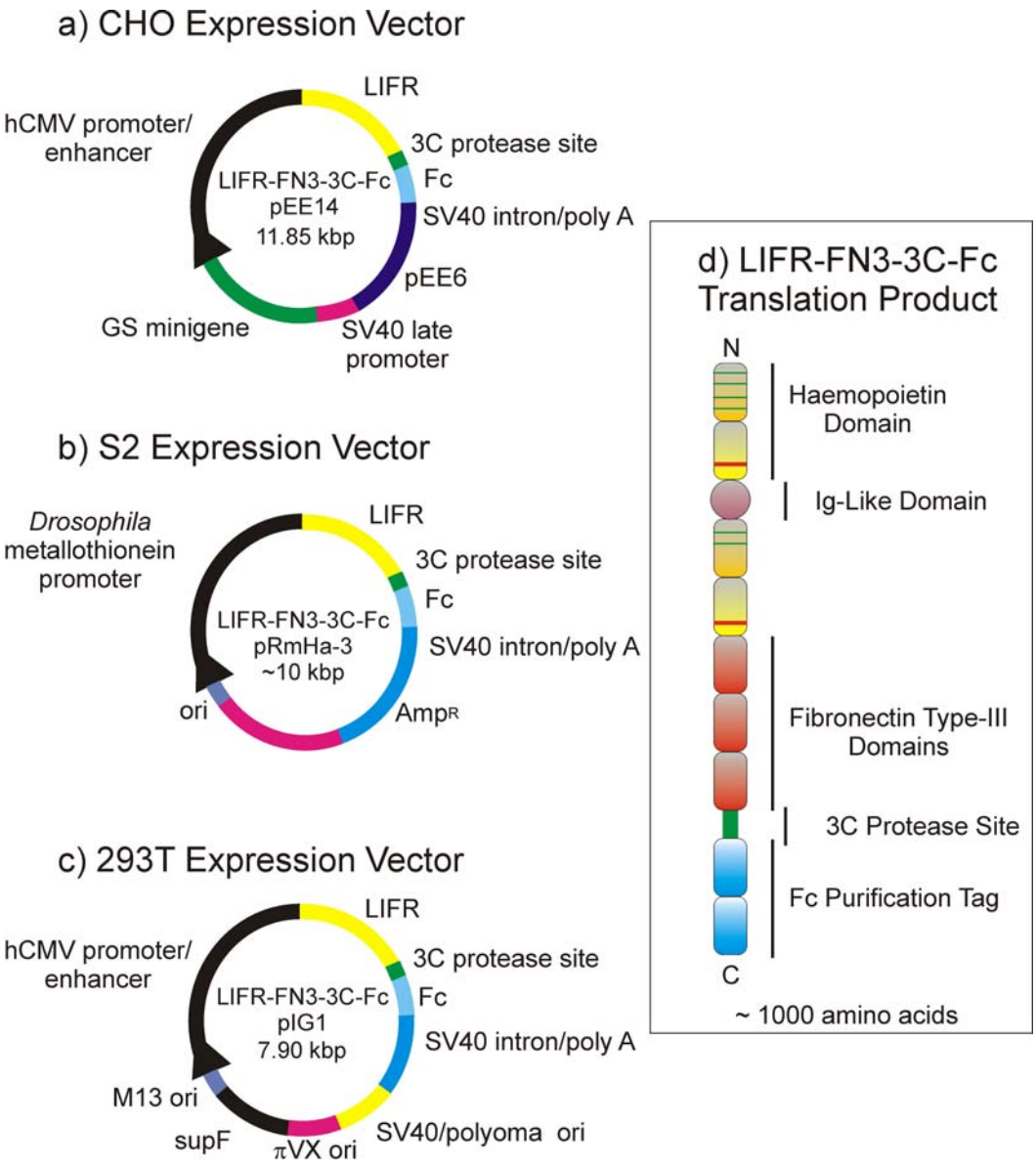


Figure 4-2

Diagrammatic representation of a) pEE14 vector (Bebbington, 1991), b) pRmHa-3 vector (Bunch *et al.*, 1988) and c) pIG-1 vector (Simmons, 1992) used for expression of soluble LIFR-FN3-3C-Fc fusion in CHO cells, *Drosophila* S2 cells and 293T fibroblast cells, respectively. Panel d) shows the diagrammatic domain representation of the translation product (LIFR-FN3-3C-Fc) consisting of the two N-terminal haemopoietin domains of LIFR an intervening Ig-like domain and the three C-terminal fibronectin type III (FN) domains. An Fc tag is included for Protein-A affinity purification and a 3C protease site is included for cleavage of fusion protein. All constructs kindly provided by K. Chobatova, CRC Growth Factors Group. Birmingham.

4.2.1 Stable Expression in CHO Cells

The mammalian expression vector pEE14 was used (Bebbington, 1991; Cockett *et al.*, 1990); see Figure 4-2) for expression of soluble LIFR-FN3-3C-Fc fusion protein. Expression was driven by the human cytomegalovirus (hCMV) promoter in host Chinese hamster ovary K1 cells (CHO-K1). This vector contains the glutamine synthetase (GS) minigene for amplification and clonal selection. GS is an allosteric enzyme required for the production of glutamine from glutamate and this enzyme is inhibited by L-methionine sulphoximine (MSX) which is a transition state analogue of the reaction:

$$\text{Glutamate} + \text{Ammonia} + \text{ATP} \Leftrightarrow \text{Glutamine} + \text{ADP} + P_i$$

CHO cells transfected with the GS minigene do not require glutamine in the growth media, providing sufficient glutamate is present. However, GS is essential for cell survival when glutamine is absent from the media and full MSX inhibition in this situation is lethal. Therefore, MSX treatment encourages an increase in the copy number of the GS gene when cells are cultured in the absence of glutamine, so effectively co-amplifying any associated genes. pEE14-LIFR-FN3-3C-Fc constructs were kindly provided by K. Chobotova, CRC Growth Factors Group. Birmingham.

4.2.1.1 CHO Cell Culture

pEE14-LIFR-FN3-3C-Fc transfected CHO-K1 cells were cultured in GMEM-S medium (Advanced Protein Products Ltd.) supplemented with 10% dialysed foetal calf serum (FCS; Gibco), 50 IU/ml penicillin, 50 μg/ml streptomycin and 50 μM MSX (Sigma) (see Appendix III). Cultures were maintained at 37°C in a humidified atmosphere of 5 % CO_2 in air. Cultures were initially seeded at 0.5 x 10^6 cells/ml and maintained by splitting each flask 1 in 5 on alternate days (or when confluent). Cells were harvested by trypsinisation as detailed in 4.2.1.2 below. All operations involving cell-lines were carried out aseptically in a

class II microbiological flow cabinet. Cells were routinely tested for mycoplasma using an enzyme-linked immunosorbant assay (ELISA) kit (Boehringer Mannheim).

4.2.1.2 Trypsinisation

Growth medium was removed by aspiration and the monolayer washed with 10 ml of PBS. The PBS was removed and 2 ml of 0.05 % (w/v) trypsin, 0.02 % (w/v) EDTA solution (Sigma) was added with gentle rocking to ensure even distribution. The flask was stoppered and incubated at 37°C for 2 minutes. The cells were dislodged by banging the flask on the bench several times. Fresh medium was added and the resulting cell suspension counted using a haemocytometer. The required number of cells was harvested by centrifugation at 450 x g for 5 minutes.

4.2.1.3 Preservation of Cells by Freezing

Typically, 2 x 10^8 cells were harvested by trypsinisation as in 4.2.1.2 and resuspended in 1.0 ml of dialysed FCS (Gibco) containing 10 % dimethyl sulphoxide (BDH). Aliquots of cell suspension were placed in cryo tubes embedded in polystyrene to ensure slow and even freezing. The tubes were placed at -70°C overnight and then transferred to liquid nitrogen for long-term storage. Re-establishment of frozen cell-lines was carried out by rapidly thawing aliquots in a 37°C water bath followed by layering of the cells onto 10 ml of pre-warmed GMEM-S medium (Advanced Protein Products Ltd.) supplemented with 10% dialysed FCS (Gibco), 50 IU/ml penicillin, 50 μg/ml streptomycin and 2 mM L-glutamine. The cells were harvested by centrifugation at 450 x g for 5 minutes and then resuspended in 10 ml of GMEM-S as above and transferred to a 75 cm^2 tissue culture flask. When confluent the cells were maintained as above.

4.2.1.4 Protein Expression

Medium was removed from confluent monolayers of cells by aspiration and replaced with 50 ml of serum free ULTRA-CHO medium (BioWhittaker). Flasks were incubated for up

to 5 weeks at 37°C in a humidified atmosphere of 5 % CO_2 in air. Samples were taken aseptically at weekly intervals and the expression of LIFR-FN3-3C-Fc monitored by means of ELISA as detailed in 4.2.5.1 below.

Different cell culture vessels were tested for optimal cell growth and protein expression including a small-scale stirred-tank bioreactor system (New Brunswick Scientific), roller bottles (Falcon) and large-scale tissue culture flasks (Costar). The use of sodium butyrate induced protein expression was also investigated (D'Anna *et al.*, 1980; Gorman *et al.*, 1983; Reeves and Cserjesi, 1979). Typically, cells were incubated as above in ULTRA-CHO medium (BioWhittaker) supplemented with 2 mM sodium butyrate (Sigma).

Supernatants were harvested and phenylmethylsulphonyl fluoride (PMSF) and EDTA added to a final concentration of 1 mM. Samples were centrifuged at 2000 x g for 30 minutes or filtered through Millex AP50 pre-filters (Millipore). Samples were kept at 4°C for short-term storage or -20°C for long-term storage. Previously frozen samples were centrifuged at 5000 x g for 10 minutes at 4 °C. Finally, supernatants were passed through 0.45 μm membrane filters (Millipore) and the samples adjusted to pH 7.0.

4.2.1.5 Affinity Purification and Cleavage of Fc Fusion Proteins

The chromatography support (Bio-Rad PolyPrep) was packed with 1 ml of Protein-A sepharose Fast Flow (Pharmacia) per litre of supernatant to be processed. The column was washed with 20 column volumes of PBS and the supernatant passed through the column under gravity at 4°C followed by 20 column volumes of PBS. Fc fusion proteins were cleaved, *in situ*, using 3C-His_6 protease as detailed in Chapter 3.

4.2.2 Stable Expression in Schneiders *Drosophila* Cell Line 2 (S2)

Schneiders *Drosophila* cell line 2 (S2) is an epithelial cell line established from 20-24 hour embryos of *Drosophila melanogaster* (Schneider, 1972). These cells grow as loosely adherent

monolayers or in suspension at temperatures between 16°C and 28°C, with the optimum lying between 22°C and 24°C. The expression plasmid pRmHa-3 (Bunch *et al.*, 1988) is utilised containing the inducible *Drosophila* metallothionein promoter (see Figure 4-2). In this system the promoter is activated by elevated amounts of the heavy metal ions Cu^{2+}, Zn^{2+} or Cd^{2+}; copper chloride was used in this study as it is less toxic and does not induce a heat-shock response (Bunch *et al.*, 1988). pRmHa-3-LIFR-FN3-3C-Fc constructs were kindly provided by K. Chobotova, CRC Growth Factor Group. Birmingham.

4.2.2.1 S2 Cell Culture and Protein Expression

Cell culture was essentially carried out as in 4.2.1.1 above, except, pRmHa-3-LIFR-FN3-3C-Fc transfected S2 cells were grown in SF900 medium (Gibco) supplemented with 5 % insect tested serum (Gibco), 50 IU/ml penicillin, 50 μg/ml streptomycin and 300 μg/ml hygromycin B (Sigma). Typically, 500 ml Erlenmeyer flasks were seeded at a cell density of $1x10^6$ cells/ml and subsequent protein expression carried out in 100 ml of growth medium containing no hygromycin B selection and serum concentrations ranging from 0-1 %. Flasks were incubated with gentle shaking (~100 RPM) at 27°C for 24 hours. Protein expression was induced with copper chloride to a final concentration of 0.5 mM. Cells were counted on a daily basis and expression of protein monitored by means of SDS-PAGE and western blot.

4.2.3 Transient Expression in 293T Cells

293T cells (293/tsA1609neo; Graham *et al.*, 1977) are a human epithelial/embryonal kidney cell line transformed with the SV40 large T antigen, so allowing high levels of transient expression from plasmids containing a polyoma origin of replication. The transient expression vector pIG-1 (Simmons, 1992) containing the hCMV enhancer and promoter was used. This vector contains both an SV40 origin and polyoma origin so allowing replication of this vector in polyoma transformed cells such as COS and 293T

cells. The remaining elements of this vector allow replication and drug selection in *E. coli*. The supF element encodes a tRNA that suppresses amber stop codons in ampicillin and tetracycline resistance genes carried on a stable episome, termed P3, found in the *E. coli* strain MC101. pIG-1-LIFR-FN3-3C-Fc constructs were kindly provided by K. Chobotova, CRC Growth Factor Group. Birmingham.

4.2.3.1 pIG-1 DNA Production and Purification

E.coli strain MC101/P3, previously transfected with pIG-1-LIFR-FN3-3C-Fc as in chapter 2, were grown in 4 L of LB medium containing 50 μg/ml ampicillin and 12 μg/ml tetracycline for 16-10 hours. The cells were harvested by centrifugation at 6000 x g for 15 minutes at 4°C and plasmid extracted using a Maxi Prep™ Kit (Qiagen). The purified pIG-1-LIFR-FN3-3C-Fc was triple digested with *Bam*HI, *Bgl*II and *Hind*III as in Table 2-4 below for confirmation of identity. The reaction mixture was incubated at 37°C for 2 hours and the digested samples were run on a 1 % agarose-TAE gel.

Volume (μl)	Component
1.00	pIG-1-LIFR-FN3-3C-Fc DNA
1.00	10x Restriction Buffer B
0.50	*Bam*HI (10 units/ml)
0.50	*Hind*III (10 units/ml)
0.25	*Bgl*II (40 units/ml)
6.75	Sterile Distilled Water
10.00	Total

Table 4-1

Table showing composition of restriction enzyme reaction mixture used to confirm identity of pIG-1-LIFR-FN3-3C-Fc.

4.2.3.2 293T Cell Culture

Cell culture was essentially carried out as in 4.2.1.1 above, except, wildtype 293T cells were grown in DMEM medium (Gibco) supplemented with 10 % FCS (LabTech International), 50 IU/ml penicillin, 50 μg/ml streptomycin, 1 mM pyruvate, 2 mM glutamate and 4500

mg/ml glucose. Typically, 162 cm^2 tissue culture flasks (Costar) containing 30 ml of medium were seeded at a cell density of 5 x 10^4 cells/ml and grown to ~90 % confluence at 37°C in a humidified atmosphere of 5 % CO_2 in air. Cells were dislodged from the growth surface by banging the flask and the cells harvested by centrifugation at 450 x g for 5 minutes. Cells were carefully resuspended in 20 ml of medium using a Pasteur pipette and then diluted 1 in 10 into fresh tissue culture flasks containing 18 ml of medium. Cells were grown for a further 24 hours to ~50-70 % confluence (sub-confluent) and then transfected with pIG-1-LIFR-FN3-3C-Fc as in 4.2.3.3 below.

4.2.3.3 Transfection of 293T Cells

DNA was sterilised by precipitation in 2 volumes of ethanol followed by centrifugation at 15000 x g. The pellet was resuspended in sterile filtered 2M calcium chloride. As a final insurance of sterility the DNA preparation was centrifuged at 15000 x g for 15 minutes in a sterile Eppendorf tube. Transfection was carried out in 162 cm^2 tissue culture flasks (Costar) using 120 μg of DNA per flask. Typically, 0.48 ml of pIG-1-LIFR-FN3-3C-Fc DNA at a concentration of ~1.0 mg/ml was added to 1.24 ml of 2M calcium chloride and 10.28 ml of sterile distilled water. This solution was added drop wise, over a period of several minutes, to 12 ml of 1.6 % (w/v) sodium chloride, 1.2 % (w/v) HEPES and 0.04 % (w/v) disodium hydrogen phosphate (HBS)(see Appendix III). The precipitate was inverted and 6 ml pipeted into each flask of sub-confluent cells followed by mixing. The precipitate was left on the cells for 16 hours and the monolayer washed with PBS before adding 30 ml of serum free ULTRA-CHO medium (BioWhittaker). The cells were incubated at 37°C in a humidified atmosphere of 5 % CO_2 in air for 4-8 days. For large-scale preparations ~9 mg of pIG-1-LIFR-FN3-3C-Fc DNA was transfected into 72 sub-confluent 162cm^2 tissue culture flasks and expression carried out for 1 week after which medium was harvested and fresh medium added for a second week of protein expression. Expression of protein was monitored by means of SDS-PAGE and western blot.

4.2.4 Deglycosylation of LIFR

The use of the enzyme peptide-N-glycosidase F (PNGase F; Tarentino *et al.*, 1990) was investigated as a means of deglycosylation as LIFR-FN3 had previously proved recalcitrant to other means of deglycosylation tested (see 4.4 for details). This enzyme functions by cleaving between the innermost N-acetylglucosamine and asparagine residues of high mannose, hybrid, and complex oligosaccharides of N-linked glycoproteins. Two variants of PNGase F were tested under non-denaturing conditions; non-recombinant material isolated from *Flavobacterium meningosepticum* (New England Biolabs) and recombinant protein expressed in *E.coli* (Barsomian *et al.*, 1990; kindly provided in pMal vector for expression as a maltose binding protein (MBP) by M. A. Hall, CRC Growth Factors Group. Birmingham.)

4.2.4.1 PNGase F Treatment

Samples of LIFR-FN3 were dialysed using a microdialysis chamber (Sigma) through 12-14 kDa exclusion limit membrane (Gibco) against PNGase F (G7) buffer consisting of 0.05 M sodium phosphate, 10 % NP-40, pH 7.5 for 5 hours at 4^{o}C. Non-recombinant PNGase F (New England Biolabs) and recombinant PNGase F-MBP (M. A. Hall, CRC Growth Factors Group, Birmingham) were incubated with glycosylated LIFR-FN3-3C-Fc as detailed in Table 4-2 below. Deglycosylation was carried out in G7 buffer at room temperature or 37^{o}C for between 1 and 24 hours. Deglycosylation was monitored by SDS-PAGE and western blot analysis (see Chapter3).

	Tube / Reaction							
Component	1	2	3	4	5	6	7	8
LIFR-FN3	—	—	17.0 µg	17.0 µg	17.0 µg	17.0 µg	17.0 µg	17.0 µg
PNGase F	30.6 µg (NR)	9.4 µg (R)	—	30.6 µg (NR)	9.4 µg (R)	18.8 µg (R)	30.6 µg (R)	—
x10 G7 Buffer	—	—	—	5.0 µl	5.0 µl	5.0 µl	5.0 µl	5.0 µl
Distilled Water	41.5 µl	41.5 µl	40.0 µl	26.5 µl	26.5 µl	17.9 µl	7.2 µl	35.0 µl
Total Volume	50.0 µl	50.0 µl	50.0 µl	50.0 µl	50.0 µl	50.0 µl	50.0 µl	50.0 µl

Table 4-2

Table showing PNGase F reaction mixtures used for deglycosylation of LIFR-FN3 under non-denaturing conditions. Deglycosylation was carried out at 37 ^{o}C in G7 buffer consisting of 0.05 M sodium phosphate, 10 % NP-40, pH 7.5 for between 1 and 24 hours. Non-recombinant PNGase F (NR) was used at a concentration of 3.6 mg/ml and recombinant PNGase F (R) at a concentration of 1.1 mg/ml. LIFR-FN3 was used at a concentration of 1.7 mg/ml.

4.2.4.2 Production of Recombinant PNGase F

Typically, 400 ml of sterile LB-amp (see Appendix I) supplemented with 0.3 % (w/v) glucose was inoculated with a colony of pMal-PNGase F-MBP *E.coli* (Kindly provided by M. A. Hall and Dr. K. Hudson, CRC Growth Factors Group, Birmingham.). The culture was incubated at 25^{o}C with shaking for 24 hours. The culture was diluted 100-fold in LB-amp and grown to OD_{600}~0.6 and IPTG added to a final concentration of 0.3 mM. The culture was grown for a further 4 hours at 25^{o}C with shaking. The cells were harvested by centrifugation at 5000 x g for 5 minutes and the pellet resuspended in 200 ml of lysis buffer consisting of 20 % sucrose and 30 mM Tris-HCl, pH 8.0 (see Appendix III). EDTA was added to a final concentration of 1 mM and the suspension incubated at room temperature with shaking for 10 minutes. The cells were harvested by centrifugation and resuspended in 200 ml of 5 mM magnesium sulphate, followed by incubation on ice for 10 minutes. The lysate was centrifuged at 6000 x g for 10 minutes and 20 mM Tris-HCl, pH

8.0 added. The PNGase F was purified by affinity chromatography as detailed in 4.2.4.3 below.

4.2.4.3 Affinity Purification of Recombinant PNGase F

The chromatography support (Bio-Rad PolyPrep) was packed with 1 ml of amylose resin (Pharmacia) and washed with 5 ml of PNGase F column buffer containing 0.2 M NaCl, 20 mM Tris-HCl, 10 mM mercaptoethanol, pH 8.0 (see Appendix III). The supernatant was passed through the column under gravity at 4°C and the column washed with 5 ml of PNGase F column buffer. Finally, protein was eluted with 0.5 ml aliquots of PNGase F elution buffer consisting of 0.2 M NaCl, 20 mM Tris-HCl, 10 mM mercaptoethanol and 10 mM maltose, pH 8.0 (see Appendix III). Protein fractions were analysed by SDS-PAGE and Coomassie assay.

4.2.5 Methods of Analysis

4.2.5.1 ELISA

MaxiSorp 96 well plates (Nunc) were coated with 1 μg/ml mouse anti-human IgG1 Fc monoclonal antibody (Pierce) diluted in PBS. The plates were covered and left for 2 hours at room temperature. The antibody was discarded and the plate then blocked with 1 % (w/v) BSA in PBS for 1 hour followed by two washes with 0.05 % (v/v) Tween-20 in PBS. Standard curves were constructed using 100 μl/well duplicates of human IgG1 at 500, 375, 250, 125, 62, 31, 16 and 8 ng/ml. Test samples were added at 100 μl/well in duplicate and the plate left at room temperature for 1 hour. The plate contents were discarded and the plate then washed twice with 0.05 % (v/v) Tween-20 in PBS. Horseradish peroxidase conjugated polyclonal goat anti-human IgG1 Fc (Pierce) was added at a dilution of 1:5000 in 1 % (w/v) BSA in PBS and the plate incubated at room temperature for 1 hour. The plate contents were discarded and the plate then washed twice with 0.05 % (v/v) Tween-20 in PBS followed by two washes with PBS. Substrate

development was carried out using 8 mg of 1,2-phenylenediamine (OPD) (Sigma) as chromogen dissolved in 12 ml of citric acid phosphate buffer consisting of 0.035 M citric acid, 0.067 M disodium hydrogen phosphate, pH 5.0 (see Appendix III). Chromogen solution was activated by addition of 5 µl of hydrogen peroxide and the solution added to each well and the colour developed, in the dark, for 10-15 minutes. The reaction was terminated by the addition of 75 µl of 1 M sulphuric acid to each well and the absorbance read at 490 nm.

4.2.5.2 Western Blot Analysis

Western blots were carried out as detailed in Chapter 3 except the primary polyclonal antibody consisted of sheep anti-LIFR (Binding Site) at a dilution of 1:10000 in PBS containing 5 % (w/v) BSA and 0.1 % (v/v) Tween-20.

4.2.5.3 Iso-electric Focussing

5.5% polyacrylamide iso-electric focussing (IEF) gels were prepared using broad range ampholines in the pI range of 3-10 (Bio-Rad) as in Table 4-3 below. Gels were run on a Mini-Protean II gel apparatus (Bio-Rad).

Component	Volume
30% (w/w) Acrylamide/0.8 % (w/w) N' N'-bis-methylene-acrylamide	1.28 ml
Water	3.89 ml
50% (v/v) Glycerol	1.40 ml
Ampholines (pI 3-10)	387 µl
10% (w/v) Ammonium Persulphate	29 µl
TEMED	14 µl
Total	7.00 ml

Table 4-3
Table showing composition of 5.5% polyacrylamide IEF gels.

Samples were diluted 1:1 in loading buffer consisting of 115 µl water, 60 µl 50% (v/v) glycerol and 30 µl ampholines. Separation was carried out for approximately 3 hours at 150 V (15 mA maximum) at 4°C. Anolyte and catholyte consisted of 0.100 M acetic acid and 0.02 M sodium hydroxide, respectively. IEF standards containing phycocyanin (pI 4.5 and 4.8), equine myoglobin (pI 6.8), human haemoglobin A (pI 7.1) and human haemoglobin C (pI 7.5) (Bio-Rad) were run along with a control sample of OSM.

4.2.5.4 Dynamic Light Scattering

Samples of LIFR-FN3-3C-Fc were filtered through a 0.1 µm membrane (Whatman) and injected into a DynaPro-LS Sizing Instrument (Protein Solutions Inc.). Typically, 0.5 ml of 2 mg/ml sample was injected and measurements taken over a 10 minute period at 20°C.

4.2.5.5 Receptor Competition and Titration Assays

Receptor competition and titration assays were carried out as detailed in chapter 3 with the exceptions that Protein-G was used in place of Protein-A as better signal to noise ratios were reported (K. Hudson, CRC Growth Factor Group. Birmingham, personal communication). For titration assays biotinylated LIF was used at saturating concentrations of 1000 ng/ml and FGFR2 included as control. For competition assays 200 ng/ml of biotinylated LIF was used as competent and OSM included as control.

4.2.6 Production, Purification and Characterisation of LIFR-LIF Complex

4.2.6.1 Expression and Purification of LIF

LIF was expressed and purified as a GST fusion in a similar manner to OSM in Chapter 3 (Robinson *et al.*, 1994). Typically, 6 L of culture was grown and the fusion protein purified on glutathione sepharose. Cleavage of the fusion protein was carried out using 105 units of thrombin (Sigma) in 150 mM sodium chloride, 2.5 mM calcium chloride and 50 mM Tris-HCl, pH 8.0 for 16 hours at room temperature.

4.2.6.2 Analysis of LIFR-LIF Complex

Affinity purified LIF and LIFR-FN3 were mixed in molar ratios of 10 to 1 and incubated at room temperature for 1 hour. Samples were centrifuged at 5000 x g to remove aggregates and the supernatant applied to a Superose-6 column (24 ml bed volume, 10 x 300 mm, separation range 5-5000 kDa; Pharmacia). Columns were typically run at a flow rate of 2 ml/min in 150 mM sodium chloride, 50 mM Tris-HCl, pH 7.5 on a Biosys 510 FPLC (Beckman). Columns were calibrated by running a cocktail of protein markers consisting of 0.5 mg each of cytochrome C (12.4 kDa), carbonic anhydrase (29 kDa), albumin bovine serum (66 kDa), alcohol dehydrogenase (150 kDa) and blue dextran (2000 kDa).

4.3 Results of Expression, Purification and Characterisation of LIFR

4.3.1 Stable Expression in CHO Cells

Expression levels of LIFR-FN3-3C-Fc in CHO cells proved exceptionally low as assessed by ELISA as shown in Figure 4-3 below. Maximal levels of expression of ~1 mg/L were typically obtained after 3 weeks incubation in the presence of 2 mM sodium butyrate. Sodium butyrate was shown to increase levels of expression by upto 30 %. All methods of scale-up, including roller bottles, cells factories and stirred-tank bioreactor systems, proved unsuccessful with cells showing poor adherence and protein expression. The inclusion of MSX was shown to be essential for maintenance of clonal selection over many passages although no increase in the levels of protein expression was observed on MSX treatment.

Soluble glycosylated LIFR-FN3-3C-Fc was successfully purified using affinity purification on Protein-A sepharose at an apparent molecular weight of ~140 kDa as assessed by SDS-PAGE (data not shown). Soluble LIFR-FN3 was successfully liberated on 3C protease treatment at an apparent molecular weight of ~116 kDa and preliminary deglycosylation with non-recombinant PNGase F resulted in a species with an apparent molecular weight of ~90 kDa as assessed by SDS-PAGE (data not shown). The identity of LIFR-FN3 was confirmed by N-terminal sequencing although all attempts at mass spectroscopy failed due to lack of ionisation.

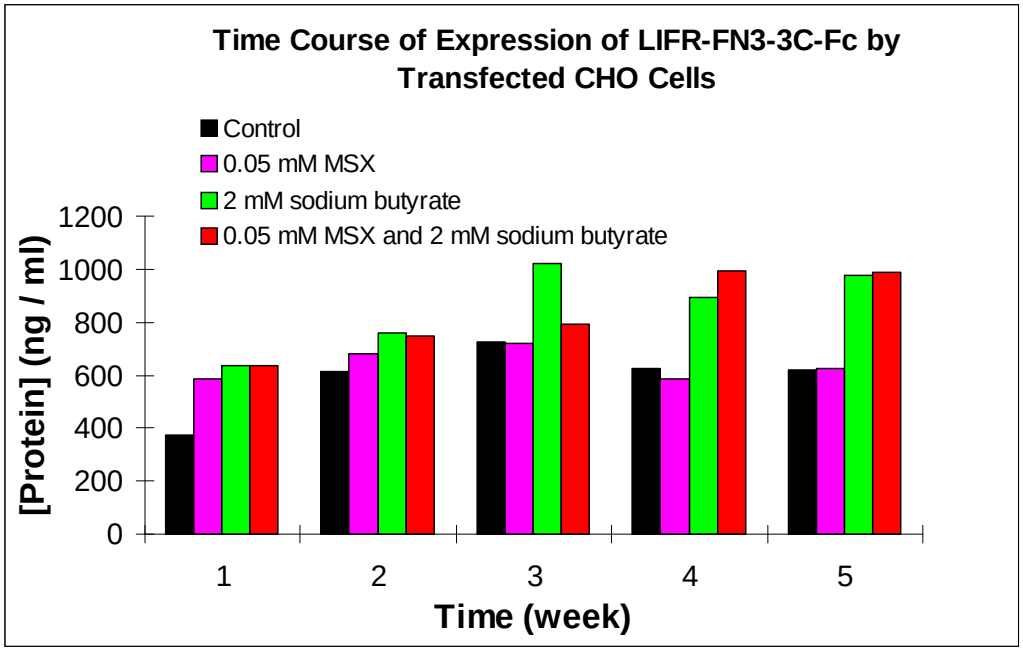


Figure 4-3

Graph showing levels of expression of LIFR-FN3-3C-Fc by stably transfected CHO cells. Cells were grown as confluent monolayers in 162cm^2 tissue culture flasks in 30 ml of ULTRA-CHO medium (BioWhittaker). Samples of medium were taken at weekly intervals for analysis by means of ELISA. Protein concentrations were determined by comparison with human IgG1 standards. Control samples represent protein expressed in ULTRA-CHO medium and other data show protein expressed in medium supplemented with 0.05 mM L-methionine sulphoximine (MSX) and 2 mM sodium butyrate.

4.3.2 Stable Expression in Schneiders *Drosophila* Cell Line 2 (S2)

No induction of LIFR-FN3-3C-Fc expression was detected using ELISA, Protein-A immunoprecipitation or western blot and samples typically contained a protein with an apparent molecular weight of ~46 kDa, compared with control LIFR-FN3-3C-Fc samples running at ~140 kDa (data not shown). No cross reactivity was observed between LIFR antibodies and samples of control expression medium containing serum, so suggesting the production of truncated translation products. Samples of S2 conditioned medium were shown to be inactive in ligand binding assays as no binding to LIF or OSM was observed, compared with control samples of supernatant containing active gp130-Fc or LIFR-Fc which showed sigmoidal dose dependant binding.

4.3.3 Transient Expression in 293T Cells

4.3.3.1 pIG-1 DNA Production and Purification

Figure 4-4 below shows a 1 % agarose gel of *Bam*HI, *Bgl*II and *Hind*III digestion of pIG-1-LIFR-FN3-3C-Fc. Lane 5 shows the five main marker bands of the λ *Hind*III digest, running at 9146, 6882, 2322, 2027 and 564 base pairs. Lane 1 contains *Bam*HI-*Hind*III digested DNA which shows a unique drop out at ~2505 bp, lane 2 contains *Bgl*II-*Hind*III digested DNA which shows a characteristic drop out at ~1752 bp and lane 3 contains *Bam*HI-*Bgl*II digested DNA which shows a unique drop out at ~753 bp. Lane 4 contains undigested pIG-1-LIFR-FN3-3C-Fc DNA with one main band running at ~7900 bp.

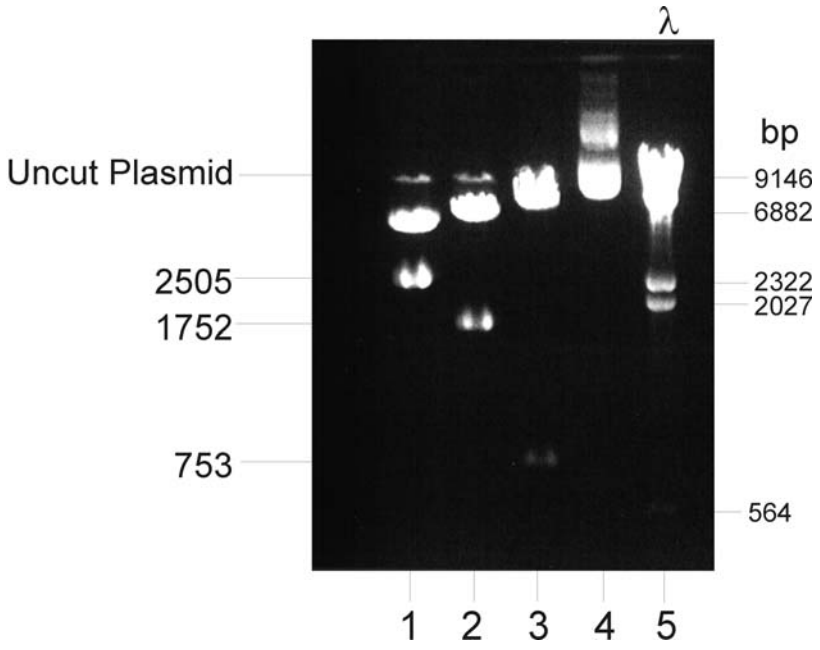


Figure 4-4

Figure showing 1 % agarose gel of *Bam*HI, *Bgl*II and *Hind*III digestion of pIG-1-LIFR-FN3-3C-Fc DNA. Lane 5 shows the five main marker bands of the λ *Hind*III digest, running at 9146, 6882, 2322, 2027 and 564 base pairs. Lanes 1, 2 and 3 contain *Bam*HI-*Hind*III, *Bgl*II-*Hind*III and *Bam*HI-*Bgl*II digested samples of pIG-1-LIFR-FN3-3C-Fc DNA, respectively. Lane 4 contains undigested pIG-1-LIFR-FN3-3C-Fc DNA.

4.3.3.2 SDS-PAGE

Figure 4-5 below shows a 12 % polyacrylamide gel of the affinity purification of LIFR-FN3-3C-Fc expressed transiently in 293T cells. Lane 1 shows a sample of Protein-A sepharose after passage of conditioned 293T supernatant and one main band is evident at ~140 kDa. Lane 2 shows a sample of Protein-A sepharose after cleavage with 3C protease and one main band is seen at ~35 kDa. Lanes 3 and 4 show unbound material and a wash fraction, respectively. Lanes 5, 6 and 7 show the first three eluted fractions of LIFR-FN3 following 3C protease cleavage with one main band running at ~116 kDa. Lane 8 contains high range molecular weight markers running at 220, 97, 66, 46, 30, 21 and 14 kDa (Amersham).

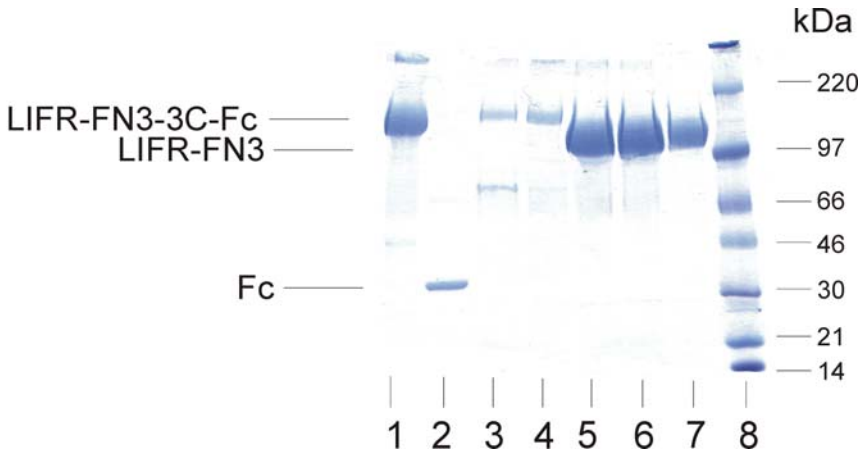


Figure 4-5

Figure showing 12 % polyacrylamide gel of the expression and affinity purification of LIFR-FN3-3C-Fc. Lane 1 contains Protein-A sepharose after passage of 293T supernatant. Lane 2 shows the Protein-A sepharose after cleavage with 3C protease. Lane 3 and 4 show unbound material and a wash step, respectively. Lane 5, 6 and 7 show the first three eluted fractions of LIFR-FN3 following 3C protease cleavage. Lane 8 contains high range molecular weight markers running at 220, 97, 66, 46, 30, 21 and 14 kDa (Amersham).

4.3.3.3 Western Blot Analysis

Figure 4-6 below shows a western blot of LIFR-FN3-3C-Fc samples at various stages of the purification scheme. Lane 1 contains conditioned 293T supernatant and no bands are visible. Lane 2 contains a sample of Protein-A sepharose after passage of conditioned 293T supernatant and three major bands are observed at ~220, ~140 and ~46 kDa. Lane 3 contains a sample of Protein-A sepharose wash and no bands are observed. Lane 4 contains recombinant PNGase F and lane 7 contains non-recombinant PNGase F (New England Biolabs) and no cross reactivity with the LIFR antibody is observed. Lanes 5 and 6 contain samples of Protein-A sepharose bound LIFR-FN3-3C-Fc after treatment with recombinant PNGase F and non-recombinant PNGase F, respectively. Comparison of the ~140 kDa band in lanes 2, 5 and 6 suggests a slight reduction in molecular weight following PNGase F treatment, with the non-recombinant variant displaying a tighter band so suggesting more complete deglycosylation. Lane 8 contains 3C-His_6 protease and no cross reactivity with the LIFR antibody is observed. Lanes 9 and 10 contain eluted fractions of LIFR-FN3 after 3C protease treatment and a single diffuse band is observed at ~116 kDa. Lane 11 contains a sample of Protein-A sepharose after elution of LIFR-FN3 and no bands are observed suggesting complete elution of LIFR. The positions of the broad range molecular weight rainbow markers is noted at 220, 97, 66, 46 and 30 kDa (Bio-Rad).

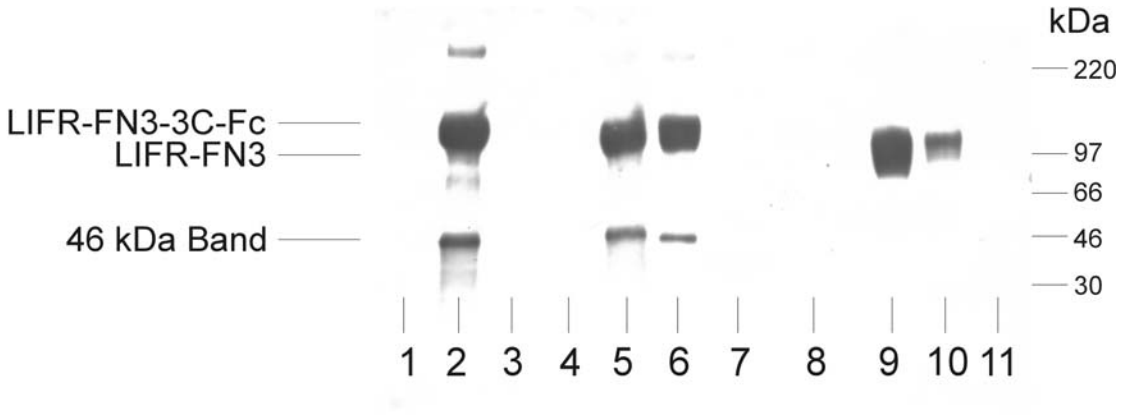


Figure 4-6

Figure showing western blot of LIFR-FN3-3C-Fc samples at various stages of the purification scheme. 12 % polyacrylamide gel is shown probed with anti-LIFR antibody (Binding Site). Lane 1 contains conditioned 293T supernatant and lane 2 contains Protein-A sepharose after passage of supernatant. Lane 3 contains Protein-A sepharose wash. Lane 4 contains recombinant PNGase F and lane 5 contains Protein-A sepharose bound LIFR-FN3-3C-Fc after treatment with recombinant PNGase F. Lane 6 contains Protein-A sepharose bound LIFR-FN3-3C-Fc after treatment with non-recombinant PNGase F (New England Biolabs) and lane 7 contains non-recombinant PNGase F. Lane 8 contains 3C-His_6 protease and lanes 9 and 10 contain eluted fractions of LIFR-FN3 after 3C protease treatment. Lane 11 contains Protein-A sepharose after elution of LIFR-FN3. The positions of the broad range molecular weight rainbow markers are noted at 220, 97, 66, 46 and 30 kDa (Bio-Rad), although no bands are visible on the blot.

4.3.3.4 Iso-electric Focussing

Figure 4-7 below shows a western blot of 5.5% polyacrylamide IEF gel of various samples of LIFR-FN3-3C-Fc. Lane 1 contains 3C protease cleaved LIFR-FN3 and no major bands are observed. Lane 2 contains acid eluted LIFR-FN3-3C-Fc which again shows no major bands and lane 3 contains Protein-A bound LIFR-FN3-3C-Fc which shows one major band at a pI ~5-6; compared with the theoretical pI of the LIFR-FN moiety at 7.2. Lane 4 contains non-recombinant PNGase F treated LIFR-FN3-3C-Fc and no major bands are observed. Lane 5 contains a control sample of OSM in which no cross reactivity to the LIFR antibody is observed. Lane 6 shows IEF standards containing phycocyanin (pI 4.5 and 4.8), equine myoglobin (pI 6.8), human haemoglobin A (pI 7.1) and human haemoglobin C (pI 7.5) (Bio-Rad) which all show single bands due to the luminescent detection method used.

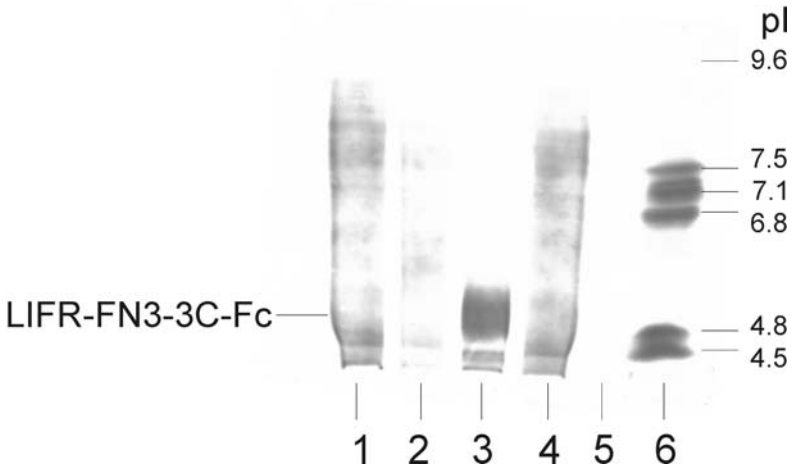


Figure 4-7

Figure showing western blot of 5.5% polyacrylamide IEF gel prepared using broad range ampholines (pI 3-10)(Bio-Rad). LIFR antibody (Binding Site) was used with ECL (Amersham) chemoluminescent detection. Lane 1 contains 3C protease cleaved LIFR-FN3. Lane 2 contains acid eluted LIFR-FN3-3C-Fc and lane 3 contains Protein-A bound LIFR-FN3-3C-Fc. Lane 4 contains PNGase F (New England Biolabs) treated LIFR-FN3-3C-Fc and lane 5 contains a control sample of OSM. Lane 6 shows IEF standards containing phycocyanin (pI 4.5 and 4.8), equine myoglobin (pI 6.8), human haemoglobin A (pI 7.1) and human haemoglobin C (pI 7.5) (Bio-Rad).

4.3.3.5 Dynamic Light Scattering

Figure 4-8 below shows samples of deglycosylated LIFR-FN3 analysed on a DynaPro-LS Sizing Instrument (Protein Solutions Inc.) at various time points. Panel (a) shows the sample at 0 minutes, directly after filtering through a 0.1 μm membrane filter, and one major species at a hydrodynamic radii of ~1 x 10^{-2} μm is observed. Panel (b) shows the same sample after 5 minutes incubation at 20^{o}C and these data display one major species at ~1 x 10^{-2} μm with the peak being split into a second minor species at ~1.5 x 10^{-1} μm. Panel (c) shows the sample after 10 minutes incubation and two distinct species at hydrodynamic radii of ~1 x 10^{-2} and ~2.9 x 10^{-1} μm are observed. Samples analysed after ~6 hours (data not shown) have the same profile as panel (c), suggesting equilibrium is reached after ~10 minutes.

The data in panel (a) successfully models as a monodisperse population with one major species at a hydrodynamic radius of 1.03 x 10^{-2} μm (10.3 nm) and an estimated molecular weight of ~84 kDa. The estimated molecular weight was calculated by comparison with the hydrodynamic radii of standard g;lobular proteins of known mass. This compares favourably with the theoretical molecular weight of LIFR-FN3 of ~94 kDa. Comparison of panels (a), (b) and (c) suggests that the system is in a state of dynamic equilibrium with large-scale aggregates (>0.1 μm) present in the solution along side a population at ~84 kDa. On filtering the aggregates are removed and over a period of 10 minutes the aggregates are reformed so re-establishing the position of equilibrium.

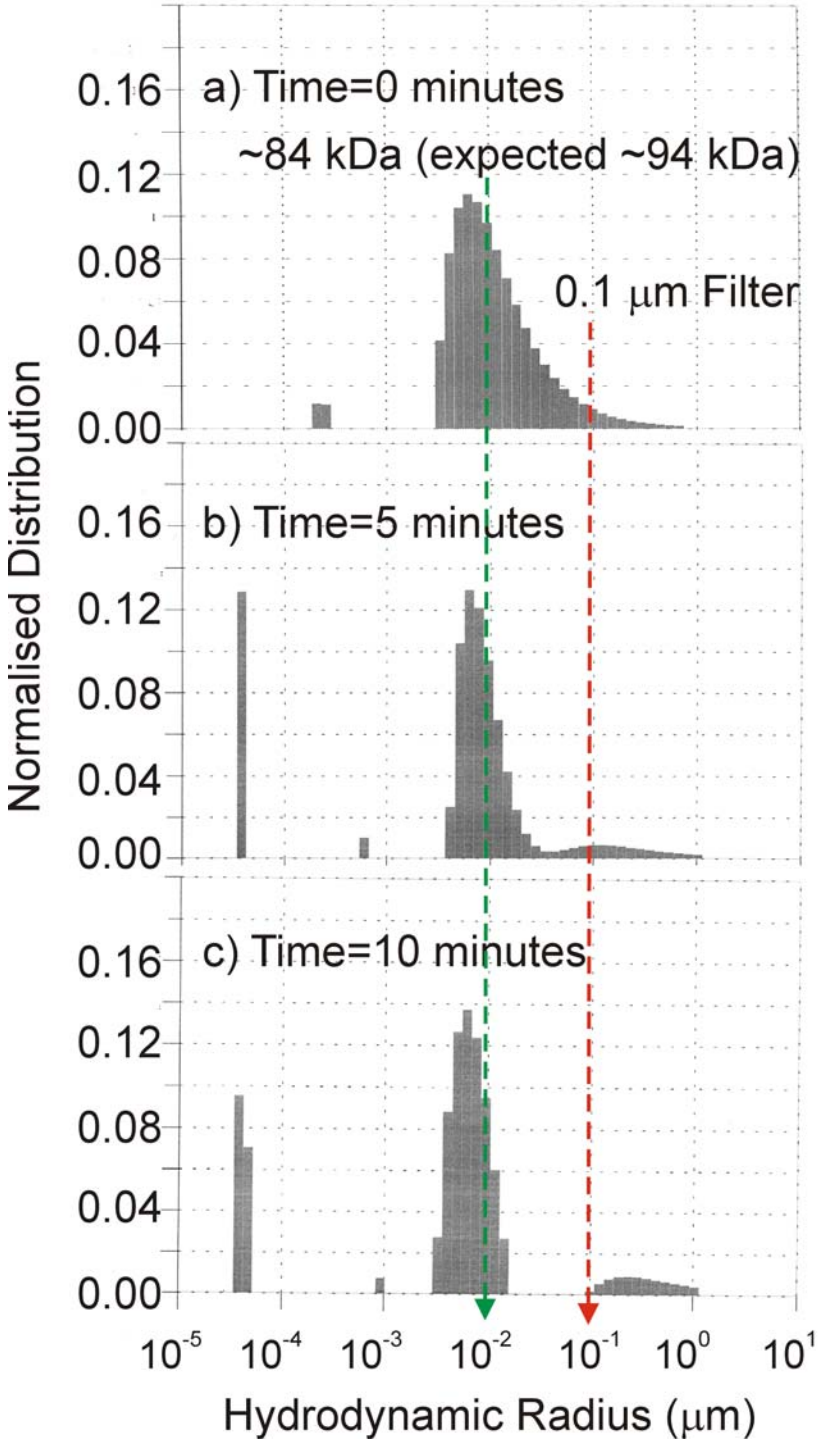


Figure 4-8

Figure showing dynamic light scattering data of deglycosylated LIFR-FN3. Samples of 293T-expressed LIFR-FN3 were deglycosylated overnight with non-recombinant PNGase F (New England Biolabs) and analysed on a DynaPro-LS Sizing Instrument (Protein Solutions Inc.). Typically, 0.5 ml of 2 mg/ml sample was injected through a 0.1 μm filter and measurements taken over a 10 minute period at 20°C. Sample shown at (a) 0 minutes, (b) 5 minutes and (c) 10 minutes after injection. Red dotted line represents hydrodynamic cut off of 0.1 μm filter. Green dotted line represents monodisperse population modelled at 10.3 nm with an estimated molecular weight of ~84 kDa.

4.3.3.6 Deglycosylation of LIFR with PNGase F

Figure 4-9 below shows a western blot of 12 % polyacrylamide gel of LIFR-FN3 samples from various deglycosylation reactions using PNGase F. Lane numbering corresponds to reactions setup in Table 4-2. Lanes 1 and 2 contain control samples of non-recombinant and recombinant PNGase F, respectively, and it can be seen that no cross reactivity to the LIFR antibody is observed. Lanes 3 and 8 contain control samples of LIFR-FN3 incubated with and without G7 buffer, respectively, and single distinct bands at ~116 kDa are observed. Lane 4 contains 17 μg of LIFR-FN3 incubated with 30.6 μg of non-recombinant PNGase F (New England Biolabs) and a single band at ~90 kDa is observed. Lanes 5, 6 and 7 contain 17 μg of LIFR-FN3 incubated with 9.4, 18.8 and 30.6 μg of recombinant PNGase F, respectively, and one main bands is observed at ~90 kDa. Lane 9 contains broad range molecular weight rainbow markers running at 135, 80, 43, 32, 14 and 8 kDa (Bio-Rad).

Comparison of lanes 3 and 7 suggest that deglycosylation of LIFR-FN3 is occurring on incubation with higher levels of recombinant PNGase F, due to the reduction in apparent molecular weight from ~114 to ~95 kDa. This compares favourably with the theoretical molecular weight of non-glycosylated LIFR-FN3 of ~94 kDa. Lane 5 suggests that lower levels of recombinant PNGase F are less effective and as such result in a slightly heavier and more diffuse protein population. Comparison with lane 4 shows that non-recombinant PNGase F is the most effective with comparable enzyme levels demonstrating a larger band shift to ~90 kDa and a less diffuse band.

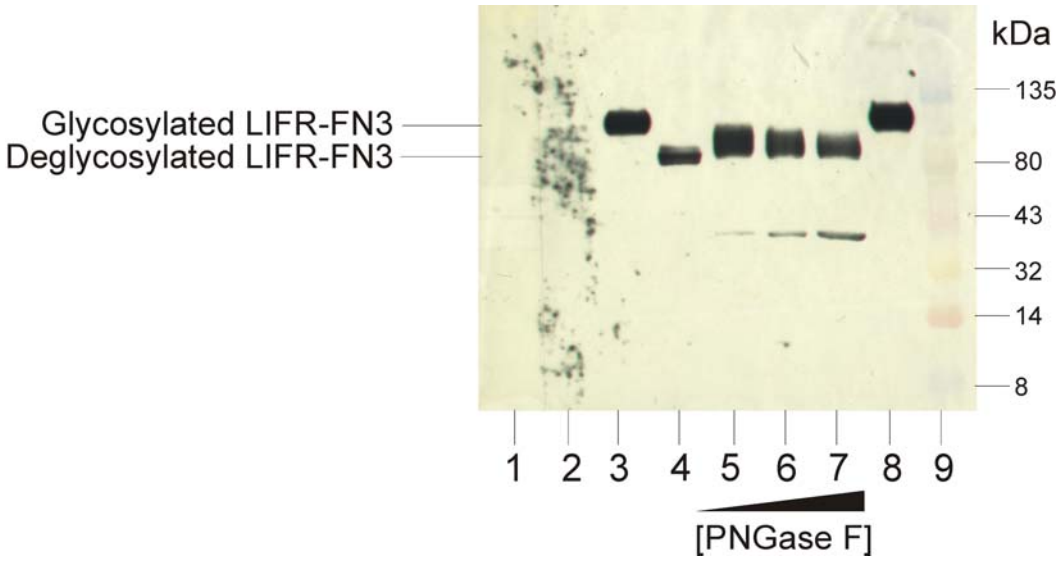


Figure 4-9

Figure showing western blot of LIFR-FN3 samples at various stages of deglycosylation using PNGase F. 12 % polyacrylamide gel is shown probed with LIFR antibody (Binding Site). Lane numbering corresponds to reactions setup in Table 4-2. Lanes 1 and 2 contain control samples of non-recombinant and recombinant PNGase F, respectively. Lanes 3 and 8 contain control samples of LIFR-FN3 incubated with and without G7 buffer, respectively. Lane 4 contains 17 μg of LIFR-FN3 incubated with 30.6 μg of non-recombinant PNGase F (New England Biolabs). Lanes 5, 6 and 7 contain 17 μg of LIFR-FN3 incubated with 9.4, 18.8 and 30.6 μg of recombinant PNGase F, respectively. Lane 9 contains broad range molecular weight rainbow markers running at 135, 80, 43, 32, 14 and 8 kDa (Bio-Rad).

4.3.3.7 Receptor Competition and Titration Assays

Figure 4-10 below shows graphs of (a) receptor titration and (b) receptor competition assays of ligand binding to LIFR-FN3-3C-Fc. Graph (a) demonstrates that the concentration range of ligand used leads to saturation of the LIFR-FN3-3C-Fc receptor under the conditions employed. This also shows that specific binding is occurring due to the sigmoidal increase in absorbance with increasing concentration of control human biotinylated LIF. Binding is not observed to purified/deglycosylated material or replenished 293T supernatant, as both show background responses similar to the control FGFR2 sample. Binding of human biotinylated LIF to 293T expressed LIFR-FN3-3C-Fc shows an EC_{50} value of 12.1 ng/ml which compares favourably to the control sample of CHO cell expressed material which displays an EC_{50} value of 10.4 ng/ml.

From graph (b) it can be seen that increasing the concentration of human LIF causes a sigmoidal decrease in the absorbance reading due to competition with biotinylated human LIF. Little or no fall in absorbance is demonstrated by the control samples of murine LIF or human OSM-196 suggesting that no competition is taking place. Binding of human LIF to 293T expressed LIFR-FN3-3C-Fc displays an EC_{50} value of 166.4 ng/ml.

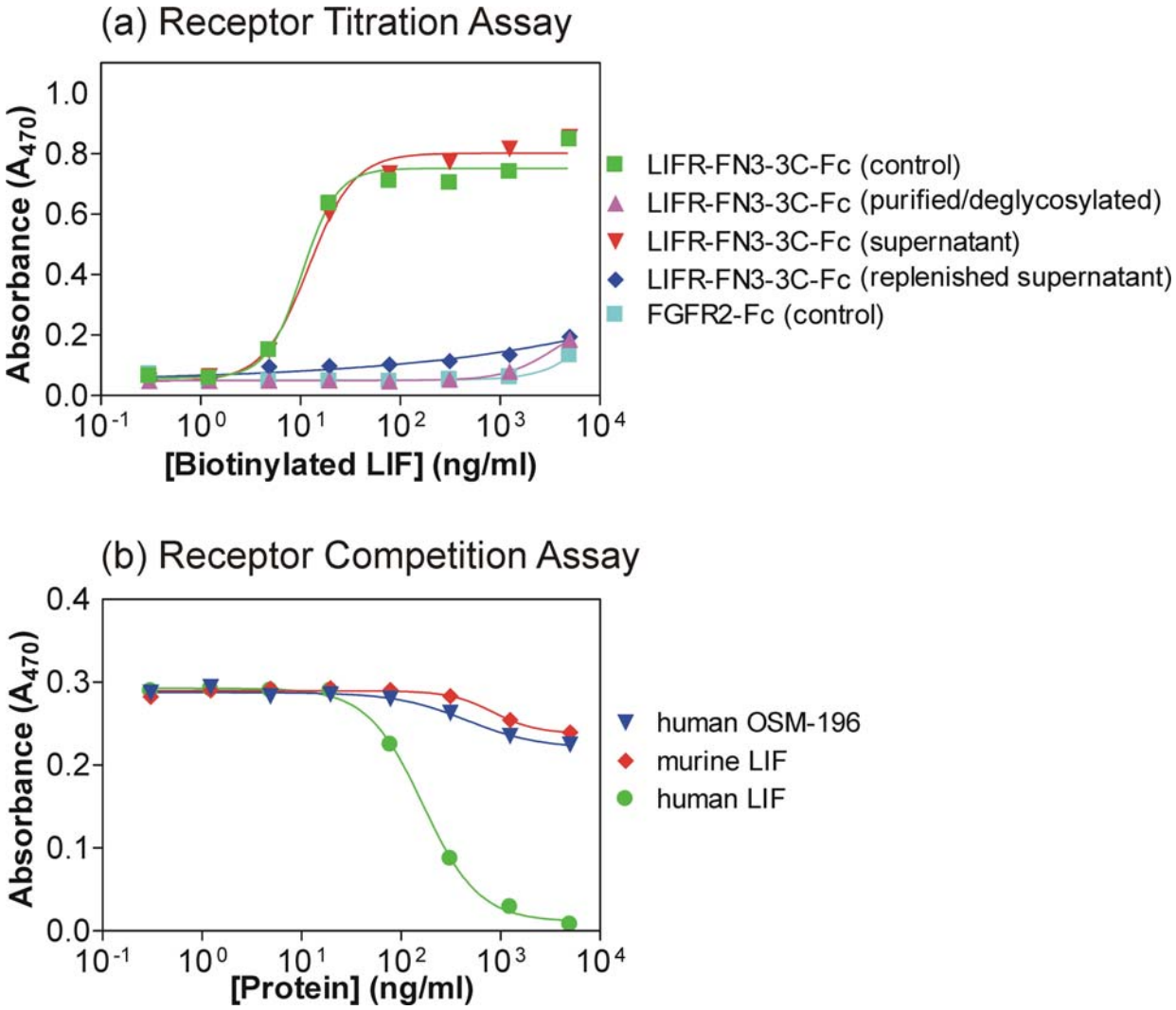


Figure 4-10

Figure showing graphs of (a) receptor titration and (b) receptor competition assays of ligand binding to 293T expressed murine LIFR-FN3-3C-Fc. Panel (a) shows binding of human biotinylated LIF to various samples of LIFR-FN3-3C-Fc including control (green; EC_{50} 10.4 ng/ml), purified/deglycosylated (pink), 293T supernatant (red; EC_{50} 12.1 ng/ml) and replenished 293T supernatant (blue). Panel (b) shows competitive binding of various ligands to murine LIFR-FN3-3C-Fc in the presence of 200 ng/ml human biotinylated LIF including human LIF (green; EC_{50} 166.4 ng/ml), murine LIF (red) and as control human OSM-196 (blue).

4.3.3.8 Production and purification of LIFR

Ion-exchange chromatography of LIFR-FN3 proved unsuccessful using both mono-Q and mono-S columns in the pH range 4-10 (theoretical pI 7.2, observed pI 5-6). No specific binding of protein was observed and only aggregates were recovered from the column at high salt concentrations.

Figure 4-11 below shows FPLC elution profile of Superose-6 analytical grade gel filtration column with panel (a) showing the elution profile of molecular weight standards with five main peaks running at 2000, 150, 66, 29 and 12 kDa. Panel (b) shows the elution profile of affinity purified LIF in which one main peak is observed at ~91 minutes (~36 kDa) and comparison with lane 6 of panel (e) shows that this peak runs at ~25 kDa on denaturing SDS-PAGE. Inspection of panel (c) shows that LIFR-FN3 is eluted as one main peak at ~67 minutes (~912 kDa) and comparison with lane 3 of panel (e) shows that this species runs as one major band on denaturing SDS-PAGE at a molecular weight of ~116 kDa. Panel (d) shows the putative LIF-LIFR-FN3 complex is eluted as a single species at ~62 minutes (~1622 kDa) and comparison with lane 2 of panel (e) shows that this peak runs as two major bands on denaturing SDS-PAGE at molecular weights of ~25 kDa and ~116 kDa. Panel (e) shows SDS-PAGE of pooled and concentrated fractions corresponding to peaks underlined by numbered red bars. Lane 1 contains void volume eluate in which no protein bands are observed and lane 2 contains the putative LIF-LIFR-FN3 complex peak eluted at ~62 minutes in which two bands at molecular weights of ~25 kDa and ~116 kDa can be seen. Lane 3 contains the LIFR-FN3 peak eluted at ~67 minutes and a single band at ~116 kDa can be observed. Lanes 5, 6 and 7 contain the LIF peak eluted at ~91 minutes and a single band is observed at ~25 kDa. Lanes 8, 9 and 10 contain control samples of putative LIFR-FN3-LIF complex, LIF and LIFR-FN3, respectively. Lane 11 contains broad range molecular weight rainbow markers running at 135, 80, 43, 32 and 14 kDa (Bio-Rad).

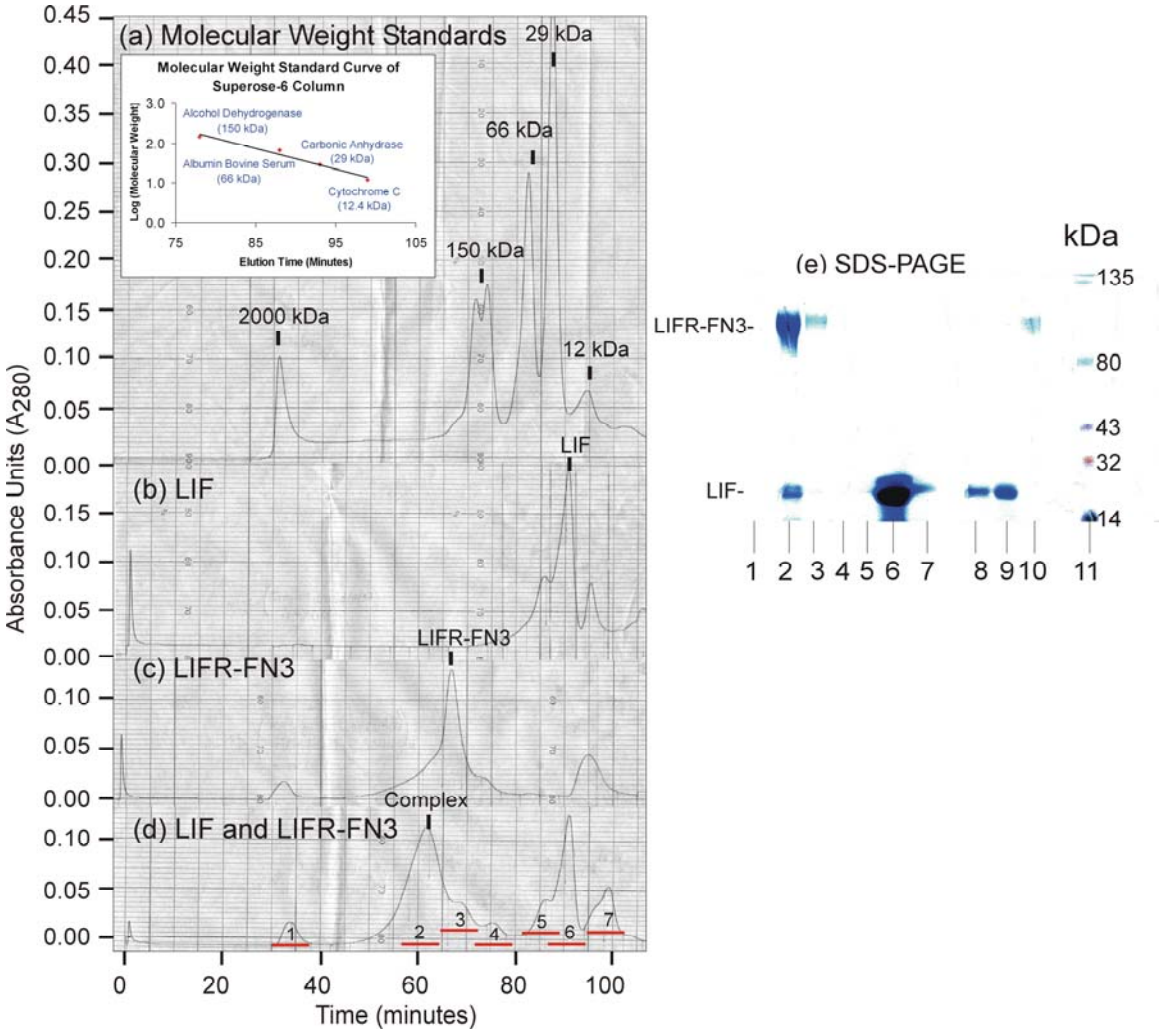


Figure 4-11

Chromatogram showing FPLC elution profile of Superose-6 analytical grade gel filtration column with (a) molecular weight standards, (b) affinity purified LIF, (c) affinity purified deglycosylated LIFR-FN3 and (d) putative LIF-LIFR-FN3 complex. 100 µl samples were loaded at a flow rate of 2 ml/min and the absorbance monitored at 280 nm. Complex was formed by mixing LIF and LIFR-FN3 in molar ratios of 10:1. LIF is eluted as one main peak at ~91 minutes (~35.6 kDa) and LIFR-FN3 is eluted as one main peak at ~67 minutes (>150 kDa). The putative LIF-LIFR-FN3 complex is eluted as a single peak at ~62 minutes (>150 kDa). SDS-PAGE (e) of pooled fractions corresponding to peaks underlined by numbered red bars. Lane 1 contains a sample of void volume eluate and lane 2 contains putative LIF-LIFR-FN3 complex eluted at ~62 minutes. Lane 3 contains LIFR-FN3 eluted at ~67 minutes. Lanes 5, 6 and 7 contain samples of LIF eluted in peaks around 91 minutes. Lanes 8, 9 and 10 contain control samples of putative complex, LIF and LIFR-FN3, respectively. Lane 11 contains broad range molecular weight rainbow markers running at 135, 80, 43, 32 and 14 kDa (Bio-Rad).

4.4 Discussion of Expression, Purification and Characterisation of LIFR

The aim of this project was to expression and purify high levels of soluble LIFR-FN3 protein suitable for crystallographic studies; such requirements are typically 10 mg or greater at a purity of 90 % or better. This was not achieved due to several technical problems although critical steps towards this goal have been established.

Expression in CHO cells resulted in very low protein yields as assessed by ELISA, as shown in Figure 4-3, and scale-up using a stirred-tank bioreactor system and roller bottles proved unsuccessful. Poor cell growth and adherence was observed, compared with wildtype cells, so suggesting that LIFR-FN3-3C-Fc is affecting cellular morphology and phenotype. Typically ~1 mg/L of glycosylated LIFR-FN3-3C-Fc was obtained after 3-4 weeks incubation and as such this system was deemed financially and technically unfeasible for large-scale protein production. Sufficient material was produced using this system for the generation of antibodies and development of preliminary purification protocols. Possible improvements to the pEE14 expression system include the use of alternative promoter-enhancer elements such as the adenovirus major late promoter in combination with the SV40 enhancer (Wong *et al.*, 1985).

Figure 4-3 demonstrates the addition of sodium butyrate to CHO cell growth medium increases the levels of expression of LIFR-FN3 by upto 30%. Sodium butyrate has previously been shown to improve expression from recombinant plasmids in mammalian cells due to increases in both enhancer-dependant transcription and stable transformation efficiencies (Gorman *et al.*, 1983). These effects have been shown to occur in a wide variety of mammalian cells including Friend erythroleukaemic cells (Reeves and Cserjesi, 1979) and CHO cells (D'Anna *et al.*, 1980).

The concentration-dependant effects of sodium butyrate in CHO cells are thought to involve inner histone acetylation and histone H1 phosphorylation (D'Anna *et al.*, 1980). Both of these events function in controlling high order chromatin structure and as such regulate chromatin transcriptional activity. This does not appear to be a protein-specific or cell-specific phenomenon, so the use of sodium butyrate may prove to be important in the development of other mammalian expression systems.

Expression using *Drosophila* S2 cells proved unsuccessful as no induction of LIFR-FN3-3C-Fc expression was detected using ELISA, Protein-A immunoprecipitation or western blot. Samples typically contained a protein with an apparent molecular weight of ~46 kDa as assessed by western blot, compared with control LIFR-FN3-3C-Fc samples running at ~140 kDa. Samples of S2 conditioned medium were shown to be inactive in ligand binding assays and showed no binding to LIF or OSM. The nature of the ~46 kDa product is unknown although it is likely that this protein is a truncated version of LIFR-FN3-3C-Fc, due to its specific reactivity towards LIFR antibodies.

Transient expression in 293T cells proved the most successful system tested for large-scale protein production. Large quantities of pIG-LIFR-FN3-3C-Fc DNA were routinely prepared from *E. coli* strain MC101/P3, with a triple *Bam*HI, *Bgl*II and *Hind*III digest (as shown in Figure 4-4) being essential for confirmation of identity and quality of DNA. Typically, ~5 mg of pIG-LIFR-FN3-3C-Fc DNA was used to transfect ~36 x 162cm^2 flask of 293T cells and protein yields of around 10 mg were achieved after 1 week incubation at 37^{o}C.

SDS-PAGE and western blot analysis as shown in Figure 4-5 and Figure 4-6, respectively, reveal one major species at an apparent molecular weight of ~140 kDa after Protein-A affinity purification of 293T conditioned supernatants; this compares favourably with the

theoretical mass of ~120 kDa. A second species was observed in western blots at ~46 kDa, presumably corresponding to a truncated translation product as observed in S2 cell conditioned supernatants. Cleavage of the fusion protein, *in situ,* with 3C protease yielded soluble LIFR-FN3 running at an apparent molecular weight of ~114 kDa. This compares with the theoretical mass of the LIFR-FN3-3C-Fc moiety of ~94 kDa; the disparity due to retarded passage through polyacrylamide due to glycosylation.

Indeed it has been predicted that LIFR is glycosylated at many of the 20 potential N-glycosylation sites (Gearing *et al.*, 1991). In order to limit microheterogeneity and conformational flexibility associated with glycoproteins, deglycosylation of LIFR-FN3 was attempted. Previous attempts at deglycosylating LIFR-FN3 using β-*N*-acetylglucosidase F (endo F; cleaves glycosidic linkage between the two core N-acetylglucosamine residues) proved unsuccessful (K. Chobotova, CRC Growth Factors Group. Birmingham). Treatment of cells with N-butyldeoxynojirimycin (Karlson *et al.*, 1993); inhibits α-glucosidase I activity thus preventing formation of complex N-linked oligosaccharides so allowing unprocessed sugars to be removed with endo H) resulted in poor levels of cell growth and expression (data not shown).

Figure 4-9 demonstrates that deglycosylation of LIFR-FN3 was successfully achieved using both recombinant (Barsomian *et al.*, 1990) and non-recombinant PNGase F (New England Biolabs). The recombinant PNGase F was shown to be active as both a native maltose-binding fusion protein and in its cleaved form. It was also demonstrated that DTT from the 3C protease cleavage step should also be removed to obtain optimal levels of deglycosylation as PNGase F activity is modulated by DTT. However, western blot analysis as shown in Figure 4-9 reveals that non-recombinant PNGase F is the most effective with comparable enzyme levels to the recombinant form resulting in the largest band shift and a less diffuse band. This suggests more complete deglycosylation and

improved homogeneity is produced by the non-recombinant preparation. Commercially supplied non-recombinant PNGase F was used for the purposes of this study although this supply is not financially feasible for large-scale deglycosylation. An alternative 'home-made' endoglycosidase preparation extracted from *Flavobacterium meningosepticum* containing endo F_1, endo F_2, endo F_3 and PNGase F activity has been demonstrated to show better activity than commercial supplies (Baker *et al.*, 1994). This extract may provide a more cost-effective solution to the problem of deglycosylation.

PNGase F functions in cleaving the N-acetylglucosaminylasparagine bond of asparagine-linked oligosaccharides to yield an aspartic acid residue, ammonia and free reducing oligosaccharide. Such large-scale mutation of asparagine to aspartate may have considerable implications for protein folding, stability, purification and crystallisation. For similar reasons PCR mutagenesis of potentially glycosylated asparagine residues has previously been avoided, as it is suspected that co-translational glycosylation may be required for correct folding and expression. Analysis of the isoelectric focussing data in Figure 4-7 suggests that this is indeed the case with only the fully glycosylated fusion protein running as one fairly tight population with a pI of between 5 and 6. Deglycosylation, cleavage or acid elution of the fusion protein results in disruption of this pI profile and gives rise to a continuous distribution of protein species with pI ranging from 4 to 8. It can be concluded that both sugars and the Fc moeity have a stabilising effect on the LIFR-FN3 protein.

Analysis of deglycosylated LIFR-FN by dynamic light scattering as shown in Figure 4-8 also suggests problems with protein stability and aggregation. These data model as a monodispersed population with a predicted molecular mass of ~84 kDa, which compares favourably with the apparent mass observed on SDS-PAGE of ~90 kDa and the theoretical mass of ~94 kDa. However, large scale aggregates (>1 μm) are observed at

equilibrium and removal of this population results in a shift in the position of equilibrium to re-establish this population over a period of 10 minutes at 20°C. Empirical observations suggest that monodispersed macromolecules are more readily crystallisable, whereas randomly aggregating or polydipersed systems, rarely, if ever, yield crystals (Ferré-D' Amaré and Burley, 1994). Therefore, prevention of aggregation is a major problem and possible remedies include incubation of receptor with excess ligand before both Fc cleavage and deglycosylation. Further to this idea the use of coexpression systems could be investigated for the expression of both LIF and LIFR-FN3-3C-Fc; an example of this is the double cistronic system used for the expression of nuclear receptor partners which resulted in increased solubility and biological activity (Li *et al.*, 1997).

Receptor titration assays as shown in Figure 4-10(a) also demonstrate that purified and deglycosylated LIFR-FN3-3C-Fc is inactive as dose dependant binding of biotinylated LIF is not observed. This again compounds the hypothesis that glycosylation is required for correct folding and activity. Receptor titration assays were also used to show that transfected monolayers of 293T cells could not be used for a second round of expression as replenished medium, although containing equivalent levels of Fc fusion as assessed by ELISA, contained inactive protein. This is presumably due to misfolding and proteolysis associated with cellular death over prolonged periods of growth. Receptor competition assays as shown in Figure 4-10(b) reveal that 293T-expressed LIFR-FN3-3C-Fc demonstrates comparable competition to control samples of LIFR-FN3-3C-Fc, with EC_{50} values of 12.1 and 10.4 ng/ml, respectively.

Purification of deglycosylated LIFR-FN3-3C-Fc using ion exchange resulted in aggregation so gel-filtration was performed as in Figure 4-11. Three single peaks corresponding to LIF (~36 kD), LIFR-FN3-3C-Fc (~912 kDa) and putative complex (~1622 kDa) were observed. The putative complex peak was shown to contain both

receptor and ligand components on denaturing SDS-PAGE; although limited information concerning the state or stoichiometry of the complex can be drawn due to the non-globular arrangement of the assemblies involved and their large Stokes' radii. However, a peak shift was observed on inclusion of ligand so suggesting some form of receptor-ligand interaction. Previous studies suggest that LIF forms a high affinity heterotrimeric complex with gp130 and LIFR (Zhang *et al.*, 1997b) so implying that all components of the assembly should be included.

In summary it can be seen that the protocols developed allow the expression, deglycosylation and preliminary purification of LIFR-FN3 for use in crystallographic studies. However, further developments will be required to gain greater quantities of protein and problems with removal of the fusion moiety and oligosaccharides still remain. Other expression systems available for investigation include glycosylation mutant yeast of the strain *Pichia pastoris* and the use of baculovirus expression systems (Lucklow and Summers, 1988).

Alternative methods of obtaining glycoproteins with minimal heterogeneity include the use of mutant CHO cells with multiple glycosylation defects (Stanley, 1987). One such group of mutant CHO cells include the lectin resistant strains (Lec^R), which grow readily in culture despite the drastic tailoring of their surface carbohydrates (Stanley and Carver, 1977). Also more rigorous analysis of the carbohydrate levels will need to be undertaken using stains such as thymol-sulphuric acid or Fuschin basic.

The methods discussed above are the first steps towards structural analysis of the LIFR-LIF-gp130 interaction. Such information will provide further insights into the molecular basis of signal-transduction by the haemopoetin cytokines and the functional epitopes involved.

5. Chapter 5 - Structure Determination of Oncostatin M

The aims of this chapter include the growth of OSM crystals and determination of the X-ray structure. Crystal growth, data collection, data processing and analysis of the quality of the structure are reported. Succesful structure determination using MAD techiniques are reported and initial attempts at molecular replacement are not discussed. A detailed description of the OSM structure and implications for receptor binding is presented in Chapter 6.

The theory and practice of structure determination using X-ray crystallography is covered in detail in texts such as (Drenth, 1994), (Rhodes, 1993) and (Blundell and Johnson, 1976). An overview of the processes involved is shown in Figure 5-1 below.

5.1 Outline of Methods Employed

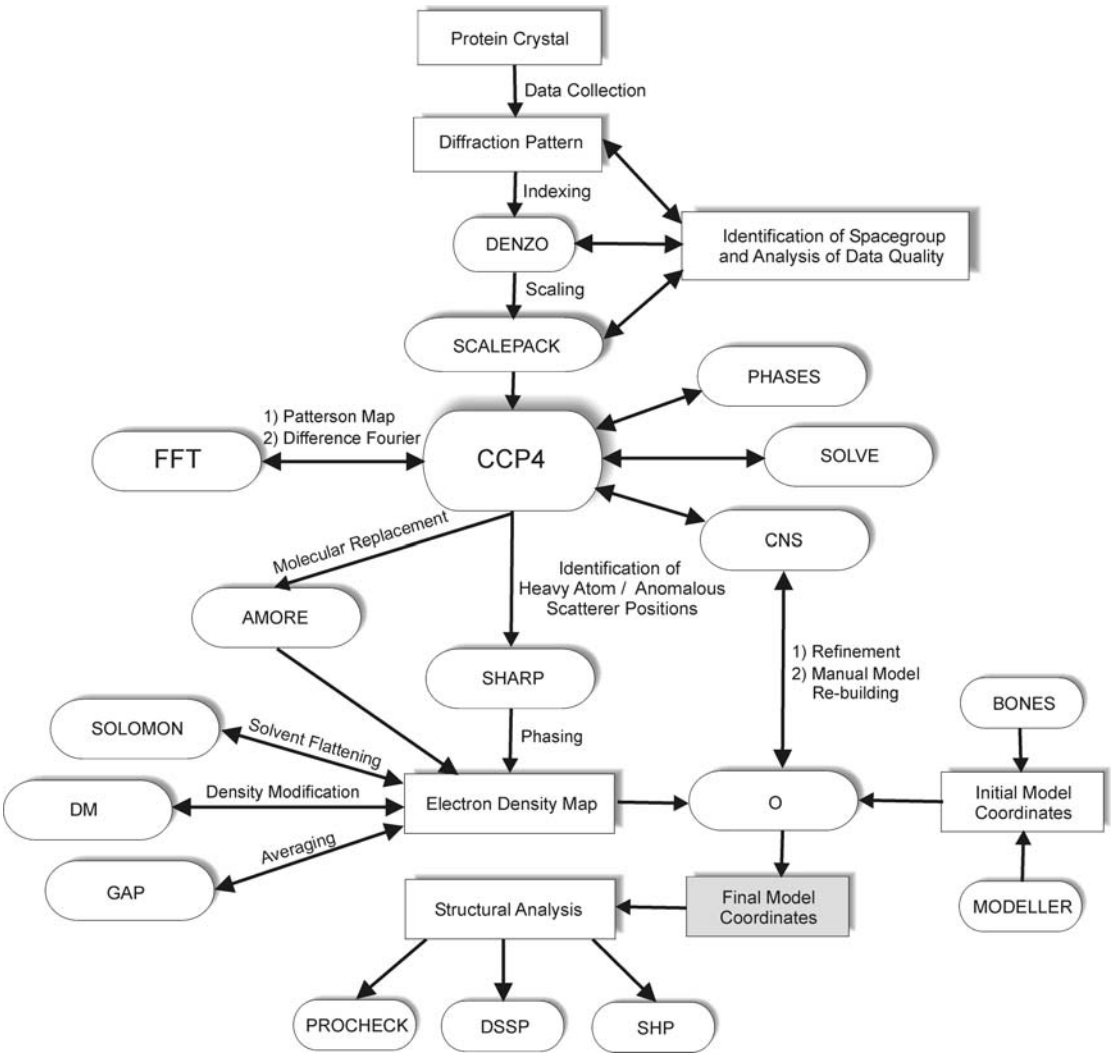


Figure 5-1
Flowchart detailing methods and techniques employed in the course of chapter 5. Programs used represented by rounded boxes.

5.2 Methods of Structure Determination

5.2.1 Crystallisation

Samples of OSM-196, Δ188_OSM and Se-Met-OSM were purified on a Sephadex-75 column in 150 mM sodium chloride, 50 mM Tris-HCl, pH 7.0 (see Chapter 3). The protein was concentrated to between 5 and 14 mg/ml in Centricon 3 kDa cut off filters (Amicon). Protein concentrations were estimated by UV spectrometry using the equation shown in Appendix IV.

Crystallisation conditions were initially screened using the hanging drop vapour diffusion method. Screening of precipitant solutions was carried out using sparse matrix sampling screens (Jancarik and Kim, 1991), ammonium sulphate screens and polyethylene glycol (PEG) screens at 22^{o}C and 4^{o}C. Typically, 0.5-2 μl of protein solution was pipeted onto a siliconised 22 mm glass coverslip (BDH) so forming a spherical drop. An equal volume of precipitant solution was pipeted onto the protein and the drop then placed over a well of a linbro plate (ICN) containing 0.5-1 ml of corresponding precipitant. The well was sealed using high-pressure vacuum grease (Dow Corning) and the plate left to equilibrate. Drops were monitored microscopically on a daily basis for the first week, followed by weekly inspections thereafter.

Conditions producing crystals or fine structured precipitates were investigated further. Finer grid screens were setup around these conditions in an attempt to improve crystal size. Macro- and micro-seeding (Stura and Wilson, 1992) were also carried out. Sitting-drop crystallisation using micro-bridges (Harlos, 1992) was investigated as an alternative to hanging drop conditions in an attempt to manipulate equilibration rates. Additive screens, different temperatures, fine pH sampling (McPherson, 1995) and crystallisation using an oil barrier (Chayen, 1997) were also tested. Crystallisation trials of Se-Met derivatised OSM

samples were also carried out. Samples were treated as above with the exception that duplicate trays were setup using protein samples with and without DTT to 1 mM, in an attempt to alleviate oxidation.

5.2.2 Crystal Mounting and Cryocrystallography

Solutions of methyl pentanediol (MPD), glycerol, dioxane and PEG 8000 were tested as cryoprotectants (Petsko, 1975). Mother liquor solutions were prepared containing 5, 10, 15, 20 and 25 % of potential cryoprotectant. Crystals were harvested using a fibre loop (Teng, 1990) and transferred to 2 µl drops of cryoprotectant solution for ~5 minutes and then mounted in a nitrogen stream (Oxford Cryosystems) at 100 K. Data were collected as in 5.2.3 below. Diffraction patterns were analysed for characteristic ice-rings at ~4 Å and the quality of crystalline disorder were also assessed.

Alternatively, direct flash cooling of crystals in liquid propane and liquid nitrogen was tested. Crystals were transferred from cryoprotectant solution, as above, and plunged directly into aliquots of liquid propane or liquid nitrogen. Crystals were transferred swiftly and directly into a nitrogen stream (Oxford Cryosystems) and data collected as in 5.2.3 below.

Crystals were also mounted for room temperature data collection in 0.7 mm quartz capillaries (Glas, Germany). Excess mother liquor was removed to prevent movement of crystals. Capillaries were sealed using Bees wax and then mounted on a goniometer head for data collection.

5.2.3 Data Collection

Initial native data were collected from a single capillary mounted crystal at room temperature on beam line BL6A2 (Sakabe, 1991) of the Photon Factory, Tsukuba, Japan. Data to 2.5 Å were collected using 1.0 Å wavelength radiation with 0.1 mm collimation.

The data collection strategy employed the Weissenberg method and involved 3.5° oscillations with a coupling constant of 1.5 °/mm. Diffraction data were recorded on BASIII imaging plates (200 x 400 mm) and scanned off line with a Fuji BA100 IP scanner. Exposure times were typically 240 seconds and the crystal to detector distance was set to 429.7 mm over a helium filled path. Data were collected at a temperature of 14°C. Typical beam currents were around 300 mA.

Native and multiple wavelength anomalous dispersion (MAD) data were collected from single propane frozen crystals on beam line BL19 of the European Synchrotron Radiation Facility (ESRF), Grenoble, France. MAD data were collected at wavelengths of 1.00914 Å, 0.826516 Å and 1.00669 Å, to a resolution of 2.6 Å. Data were collected in a nitrogen stream at 100 K and data recorded as 1° oscillations using a MAR 345 image plate detector (MAR Research, Hamburg, Germany) in 18 cm mode. Exposure times were typically 40 seconds with the crystal to detector distance set to 225 mm. Native data were collected at the same beamline at a wavelength of 0.98 Å to a resolution 2.2 Å. Data were collected in a nitrogen stream at 100 K and data recorded as 1° oscillations using a MAR 345 image plate detector (MAR Research, Hamburg, Germany) in 30 cm mode. Exposure times were typically 60 seconds and the crystal to detector distance was set to 225 mm.

Heavy atom derivative screening was carried out using 'in house' radiation and synchrotron radiation at station 9.6 of the Synchrotron Radiation Source (SRS), Daresbury, U.K. 'In house' data were collected using CuKα X-rays (1.54 Å) produced by a rotating anode (Rigaku RU2000) fitted with a graphite monochromator. Data were collected at 100 K or at room temperature. Data were recorded as 1° oscillations using a 18 cm MAR image plate detector (MAR Research, Hamburg, Germany). Exposure times were typically 2000 seconds and the crystal to detector distance was set to 200 mm. Data were collected at station 9.6 of the SRS using a wavelength of 0.87 Å. Data were collected at room

temperature as 2° oscillations using a charged coupled device (CCD) detector (Advanced Detector Systems Corporation, USA). Exposure times were typically 40 seconds with the crystal to detector distance set to 300 mm.

5.2.4 Data Processing

Data were autoindexed, integrated and corrected for Lorentz and polarizing effects with the program DENZO (Otwinowski and Minor, 1997). Scaling, merging and addition of partially recorded reflections was carried out using the program SCALEPACK (Otwinowski and Minor, 1997). General processing and file manipulation was carried out using the CCP4 suite of programs (Collaborative Computational Project, 1994). SCALEPACK files were converted to CCP4 format using the program SCALEPACK2MTZ and intensities were converted to amplitudes using the program TRUNCATE (French and Wilson, 1978).

5.2.5 Heavy Metal Derivatisation

Heavy metal derivative screening was carried out as detailed in Table 5-1 below. Heavy metal solutions were prepared at twice the required concentration in mother liquor consisting of 30 % (w/v) PEG 35000, 0.25 M sodium acetate, pH 7.5. Solutions were left overnight, with agitation, to equilibrate. Crystals were transferred to fresh 2 µl drops of mother liquor and 2 µl of heavy metal solution was added. Drops were sealed over 0.5 ml reservoirs of mother liquor and the crystals incubated at room temperature for between 1 and 72 hours.

Heavy Metal Reagent	Concentration	Duration of Soak
Trimethyllead Acetate (TA)	20-50 mM	16-72 hours
Lead Acetate (LA)	5 mM	1-24 hours
Potassium Mercuric Iodide (PMI)	2 mM	16-24 hours
Ethylmercury Phosphate (EMP)	1-40 mM	16-24 hours
Uranyl Acetate (UA)	1-5 mM	1-72 hours
Uranyl Nitrate (UN)	1-5 mM	1-72 hours
Cis-dichlorodiamine Platinum II (CP)	5 mM	16-72 hours
Mercury meso-(2,3-dimethoxytetramethylene)acetate (BAKERS)	5 mM	16-72 hours
Tetrakis(acetomercuric)methane (TAMM)	5 mM	16-72 hours
Sodium Gold Chloride (SGC)	5 mM	1-72 hours
Di-potassium Platinum Nitrate (PN)	5 mM	1-72 hours

Table 5-1

Table showing summary of heavy metal reagents used for derivative screening of Δ188_OSM. All soaks were carried out for between 1 and 72 hours in equilibrated mother liquor containing 30 % (w/v) PEG 35000, 0.25 M sodium acetate, pH 7.5 and corresponding heavy metal.

The state of derivatisation was analysed by collecting ~45^{o} of data (as 1^{o} or 2^{o} oscillations) and manual inspection of anomalous difference Patterson and Fourier maps, calculated using the program FFT (Ten Eyck, 1973). Maps were contoured using the program GROPAT (Esnouf, 1993).

5.2.6 Multiple Wavelength Anomalous Dispersion

MAD data were collected to a resolution of 2.6 Å from a single EMP derivatised propane cooled crystal on beam line BL19 at the ESRF, Grenoble, France (see 5.2.3). The mercury LIII edge of the derivative crystal was characterised by a fluorescence scan and three wavelengths were selected for data collection. The first wavelength was selected just below the LIII edge at a wavelength of 1.00669 Å (f '=-14.8 e^-, f "=10.1 e^-, E=12.3163 keV). The second wavelength was selected at the inflexion point at a wavelength of 1.00914 Å (f '=-20.3 e^-, f "=10.2 e^-, E=12.2863 keV). A remote wavelength was also selected above the absorption edge at 0.826516 Å (f '=-7.4 e^-, f "=11.8 e^-, E=15.0009 keV). Data were collected as 1^o oscillations over the same 90^o phi range for each wavelength.

All data was autoindexed and processed as previously (see 5.2.4) except scaling and merging of MAD data at each wavelength was carried out separately, with preservation of Bijvoet pairs, using the program SCALEPACK (Otwinowski and Minor, 1997). Structure factor amplitudes were calculated using the program TRUNCATE (French and Wilson, 1978) and the MAD data were scaled, with reference to the inflexion data, using the program SCALEIT (Howell and Smith, 1992). Finally, normalised anomalous scattering magnitudes were estimated, to aid in analysis of anomalous scatterer positions, using the program REVISE (Woolfson *et al.*, 1997).

Mercury positions were determined by manual inspection of anomalous difference Patterson maps, calculated using the program FFT (Ten Eyck, 1973). Confirmation of the mercury positions was carried out using the programs RSPS and SOLVE (Terwilliger and Berendzen, 1999). The positions of the mercury sites were refined using the program SHARP (de La Fortelle and Bricogne, 1997).

5.2.7 Molecular Replacement

The refined MAD determined structure was used as a search model for molecular replacement of subsequently collected native data using the program AMoRe (Navaza, 1994). Residues 3-134 and 156-187 of the model were used in the search with water molecules removed. Reflections were first sorted using the SORTING routine and the molecule translated so that the centre of gravity is at the origin, using the TABLING routine. The search model was placed in the symmetry free spacegroup P1 to minimize cross vectors and the cross-rotation function was carried out in the resolution range 12-3.5 Å, using the routine ROTING. Having determined the orientation of the molecules in the asymmetric unit the translation parameter was calculated to orientate the molecules relative to the symmetry axis. This was carried out using the routine TRAING in the resolution range 12-3.5 Å. Rigid body refinement of the solution was carried out using the routine FITTING and the resulting orientation matrix used to transform the output of the TABLING routine using the program PDBSET.

5.2.8 Map Calculation and Model Refinement

MAD phases were calculated to 2.6 Å resolution using the program SHARP (de La Fortelle and Bricogne, 1997). Phase improvement was carried out by solvent flattening, using cyclical phase extension for 160 cycles starting from 4.0 Å to a final resolution of 2.6 Å using the program SOLOMON (Abrahams and Leslie, 1996). Structure factors and phases were calculated using the program SFALL (Agarwal, 1978) and phase combination carried out using SIGMAA (Read, 1986). Initial $2|F_o|-|F_c|$ maps were calculated using data in the range 20.0-2.6 Å using the program FFT (Ten Eyck, 1973). All subsequent refinement and map calculations were carried out using the program CNS-SOLVE (Brunger *et al.*, 1998).

An homology model of Δ188_OSM (see Chapter 2) containing residues 1-188 of mature OSM was used as an initial model. Rigid body refinement of the model was carried out using native data between 20.0-2.6 Å. Helices were treated as separate rigid bodies and 40 cycles of maximum likelihood rigid body refinement were carried out. Low weighting of the X-ray terms (W_a=2.4) was initially used, to maintain close to ideal stereochemistry, and 10 % of the data was flagged for use as a free-R (Brunger, 1992) test set. Positional refinement using maximum likelihood was carried out for 200 cycles using data in the range 20.0-2.6 Å followed by simulated annealing for 1000 cycles, using torsion dynamics with slow cooling from 4000 K.

Manual rebuilding was carried out using the interactive graphics program O (Jones *et al.*, 1991) with reference to the original solvent flattened $2|F_o|-|F_c|$electron density maps. Refined coordinates were output and $2|F_o|-|F_c|\alpha_{calc}$ maps and $|F_o|-|F_c|$ difference maps calculated using the program CNS-SOLVE (Brunger *et al.*, 1998) using data in the range 20.0-2.6 Å. Further rebuilding was carried out with reference to the $2|F_o|-|F_c|\alpha_{calc}$ maps and $|F_o|-|F_c|$ difference maps. For regions of ambiguous density composite omit maps were calculated using 5 % omissions of data. After five rounds of such rebuilding and refinement, bulk solvent correction was applied with a solvent density of 0.29 $e^-/Å^3$ and a B-factor of 31.1 $Å^2$. Anisotropic B-factor correction was also applied on data in the range 6.0-2.2 Å. In subsequent rounds of refinement all data in the resolution range 20.0-2.2 Å were included. A further ten rounds of refinement and manual rebuilding was carried out with the inclusion of grouped B-factor refinement and subsequent individual B-factor refinement.

In the final stages of refinement bound water molecules were built into the model using an automated peak search utility in the program CNS-SOLVE (Brunger *et al.*, 1998). Waters were built into peaks above 1.5 σ at a distance of 2.6-3.2 Å from the nearest oxygen or

nitrogen atoms. A second round of water addition and refinement was carried out at a peak search level of 1.0 σ. Further waters were added manually through inspection of $2|F_o|-|F_c|\alpha_{calc}$ maps and $|F_o|-|F_c|$ difference maps. In a final round of refinement the X-ray weighting term was increased (W_a=10) so relaxing the restraint on stereochemical terms.

5.2.9 Analysis of Model

The stereochemistry of the final model was assessed using the program PROCHECK (Laskowski *et al.*, 1993). Secondary structure assignments and solvent accessibilities were calculated using the programs DSSP (Kabsch and Sander, 1983), HERA (Hutchinson and Thornton, 1990) and NACCESS (Hubbard and Thornton, 1993). Structural superpositions were performed using the program SHP (Stuart *et al.*, 1979). Molecules were drawn and rendered using the programs BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

5.3 Results of Structure Determination

5.3.1 Crystallisation

Crystal Screen I and Crystal Screen II (Hampton Research) produced no crystals in all samples tested at 22°C or 4°C. Δ188_OSM samples displayed high solubility in all screens with Se-Met-OSM and OSM-169 displaying heavy precipitation in ~70 % of all conditions.

Initial crystals were obtained within 1-3 days at 22°C using the hanging drop vapour diffusion method with a custom salt/PEG screen. Crystals formed in drops consisting of 2 µl of protein solution at ~8 mg/ml plus 2 µl of reservoir solution consisting of 30 % (w/v) PEG 35000 and 0.25 M sodium acetate, pH 7.5. Crystal size was improved by decreasing the ratio of precipitant to protein. The largest crystals measuring approximately 0.2 x 0.2 x 0.1 mm^3 were obtained from drops consisting of 0.5 µl of precipitant and 1 µl of protein at 8 mg/ml (see Figure 5-2 (a)). The crystals showed ordered Bragg diffraction to 2.2 Å when exposed to synchrotron radiation (see Figure 5-2 (c)). Preliminary processing showed that the crystals belong to the orthorhombic spacegroup $P2_12_12_1$.

A second crystal form was obtained over a period of several months from 30 % (w/v) PEG 4000, 0.1 M Tris-HCl and 0.2 M sodium acetate, pH 8.5 using the hanging drop vapour diffusion method. Crystals measuring approximately 0.1 x 0.1 x 0.1 mm^3 were obtained from drops consisting of 1 µl of precipitant and 1 µl of protein solution at 14 mg/ml (see Figure 5-2 (b)). The crystals displayed ordered Bragg diffraction to 2.5 Å when exposed to synchrotron radiation. Preliminary processing showed the crystals belong to the spacegroup I222 or $I2_12_12_1$.

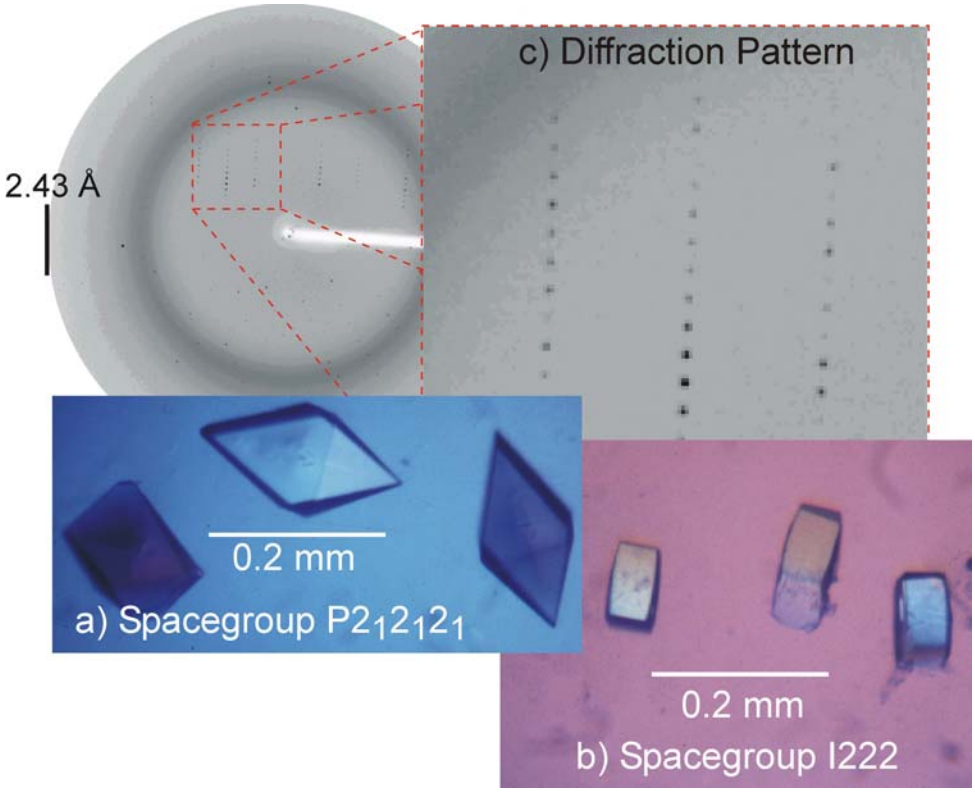


Figure 5-2

Figure showing (a) $P2_12_12_1$ and (b) I222 crystal forms of Δ188_OSM with bar representing 0.2 mm. Panel (c) shows resulting diffraction pattern of the $P2_12_12_1$ crystal form with bar showing high resolution limit of detector at 2.43 Å. Data to 2.6 Å were collected from single propane cooled crystals on beam line BL19 of the ESRF, Grenoble, France. The I222 crystal form was grown in 30 % (w/v) PEG 4000, 0.1 M Tris-HCl and 0.20 M sodium acetate, pH 8.5. The $P2_12_12_1$ form was grown in 30 % (w/v) PEG 35000 and 0.25 M sodium acetate, pH 7.5. Crystals were grown from protein solution at 8 mg/ml to 14 mg/ml, using hanging drop vapour diffusion.

5.3.2 Crystal Mounting and Cryocrystallography

Optimum flash cooling of the $P2_12_12_1$ crystal form was carried out as follows; crystals were removed from the mother liquor and briefly passed through a cryoprotectant solution consisting of a 2µl drop of 30 % (w/v) PEG 35000 and 0.25 M sodium acetate, pH 7.5 overlaid with 2µl of 30 % (w/v) PEG 8000 and 0.25 M sodium acetate, pH 7.5. Crystals were cooled directly in liquid propane and placed in a nitrogen stream at 100 K and data recorded.

5.3.3 Data Collection and Processing

The two crystal forms of Δ188_OSM showed ordered Bragg diffraction when exposed to synchrotron radiation. Statistics relevant to data collection are shown below in Table 5-2. Initial processing with the program DENZO (Otwinowski and Minor, 1997) suggested the crystals belong to the primitive orthorhombic pointgroup 222. Scaling and addition of partial reflections was carried out using the program SCALEPACK (Otwinowski and Minor, 1997). The pattern of systematically absent odd reflections on indices *h*00, 0*k*0 and 00*l*, suggested that the spacegroup has three 2-fold screw axes, so data were finally merged in the spacegroup $P2_12_12_1$. Low or negative $I/\sigma(I)$ values for all systematic absences, in the resultant SCALEPACK output, confirmed $P2_12_12_1$ as the correct spacegroup assignment. The crystals have unit cell dimensions of a=35.8 Å, b=53.1 Å, c=106.8 Å and $\alpha=\beta=\gamma=90^o$ (see Table 5-2). The volume of the unit cell is approximately 203025 Å^2 and the theoretical molecular weight of the protein is 21.4 kDa. Assuming a standard specific volume of 1.23 cm^3/g this suggests one molecule per asymmetric unit and a crystal solvent content of ~48.8 % (Matthews coefficient, V_m=2.4 Å^3/Da). From Table 5-2 it can be seen that native data collected from these crystals is of good quality with an $I/\sigma(I)$ value in the outer resolution shell (2.25-2.20 Å) of 4.3. Completeness in the outer resolution shell is 98.2 %.

The second crystal form was processed in the body centred orthorhombic spacegroup I222 with unit cell dimensions *a*=52.9 Å, *b*=62.0 Å, *c*=109.4 Å and α=β=γ=90°(see Table 5-2). Spacegroups I222 and $I2_12_12_1$ are indistinguishable by diffraction intensities and correct assignment as I222 was only confirmed during molecular replacement (see 5.3.6). The volume of the unit cell is approximately 358810 Å^2 and the theoretical molecular weight of the protein is 21.4 kDa. Assuming a standard specific volume of 1.23 cm^3/g this suggests one molecule per asymmetric unit and a crystal solvent content of ~41.4 % (Matthews coefficient, V_m=2.1 Å^3/Da). From Table 5-2 it can be seen that native data collected from these crystals is of poorer quality than the $P2_12_12_1$ crystal form with an I/σ(I) value in the outer resolution shell (2.56-2.50 Å) of 2.4. Completeness in the outer resolution shell is 88.8 %.

	Native	Native	EMPλ_{peak}	EMP$\lambda_{inflexion}$	EMPλ_{remote}
Data Collection Site	Photon Factory	ESRF (BL14)	ESRF (BL14)		
Data Collected	100°	100°	100°		
Temperature	295 K	100 K	100 K		
Wavelength	1.000000 Å	0.979273 Å	1.00914 Å	0.826516 Å	1.00669 Å
Spacegroup	I222	$P2_12_12_1$	$P2_12_12_1$		
Unit Cell Dimensions	a=52.9 Å b=62.0 Å c=109.4 Å α=β=γ=90°	a=35.8 Å b=53.1 Å c=106.8 Å α=β=γ=90°	a=35.9 Å b=53.3 Å c=106.7 Å α=β=γ=90°		
Resolution Range	20-2.5 Å	25-2.2 Å	25-2.6 Å		
Total Number of Reflections	15637	43901	24040	26223	24159
Number of Unique Reflections	5544	10513	11650	12069	11691
Redundancy	2.8	4.2	2.1	2.2	2.1
Completeness	85.2%	96.5% (98.2%)	96.4% (63.1%)	99.9% (100%)	96.8% (66.7%)
$<I/\sigma(I)>$	9.2 (2.4)	13.2 (4.3)	12.1 (2.9)	12.5 (4.7)	12.3 (3.3)
R_{merge}	9.9%	8.3% (37.6%)	4.6% (18.7%)	4.6% (16.4%)	4.3% (17.3%)

Table 5-2

Table showing data collection and processing statistics for Δ188_OSM. EMPλ denotes MAD data collected from ethylmercuryphosphate (EMP) derivatised Δ188_OSM, with EMP$\lambda_{inflexion}$, EMPλ_{remote} and EMPλ_{peak} representing inflexion, remote and peak data, respectively. $R_{merge}=\Sigma|I-<I>|/\Sigma<I>$. Redundancy=Total Number of Reflections/Number of Unique Reflections. Values in parentheses correspond to the highest resolution shell (2.66–2.60 Å for EMPλ, 2.25–2.20 Å for native $P2_12_12_1$ and 2.56–2.50 Å for native I222).

5.3.4 Heavy Metal Derivatisation

Initial heavy atom derivative trials failed due to the high reactivity of mercury and uranyl compounds towards the OSM crystals (see Table 5-3). Even at low concentration UA, UN, BAKERS, TAMM and PMI soaks destroyed crystal instantaneously. Platinum salts presumably disrupted crystalline disorder with PN and CP soaked crystals showing no diffraction. Lead salts such as TA and LA showed no reactivity towards OSM crystals,

even at higher concentrations, as no derivatisation was observed. One mercury derivative was identified by soaking crystals in equilibrated drops of mother liquor containing 14 mM ethylmercuryphosphate (EMP) for 16-24 hours.

Heavy Metal Reagent	Target Residues	Comments
Potassium Mercuric Iodide (PMI)	Gly/Asn/Met/His	Crystal Destroyed
Mercury meso-(2,3-dimethoxytetramethylene)acetate (BAKERS)	Met/His/Cys	Crystal Destroyed
Tetrakis(acetomercuric)methane (TAMM)	His/Cys	Crystal Destroyed
Uranyl Acetate (UA)	Glu/Asp	Crystal Destroyed
Uranyl Nitrate (UN)	Glu/Asp	Crystal Destroyed
Sodium Gold Chloride (SGC)	Cys/His	Crystal Destroyed
Cis-dichlorodiamine Platinum (II) (CP)	Met/His/Cys	No Diffraction
Di-potassium Platinum Nitrate (PN)	Met/His/Cys	No Diffraction
Trimethyllead Acetate (TA)	Glu/Asp	No Binding
Lead Acetate (LA)	Glu/Asp	No Binding
Ethylmercury Phosphate (EMP)	Met/His/Cys	Peaks in Patterson

Table 5-3

Table showing summary of heavy metal reagents used for derivative screening of Δ188_OSM. All soaks were carried out for between 16 and 72 hours in equilibrated mother liquor containing 30 % (w/v) PEG 35000, 0.25 M sodium acetate, pH 7.5 and corresponding heavy metal reagent. Target residues represent sites of attachment previously reported for each heavy metal reagent (Blundell and Johnson, 1976).

An EMP derivative was originally identified using SRS data and a full MAD experiment was subsequently carried out at the ESFR. Derivatisation was confirmed by manual inspection of anomalous difference Patterson maps and anomalous difference Fourier maps as shown in Figure 5-3 and Figure 5-4 below. Inspection of Harker sections of the anomalous difference Patterson map reveals two major sites for the EMP derivative at refined fractional coordinates of 0.22, 0.14, 0.20 (site 1) and 0.94, 0.00, 0.13 (site 2). Inspection of anomalous difference Fourier maps, calculated using phases generated by the two major sites, revealed a third minor site at refined fractional coordinates of 0.32, 0.33,

0.24 (site 3). The same two major sites were also identified using an automated procedure within the program SOLVE (Terwilliger and Berendzen, 1999). All three sites showed good refinement using the program SHARP (de La Fortelle and Bricogne, 1997) with occupancies of 23 %, 19 % and 16 % for site 1, site 2 and site 3, respectively. Analysis of the final refined model showed the heavy atom binding sites correspond to residues His 48, His 171 and His 178 which are all solvent exposed (see Figure 5-5). Similar analysis of the final refined model reveals that the four methionine residues in OSM are buried and as such may affect folding of derivatised material (see Figure 5-5).

5.3.5 Multiple Wavelength Anomalous Dispersion

The mercury LIII absorption edge (observed in ideal systems at λ=1.00910 Å and 12.2858 keV) was characterised using a fluorescence scan as shown below in Figure 5-6. This shows a characteristically broad f ” peak often associated with mercury derivatised proteins.

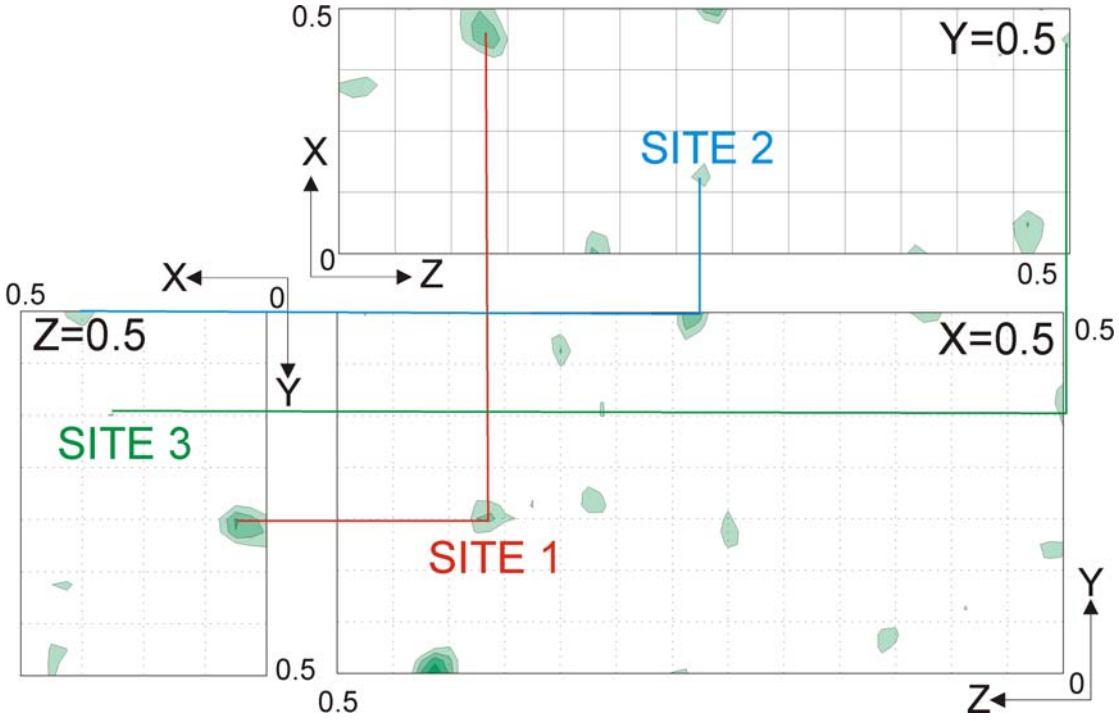


Figure 5-3

Figure showing Harker sections of anomalous difference Patterson map of EMP derivatised Δ188_OSM produced using the program FFT (Ten Eyck, 1973). Sections at X=0.5, Y=0.5 and Z=0.5 are shown contoured from 2 σ in 1 σ steps using the program GROPAT (Esnouf, 1993). Data in the resolution range 10-3 Å were used. The two major sites are denoted as site 1 and site 2 at refined fractional coordinates of 0.22, 0.14, 0.20 and 0.94, 0.00, 0.13, respectively. A third minor site identified through inspection of anomalous difference Fouriers is denoted as site 3 at refined fractional coordinates of 0.32, 0.33, 0.24. Corresponding peaks in each section are shown with connecting lines.

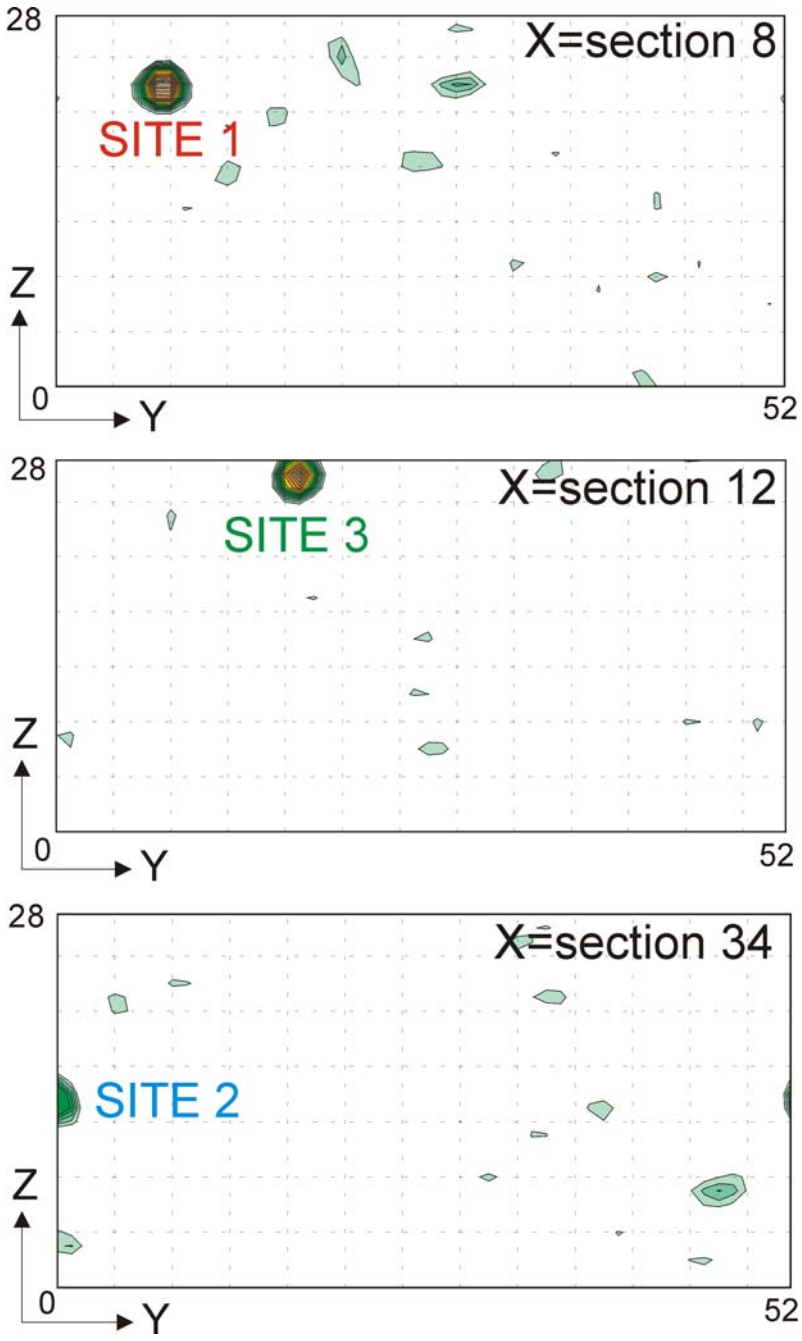


Figure 5-4

Figure showing anomalous difference Fourier map of EMP derivatised Δ188_OSM. Map produced using phases calculated from site 1 and site 2 with data in the range 10-3 Å using the program FFT (Ten Eyck, 1973). Sections in X=1, Y=0.5 and Z=0.5 are shown contoured from 2 σ in 1 σ steps using the program GROPAT (Esnouf, 1993). The two major sites are denoted as site 1 and site 2 at refined orthogonal coordinates of 7.86 Å, 7.43 Å, 21.37 Å and 33.60 Å, 0.00 Å, 13.89 Å, respectively. A third minor site is denoted as site 3 at refined orthogonal coordinates of 11.44 Å, 17.51 Å, 25.65 Å.

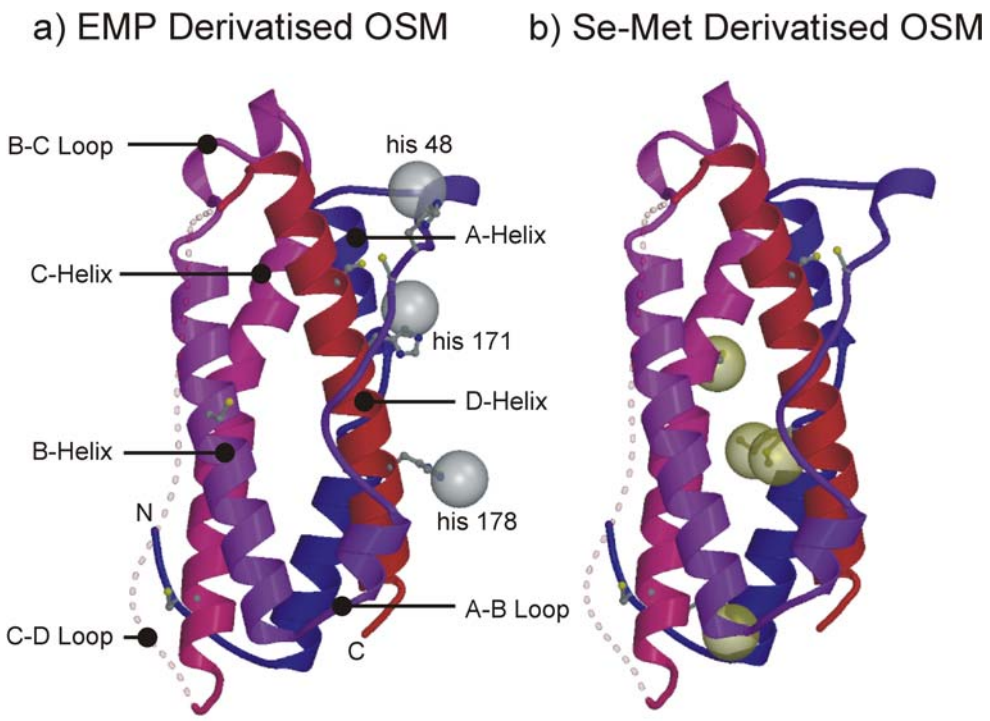


Figure 5-5

Figure showing final refined model of Δ188_OSM with bound (a) EMP represented by silver transparent spheres and (b) hypothetical Se-Met substitutions represented by gold transparent spheres. Transparent dotted section represents loop region as observed in LIF. Heavy atom positions were found by manual inspection of anomalous difference Patterson maps and anomalous difference Fourier maps. Refinement of heavy atom positions was carried out using SHARP (de La Fortelle and Bricogne, 1997). Residues His 48, His 171 and His 178 responsible for heavy metal binding are shown as ball and stick objects. Secondary structure is coloured blue at the N-terminus through to red at the C-terminus. Figure produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

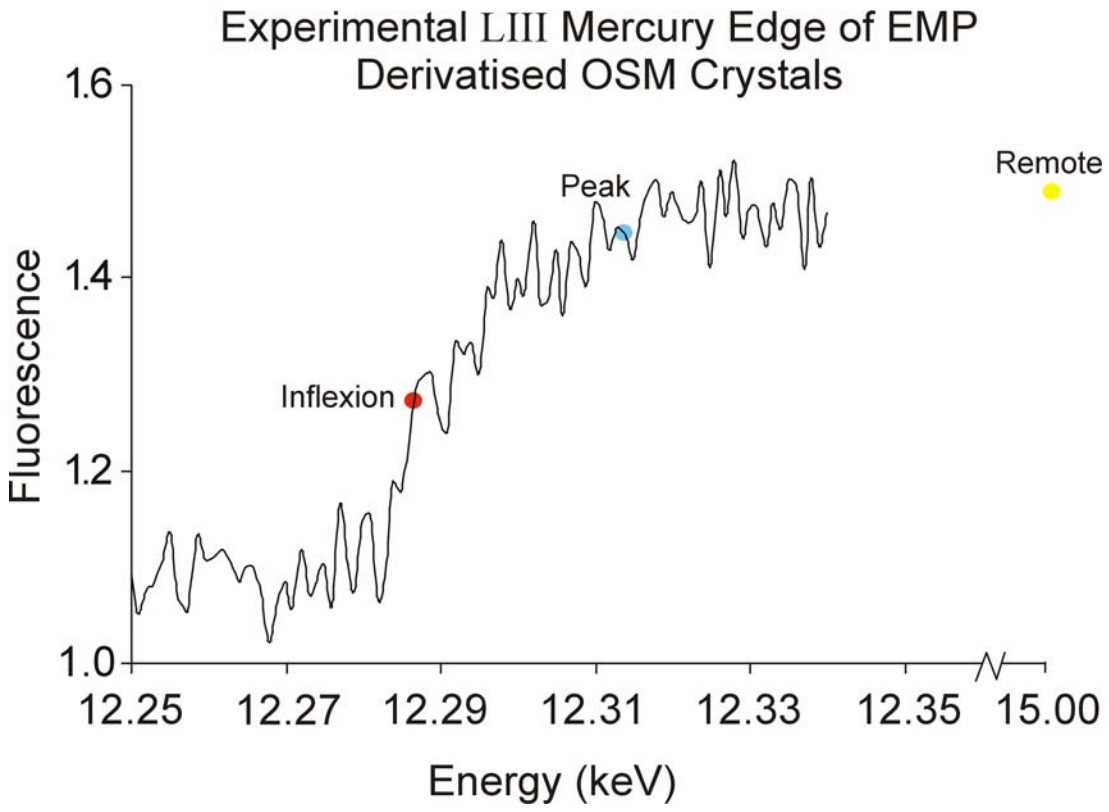


Figure 5-6

Figure showing fluorescence scan of an EMP derivatised Δ188_OSM crystal. Energy region corresponds to mercury LIII absorption edge. Data were collected from a single EMP derivative propane cooled $P2_12_12_1$ crystal on beam line BL19 at the ESRF, Grenoble, France. Three wavelengths selected for data collection are shown as peak (λ=1.00669 Å, f '=-14.8 e^-, f "=10.1 e^-, E=12.3163 keV), inflexion (λ=1.00914 Å minimum f '=-20.3 e^-, f "=10.2 e^-, E=12.2863 keV) and remote (λ=0.826516 Å f '=-7.4 e^-, f "=11.8 e^-, E=15.0009 keV). All values are experimental values prior to refinement.

From Table 5-2 it can be seen that MAD data collected from EMP derivatised crystals is of good quality with an $I/\sigma(I)$ value in the outer resolution shell (2.66-2.60 Å) greater than 2.9 for all wavelengths. Completeness in the outer resolution shell is 63.1 % or better.

Initial maps based on MAD phases and F_{obs} from the remote data set (figure of merit=0.46, phasing power=1.84) proved untraceable. Various phasing packages were tried including SOLVE (Terwilliger and Berendzen, 1999), PHASES (Furey and Swaminathan, 1996) and SHARP (de La Fortelle and Bricogne, 1997). Hendrickson-Lattman coefficients from an independent native-remote single isomorphous replacement with anomalous scattering (SIRAS) experiment were combined with the MAD phases using SHARP. Maps based on the combined phases and F_{obs} from the remote data set (figure of merit=0.51, phasing power=1.90) showed improved main chain density. A summary of phasing statistics and the programs used is shown below in Table 5-4.

Phasing Package	Experiment	Dispersive[1]/Bijvoet[2] Difference $<\|\Delta F\|>/<FP>$	Figure of Merit $\|F(hkl)best\|/\|F(hkl)\|$	Phasing Power $\Sigma n\|FH\|/\Sigma n\|E\|$
SOLVE	MAD	5.3 %	0.42	-
PHASES	MAD	5.0 %	0.23	1.33
SHARP	MAD	9.9 % (peak-inflexion) [1]* 9.7 % (inflexion-inflexion) [2]* 10.6 % (remote-inflexion) [1]*	0.46	1.84
	SIRAS	16.3 % (remote-native) [1]*	0.12	0.95
	MAD-SIRAS	-	0.51	1.90

Table 5-4

Table showing statistics for MAD phasing of EMP derivatised Δ188_OSM crystals. * denotes values calculated using program SCALEIT (Howell and Smith, 1992). [1] and [2] denote dispersive or Bijvoet differences. All data was scaled relative to inflexion point data except for SIRAS experiment in which data was scaled to native data. SIRAS denotes independent phasing experiment between native and remote MAD data. MAD-SIRAS denotes combined phases from MAD and SIRAS experiment.

5.3.6 Molecular Replacement

Data were initially processed in the spacegroup $I2_12_12_1$ and this yielded strong peaks in the cross-rotation function, but no peaks in the translation function, so suggesting an alternate spacegroup. Cross-rotation functions of data processed in the spacegroup I222 produced a clear solution with a correlation coefficient of 21.5 %. From Table 5-5 below it can be seen that this solution shows a 9 % improvement in correlation coefficient compared to the next best solution. Solution 1 of the cross-rotation function was processed by the translation function and Table 5-6 below shows a clear solution with a correlation coefficient of 55.0 % and an R-factor of 43.5 %. These values compare favorably with the next best solution, which has a lower correlation coefficient of 41.9 %, and a higher R-factor of 49.6 %. Rigid body refinement resulted in an increase in the correlation coefficient to a value of 64.5 % and a drop in the R-factor to a value of 40.5 % or all data in the resolution range 12.0-3.5 Å.

Cross-Rotation Solution	Eulerian Angle (°) α	β	γ	Correlation Coefficient (%)
1	6.8	77.2	153.8	21.5
2	172.4	76.5	217.2	12.5
3	170.6	90.0	326.5	12.5
4	117.2	75.0	207.8	11.8
5	147.3	82.3	184.9	11.3

Table 5-5

Table showing top 5 results after cross-rotation search of molecular replacement of I222 data. Cross-rotation carried out using program AMoRe (Navaza, 1994).

Translation Solution	Eulerian Angle (°) α	β	γ	Translation (Fractional) x	y	z	Correlation Coefficient (%)	R-factor (%)
1	6.8	77.2	153.8	0.160	0.162	0.311	55.0	43.5
2	6.8	77.2	153.8	0.160	0.457	0.310	41.9	49.6
3	6.8	77.2	153.8	0.160	0.067	0.311	40.0	50.3
4	6.8	77.2	153.8	0.159	0.271	0.313	37.7	50.7
5	6.8	77.2	153.8	0.360	0.043	0.000	33.4	53.3

Table 5-6

Table showing top 5 results after cross-rotation and translation search of molecular replacement of I222 data. Cross-rotation and translation carried out using program AMoRe (Navaza, 1994).

Refined Solution	Eulerian Angle (°) α	β	γ	Translation (Fractional) x	y	z	Correlation Coefficient (%)	R-factor (%)
1	5.3	76.8	154.2	0.159	0.156	0.309	64.5	40.5
2	5.6	77.8	154.6	0.160	0.454	0.310	50.1	47.8
3	5.9	77.8	154.3	0.158	0.065	0.310	47.7	49.3
4	6.2	77.3	154.1	0.160	0.267	0.310	44.8	49.4
5	4.6	77.1	154.5	0.409	0.130	0.043	43.0	51.5

Table 5-7

Table showing top 5 results after refinement of cross-rotation and translation matrix of molecular replacement solution. Cross-rotation, translation and rigid body refinement carried out using program AMoRe (Navaza, 1994).

5.3.7 Map Calculation and Model Refinement

Initial electron density maps based on combined MAD and SIRAS phases showed poor side chain density although the main chain density was generally contiguous. Solvent flattening improved the quality of density and further improvements were seen during the course of refinement as shown below in Figure 5-7.

Crystallographic-R and free-R (Brunger, 1992) values were monitored throughout the course of refinement, as were deviations from stereochemical parameters as defined by Engh and Huber (Engh and Huber, 1991). Initial rigid body refinement resulted in a model with crystallographic-R value of 48 % and a free -R value of 52 %. Low weighting of the X-ray terms resulted in tight stereochemistry with root mean squared deviations (RMSD) from ideality in bond lengths typically less than 0.008 Å. After 15 rounds of manual rebuilding and refinement and the inclusion of 116 waters the current model has a crystallographic-R value of 20.5 % with a free-R value of 26.1 % for all data between 20.0-2.2 Å. Stereochemistry is typified by RMSD in bond lengths of 0.0147 Å and RMSD in bond angles of 1.7997°. Statistics relevant to model refinement are shown below in Table 5-8.

The current model for the I222 structure contains 35 waters and has been refined to a crystallographic-R value of 20.8 % and a free-R value of 28.9 % for all data between 20.0-2.5 Å. Stereochemistry is typified by RMSD in bond lengths of 0.0129 Å and RMSD in bond angles of 1.6112°. Statistics relevant to the final models are shown below in Table 5-8.

Spacegroup	$P2_12_12_1$	I222
Resolution Range	20–2.2 Å (2.28–2.20 Å)	20–2.5 Å (2.59–2.50 Å)
R_{cryst}	20.5% (27.1%)	20.8% (27.8%)
R_{free}	26.1% (33.6%)	28.9% (31.5%)
Completeness		
Working Set	86.1% (81.0%)	75.7% (75.4%)
Test Set	9.9% (8.6%)	9.0% (6.6%)
Number of Reflections (F>0)		
Working Set	9410 (870)	4927 (481)
Test Set	1082 (92)	583 (42)
Number of Non-Hydrogen Atoms		
Protein	1319	1219
Water	116	35
RMSD		
Bond Lengths	0.0147 Å	0.0129 Å
Bond Angles	1.7997°	1.6112°
Dihedrals	19.8882°	20.6155°
Improper	1.0381°	1.1047°
Average B-factors		
Main Chain	27.1 Å²	38.5 Å²
Side Chain	32.4 Å²	42.9 Å²
Water	38.4 Å²	41.5 Å²
All Atoms (Range)	30.5 Å² (6.3–84.9 Å²)	40.8 Å² (10.2–84.4 Å²)

Table 5-8

Structural refinement and model statistics for the two crystal structures of Δ188_OSM. $R_{cryst}=\Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|$. R_{free} is as for R_{cryst} but calculated using a 10% test set of reflections excluded from the refinement. Values in parentheses correspond to the highest resolution shell (2.28–2.20 Å for $P2_12_12_1$ and 2.59–2.50 Å for I222).

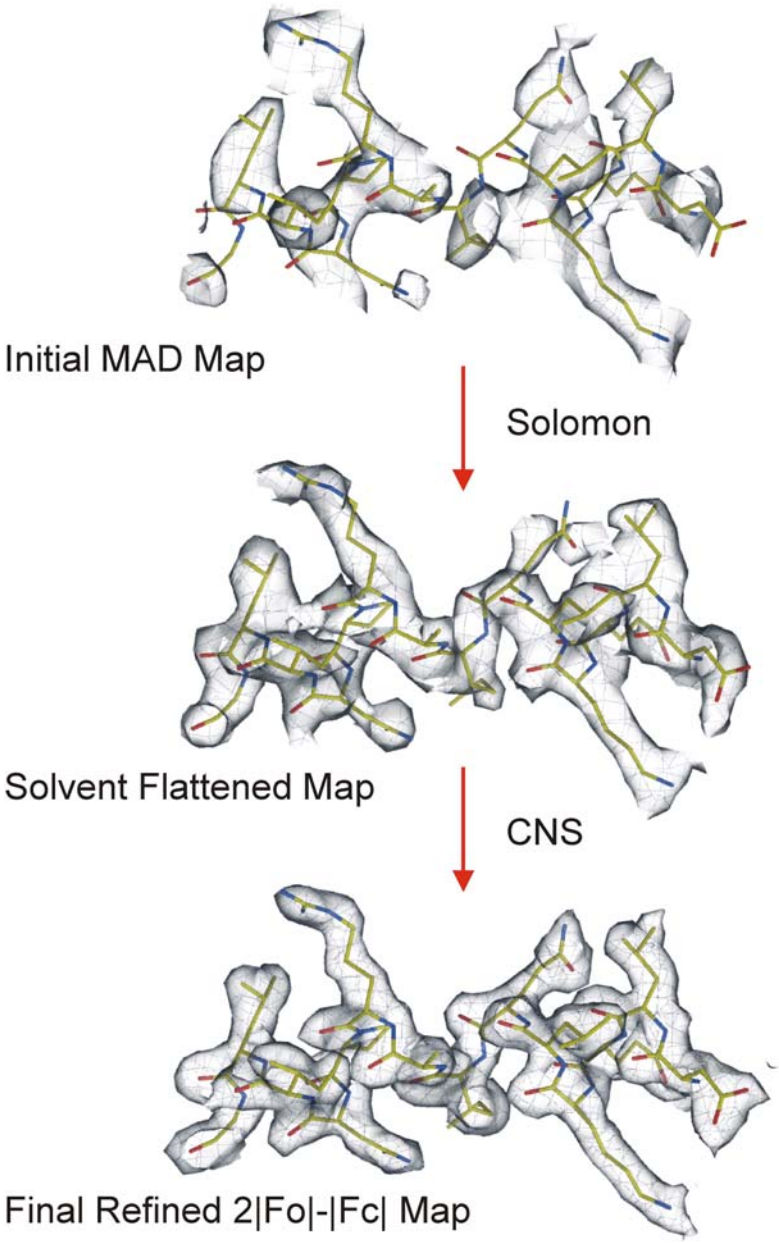


Figure 5-7

Figure showing the course of refinement of the Δ188_OSM model. All maps are contoured at 1σ. Region shown corresponds to the helix C running from residues 106 to 119. Maps were produced from MAD data using SHARP (de La Fortelle and Bricogne, 1997) followed by solvent flattening with SOLOMON (Abrahams and Leslie, 1996) and refinement with CNS-SOLVE (Brunger *et al.*, 1998). Final refined coordinates are shown in each case. Figure was produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

5.3.8 Analysis of Model

In the $P2_12_12_1$ structure residues 4-134 and 156-187 are well ordered, with residues 1-3 at the N-terminus and residues 135-155 of the C-D loop being disordered with no interpretable electron density. Electron density is good throughout the course of the main chain and a representative portion of electron density is shown below in Figure 5-8. Residue numbering corresponds to that of the human native mature protein (GenBank accession number M27286; (Malik *et al.*, 1989).

In the I222 structure residues 6-94, 104-132 and 158-187 are well ordered. Residues 1-5 at the N-terminus, residues 95-103 in the B-C loop and residues 133-157 of the C-D loop are disordered and display no interpretable electron density. Electron density is good throughout the course of the main chain and a comparison of the electron density for the two structures is shown below in Figure 5-8.

Reasons for the poor electron density at the N-terminus and inter helical loop regions are readily apparent on inspection of the B-factors for the $P2_12_12_1$ structure as shown in Figure 5-9 below. In both the $P2_12_12_1$ and I222 structures, the N-terminal loop region preceding Leu 10 shows a sharp increase in main chain B-factor values, with both rising above the mean values of 27 $Å^2$ and 39 $Å^2$, respectively. Similarly, the C-terminal region of helix C from Glu 130 and the N-terminal region of helix D before Phe 160 show a similar increase. Also, higher than average main chain B-factors are observed for both structures between residues Glu 90 and Lys 110. The level of thermal motion in these regions suggests a high degree of flexibility within the crystal, which therefore gives rise to multiple conformations and ambiguous electron density in these regions.

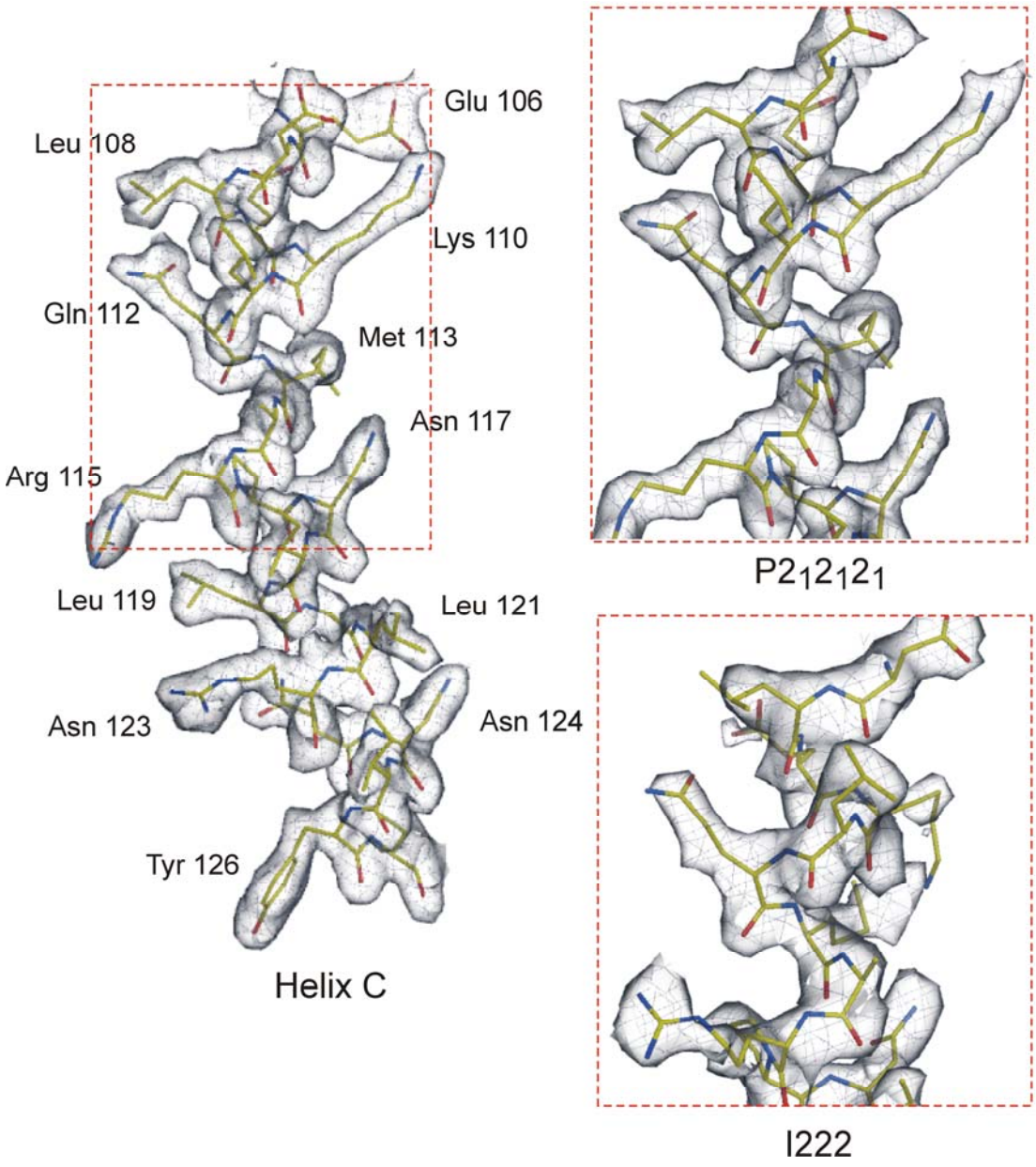


Figure 5-8

Figure showing the refined co-ordinates and $2|F_o|-|F_c|\alpha_{calc}$ electron density maps of Δ188_OSM contoured at 1 σ. Main panel shows region of $P2_12_12_1$ map corresponding to helix C running from residues 106 to 126. Boxed section shows corresponding region of map for each of the two spacegroups. The I222 map is to a resolution of 2.5 Å and the $P2_12_12_1$ map to a resolution of 2.2 Å. Maps were produced from MAD data using SHARP (de La Fortelle and Bricogne, 1997) followed by solvent flattening with SOLOMON (Abrahams and Leslie, 1996) and refinement with CNS-SOLVE (Brunger *et al.*, 1998). Figure was produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

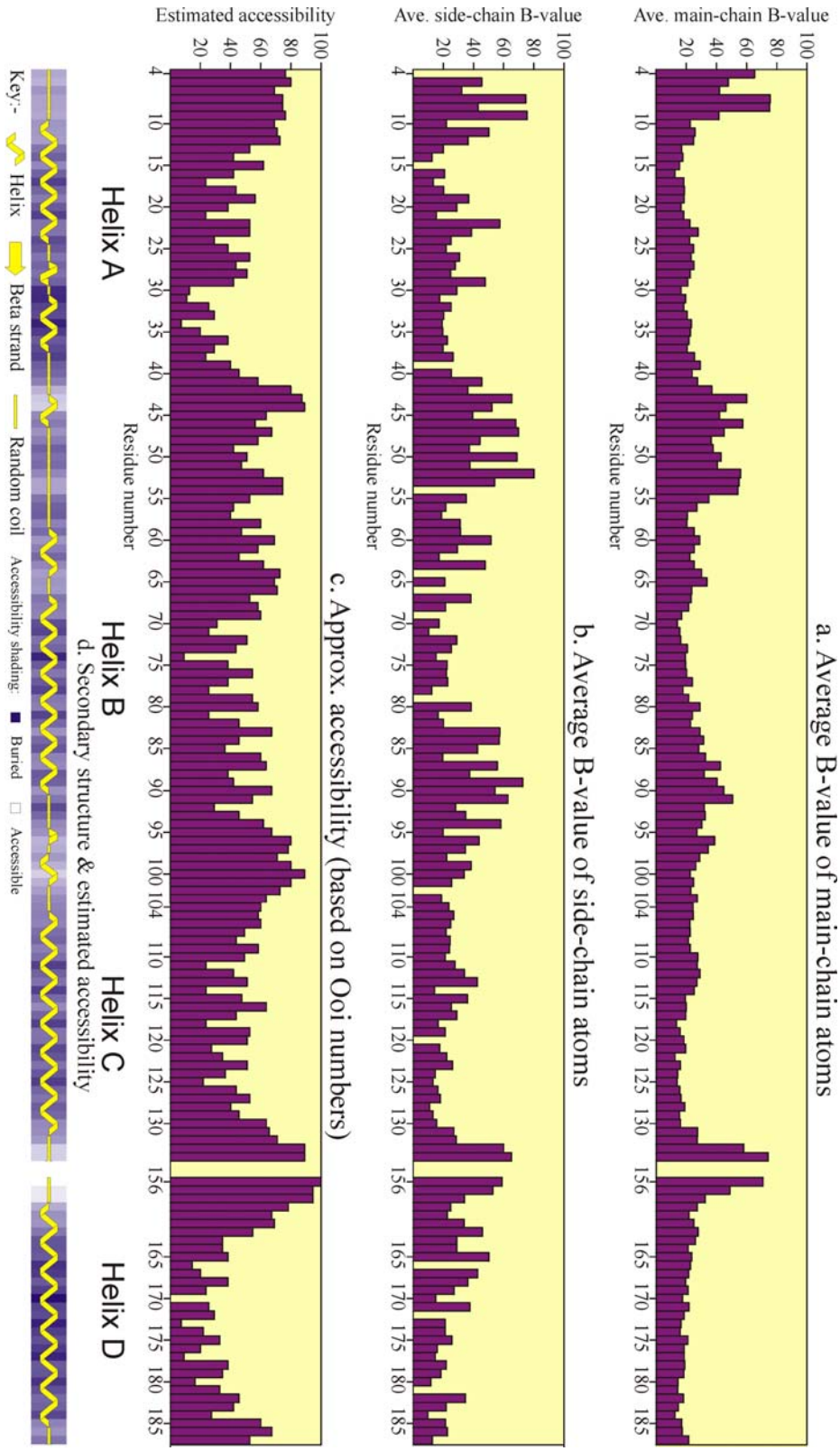


Figure 5-9

Figure showing a) average main chain B-factor, b) average side chain B-factor, c) solvent accessibility and d) secondary structure assignment by residue for the entire Δ188_OSM sequence. Data shown for $P2_12_12_1$ structure. Figure produced using PROCHECK (Laskowski *et al.*, 1993).

Analysis of the Ramachandran plot shown in Figure 5-10 below shows that the stereochemistry of the molecule is good with no residues lying in energetically disallowed regions, and only Arg 67 of the I222 structure falling in a generously allowed region. No clear density was observed for Arg 67, so suggesting high levels of flexibility for this side chain.

For the $P2_12_12_1$ structure 94.3 % of residues lie in the most favoured regions (A, B, L) and 5.7 % lie in additionally allowed regions (a, b, l, p). No residues lie in generously allowed (~a, ~b, ~l, ~p) or disallowed regions.

For the I222 structure 88.1 % of residues lie in the most favoured regions and 11.1 % lie in additionally allowed regions. No residues lie in disallowed regions and only 0.8 % of residues fall in generously allowed regions

A number of residues including Gln 20, Asp 22, Asp 26, Glu 51, His 83, Asp 87, Asn 123, Lys 163, Arg 162, and Arg 84 appeared to have alternate side-chain conformations. Refinement of conformers was not successful, presumably due to the low occupancies (typically <20 %) of the minor conformations and the relatively low resolution of data.

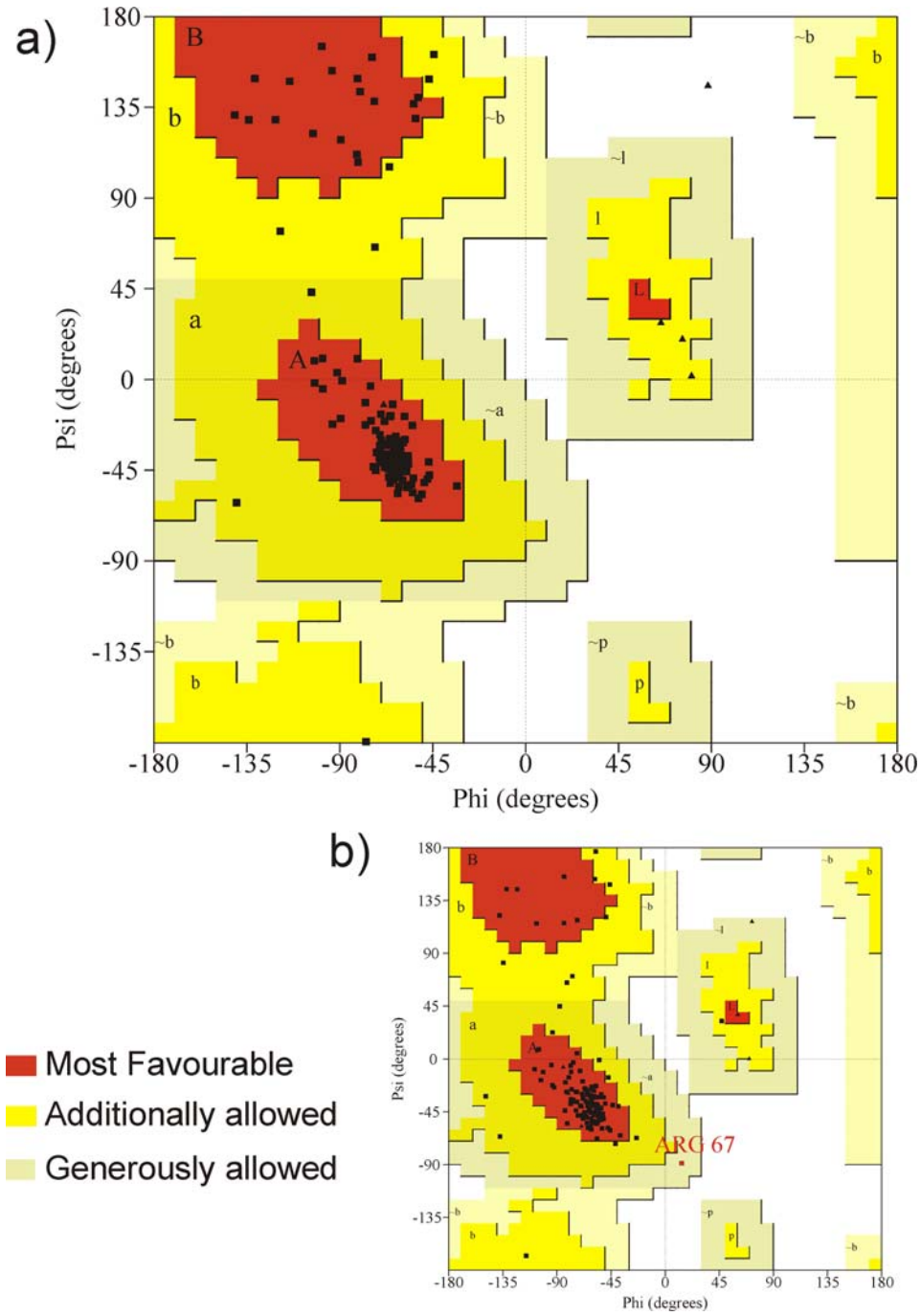


Figure 5-10

Figure showing Ramachandran plot of the stereochemistry observed in a) $P2_12_12_1$ model and b) I222 model. For the $P2_12_12_1$ model 94.3 % of residues lie in the most favoured regions (A, B, L) and 5.7 % lie in additionally allowed regions (a, b, l, p). No residues lie in generously allowed (~a, ~b, ~l, ~p) or disallowed regions. For the I222 structure 88.1 % of residues lie in the most favoured regions and 11.1 % lie in additionally allowed regions. No residues lie in disallowed regions and only 0.8 % of residues fall in generously allowed regions. Figure was produced using the program PROCHECK (Laskowski *et al.*, 1993).

5.4 Discussion of Structure Determination

Δ188_OSM was the only sample that produced crystals. Both OSM-196 and Se-Met-OSM variants showed exceptionally poor solubility and aggregation. Crystals used for structure determination were grown from PEG 35000 and this is the first reported use of this reagent for the growth of protein crystals.

Se-Met derivatised OSM, displayed exceptionally poor solubility presumably due to the reduced solubilty of Se-Met compared to Met. Indeed, the C-Se bond is 0.14 Å longer than the C-S bond and this leads to greater polarity, weaker bonds and higher reactivity. Selenides are more susceptible to oxidation than the corresponding sulphide and the resultant selenoxides themselves act as mild oxidising agents. Inclusion of reducing agents to the preparation was not sufficient, so suggesting further problems with protein folding. This is confirmed on inspection of the final refined model, which shows that all four methionine residues in the sequence are buried in the core of the protein and all display solvent accessibilities of less than 50 % (see Figure 5-5 (b)).

Flash cooling of the crystals proved problematic, presumably due to the high viscosity of mother liquor solutions and the associated lack of heat transfer. Passing crystals through solutions of lower molecular weight PEG aleviated the problem, so allowing MAD analysis of single crystals to be carried out. An EMP derivative of OSM was produced and this was shown to bind at two major and one minor site (see Figure 5-5 (a)). Although the initial identification of the derivative was achieved with room temperature SIRAS data, crystal decay and non-isomorphism underscored the necessity for the single crystal MAD experiment. A fluorescence scan was used to characterise the mercury LIII edge of the derivative and a traditional three wavelength MAD experiment carried out. Phases based on this experiment were supplemented by the addition of further phasing from the independent SIRAS experiment and the resulting maps showed improved quality.

The current model has a crystallographic-R value of 20.5 % with a free-R value of 26.1 % for all data between 20.0-2.2 Å. A second crystal form belonging to the spacegroup I222 was subsequently solved using molecular replacement and this structure has been refined to a crystallographic R-value of 20.8 % and an R-free value of 28.9 % for all data between 20.0-2.5 Å.

The final model shows significant deviations to the initial model generated using Modeller, with an RMSD value of 2.40 over 151 structurally equivalent Cα. This suggests possible reasons for the failure of molecular replacement techniques; although solutions using the refined model coordinates of human LIF and murine LIF also failed.

6. Chapter 6 – Structure Description, Comparison and Functional Implications

The aims of this chapter include the description of the OSM structure and comparisons with other members of the haemopoietin cytokine family. Detailed analysis of the structure is reported along with implications for receptor binding and species specificity. The crystal structure solved in the spacegroup $P2_12_12_1$ is used for analysis (unless stated otherwise), due to the higher quality of experimental data. A comparison of the I222 and $P2_12_12_1$ structures is also given. Finally, general structural themes pertinent to the haemopoietin cytokine network are also discussed along with functional dissection of the site 1, 2 and 3 receptor binding epitopes as depicted schematically in Figure 6-0 below.

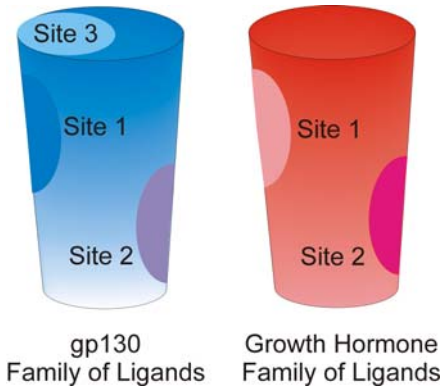


Figure 6-0
Schematic representation of the site 1, 2 and 3 receptor binding site motifs as observed in the gp130 and growth hormone family of ligands. The 4-helix bundle is depicted as a cylinder and the regions responsible for receptor binding are highlighted.

6.1 Structure Description

As predicted from sequence analysis and homology models (Bazan, 1991), OSM conforms to the up-up-down-down 4-helix bundle topology, as observed for other members of the haemopoietin cytokine family whose crystal structures have been determined. Figure 6-1 and Figure 6-2 below show Cα stereodiagrams and ribbon representations of OSM, respectively.

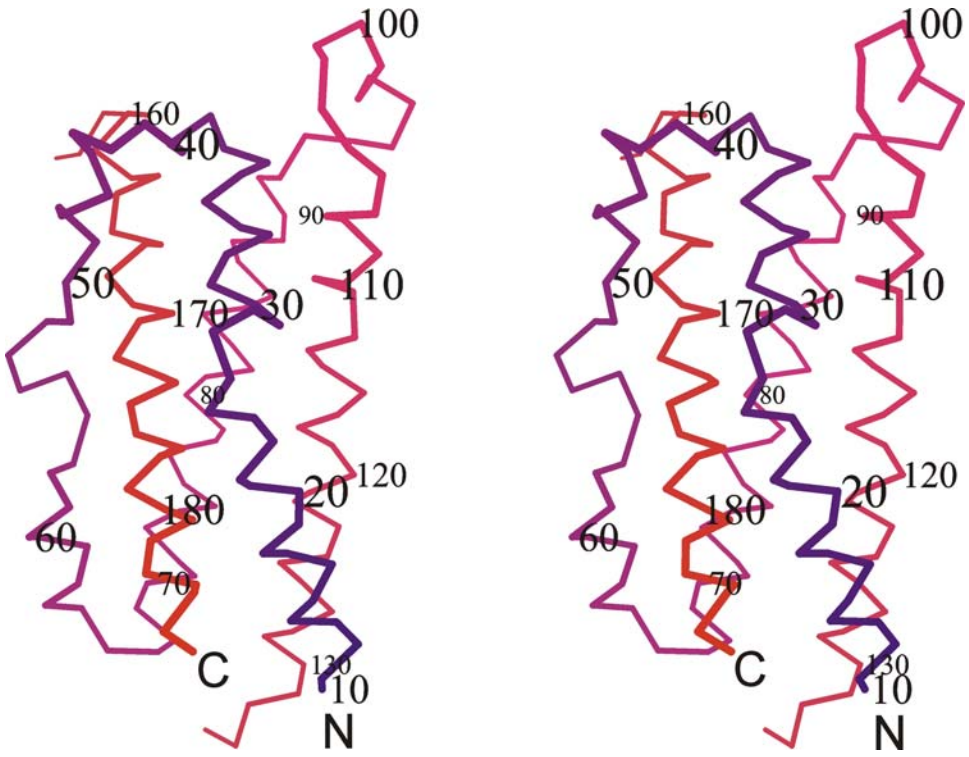


Figure 6-1

Figure showing Cα stereodiagram of OSM coloured from blue at the N-terminus through to red at the C-terminus. Every tenth residue is numbered. Figure produced using the program MOLSCRIPT (Kraulis, 1991).

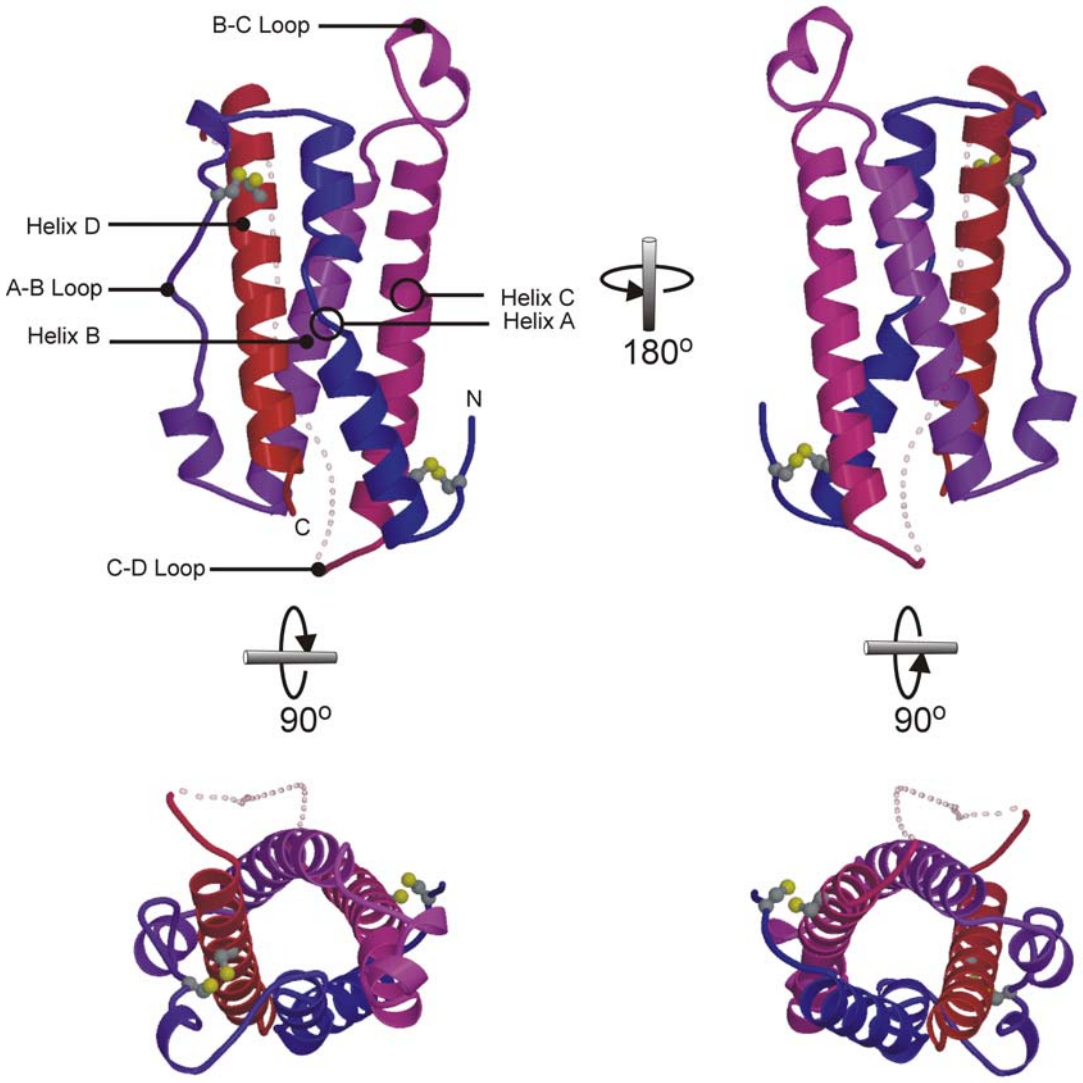


Figure 6-2

Figure showing orthogonal projections of the ribbon representation of OSM coloured from blue at the N-terminus through to red at the C-terminus. Disulphide bonds are represented as ball and stick objects with the sulphur atoms highlighted by yellow spheres. Transparent dotted section represents loop region as observed in LIF. Position of open circles on helix A and C denotes location of kinked regions. Figure produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

OSM is a compact barrel shaped molecule with dimensions of approximately 20 Å x 27 Å x 56 Å. The basic fold of OSM consists of 4 main helical regions (helix A, residues 10-37; helix B, residues 67-90; helix C, residues 105-131; helix D, residues 159-185) connected by two long overhand loops (A-B loop, residues 38-66; C-D loop, residues 132-158) and one short loop (B-C loop, residues 91-104). The secondary structure assignment and hydrogen bonding pattern is shown below in Figure 6-3.

The overall arrangement of helices in this fold may be considered as two pairs of anti-parallel helices that pack together, with helices A and D forming one pair and helices B and C forming the other. Helices B and C pack in a classical fashion, with a skew angle between the helical axis of around 20^{o}, as observed in other members of the long chain 4-α-helix cytokines. Helix C is the longest helix consisting of 26 residues that run the length of the molecule with helix B being the shortest helix of only 23 residues in length. Figure 6-4 (b) below shows that helix C exhibits a slight kink towards the N-terminus of its length, between residues Pro 116 and Gln 112, due to disruption of the classic hydrogen bonding pattern observed for α-helices. This disruption is in part due to substitute hydrogen bonds, which are formed with two tightly bound water molecules.

Helices A and D pack in similar fashion although a more pronounced kink is observed in helix A (see Figure 6-4 (a)). This serves to allow closer packing of the A:D and B:C helical pairs, so forming a tighter solvent inaccessible core. The kink in helix A is caused by a break in the helical conformation, as residues Gln 25 and Leu 30 adopt a hydrogen bonded turn conformation, and also four tightly bound water molecules disrupt the classic hydrogen bonding pattern. Residues Thr 27 to Ile 37 of helix A adopt a 3_{10}-helix conformation.

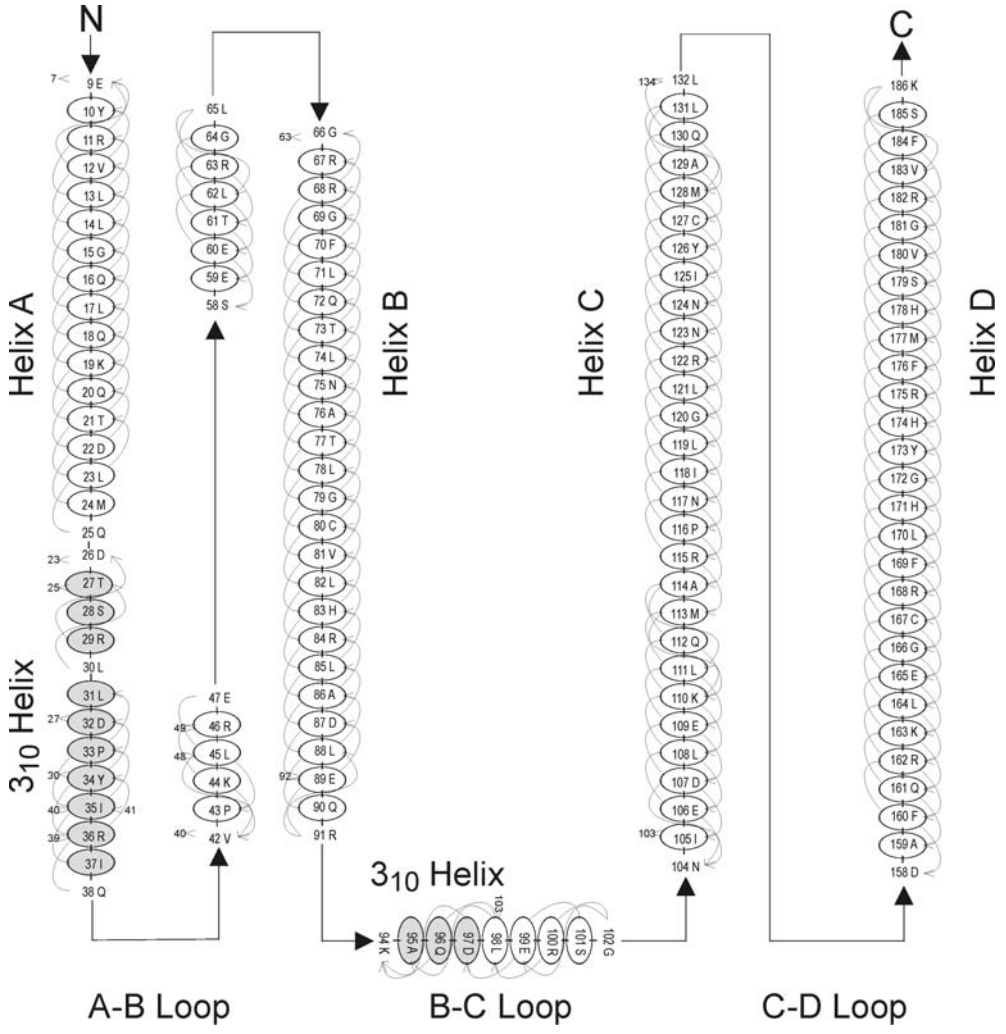


Figure 6-3

Figure showing secondary structure and hydrogen bonding pattern of $P2_12_12_1$ OSM. Figure produced using DSSP (Kabsch and Sander, 1983) and HERA (Hutchinson and Thornton, 1990). Residues represented as open ovals are in a α-helical conformation and residues represented as shaded ovals are in a 3_{10}-helical conformation. All other residues are in an extended loop conformation, except for residues 1-3 at the N-terminus and residues 135-155 of the C-D-loop for which no appreciable density was observed. Arrows represent hydrogen bonding pattern.

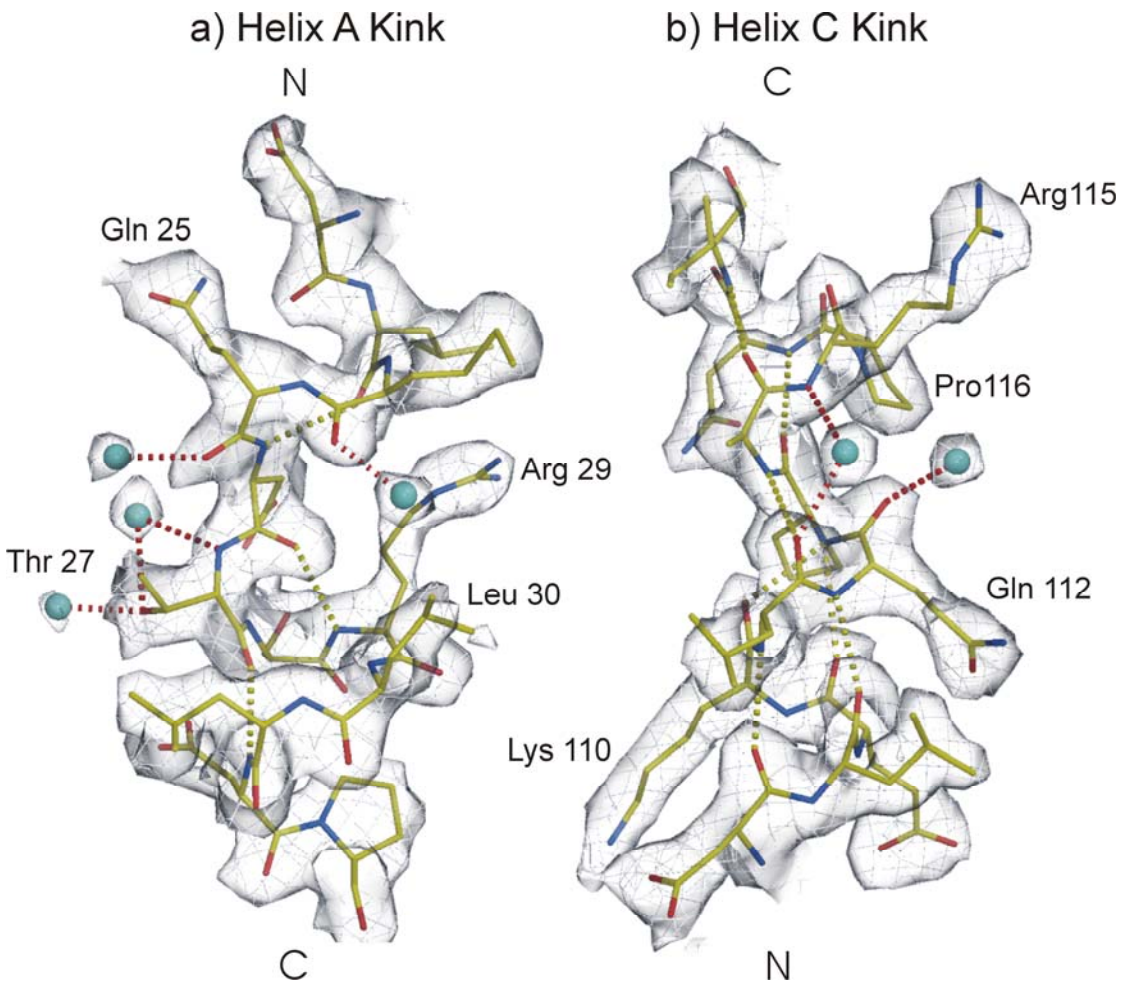


Figure 6-4

Figure showing kinked regions of a) helix A and b) helix C. Refined 2.2 Å $2|F_o|-|F_c|$ electron density map is shown contoured at 1σ. Main chain hydrogen bonds are shown as yellow dashed lines and hydrogen bonds to waters are shown as red dashed lines. Bound water molecules as shown as sky blue spheres. Map was produced from MAD data using SHARP (de La Fortelle and Bricogne, 1997) followed by solvent flattening with SOLOMON (Abrahams and Leslie, 1996) and refinement with CNS (Brunger *et al.*, 1998). Figure was produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

Analysis of Figure 6-5 (a) and Figure 6-5 (c) below demonstrates that close packing of the A:D and B:C helical pairs results in a tight solvent inaccessible core with the outer faces of the helices having many charged residues and the inner faces consisting of many hydrophobic residues. This solvent inaccessible core of the protein is composed of two groups of aromatic stacking interactions in which Phe 56, Tyr 173, Phe 169 and Phe 176 form one group and Phe 70, Phe 185 and Trp 187 form the second. This distribution emphasises the hydrophobic nature of helix D, which donates all of the aromatics to these groupings with the exception of Phe 56 (A-B loop) and Phe 70 (B helix). Helix D also shows a clustering of around seven positively charged residues on its outer surface which presumably affects the surface electrostatic potential of site 2 interactions in a similar fashion to that proposed for CNTF (Panayotatos *et al.*, 1995); see 6.3.3.1 for discussion) and GH (Cunningham and Wells, 1993). In contrast to the large number of hydrophobic interactions stabilising the fold, no hydrogen bonds between the main helices are observed (see Figure 6-3).

Analysis of B-factor values in Figure 6-5 (b) shows that the tight packaging of the helices also results in a rigid helical bundle, with only the loop regions and the C-terminus end of helix B exhibiting any motion. Average B-factors are 30.5 $Å^2$ for the $P2_12_12_1$ structure (see Figure 6-8 for comparison with I222 structure).

The N-terminal loop preceding helix A (Glu 4 to Gly 9) is tethered to the N-terminus of helix C by one of the two disulphide bridges spanning Cys 6 and Cys 127. In a similar fashion, the second disulphide bridge, between Cys 49 and Cys 167, anchors the start of the A-B loop to the N-terminal region of helix D. The A-B loop consists of two separate α-helical regions running between Pro 43 to Arg 46 and Glu 59 to Gly 64. Residues between the two helical regions of the A-B loop adopt an extended conformation and pack closely against helix D, with Phe 56 also contributing to the hydrophobic molecular core. The B-C loop contains a second small region of 3_{10}-helix running between residues Ala 95 and Asp 97 with a traditional α-helical conformation being adopted for the rest of the helix to residue Ser 101.

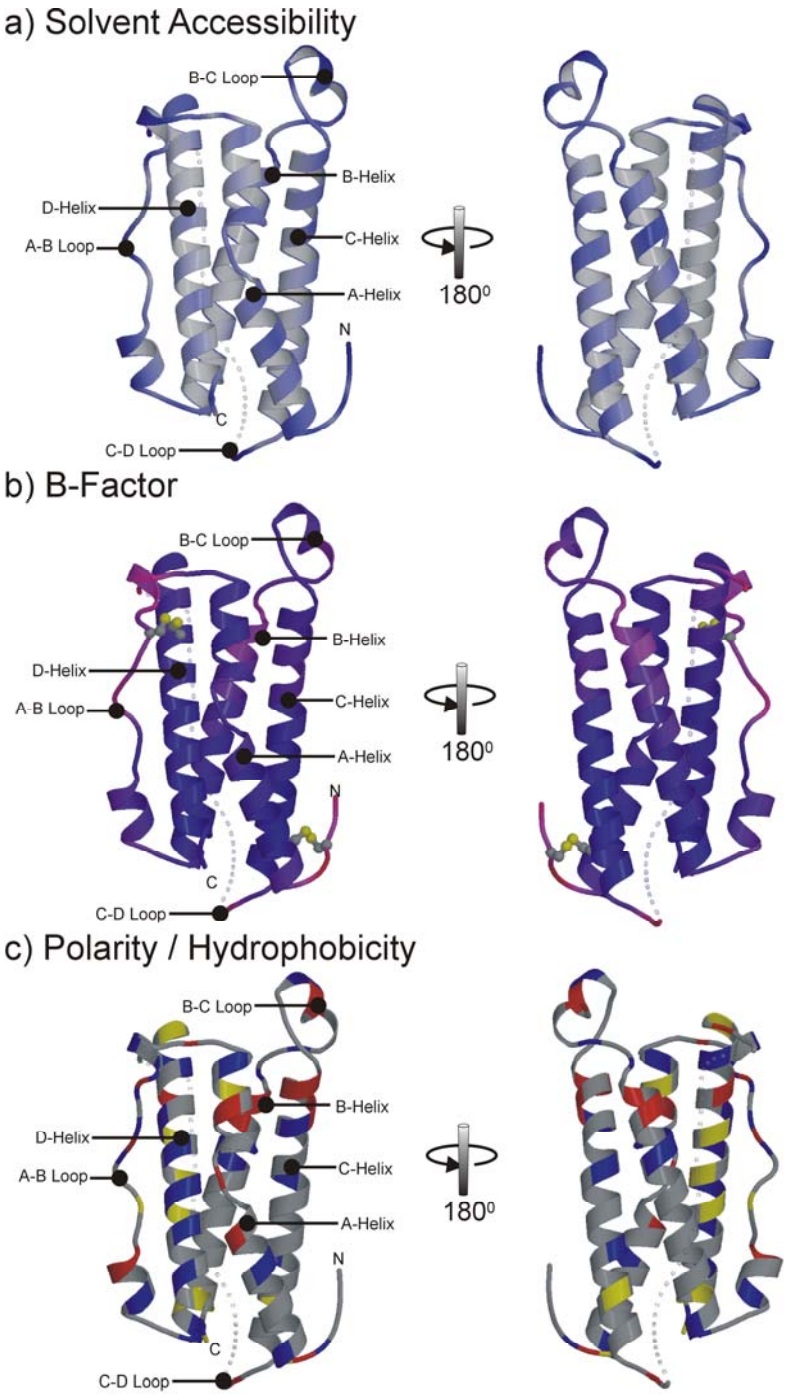


Figure 6-5

Figure showing ribbon representation of OSM with Cα coloured according to (a) relative solvent accessibility, (b) refined B-factors and (c) charge and hydrophobicity. Relative solvent accessibility values are coloured from the minimum value of 0 % to the maximum value of 100 % from grey to blue, respectively. B-factors are coloured from the minimum value of 6.3 Å^2 to the maximum value of 84.9 Å^2 from blue to red, respectively (average B-factor 30.5 Å^2). Positively charged residues (Arg, Lys or His) are coloured blue, negatively charged residues (Asp or Glu) red and hydrophobic residues (Trp, Tyr or Phe) yellow. Figure produced using NACCESS (Hubbard and Thornton, 1993), BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

6.2 Structural Comparisons

The helical cytokine family has previously been classified in terms of the helix length; short chain or long chain (Nicola, 1994). Short chain helical cytokines analysed structurally include IL-5 (Milburn *et al.*, 1993), M-CSF (Pandit *et al.*, 1992), IFN-γ (Ealick *et al.*, 1991; Samudzi *et al.*, 1991) and IL-3 (Feng *et al.*, 1996) which are dimeric, and GM-CSF (Diederichs *et al.*, 1991; Walter *et al.*, 1992b), IL-2 (Bazan, 1992) and IL-4 (Powers *et al.*, 1993; Walter *et al.*, 1992b; Wlodawer *et al.*, 1992) which are monomeric (see Chapter 1 for overview).

Long chain helical cytokines resolved structurally include GH (Abdel Meguid *et al.*, 1987; de Vos *et al.*, 1992), Epo (Livnah *et al.*, 1998; Livnah *et al.*, 1996; Syed *et al.*, 1998), Leptin (Zhang *et al.*, 1997a), GCSF (Aritomi *et al.*, 1999; Hill *et al.*, 1993; Lovejoy *et al.*, 1993), IL-10 (Zdanov *et al.*, 1995), IFN-β (Senda *et al.*, 1995) and IFN-τ (Radhakrishnan *et al.*, 1999) with the later two adopting a five helix bundle topology.

6.2.1 Haemopoietin Cytokine Family

The long chain helical cytokine family can be further sub-divided into those using the gp130 receptor or the interferon receptor and this present structure represents the fourth member of the gp130 cytokine family to be solved structurally. Members of the gp130 family of cytokines delineated structurally so far include IL-6 (Somers *et al.*, 1997; Xu *et al.*, 1997), CNTF (McDonald *et al.*, 1995) and LIF (Robinson *et al.*, 1994 [X-ray]; Purvis and Mabbutt, 1997 [NMR]). For the purpose of this study medium to high-resolution X-ray structures will be used for comparative purposes with coordinates deposited in the PDB, except for GCSF-GCSFR coordinates kindly provided by K. Morikawa, Biomedical Engineering Research Council, Osaka, Japan.

Analysis of root mean squared deviation (RMSD) values as in Table 6-1 below confirms that, structurally OSM belongs to the gp130 family of cytokines. Murine LIF shows

greatest similarity to OSM with an RMSD value of 2.10 Å over 145 structurally equivalent residues and a sequence identity of 21 %. Human LIF is close behind with an RMSD value of 2.27 Å over 138 structurally equivalent residues and a sequence identity of 20 %. Human IL-6 and human CNTF are the next best matches with RMSD values typically greater than 2.40 Å and a sequence identity typically greater than 12 %. Human OSM shows more distant structural similarity to the growth hormone family of helical cytokines with RMSD values typically greater than 2.80 Å and a sequence identity typically less than 9 %. Within the growth hormone family the greatest similarity is shared with Epo with an RMSD value of 2.80 Å, although sequence identity is exceptionally low at only 6 %. However, OSM clearly shows greater similarity to GCSF with an RMSD value of 2.80 Å over 127 equivalencies and a sequence identity of 19 %.

Family	Protein	Total Number of Residues	Number of Structurally Equivalent Residues	Number of Sequence Identities within Equivalents	RMSD (Å)
gp130 Family	Murine LIF	172	145	31 (21 %)	2.10
	Human LIF	159	138	28 (20 %)	2.27
	Human IL-6	157	127	22 (17 %)	2.40
	Human CNTF	150	120	14 (12 %)	2.95
	Human OSM (I222)	149	147	147 (100 %)	0.72
	Human GCSF	172	127	25 (19 %)	2.65
Growth Hormone Family	Human Epo	154	120	7 (6 %)	2.80
	Human GH	186	120	9 (8 %)	3.08
	Human Prl	186	133	12 (9 %)	3.17

Table 6-1

Table showing structural comparison of Cα positions ranked according to RMSD values of human OSM with various other helical cytokines of the gp130 and growth factor related families. Structural superpositions and RMSD values calculated using the program SHP (Stuart *et al.*, 1979). GCSF-GCSFR coordinates kindly provided by K. Morikawa, Biomedical Engineering Research Council, Osaka, Japan. All other co-ordinates from PDB.

Structural superposition of human OSM onto human LIF, human IL-6, human CNTF, human GCSF and human Epo are shown below in Figure 6-6. To a first approximation the length and orientation of the four main helical regions is conserved between the five molecules. However, several subtle differences are observed. Starting at the N-terminus the loop preceding helix A is anchored to helix C by two disulphides in LIF, whereas only one disulphide is present in this region in OSM. These anchoring disulphide bridges are not present in IL-6, CNTF, Epo or GCSF. In the IL-6 structure this loop region deviates away from the body of the molecule and in the GCSF structure the loop passes over the end of the helical bundle, with the side chain of Leu 10 packing between the helical bundle. CNTF is unique in having no disulphide bridges in the structure, and the lack of an anchoring N-terminal disulphide linkage results in high flexibility for the first 10 residues, although a similar course to both OSM and LIF is suggested.

Helix A adopts a similar course in all structures although OSM differs from others in the severity of the kink. Both CNTF and LIF have minor kinks due to disruption of the main chain hydrogen bonding around residues Thr 32 and Ser 36, respectively (McDonald *et al.*, 1995; Robinson *et al.*, 1994). Both these residues bind waters through the Oγ1 atom and a similar pattern is observed in OSM through an equivalent residue Thr 27 (see Figure 6-4 (a)). Helix A of Epo differs from all the other structures in being approximately one turn shorter at both the N- and C-terminus. In OSM the C-terminal section of helix A adopts a 3_{10} helical conformation; a secondary structural feature not observed in any of the other structures. The unique use of 3_{10} helix is also seen in the N-terminal portion of the B-C loop.

The A-B loop packs against the *external* face of helix D approximately half way along its length. Similar paths are observed in OSM, LIF, IL-6 and CNTF. Epo, however, differs markedly in this region to all the other structures as the C-terminal region of the A-B loop

adopts a path at approximately right angles to that of the other structures, so crossing the *internal* face of helix D at its N-terminus. The altered path of Epo is mainly due to the shorter length of the A-B loop and a sharp turn stabilised by a disulphide bridge between residues 29 and 33. In a similar fashion the A-B loop of GCSF adopts an altered path around the *external* face of helix D at its N-terminus. The altered path again due in part to a disulphide bond stabilised turn between residues 37 and 43 and the insertion of an extra helical segment. In common, both LIF and OSM, have a disulphide bridge anchoring the first part of the A-B loop to the top of helix D. In comparison a similar region of the A-B loop of IL-6 is anchored to helix B. No density is present for the A-B loop of CNTF although a similar course is suggested. A major distinction of the OSM structure is the presence of two helical regions in the A-B loop; one running from the C-terminus of helix A and the second running into the N-terminus of helix B. In comparison, LIF and GCSF contain only the first helical region running from the C-terminus of helix A and IL-6 and Epo contain only the second running into the N-terminus of helix B.

Both helices B and C match closely between the structures with LIF, IL-6, Epo and GCSF all having short interconnecting B-C loop regions. However, this loop region in OSM has a 5 residue insertion that forms just over 1 turn of helix that protrudes from the top of the helical bundle. This feature has so far not been observed in any other members of the gp130 cytokine family, although similar helical features are observed in GH and short chain cytokines such as GM-CSF and IL-4. The N-terminus of helix C is approximately 1 turn shorter in both Epo and GCSF.

The C-D loop has not been fully modelled into OSM and CNTF, although the orientation of residues preceding this loop suggests a similar route is taken to that of LIF, IL-6 and CNTF in which the loop passes over helix B approximately half way along its length. Deviation from this pattern is observed in Epo as the C-terminal section of the loop

protrudes away from the main helical bundle as a series of residues adopt a turn conformation. GCSF, IL-6 and LIF all show similar termination positions of helix D to within 2 or 3 residues.

Helix D shows similar positioning in all structures, although that of Epo is ~2 turns shorter and that of CNTF is ~2 turns longer with the extended section packing towards the N-terminus of helix A, presumably in a similar fashion to non-truncated wildtype OSM.

In summary, features unique to OSM include the helical B-C loop region which forms an extended 'ear' protruding from the main body of the helical bundle and two sections of helix in the A-B loop. The severity of the kink in helix A is also unique to OSM along with the use of more tightly coiled 3_{10} helix. Together these subtle but none the less unique features may function in modulation of receptor-ligand interactions in the gp130-LIFR system.

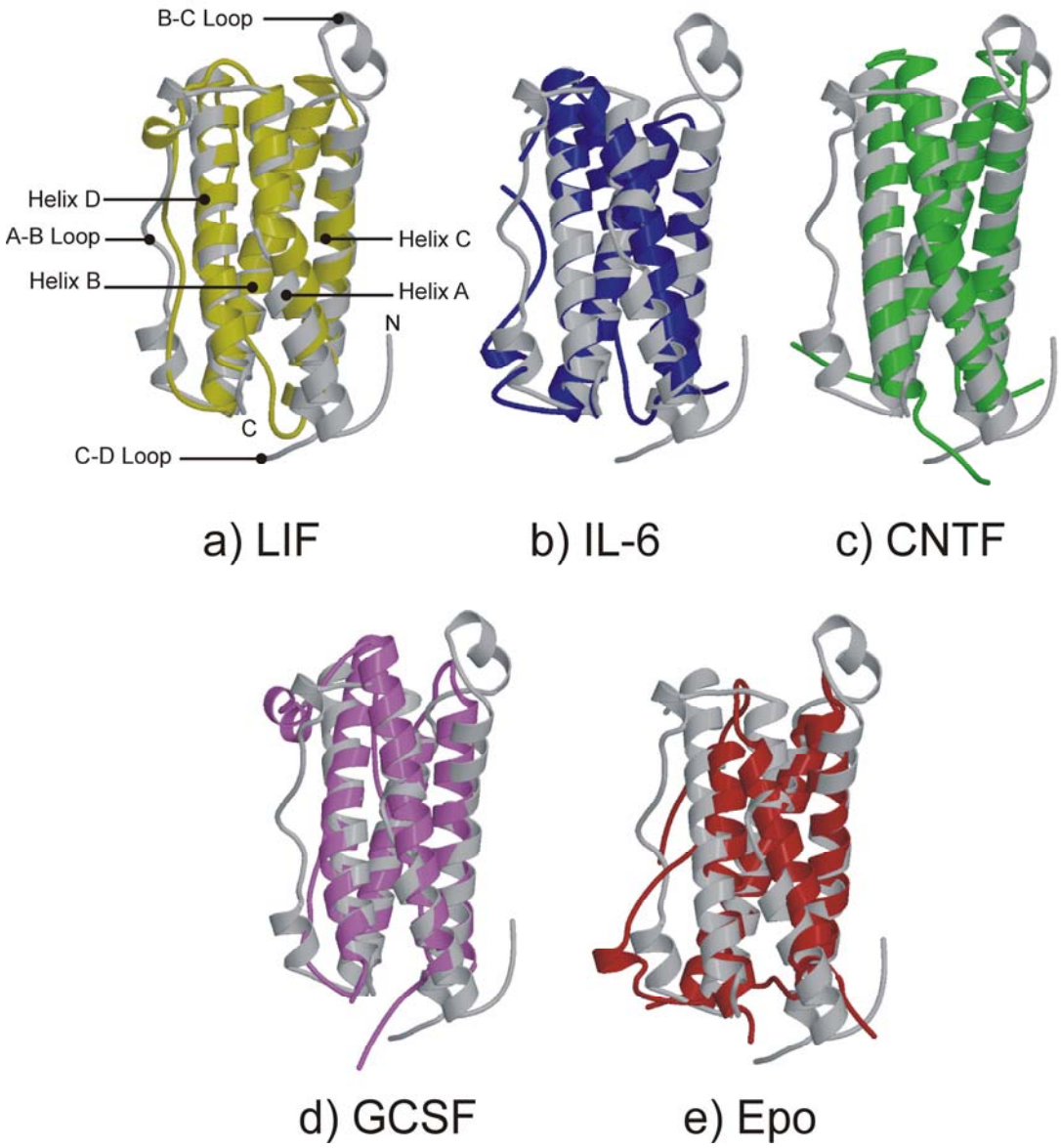


Figure 6-6

Figure showing pairwise structural superpositions of (a) human LIF (yellow), (b) human IL-6 (blue; Somers *et al.*, 1997), (c) human CNTF (green; McDonald *et al.*, 1995), (d) human GCSF (purple; Aritomi *et al.*, 1999; Hill *et al.*, 1993) and human Epo (red; Livnah *et al.*, 1996) with OSM (grey). GCSF coordinates kindly provided by K. Morikawa, Biomedical Engineering Research Council, Osaka, Japan. Structural superpositions were produced using SHP (Stuart *et al.*, 1979) and figures were produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

6.2.2 Haemopoietin Receptor Family

The structural similarity of the OSM-gp130 system to that of the Epo-EpoR and GCSF-GCSFR system is further reflected on analysis of the structural similarities of gp130, EpoR and GCSFR as shown below in Table 6-2. This suggests that gp130 shows greatest structural similarity to GCSFR with an RMSD value of 2.52 Å, and a sequence identity of 29 % over 189 structurally equivalent residues. Substantially similarity has also been reported between gp130 and EpoR (Bravo *et al.*, 1998) with an RMSD value of 2.52 Å over 177 equivalencies.

Family	Protein	Total Number of Residues	Number of Structurally Equivalent Residues	Number of Sequence Identities within Equivalents	RMSD (Å)
Haemopoietin Receptor	Human GCSFR	213	189	54 (29 %)	2.52
	Human EpoR	217	177	36 (20 %)	2.52
	Human PrlR	197	171	33 (19 %)	3.03
	Human GHR	192	164	31 (19 %)	3.07

Table 6-2

Table showing structural comparison ranked according to RMSD values of Cα superpositions for human gp130 with various other receptors of the haemopoietin family. GCSF-GCSFR coordinates (Aritomi *et al.*, 1999) kindly provided by K. Morikawa, Biomedical Engineering Research Council, Osaka, Japan. All other coordinates are from the PDB. Structural superpositions and RMSD calculated using the program SHP (Stuart *et al.*, 1979).

Figure 6-7 below structural superposition of human gp130 onto human GCSFR, human EpoR, human GHR and human PrlR. The overall two domain alignment with EpoR has an RMSD value of 2.52 Å, with 177 structurally equivalent residues. These relatively poor superposition values are mainly due to differences in the tilt angle, which is defined as the angle between the long axes, running approximately parallel to the β-strands of the two domains. EpoR has a tilt angle of ~90^{o} in common with other class 1 haemopoietin receptors, whereas gp130 has a tilt angle of ~70^{o}. This ~20^{o} difference results in poor

individual domain superposition, but allows analysis of the E-F and B-C loops which are on the N- and C-terminal domains, respectively. These regions are thought to contain the key binding determinants of gp130. Indeed, Phe 93 and Phe 205 of EpoR which are major binding determinants (Middleton *et al.*, 1999) are structurally equivalent to Phe 169 and Val 230 of gp130, respectively. Mutagenesis has revealed the importance of Phe 169, Val 230 of the B-C loop and Gly 286 and Lys 285 of the F-G loop (Horsten *et al.*, 1997) as discussed in section 6.3.4.

The two fibronectin type III-like cytokine binding domains have the same basic structural scaffold in both gp130 and EpoR consisting of a β-sandwich formed from a three-strand (A, B, E) and four-strand (C, C', F, G) β-sheet. The C-terminal domains show good agreement between the two structures with an RMSD value of 1.0 Å for 81 structurally equivalent residues. This good agreement is mainly attributed to the short length of the β-strands (Bravo *et al.*, 1998). However, the N-terminal domains show a poorer agreement within the core framework with an RMSD value of 1.0 Å, for only 53 structurally equivalent residues. This discrepancy is mainly attributed to a shift in the angle of β-strands C, F and G of the C, C', F, G β-sheet (Bravo *et al.*, 1998).

Figure 6-7 below shows the structural superposition of human gp130 onto human GCSFR ((Aritomi *et al.*, 1999). The overall two domain alignment has an RMSD value of 2.52 Å, with 189 structurally equivalent residues and a sequence identity of 29 %. Again these relatively poor superposition values are mainly due to ~25^0 differences in the tilt angle between the N- and C-terminal domains. In gp130, but not GCSFR or EpoR, the interdomain angle is tightly restrained by main-chain hydrogen bonds between the WSXWS motif and the interdomain linker strand (Bravo *et al.*, 1998), so suggesting that the WSXWS motif may play a role in modulating the interdomain orientation of haemopoietin receptors. Mutational analysis has revealed that in contrast to the even split between the -N

and –C terminal domains the major binding determinants of GCSFR reside on the B-C and F-G loops and involve residues Ile 239 and Arg 288, respectively (Layton *et al.*, 1997). This use of residues in the C-terminal B-C and F-G loops at the expense of residues in the N-terminal E-F loop is similar to that observed in gp130, as discussed above. Structurally equivalent residues to Ile 239 and Arg 288 of GCSFR are Lys 281 and Ile 231 of gp130, respectively, and these residues may indeed prove to be important epitopes for ligand binding.

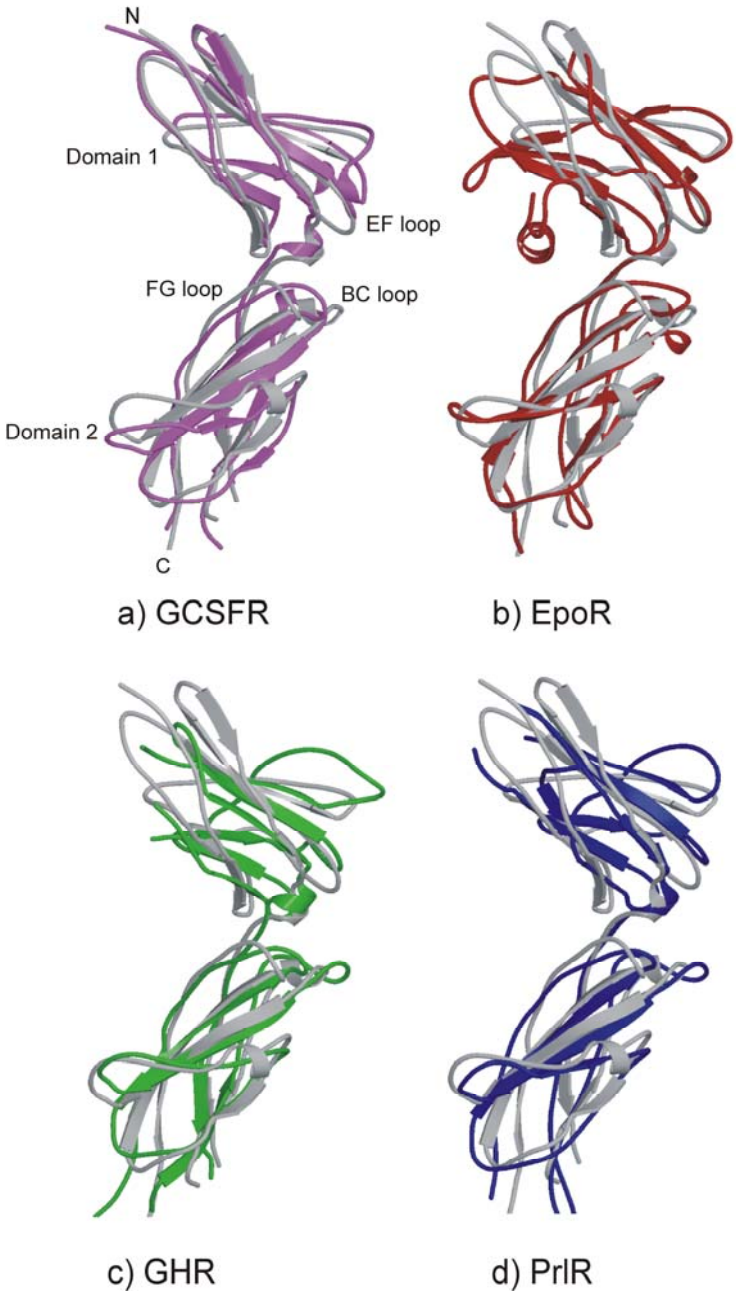


Figure 6-7

Figure showing structural superpositions of human GCSFR (pink; Aritomi *et al.*, 1999), human EpoR (red; Livnah *et al.*, 1996), human GHR (green; de Vos *et al.*, 1992) and human PrlR (blue; Somers *et al.*, 1994) with gp130 (grey). The E-F loop located on domain 1 typically contains a solvent exposed aromatic residue important for ligand binding. Two other potential binding sites are located on the F-G and B-C loops of domain 2. Structural superpositions produced using SHP (Stuart *et al.*, 1979) and figure produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

6.2.3 $P2_12_12_1$ and I222 Structures

Structural superposition as shown in Table 6-1 and Figure 6-8 below reveals that the I222 crystal structure of OSM shows a high degree of structural similarity to the $P2_12_12_1$ crystal structure of OSM. The two molecules display a RMSD in Cα positions of 0.72 Å over 147 structurally equivalent residues. Side chain conformations all lie in close agreement with the exception of residues in the flexible N-terminus preceding Leu 13 and large flexible solvent exposed arginine residues (e.g. Arg 50, 52, 63, 67 and 84) and lysine residues (e.g. Lys 94 and 110) which all show conformational flexibility between the two structures. Three clusters of residues grouped around Glu 89, Lys 110 and Tyr 126, also exhibit conformational flexibility about the χ_1 angle. These clusters are located on helices B and C and residues from each are involved in crystal contacts in one or both of the structures as shown in Figure 6-9 below. It is also interesting to note that this region surrounding the B-C loop is structurally equivalent to site 4 of IL-6, which is postulated to be involved in IL-6 dimerisation as discussed in section 6.3.2.4. The conformational flexibilty of side chains and the formation of crystal contacts in this region suggests OSM may also have a propensity for dimer formation; an observation which is supported by this thesis (see Chapter 3) and other studies (Staunton *et al.*, 1998).

All helices and loops adopt similar positions in the two structures and major deviations between the two all lie in flexible regions at the end of helices and loop regions. Comparison of the B-factor values as shown below in Figure 6-8 shows that the two structures differ markedly in their flexibility. B-factors of the $P2_12_12_1$ structure range from a minimum value of 6.3 Å^2 to a maximum value of 84.9 Å^2, with an average value of 30.5 Å^2. Whereas, the I222 structure displays higher B-factors ranging from a minimum value of 10.2 Å^2 to a maximum value of 84.4 Å^2, with a higher average value of 40.8 Å^2. The higher average B-factor observed in the I222 structure is presumably due to the fact that

data was collected at room temperature (295 K) compared with the $P2_12_12_1$ data that was collected under cryo-conditions (100 K). As discussed previously, higher than average B-factor values at the N- and C-terminus of helix C result in flexibility and hence disordered density for both the C-D and B-C loops of the I222 structure. Other areas of high mobility include the N-terminus preceding Leu 13 and the second helical section of the A-B loop running into helix B.

Figure 6-9 below shows that the crystal contact points for the two structures also differ with eight main contact points in the I222 structure (8 molecules per unit cell) and four main contact points in the $P2_12_12_1$ structure (4 molecules per unit cell). Panel (a) shows crystal contact points in the $P2_12_12_1$ structure with residues shown in red corresponding to the N-terminal loop and the N-terminus of helix B and the C-terminus of helix C. Contacts shown in blue correspond to the second helical region at the C-terminus of the A-B loop and the C-terminus of helix D. Contacts shown in green lie on the N-terminus of helices A and C and the final contact site shown in yellow resides on the B-C loop and the first helical portion at the N-terminus of the A-B loop. Comparison with the crystal contact sites in the I222 structure shown in panel (b) reveals that no particular region is favoured for crystal contact and most exposed regions of loop and helix play a role.

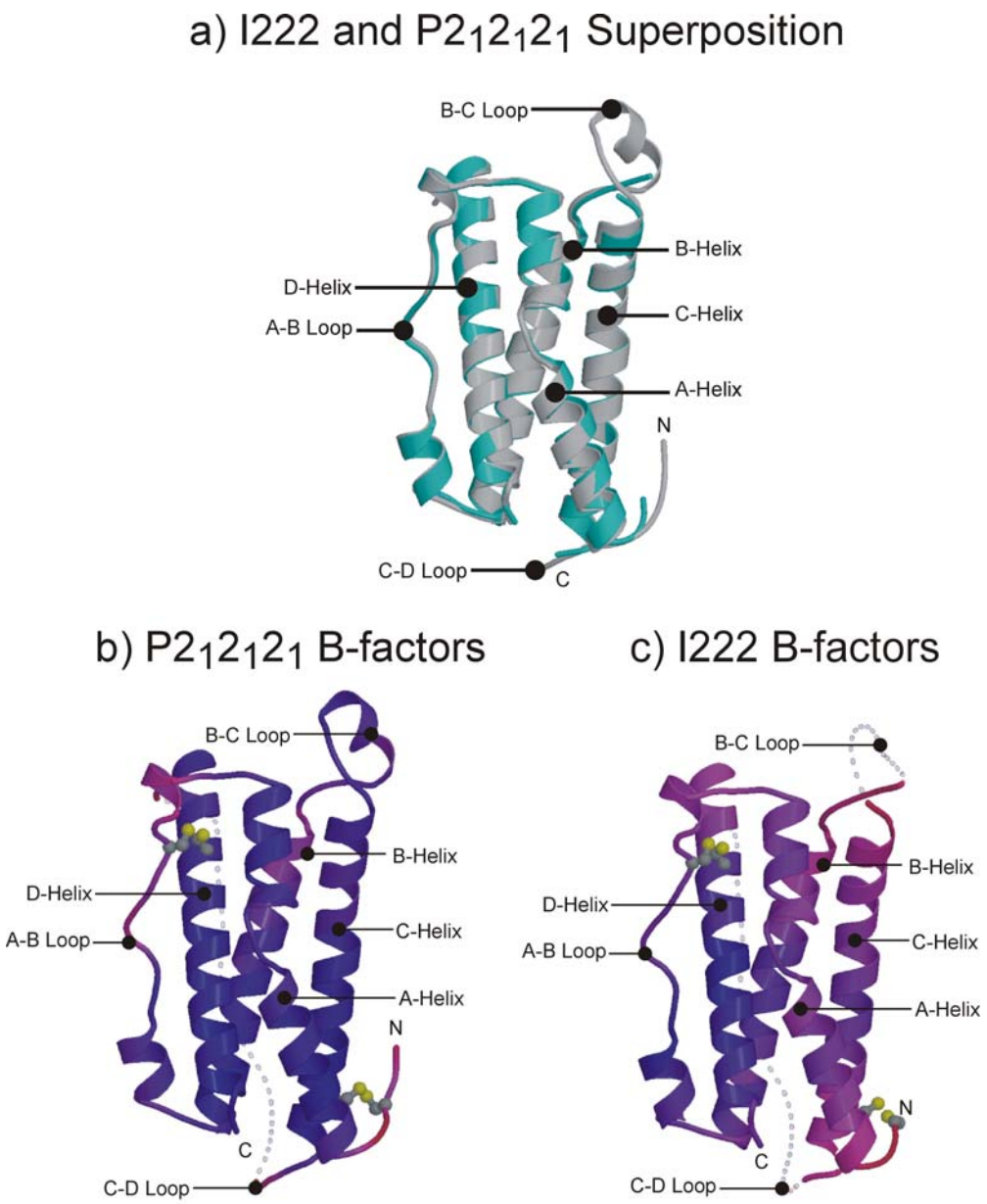


Figure 6-8

Figure showing (a) structural superpositions of $P2_12_12_1$ OSM (grey) and I222 OSM (cyan) and B-factors of (b) $P2_12_12_1$ and (c) I222 structures. Cα positions show 0.72 Å RMSD over 147 structurally equivalent residues. Structural superpositions were produced using SHP (Stuart *et al.*, 1979). Panel (b) and (c) show comparison of refined B-factor values for crystal structures of OSM. In (b) B-factors are coloured from the minimum value of 6.3 Å^2 to the maximum value of 84.9 Å^2 from blue to red, respectively (average B-factor 30.5 Å^2). In (c) B-factors are coloured from the minimum value of 10.2 Å^2 to the maximum value of 84.4 Å^2 from blue to red, respectively (average B-factor 40.8 Å^2). Figure produced BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

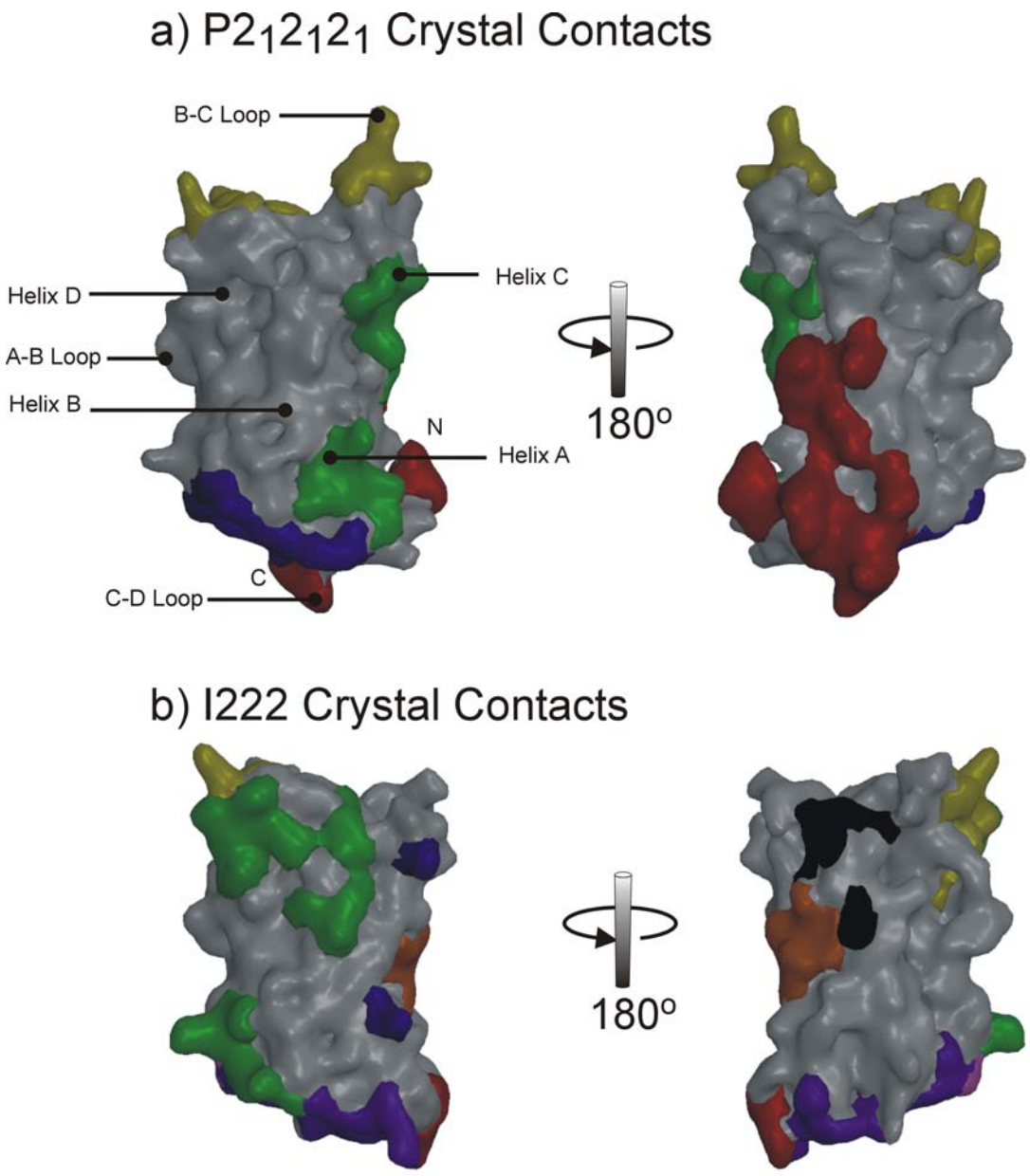


Figure 6-9

Figure showing crystal contact points observed for $P2_12_12_1$ and I222 OSM. OSM molecules are shown with a solvent accessible contoured surface produced using the program VOLUMES (R. Esnouf, Oxford, U.K.). The four crystal contact points for the $P2_12_12_1$ structure are shown as red, blue, green and yellow clusters on the protein surface. The eight crystal contact points for the I222 structure are shown as red, blue, green, yellow, pink, purple, orange and black clusters on the protein surface. Figure produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

6.3 Structure-Function Relationship Comparisons

Considerable mutagenesis work has been carried out on the gp130 family of cytokines that reveals, in common with GH, several discrete sites on the molecules involved in interactions with receptor components. Sites 1 and 2 are in broadly similar locations of the molecule to receptor binding sites on GH, while a third site has been identified that enables an extra interaction not seen in the GH-GHR system. The structure-function relationships of LIF (Hudson *et al.*, 1996; Layton *et al.*, 1994a; Layton *et al.*, 1994b; Owczarek *et al.*, 1993) and CNTF (Di Marco *et al.*, 1996; Inoue *et al.*, 1995; Panayotatos *et al.*, 1995) are the most fully understood due to the availability of crystal structures. A summary of the structural and mutagenesis data for each of these molecules will now be discussed. IL-6 and IL-11 are of less importance to this study, as they do not participate in LIFR binding. However, the crystal structure of IL-6 has been solved and mutagenesis data are available (Simpson *et al.*, 1997b) so comparisons can also be made. Discussion of the structure-function relationship for the Epo and GCSF systems will also be discussed due to the structural similarity with the OSM and gp130 system.

6.3.1 Leukemia Inhibitory Factor

As discussed previously in 6.2.1 human OSM exhibits close structural similarity to human LIF with RMSD in Cα positions of 2.27 Å for 138 structurally equivalent residues with a sequence identity of ~20 %. Through the use of native gel electrophoresis, gel-filtration chromatography, sedimentation equilibrium analysis, surface plasmon resonance spectroscopy and chemical cross-linking it has been shown that LIF forms a heterotrimeric complex with the extracellular portions of LIFR and gp130 in a stoichiometry of 1:1:1 (Zhang *et al.*, 1997b). The complex is stabilised by site 1 and site 3 interactions with LIFR, and this system differs from that of hGH in that site 3 interactions predominate energetically. Site 3 interactions facilitate LIFR binding through a recently identified LIGHT motif in the Ig-like domain of LIFR (Chobotova and Heath, 2000).

6.3.1.1 Site 1 Interactions

Alanine scanning mutagenesis of murine LIF (Hudson *et al.*, 1996) has revealed a site analogous to site 1 of hGH, responsible for LIFR binding, centralised around Val 175 with contributions from Asp 57, Lys 58, Lys 170 and Ala 174. These residues map to the C-terminus of helix D and the A-B loop as shown in blue in Figure 6-10 (a) below.

Structurally equivalent residues in OSM are Arg 46, Glu 47, His 174, His 178 and Ser 179 as shown in Figure 6-10 (b) below, although the function of these residues in OSM is currently unknown and awaiting further mutational analysis. However, these residues, with the exception of His 174 and Ser 179, are solvent exposed in the OSM crystal structure and as such would be expected to form a similar site 1 epitope to that of LIF.

6.3.1.2 Site 3 Interactions

The residues of site 1 contribute weakly to LIFR binding and residues Phe 156 and Lys 159, located on the N-terminus of helix D as shown in red in Figure 6-10 (a) below, contribute the majority of the free energy for binding (Hudson *et al.*, 1996). Alanine mutation of these residues individually reduces binding more than 100-fold and biological activity by upto 3000-fold. Combined alanine mutagenesis eliminates biological activity completely. These two residues are solvent exposed and are conserved in all members of the gp130 family that utilise LIFR; in contrast these residues are not conserved in IL-6 and IL-11 which do not use LIFR (see section 6.4 and Figure 6-12 for sequence alignments). Surrounding Phe 156 and Lys 159 are several other residues that influence LIFR binding; this situation is analogous to that of growth hormone in which key residues are surrounded by ones of lesser importance (Cunningham and Wells, 1989).

Structurally equivalent residues in OSM are Phe 160 and Lys 163, which reside on the N-terminus of helix D as shown in red in Figure 6-10 (b) below. These two residues have been confirmed as the major LIFR binding epitope for OSM with alanine mutagenesis of

either residue resulting in a >100-fold decrease in binding to LIFR and a >10000-fold reduction in biological activity, with no effect on gp130 binding (Deller *et al.*, 2000).

Through analysis of the solution dynamics and secondary structure of murine LIF it has been suggested that the C-D loop of LIF is rigid and shows limited mobility in solution (Purvis and Mabbutt, 1997). In comparison, other cytokines utilising the LIFR unit demonstrate marked mobility in the C-D loop. For example, previous studies have shown that the C-D loop of OSM adopts a long extended conformation in solution (Hoffman *et al.*, 1996) and this present study supports this claim, as considerable disorder and mobility in this region is observed. Also, the crystal structure of CNTF (McDonald *et al.*, 1995) shows considerable flexibility in this region. Although no directly functional residues lie on the C-D loop the site 3 motif lies close on the N-terminus of helix D and as such mobility in this region could affect receptor engagement through some form of entropic component.

Interestingly OSM and CNTF, which both show high levels of flexibility in this region both show low affinity binding to LIFR, compared with LIF. This could also help account for the differential selection of LIFR or OSMR in the OSM signalling system and also the observed anomaly of high affinity binding of hLIF to mLIFR compared to mLIF binding to mLIFR (see 6.5.1).

6.3.1.3 Site 2 Interactions

Site 2 interactions with gp130 were analysed in a similar fashion using multiple simultaneous substitutions due to the weaker interactions involved (Hudson *et al.*, 1996). Residues Gln 25, Ser 28, Gln 32, Ile 121, Gly 124 and Ser 127 have all been shown to participate in gp130 binding. These residues cluster at the ends of helices A and C as shown in yellow in Figure 6-10 (a) below. Non-conservative mutations of these residues result in loss of biological action and act as antagonists of wildtype LIF action.

Structurally equivalent residues in the OSM structure are Gln 16, Lys 19, Leu 23, Asn 117, Gly 120 and Asn 123 as shown in yellow in Figure 6-10 (b) below. The importance of Gln 16, Gly 120 and Asn 123 in the formation of the signalling complex of LIFR has been confirmed through mutagenesis, which reveals that mutation of all of these residues results in an 8 to 100-fold reduction in binding to gp130 and a corresponding reduction in biological activity, with no effect on LIFR binding (Deller *et al.*, 2000).

6.3.2 Interleukin-6

In common with LIF, mutagenesis studies of IL-6 (Simpson *et al.*, 1997a) also suggest a three site receptor binding model. However, in contrast to LIF these three sites function in binding three separate receptor components.

It has been shown that sites 2 and 3 interact independently with gp130 only in the presence of IL-6R (Paonessa *et al.*, 1995). Immunoprecipitation (Paonessa *et al.*, 1995) studies have suggested that the complex contains two copies of each component and a hexameric complex with a stoichiometry of 2:2:2 has been confirmed by analytical ultracentrifugation and size exclusion chromatography (Ward *et al.*, 1994).

6.3.2.1 Site 1 Interactions

This site forms a cleft on the surface of IL-6 and all mutations that affect binding to IL-6R map to this surface, which is composed of residues in helix D and the A-B loop as shown in blue in Figure 6-10 (c) below. Residues of key importance include Arg 179, which results in 100-fold decrease in activity (Fontaine *et al.*, 1993), and Gln 175, which results in 5-fold decrease in activity (Savino *et al.*, 1993). Other solvent exposed residues implicated in binding IL-6R include Ser 176 (Savino *et al.*, 1993), Arg 182 (Lutticken *et al.*, 1991) and Phe 74.

As for the other cytokines, site 1 in IL-6 is thus implicated as a binding site for a specificity conferring receptor in the signalling complex. This region in OSM corresponds to residues Arg 182, His 178, Ser 179, Ser 185 and Arg 63 as shown in blue in Figure 6-10 (d) below. The function of these residues is currently unknown although all these residue, except Ser 179 are solvent exposed and are likely to form a site 1 epitope, which by analogy with LIF might be expected to interact with LIFR or OSMR.

6.3.2.2 Site 2 Interactions

As in LIF site 2 interacts independently with a gp130 receptor and consists of residues in helix A and helix C as shown in yellow in Figure 6-10 (c) below. Important residues in this location include Tyr 31, Ser 118 and Val 121, which are all solvent exposed.

Structurally equivalent residues in the OSM structure are Leu 23, Asn 117 and Gly 120 as shown in yellow in Figure 6-10 (d) below. All of these residues have been shown to be important in varying degrees for OSM binding to gp130 as detailed in Table 6-3.

6.3.2.3 Site 3 Interactions

In analogy with LIF site 3 consists of residues in helix D and the C-D loop; however in IL-6 both sites 2 and 3 interact independently with separate gp130 receptors. Residues important for site 3 binding to the second gp130 receptor, include the solvent exposed residues Trp 157 and Asp 160 (Paonessa *et al.*, 1995) as shown in red in Figure 6-10 (c) below.

Structurally equivalent residues in the OSM structure are Phe 160 and Lys 163 as shown in red in Figure 6-10 (d) below, and both of these residues have been shown to form the major determinant of OSM binding to LIFR as detailed in Table 6-3 below.

6.3.2.4 Site 4 Interactions

On the basis of the crystal structure of IL-6 (Somers *et al.*, 1997) it has also been predicted that a fourth site on the IL-6 exists that mediates dimer formation between two IL-6 molecules. It has been proposed that the region Glu 106 to Arg 113, shown in pink in Figure 6-10 (b) below, would interact with the same residues on an adjacent molecule, across a 2-fold axis of rotation, so stabilizing the hexameric signaling complex.

This region maps to the structurally equivalent residues Pro 93 to Gln 112 of the OSM structure, which reside on the B-C loop and N-terminal portion of helix C, as shown in pink in Figure 6-10 (b) below. As discussed previously several residues in this region, clustered around Glu 89, Lys 110 and Tyr 126, exhibit conformational flexibility between the I222 and $P2_12_12_1$ crystal structures of OSM, due to their involvement in crystal contacts in one or both of the structures. It has been reported that OSM forms low levels of dimers and trimers in solution (Staunton *et al.*, 1998) and this present study also supports this claim as detailed in Chapter 3. Further support is given by the pH dependency of the dimerisation with 0.2 % of total protein being dimerised at pH 6.0 and 20 % at pH 7.5 (Staunton *et al.*, 1998), this dependency could possibly be due to the charged character of Glu 89 and Lys 110 in the interface. This interface would provide a smaller surface area for burial than that observed in the CNTF dimer (~1038 Å) (McDonald *et al.*, 1995) and as such would explain the low levels of dimerisation observed in OSM. However, CNTF does use similar residues in the dimer interface along the face of helix C, so suggesting a common theme may prevail. However, the stoichiometry of the OSM-gp130-LIFR complex is believed to be 1:1:2 (Zhang *et al.*, 1997b) and there is no evidence to suggest any direct interaction between the OSM molecules of the assembly (Staunton *et al.*, 1998). Moreover, both IL-6 (May *et al.*, 1991) and GH (Cunningham *et al.*, 1991a) have also been reported to exist as oligomers and such complexes may be important for storage and transport of potentially toxic levels of cytokine.

6.3.3 Ciliary Neurotrophic Factor

CNTF requires three separate receptor components for signalling; CNTFR, LIFR and gp130 (Davis *et al.*, 1993). In common with LIF and IL-6, as discussed above, three sites on the ligand have been identified as important for receptor binding. CNTF, CNTFR, gp130 and LIFR have been shown to form a hexameric complex in the stoichiometry of 2:2:1:1 (De Serio *et al.*, 1995).

Using a series of cell survival assays (Inoue *et al.*, 1997; Panayotatos *et al.*, 1994) measuring specific responses to both α-receptor (CNTFR) and β-receptor (LIFR and gp130) units the three receptor binding sites of CNTF have been delineated in a similar fashion to LIF and IL-6 as discussed above.

6.3.3.1 Site 1 Interactions

In agreement with the theme introduced in LIF and IL-6, site 1 functions in binding CNTFR; a specificity conferring receptor element. Interactions are believed to involve residues Arg 25 and Arg 28 which are located on helix A, residues Gln 63 and Trp 64 which are located on the A-B loop and, residue Gln 74 located on helix B and residue Arg 177 located on helix D (Panayotatos *et al.*, 1995). These residues are all solvent exposed as shown in blue in Figure 6-10 (e) below.

Structurally equivalent residues in the OSM structure include Asp 22, Thr 27, Asn 75 and Arg 182 as shown in blue in Figure 6-10 (b) below. It is interesting to note that two of these residues are positively charged and given the previous observation (see Figure 6-5 (f)) that helix D shows an unusually large clustering of around seven positive charges, it is possible that OSM-LIFR interactions may involve an electrostatic component for modulation of binding. A similar clustering is observed in site 1 of CNTF and it has been postulated that electrostatic interactions are involved in CNTFR recruitment (McDonald *et*

al., 1995). LIF does not possess such a prominent clustering of charged residues, so suggesting that the high affinity site 3 interaction prevails.

Such clustering of electrostatic potential has previously been reported to play a role in cytokine ligand-receptor interactions; for example the use of three lysine residues in the NGF-p75NGFR interaction (Ibanez *et al.*, 1992). Similarly, studies of the electrostatics of GH-GHR contact residues has revealed that mutation of Arg residues most affect association and together contribute a factor of about 20 to the on-rate (Cunningham and Wells, 1993). These data suggests that the hormone and receptor associate by diffusion and electrostatics to form an ensemble of weak collisional complexes, and the structure of OSM suggests a similar mechanism may be utilised.

6.3.3.2 Site 2 Interactions

In continuation of the themes observed in LIF site 2 residues interact with gp130 and this region has been partially mapped to the solvent exposed residues Lys 26 and Asp 30 (Panayotatos *et al.*, 1995) which both reside on helix A as shown in yellow in Figure 6-10 (e) below.

Structurally equivalent residues in the structure of OSM include Leu 23 and Arg 29 as shown in Figure 6-10 (f) below. Analysis of Table 6-3 below reveals that mutation of Leu 23 has a small effect on both gp130 and LIFR binding although the role of the solvent exposed residue Arg 29 is unknown.

6.3.3.3 Site 3 Interactions

Site 3 in common with LIF and IL-6 consists of a conserved pair of residues, Phe 152 and Lys 155 (Inoue *et al.*, 1997), which have been shown to facilitate binding to LIFR without affecting binding to CNTFR or gp130 (Di Marco *et al.*, 1996). Both of these residues are solvent exposed and lie on the N-terminus of helix D, as shown in red in Figure 6-10 (e) below.

From sequence alignments as shown in Figure 6-12 below it is clear that Phe 152 and Lys 155 correspond to the functionally equivalent Phe 160 and Lys 163 pair of OSM and Phe 156 and Lys 159 of LIF. However, these pairs do not correspond structurally between the two molecules due to the lengthened helix D of OSM, which results in a shift of 3 residues between the pairs as shown in red in Figure 6-10 (f) below.

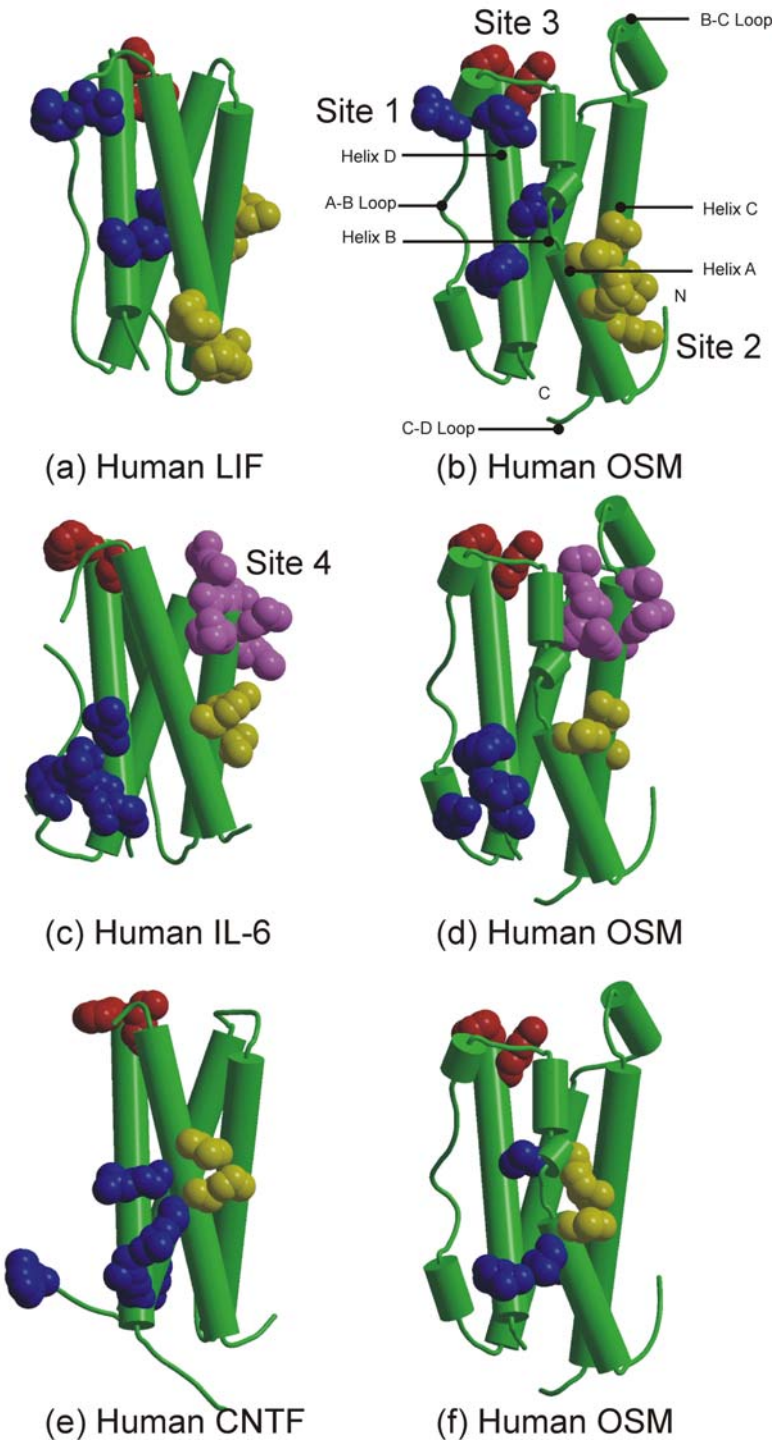


Figure 6-10

Figure showing summary of residues important for receptor binding in (a) LIF, (c) IL-6 and (e) CNTF. Structurally equivalent residues in OSM are shown in panels (b), (d) and (f). Residues involved in site 1 (blue) and site 3 (red) interactions with LIFR are shown along with residues involved in site 2 (yellow) interactions with gp130. Residues represent those deemed important from alanine scanning mutagenesis and the structurally equivalent residues in OSM were analysed using the program SHP (Stuart *et al.*, 1979). Figure produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

6.3.4 Erythropoietin

The Epo-EpoR structure (Syed *et al.*, 1998) as shown in Figure 6-11 (a) below shows considerable homology to the GH-GHR system as discussed in Chapter 1 and also to the OSM and gp130 system as assessed by the RMSD values as shown in Table 6-1 and Table 6-2. Given the wealth of mutagenesis and structural data available, this system provides an ideal model for understanding OSM-gp130 interactions. However, the wider picture involving LIFR interactions will require structural studies of complexes involving the site 3 epitope; a feature that is unique to the LIFR family of ligands.

6.3.4.1 Site 1 Interactions

Site 1 forms a central hydrophobic pocket involving Epo residues from helices A, B, D and part of the A-B loop. Epo residues Thr 44 and Asn 147 function in hydrogen bonding to Phe 93 of EpoR, which is the key binding determinant on the receptor. A hydrophobic quartet of Epo residues including Gly 151, Val 46, Phe 48 and Leu 155 form the centre of site 1 with the phenyl ring of the EpoR residue Phe 93 packing against the Cα of Epo residue Gly 151. The site 1 interface also has contributions from Epo residues Thr 44 to Phe 48 of the A-B loop with Phe 48 being in van der Waals contact with a minor determinant Met 150 of EpoR.

OSM residues structurally equivalent to the site 1 hydrophobic pocket in Epo include Ser 179, Ser 58, Glu 60 and Arg 182, so suggesting that the hydrophobic nature of this motif is not conserved between the two. This is further borne out through comparison with LIF mutagenesis data (discussed in section 6.3.1.1) which maps site 1 to residues Arg 46, Glu 47, His 174, His 178 and Ser 179, so suggesting that this site is more hydrophilic in character than observed in Epo. This is not unexpected, as the major determinant of OSM interaction with LIFR is the Phe 160 and Lys 163 pair on the ligand, as identified through alanine mutagenesis (Deller *et al.*, 2000). This site 3 motif provides most of the binding

energy with site 1 playing only a minor role. Therefore, it could be predicted that the site 3 aromatic of the Phe/Lys pair functions in binding to a hydrophobic patch on the LIFR, in a reverse situation to that of the site 1 Epo-EpoR interaction (Middleton *et al.*, 1999; Syed *et al.*, 1998). Alternatively, the basic partner of the Phe/Lys pair could function in salt bridge formation with equivalents on the LIFR. This situation is observed in the crystal structure of IL-4-IL4R complex (Hage *et al.*, 1999) in which Arg 88 of the ligand and Asp 72 of the receptor form a salt bridge with the bonding potential being enhanced by the surrounding hydrophobic environment.

Equivalents to the Phe/Lys pair are not totally conserved outside the LIFR binding family of cytokines (see Figure 6-12). For example, IL-6 has a structurally and functionally equivalent Trp 157 and Asp 160 pair (Paonessa *et al.*, 1995) and no such structural equivalents are observed for Epo. This is as would be expected, as the major binding determinant is the hydrophobic quartet of site 1 and site 3 interactions are redundant. GCSF shows a partially conserved Phe 144 and Lys 147 pair, so again suggesting homology to the gp130-LIFR family of ligands.

6.3.4.2 Site 2 Interactions

Site 2 involves Epo residues located in the A and C helices although the interactions are less extensive than site 1 and most of the interactions involve residues of the C-helix. In a similar fashion to site 1 a hydrophobic surface is produced by Epo residues Leu 5, Val 11, Tyr 15, Ser 104, Thr 107 and Leu 108, which are all within 4.5 Å of the major binding determinant of EpoR, residue Phe 93 (Middleton *et al.*, 1999; Syed *et al.*, 1998). Epo residues Arg 10, Val 11 and Arg 14 are in van der Waals contact with the minor binding determinant Met 150 of EpoR. The EpoR residue Phe 93 is structurally and functionally equivalent to the gp130 residue Phe 169 located on the E-F loop of domain 1, whereas,

the EpoR residues Met 150 and Phe 205, which have also been identified as essential binding determinants (Middleton *et al.*, 1999) have no structural equivalents in gp130.

Structurally equivalent OSM residues to the site 2 hydrophobic pocket in Epo include Leu 13, Gln 20, Met 24, Gly 120, Asn 123 and Asn 124, so suggesting that this motif is only partially conserved in hydrophobic character between the two molecules. Indeed, the side chains of the two main hydrophobic residues Leu 13 and Met 24 are buried in the OSM structure with relative solvent accessibility's of 2 and 3 %, respectively. Equivalent residues in Epo are Leu 5 and Tyr 15 with relative solvent accessibility's of 60 and 11 %, respectively. However, of these six residues alanine mutation of Gln 20, Gly 120, Asn 123 and Asn 124 has been shown to display upto 100-fold decrease on gp130 binding with little or no effect on LIFR binding (Deller *et al.*, 2000). In addition a fifth member of the OSM site 2 motif has been identified as Gln 16. This residue is structurally equivalent to the Epo residue Asp 8 which hydrogen bonds to His 153 of EpoR, which is one of the 11 hydrogen bonds in the site 2 interface.

Therefore, it can be concluded that of the 5 residues identified as the site 2 motif of OSM through mutagenesis studies, all are structurally equivalent to residues directly involved in the Epo-EpoR interface. OSM residues Gln 20, Gly 120, Asn 123 and Asn 124 are all equivalent to the series of residues in Epo that pack against the EpoR residue Phe 93. The fifth member of OSM site 2 is Gln 16, which is equivalent to an Epo residue involved in one of the hydrogen bonds in the receptor-ligand interface. The EpoR residue Phe 93 is in turn structurally and functionally equivalent to the gp130 residue Phe 169 which resides on the E-F loop of the N-terminal domain.

The two remaining major determinants in EpoR, Met 150 and Phe 205 (Middleton *et al.*, 1999) have no structural equivalents although mutation of Val 230 in the B-C loop has

been shown to affect binding in the IL-6 system (Horsten *et al.*, 1997). However, this residue is relatively buried so suggesting indirect effects on ligand binding. Searches for other regions of binding activity in gp130 have revealed that the solvent exposed residues Gly 286 and Lys 285 of the F-G loop also function in binding IL-6 (Horsten *et al.*, 1997). The F-G loop region has also been confirmed as a binding epitope in the GCSF system as discussed below (Yamasaki *et al.*, 1997). Thus, it appears that the gp130 system has diverged from the Epo related systems in that only one aromatic residue is involved in binding compared with two in Epo. This may be to allow greater conformational flexibility so allowing formation of higher order receptor-ligand signalling complexes such as the hexameric assemblies observed for IL-6 (Ward *et al.*, 1996) and CNTF (De Serio *et al.*, 1995), the tetrameric assemblies of OSM (Staunton *et al.*, 1998) (see Figure 6-11) and the trimeric assemblies of LIF (Zhang *et al.*, 1997b).

6.3.5 Granulocyte Colony Stimulating Factor

The GCSF-GCSFR (Aritomi *et al.*, 1999; Hill *et al.*, 1993); GCSF-GCSFR coordinates kindly provided by K. Morikawa, Biomedical Engineering Research Council, Osaka, Japan.)(see Figure 6-11 (b)) system provides an attractive model for comparison with the OSM and gp130 system for several reasons. Firstly, the ligand and receptor components of this system show high levels of structural similarity to OSM and gp130, respectively, as assessed by the RMSD values as shown in Table 6-1 and Table 6-2. Particularly, as discussed in section 6.2.1, the positioning and orientation of the A-B loop in GCSF more closely matches that of OSM than the Epo structure. Secondly, the Phe/Lys pair as utilised by the LIFR binding cytokines is conserved, as shown in Figure 6-12 below, and these two residues have also been shown to be critically important for binding GCSFR (Reidhaar-Olson *et al.*, 1996). Thirdly, mutations in GCSFR have revealed that residues in the B-C and F-G loops of the C-terminal domain are critical for ligand binding, in a similar fashion to gp130 and in contrast to the use of the B-C and E-F loops in Epo. Finally, as seen with

gp130, GCSF has a propensity for forming higher order complexes with stoichiometries of 1:1, 2:2 and 4:4 having been reported depending on the experiment performed (Hiraoka *et al.*, 1995; Hiraoka *et al.*, 1994; Horan *et al.*, 1996; Horan *et al.*, 1998).

However, one major draw back of the comparison is the lack of site 1 and site 3 contacts in the current crystal structure of the GCSF-GCSFR complex (Aritomi *et al.*, 1999) as shown in Figure 6-11 below. The two contact regions are described in terms of a major site, which is equivalent to site 2, and a minor site that displays novel interactions. Therefore, this system can only be used for modelling of site 2 interactions with gp130.

6.3.5.1 Major Interface

The current model of GCSF-GCSFR interaction involves a 2:2 complex consisting of two 1:1 complexes of GCSF and GCSFR related by a non-crystallographic pseudo-twofold axis. The 1:1 complex is formed by binding of the hinge region of the receptor to an epitope equivalent to site 2 on the ligand. The core of this interaction involves hydrogen bonding of Glu 19 of GCSF to Arg 288 of GCSFR; both of these residues have previously shown to be important through mutagenesis (Layton *et al.*, 1997; Layton *et al.*, 1999). Lys 19 is the structurally equivalent residue in OSM and although this residue is ~80 % solvent exposed alanine mutation showed no effect on either gp130 or LIFR binding. However, mutation of adjacent residues Gln 16 and Gln 20, which show much lower solvent accessibility's, affect gp130 binding by ~75-fold or greater as detailed in Table 6-3 below. In contrast to the Epo-EpoR interface which is predominantly hydrophobic in character the GCSF-GCSFR major interface displays an even spread of hydrophobic and polar interactions in a similar fashion to the mosaic interface observed in the IL-4-IL-4R complex (Hage *et al.*, 1998). This again compounds the similarity to the OSM system with the present mutagenesis data suggesting that this interface maps to a surface of mixed character.

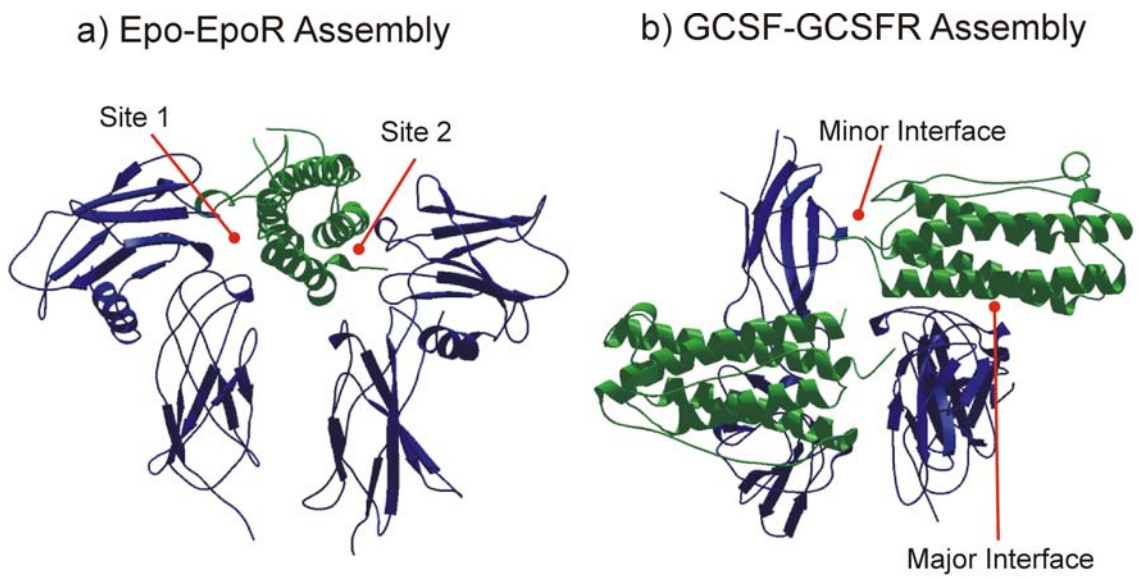
Mutational analysis has revealed that in contrast to EpoR the major interface determinants of GCSFR reside on the B-C and F-G loops of the C-terminal domain and involve residues Ile 239 and Arg 288, respectively (Layton *et al.*, 1997). Structurally equivalent residues to Ile 239 and Arg 288 of GCSFR are Ile 231 and Lys 285 of gp130, respectively, which demonstrates conservation between the two receptors, although the role of these two residues in gp130 is currently unknown. As discussed previously mutagenesis of Val 230 and Lys 285 resulted in loss of affinity between gp130 and IL-6 (Horsten *et al.*, 1997) so impying that activity resides around these residues and so suggesting a similar mode of receptor engagement between the two systems.

6.3.5.2 Minor Interface

The minor interface formed between GCSF and GCSFR is not equivalent to interactions observed in the Epo or GH system as it involves residues on the face of the C-terminal domain of GCSFR and residues at the N-terminal end of the helical bundle of GCSF; namely Ser 7 and Gln 11 of GCSF and Phe 259 and His 260 of GCSFR.

Therefore, the unusual minor site interaction of GCSF with its receptor may provide a useful basis for understanding non-standard epitopes, such as the site 3 motif of the LIFR binding family of cytokines. For example, the Ig-like domain of gp130 has been shown to bind OSM through site 3, so suggesting the formation of a tetrameric assembly of two OSM, one LIFR and one gp130 molecule as shown in Figure 6-11.

A comparison of the themes observed in the Epo, GCSF and OSM signalling systems is given below in Figure 6-11.



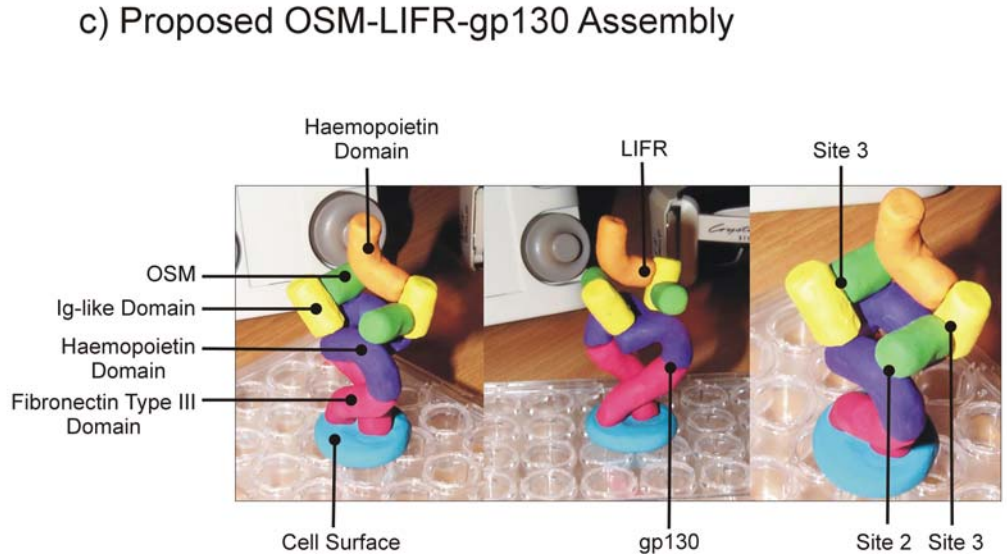


Figure 6-11

Figure showing ribbon representation of (a) Epo-EpoR (Syed *et al.*, 1998) and (b) GCSF-GCSFR (Aritomi *et al.*, 1999)(coordinates kindly provided by K. Morikawa, Biomedical Engineering Research Council, Osaka, Japan.) crystal structures. In each case the haemopoietin receptor domains are coloured in blue and the cytokines are coloured in green. Panel (c) shows Plasticene model of proposed OSM-LIFR-gp130 complex (Staunton *et al.*, 1998) observed interactions shown in hypothetical orientation. Figure produced using BOBSCRIPT (Esnouf, 1997), RASTER3D (Merritt and Murphy, 1994), Plasticene (Toys R Us) and Anglepoise lighting.

6.4 Sequence Comparisons

Despite the common framework amongst the helical cytokines essentially no sequence identity is observed across the family as a whole. However, core residues are conserved within the gp130 family of cytokines and this gives rise to low but significant levels of sequence identity and conservation as shown in Figure 6-12 below. Typically, members of the gp130 family show ~20 % pairwise identity or conservative substitution. Approximately 50 % of these conserved residues are inaccessible to solvent, and as such contribute to the molecular core so maintaining the protein fold and stability. The remaining residues are solvent exposed and as such represent potential sites of receptor engagement.

Figure 6-12 also demonstrates the conserved nature of the Phe/Lys (residues 160 and 163, respectively in OSM) pair in the LIFR binding family of cytokines. This pair is invariant over all species of OSM and is conserved between OSM, LIF and CNTF. These residues are not observed in IL-6 and are partially conserved in both Epo and GCSF, which contain a Phe/Leu and Phe/Arg pair, respectively. This suggest that GCSF and Epo have diverged from the LIFR binding family of cytokines, as they place less importance on the use of a high affinity site 3 epitope and rely solely on a classic hydrophobic site 1 and 2 interaction as first observed in the growth hormone system.

Given the importance of these residues in LIFR engagement this level of conservation again supports an evolutionary link between the molecules and together with the similar intron/exon boundaries that are observed for these cytokines suggests an homologous relationship (Bazan, 1991).

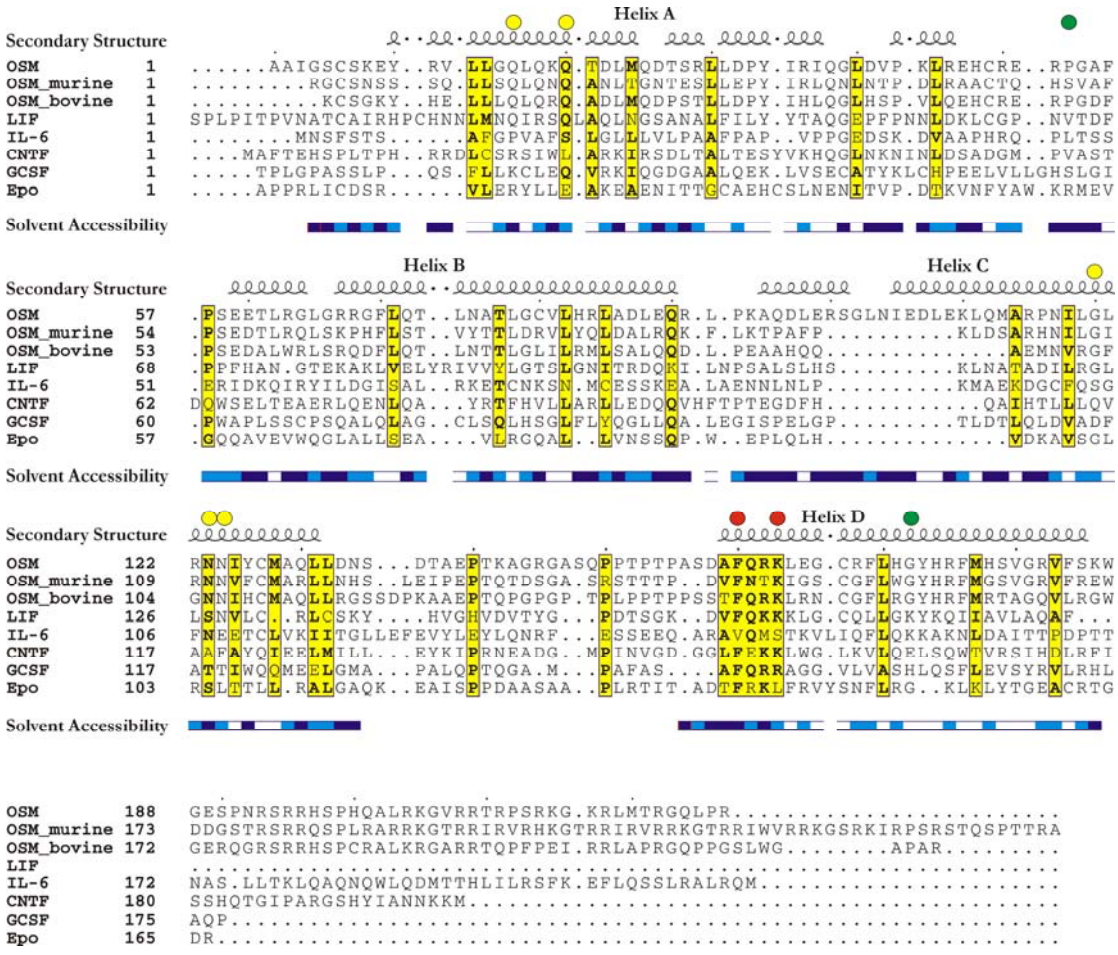


Figure 6-12

Figure showing sequence alignments of the gp130 family of cytokines and GCSF and Epo. Coils represent secondary structure assignment of OSM and blue blocks represent solvent accessibility ranging from white (~0%) through light blue (~50%) to dark blue (~100 %). Yellow circles represent site 2 residues implicated in gp130 binding and red circles represent site 3 residues implicated in LIFR binding (Deller *et al.*, 2000). Green circles represent residues structurally equivalent to residues implicated in modulating species specificity in LIF (Robinson *et al.*, 2000). Figure produced using DSSP (Kabsch and Sander, 1983) and ESPript (Gouet *et al.*, 1999).

6.5 Implications for Receptor Binding

6.5.1 Implications for Species Specificity

As discussed previously both murine LIF and human LIF interact with human LIFR, whereas, murine LIFR interacts only with murine LIF with appreciable affinity (Layton *et al.*, 1994a; Layton *et al.*, 1994b; Owczarek *et al.*, 1993). From analysis of murine-human chimeras the region responsible has been localised to residues 99-180 (Owczarek *et al.*, 1993). From analysis of the crystal structures of murine LIF (Robinson *et al.*, 1994) and human LIF (Robinson *et al.*, 2000) it has been proposed that a change in surface character of helix D, which is ideally placed to interact with LIFR, is responsible for the species anomaly.

The surface volume occupied by residue Met 64 in murine LIF is displaced by residue Lys 168 in human LIF, due to perturbation of the A-B loop (Robinson *et al.*, 2000). It is proposed that this change in surface character from hydrophobic to charged modulates the specificity of LIF towards LIFR. Although the species specificity of OSM has not been fully investigated it is plausible that a change in the surface character of helix D could be involved given the high number of charged residues residing on the face of this helix. This may prove to be another example of the importance of electrostatic interactions in ligand-receptor docking as discussed in 6.3.3.1.

Analysis of the OSM structure reveals that structural equivalents to murine LIF Met 64 and human LIF Lys 168 are human OSM residues Pro 53 and Gly 172, respectively. This precludes these residues from involvement in such a species 'switch' mechanism in OSM as glycine has no surface character and it is unlikely that a proline side-chain will be displaced due to its rigid nature. Sequence alignments reveal that Gly 172 is conserved across all species of OSM, whereas Pro 53 is replaced by serine in murine OSM as identified by green circles in Figure 6-12.

Other local or long-range conformational changes in this region are not precluded and further crystallographic studies of other OSM and LIF species variants will be required to address the problem.

6.5.2 Summary of Mutagenesis Data

Several residues have been implicated to play in a role in site 3 and site 2 interactions through alanine mutagenesis (Deller *et al.*, 2000) as shown below in Figure 6-13 and Table 6-3. All active residues have side chains that are solvent accessible to 27 % or greater. Residues Phe 160 and Lys 163 form the main interaction site for LIFR and both reside at the N-terminus of helix D. Site 2 interactions with gp130 are dominated by residues on the A and C helices, towards the N-terminal loop region.

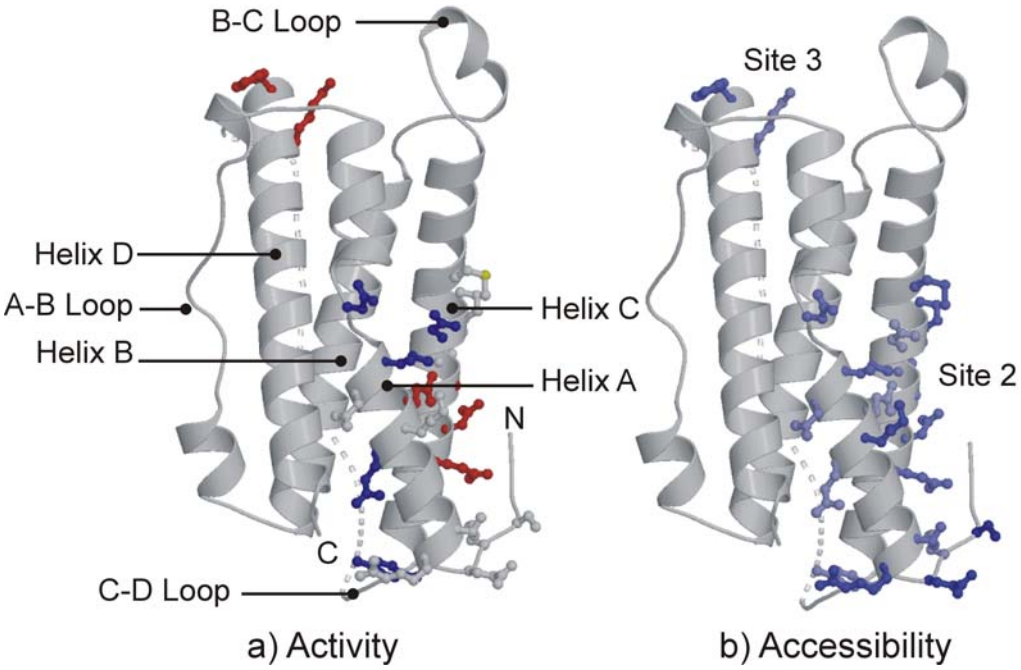


Figure 6-13

Figure showing side chains identified as major site 2 and site 3 epitopes in OSM (Deller *et al.*, 2000). Panel (a) represents activity in gp130 binding assays and biological assays for each mutated side chain: red represents >75-fold increase in activity; blue represents <75-fold activity and grey represents no activity. Panel (b) shows relative solvent accessibility's for each mutated side chain ranging from 0% (grey) through to 100 % (blue) calculated using the program NACCESS (Hubbard and Thornton, 1993). Figure produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

Mutant	Site	Location	RSA (%)	Receptor Binding Assays		Ba/F3 Cell Survival Assays		Conclusion
				gp130	LIFR	LIFR/ gp130	OSMR/ gp130	
Wildtype	—	—	—	1.0	1.0	1.0	1.0	—
F160A	3	N-terminus Helix D	65	1.1	**>100**	**>10000**	**>10000**	**Active towards LIFR only**
K163A			42	0.9	**>100**	**>100**	**>10000**	
F160A/ K163A			—	1.0	**>100**	**>10000**	**>10000**	
S7A	2	N-terminal Loop	100	1.1	0.9	0.8	0.5	No activity
K8A			8	0.8	1.1	0.8	0.5	
E9A			96	1.2	0.8	1.4	1.2	
Y10A		N-terminus Helix A	35	3.4	7.2	2.1	0.9	Limited activity towards LIFR and gp130
R11A			74	1.0	1.9	0.8	0.5	No activity
V12A			55	1.0	1.2	0.7	0.7	
Q16A			56	**75.0**	1.0	**100.0**	**125.0**	**Active towards gp130 only**
Q18A			36	4.0	5.0	1.2	1.1	Limited activity towards LIFR and gp130
K19A			79	1.0	2.0	1.4	2.0	No activity
Q20A			27	**90.0**	1.0	**20.0**	**30.0**	**Active towards gp130 only**
D22A			40	1.0	1.0	0.8	1.0	No activity
L23A			61	3.0	6.7	2.5	5.0	Limited activity towards LIFR and gp130
D26A			63	4.0	5.0	3.0	4.1	
M113A		Helix C	66	0.7	1.0	1.0	1.0	No activity
P116A			52	0.7	2.3	1.0	0.9	
N117A			29	2.0	1.0	2.6	8.0	Limited activity towards gp130 only
L119A			46	1.7	2.2	1.2	1.7	No activity
G120A			76	**>100**	1.8	**60.0**	**60.0**	**Active towards gp130 only**
G120Y				**>100**	1.8	**>10000**	**>10000**	
N123A			61	**8.0**	1.0	**3.3**	**10.0**	
N124A			31	**>100**	1.0	**>10000**	**>10000**	

Table 6-3

Table showing summary of mutagenesis data of OSM. Mutagenesis studies carried out by K. Hudson, CRC Growth Factors Group. Birmingham. All mutants were created by PCR overlap using a pGEX OSM plasmid as detailed in chapter 2. Relative solvent accessibilities (RSA) were calculated using NACCESS (Hubbard and Thornton, 1993) from the OSM coordinates. Receptor binding assays were carried out as detailed in Chapter 3 and 4 with values quoted representing mutant IC_{50}/wildtype IC_{50} ratios. Ba/F3 cell survival assays were carried out as detailed in Chapter 3 with values quoted representing mutant EC_{50}/wildtype EC_{50} ratios. Fields highlighted in bold represent important results.

6.5.3 Proposed Model of Receptor-Ligand Interaction

On the basis of structural superposition of OSM and gp130 onto the Epo-EpoR structure as discussed above and shown in Figure 6-14 below it can be seen that all residues implicated in site 2 receptor binding through mutagenesis (Deller *et al.*, 2000; Horsten *et al.*, 1997) are indeed predicted to map to the interface region. Analysis of Table 6-3 shows that alanine mutation of Gly 120 has a major effect on gp130 binding with no effect on LIFR binding and limited effect on biological activity, however, mutation to tyrosine results in both major disruption of receptor binding and biological activity. Inspection of the proposed model in Figure 6-14 below, suggests that Gly 120 is central to the site 2 epitope and as such mutation to tyrosine would result in steric hinderance between Phe 169 and/or components of the B-C loop, whereas introduction of alanine would limit such a clash. Such a scenario has also been observed in GH in which mutation of the structurally equivalent Gly 120 results in a steric block on the formation of a 1:2 ligand-receptor complex (Clackson *et al.*, 1998).

Other important residues residing on helix C of OSM are Asn 124 and Asn 123, with mutation of Asn 124 showing the greatest effect on receptor binding and biological activity. Analysis of the proposed model suggests that Asn 124 is directed towards Phe 169 and the B-C loop of the receptor, whereas Asn 123 has a higher solvent accessibility being placed on the outer face of helix A, so showing less interaction towards this region of the receptor. The only residues implicated in receptor engagement on helix A are Gln 16 and Gln 20. It is interesting to note that in the long stretch of residues mutated on this helix (i.e. residues 7-26) only this pair showed activity. Their spacing of one helical turn suggests that both residues face in a similar fashion towards the major epitope on the receptor. Clearly this model fails to account for all data, including site 1 and site 3 interactions, and further crystallographic studies involving LIFR and complexes thereof will be required.

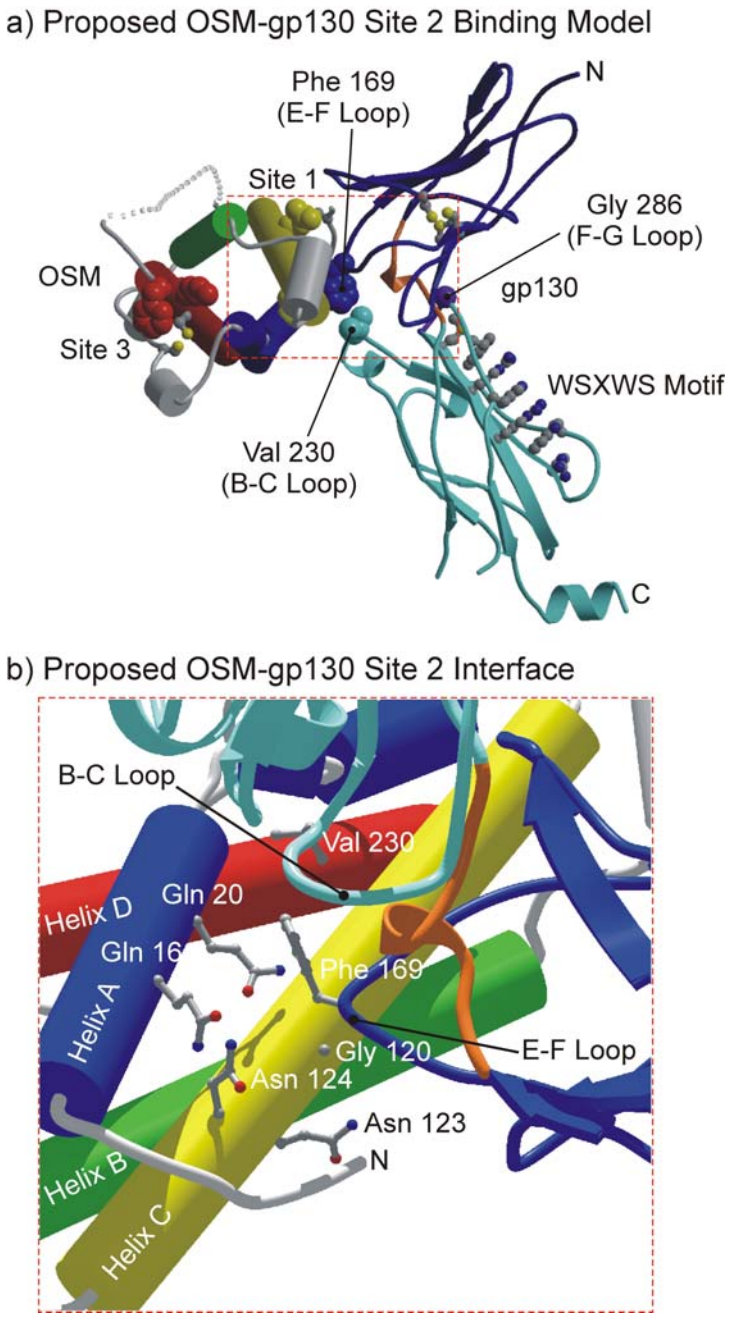


Figure 6-14

Figure showing (a) proposed OSM-gp130 site 2 binding model and (b) proposed OSM-gp130 site 2 interface, based on Epo-EpoR structure (Syed *et al.*, 1998). Panel (a) shows gp130 with the N-terminal domain coloured dark blue and the C-terminal domain coloured light blue. The WSXWS motif is shown along with the interdomain linker region in orange and the F-G loop in purple. Residues implicated in site 1 and site 3 are shown as CPK objects in yellow and red, respectively. Panel (b) shows a closeup of the interface surface with residues shown to be important through mutagenesis (Deller *et al.*, 2000; Horsten *et al.*, 1997) shown as ball and stick objects. Helices A (blue), B (green), C (yellow) and D (red) are shown along with connecting loops in grey. Superposition was carried out using the program SHP (Stuart *et al.*, 1979). Figure produced using BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994).

6.6 Summary

In conclusion it can been seen that the present structure of OSM highlights several novel features unique to the gp130 family of cytokines, whilst at the same time supporting common themes observed in other 4 helix bundle cytokines. Although exhibiting a high degree of structural homology to other helical cytokines, key features such as the helical content of both the B-C and A-B loops and the severity of the kink in helix A set OSM apart. These subtle differences are likely to be key features responsible for modulation of receptor-ligand engagement in the gp130 family which utilise otherwise homologous systems.

Structural comparison reveals degrees of similarity to the classic two site binding system utilised by GH and Epo in which hydrophobic burial of key aromatic residues of the receptor drives the interaction (de Vos *et al.*, 1992; Syed *et al.*, 1998). However, models based on these interactions suggest the use of an interface that is less hydrophobic in nature in both LIF and OSM. Indeed, mutagenesis has revealed that site 2 interactions with gp130 and site 3 interactions with LIFR dominate the binding energy and as such the classic site 1 is of less importance (Deller *et al.*, 2000; Hudson *et al.*, 1996; Layton *et al.*, 1994b). Residues comprising site 2 and site 3 of OSM are both hydrophobic and polar in nature in a similar fashion to the mosaic binding sites observed in both the GCSF (Aritomi *et al.*, 1999) and IL4 (Hage *et al.*, 1998) receptor-ligand systems. This move away from the use of hydrophobic surfaces may allow tighter modulation of fidelity, albeit at the possible expense of affinity, a situation that is ideal for the 'pick and mix' approach to receptor engagement observed in the gp130 family of cytokines (Staunton *et al.*, 1998).

The level of promiscuity exhibited by LIFR and gp130 receptors suggests that novel motifs on the receptors are required. Models of the OSM-gp130 interaction, based on the Epo system, suggest that only one hydrophobic residue of the receptor dominates the

binding energy; a claim which is so far further supported by mutagenesis data (Horsten *et al.*, 1997). It is possible that the gp130 interaction has moved away from the classic B-C and E-F loop interactions which would lead to invariant and 'rigid' orientation of the ligand between the two receptor domains to a more 'flexible' interaction involving components of the B-C loop and the F-G loop, as observed in the GCSF complex (Aritomi *et al.*, 1999; Layton *et al.*, 1999). It has also been suggested that such receptor-ligand complexes are likely to be dependent on relative domain conformation and orientation, as observed in the Epo system (Syed *et al.*, 1998) and as such this model would allow for such optimisation of orientation. Structural features such as the WSXWS motif, which effectively forms a supporting 'back bone' of gp130 may be important in defining the allowed range of orientations.

Other evidence suggests that the gp130 family of cytokines may use complex modes of binding including electrostatic (McDonald *et al.*, 1995) and entropic components (Purvis and Mabbutt, 1997). Further receptor-ligand interactions, as observed in the minor site of the GCSFR-GCSF complex (Aritomi *et al.*, 1999) and the proposed site 3 interaction between LIF and the Ig domain of LIFR (Chobotova and Heath, 2000), adds a further level of complexity, which the classic two site model of GH and Epo fails to explain. This again highlights the complexities involved in these systems and traditional mutagenesis may prove of little merit for analysis of the finer details of the active signalling complexes.

Clearly, structural studies through the use of X-ray crystallography and NMR techniques will allow finer dissection of the interactions involved. This present work represents one such study in which a further piece of structural information has been added. Further analysis of the higher order complexes of gp130, LIFR and the specific α-receptor subunits is now required to complete the understanding of signal transduction through the haemopoietin cytokine network; central to this understanding is the site 3 interaction with LIFR, which still remains the elusive target of structural studies.

7. Appendices

7.1 Appendix I – Molecular Biology Reagents

7.1.1 Lund-Bertani Medium (LB)

Bacto-tryptone	10 g
Bacto-yeast extract	5 g
NaCl	10 g

dH_20 to 1000 cm^3 followed by autoclaving for 20 min at 15 p.s.i.

7.1.2 Tris-Acetate (TAE) (x50 Stock)

Tris Base	242 g
Glacial Acetic Acid	57.1 ml
0.5 M EDTA (pH 8.0)	100 ml

7.1.3 Tris-Borate (TBE) (x5 Stock)

Tris Base	54 g
Boric Acid	27.5 g
0.5 M EDTA (pH 8.0)	20 ml

7.1.4 LB Plates

Agar	12 g
Bactotryptone	8 g
Yeast Extract	4 g
Sodium Chloride	8 g
Sodium Hydroxide (10 M)	80 µl

dH_20 to 1000 cm^3 followed by autoclaving for 20 min at 15 p.s.i.

7.1.5 Transformation Media

LB media containing:

Sucrose	20 mM
$MgSO_4$	10 mM
$MgCl_2$	10 mM

7.1.6 DNA Loading Buffer Type III (6x)

Bromophenol Blue	0.25 %
Xylene Cyanol FF	0.25 %
Glycerol	30 %

Store at 4°C

7.2 Appendix II – Protein Expression and Purification Reagents

7.2.1 SOB Medium

Bacto-Tryptone	20 g
Bacto-Yeast Extract	5 g
NaCl	0.5 g
250 mM KCl (pH 7)	10 ml
dH_20 upto 1000cm^3 followed by autoclave for 20 min at 15	
2 M $MgCl_2$ (before use)	5 ml

7.2.2 SOC Medium

Bacto-Tryptone	20 g
Bacto-Yeast Extract	5 g
NaCl	0.5 g
250 mM KCl (pH 7)	10 ml
dH_2O upto 1000cm^3 followed by autoclave for 20 min at 15	
2 M $MgCl_2$ (before use) (sterile)	5 ml
1 M Glucose (before use)	20 ml

7.2.3 TNED Buffer

Tris-base (pH 8.0)	50 mM
NaCl	150 mM
EDTA	10 mM
DTT	1 mM

7.2.4 Sodium Citrate Buffer

Citric acid (0.1 M)	46.5 ml
Sodium citrate (0.1 M)	3.5 ml

7.2.5 MTPBS

NaCl	150 mM
Na_2HPO_4	16 mM
NaH_2PO_4	4 mM

pH 7.3

7.2.6 Gel Filtration Buffer

Tris-HCl	50 mM
NaCl	150 mM

7.2.7 Reduced Glutathione

Reduced glutathione	10 mM
Tris-HCl	20 mM

pH 8.0

7.2.8 G7 PNGase F Buffer

Sodium phosphate	0.05 M
NP-40 (v/v)	10 %

7.2.9 PNGase F Column Buffer

NaCl	200 mM
Tris-HCl	20 mM
Mercaptoethanol	10 mM

pH 8.0

7.2.10 PNGase F Elution Buffer

NaCl	200 mM
Tris-HCl	20 mM
Mercaptoethanol	10 mM
Maltose	10 mM

pH 8.0

7.2.11 Citric Acid Phosphate Buffer

Citric acid	0.035 M
Disodium phosphate	0.067 M

pH 5.0

7.2.12 Phosphate Buffered Saline (PBS)

Sodium chloride	8 g
Potassium chloride	0.2 g
Disodium hydrogen	1.44 g
Potassium dihydrogen	0.24 g

dH_20 upto 1000 ml, pH 7.4

7.2.13 Acrylamide/Bis (30 % T, 2.67 % C)

Acrylamide	87.6 g
N'N'-bis-methylene-	2.4 g

dH_2O upto 300 ml filter (0.45 μm) and store at 4°C in dark

7.2.14 SDS Reducing Buffer (x1 stock)

dH_2O	3.5 ml
0.5 M Tris-HCl (pH 6.8)	1.3 ml
Glycerol	6.0 ml
10 % (w/v) SDS	3.5 ml
β-mercaptoethanol	0.5 ml
1 % (w/v) BPB	0.5 ml

7.2.15 SDS Running Buffer (x5 stock)

Tris-Base	9 g
Glycine	14 g
SDS	3 g

dH_2O upto 600 ml

7.2.16 Separating Gel (12 %)

dH_2O	3.35 ml
1.5 M Tris-HCl (pH 8.8)	2.50 ml
10 % (w/v) SDS	100 µl
30 % Acrylamide/Bis	4.0 ml
10 % (w/v) APS	50 µl
TEMED	5 µl

7.2.17 Stacking Gel (4 %)

dH_2O	6.1 ml
0.5 M Tris-HCl (pH 6.8)	2.5 ml
10 % (w/v) SDS	100 µl
30 % Acrylamide/Bis	1.33 ml
10 % (w/v) APS	50 µl
TEMED	10 µl

7.2.18 Coomassie Destain

Methanol	400 ml
Acetic Acid	100 ml

dH_2O upto 1000 ml

7.2.19 Coomassie Stain

Coomassie Blue R-250	1g
Methanol	400 ml
Acetic Acid	100 ml

dH_2O upto 1000 ml

7.2.20 Protein Loading Buffer (SDS Reducing)

dH_20	4.0 ml
0.5 M Tris-HCl (pH 6.8)	1.0 ml
Glycerol	0.8 ml
β-mercaptoethanol	0.4 ml
10 % (w.v) SDS	1.6 ml
0.05 % (w/v) BPB	0.2 ml

7.2.21 Western Blot Transfer Buffer

Tris-Base	3.03 g
Glycine	14.4 g
Methanol	100 ml

dH_20 to 1000 ml

7.3 Appendix III - Eukaryotic Expression Reagents

7.3.1 DMEM-S for Culture/Transfection of 293T Cells

DMEM Supplemented with:

FCS	10%
Penicillin	50 IU/ml
Streptomycin	50 μg/ml
Glutamine	2 mM
Pyruvate	1 mM
Glucose	4500 mg/ml

7.3.2 SF900 Media for Culture of *Drosophila* S2 Cells

SF900 Media Supplemented with:

FCS	5%
Penicillin	50 IU/ml
Streptomycin	50 μg/ml
Hygromycin-B	300 μg/ml

7.3.3 GMEM-S for Culture/Transfection of CHO Cells

GMEM-S Supplemented with:

FCS	10%
Penicillin	50 IU/ml
Streptomycin	50 μg/ml
L-methioninesulphoximine	50 μM

7.3.4 HBBS (10x)

NaCl (w/v)	1.60 %
Hepes (w/v)	1.20 %
Na_2HPO_4 (w/v)	0.04 %

7.4 Appendix IV - Useful Formulae

7.4.1 Spectrophotometric Coefficient for Calculating Protein Concentration

$$\frac{A_{280}}{[\text{Protein}](\text{mg/ml})} = \frac{(1331 \times Y) + (5650 \times W)}{M_r}$$

Equation 7-1

Equation used to determine protein concentration from absorbance readings at 280 nm (A_{280}). Measurements were made in a 1 ml cell using a Beckman DU-600 instrument. Protein concentration (mg/ml) were estimated based on the number of tyrosine residues (Y) and the number of tryptophan residues (W) in relation to the molecular weight (M_r).

7.4.2 Spectrophotometric Measurement of Oligonucleotide Concentration

$$[\text{Oligonucleotide}](\text{pM}/\mu\text{l}) = 118 \times A_{260}$$

Equation 7-2

Equation used to determine primer and oligonucleotide concentrations for a 1:40 dilution of stock oligonucleotide. Absorbance measurements at 260 nm (A_{260}) were made in a 1 ml cell using a Beckman DU-600 instrument. The factor 118 is derived from the number of bases and their extinction coefficients and is specific for each oligonucleotide.

7.4.3 Spectrophotometric Measurement of Plasmid DNA Concentration

$$A_{260}1.0 = 50\ \mu\text{g} / \text{ml}$$

Equation 7-3

Equation used to determine plasmid DNA concentrations for a 1:100 dilution of plasmid stock. Absorbance measurements at 260 nm (A_{260}) were made in a 1 ml cell using a Beckman DU-600 instrument.

Abdel Meguid, S.S., Shieh, H.S., Smith, W.W., Dayringer, H.E., Violand, B.N. and Bentle, L.A. (1987) Three-dimensional structure of a genetically engineered variant of porcine growth hormone. Proc. Natl. Acad. Sci. USA, 84, 6434-7.

Abrahams, J.P. and Leslie, A.G.W. (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta. Cryst.., D52, 30-42.

Agarwal, R.C. (1978) A new least-squares refinement technique based on the fast Fourier transform algorithm. Acta. Cryst.., A34, 791-793.

Arai, K.I., Lee, F., Miyajima, A., Miyatake, S., Arai, N. and Yokota, T. (1990) Cytokines: coordinators of immune and inflammatory responses. .

Aritomi, M., Kunishima, N., Okamoto, T., Kuroki, R., Ota, Y. and Morikawa, K. (1999) Atomic structure of the GCSF-receptor complex showing a new cytokine- receptor recognition scheme. Nature, 401, 713-7.

Baeyens, K.J., De Bondt, H.L., Raeymaekers, A., Fiers, W. and De Ranter, C.J. (1999) The structure of mouse tumour-necrosis factor at 1.4 A resolution: towards modulation of its selectivity and trimerization. Acta. Cryst., 55, 772-8.

Baggiolini, M. and Dahinden, C.A. (1994) CC chemokines in allergic inflammation. Immunol. Today, 15, 127-33.

Baker, H.M., Day, C.L., Norris, G.E. and Baker, E.N. (1994) Enzymatic Deglycosylation As a Tool For Crystallization of Mammalian Binding-Proteins. Acta. Cryst., 50, 380-384.

Baldwin, E.T., Weber, I.T., St. Charles, R., Xuan, J.C., Appella, E., Yamada, M., Matsushima, K., Edwards, B.F., Clore, G.M., Gronenborn, A.M. et al. (1991) Crystal structure of interleukin 8: symbiosis of NMR and crystallography. Proc. Natl. Acad. Sci. USA, 88, 502-6.

Balkwill, F.R. and Burke, F. (1989) The cytokine network. Immunol.Today, 10, 299-304.

Banner, D.W., D'Arcy, A., Janes, W., Gentz, R., Schoenfeld, H.J., Broger, C., Loetscher, H. and Lesslauer, W. (1993) Crystal structure of the soluble human 55 kd TNF receptor-human TNF beta complex: implications for TNF receptor activation. Cell, 73, 431-45.

Barsomian, G.D., Johnson, T.L., Borowski, M., Denman, J., Ollington, J.F., Hirani, S., McNeilly, D.S. and Rasmussen, J.R. (1990) Cloning and expression of peptide-N4-(N-acetyl-beta-D-glucosaminyl)asparagine amidase F in Escherichia coli. J. Biol. Chem., 265, 6967-72.

Baumann, H., Ziegler, S.F., Mosley, B., Morella, K.K., Pajovic, S. and Gearing, D.P. (1993) Reconstitution of the response to leukemia inhibitory factor, oncostatin M, and ciliary neurotrophic factor in hepatoma cells. J. Biol. Chem., 268, 8414-7.

Bazan, J.F. (1991) Neuropoietic cytokines in the hematopoietic fold. Neuron, 7, 197-208.

Bazan, J.F. (1992) Unraveling the structure of IL-2. Science, 257, 410-3.

Bebbington, C.R. (1991) Expression of antibody genes in nonlymphoid mammalian cells. Methods: A companion to methods in enzymology, 2, 136-145.

Blaber, M., DiSalvo, J. and Thomas, K.A. (1996) X-ray crystal structure of human acidic fibroblast growth factor. Biochemistry, 35, 2086-94.

Blow, D.M., Janin, J. and Sweet, R.M. (1974) Mode of action of soybean trypsin inhibitor (Kunitz) as a model for specific protein-protein interactions. Nature, 249, 54-7.

Blundell, T. and Johnson, L. (1976) Protein crystallography. Academic Press, New York.

Bravo, J., Staunton, D., Heath, J.K. and Jones, E.Y. (1998) Crystal structure of a cytokine-binding region of gp130. EMBO. J., 17, 1665-74.

Brown, T.J., Lioubin, M.N. and Marquardt, H. (1987) Purification and characterization of cytostatic lymphokines produced by activated human T lymphocytes. Synergistic antiproliferative activity of transforming growth factor beta 1, interferon-gamma, and oncostatin M for human melanoma cells. J. Immunol., 139, 2977-83.

Bruce, A.G., Hoggatt, I.H. and Rose, T.M. (1992) Oncostatin M is a differentiation factor for myeloid leukemia cells. J. Immunol., 149, 1271-5.

Brunger, A.T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature, 355, 472-475.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., GrosseKunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Cryst., 54, 905-921.

Bunch, T.A., Grinblat, Y. and Goldstein, L.S. (1988) Characterization and use of the Drosophila metallothionein promoter in cultured Drosophila melanogaster cells. Nucleic Acids Res., 16, 1043-61.

Cai, J., Gill, P.S., Masood, R., Chandrasoma, P., Jung, B., Law, R.E. and Radka, S.F. (1994) Oncostatin-M is an autocrine growth factor in Kaposi's sarcoma. Am. J. Pathol., 145, 74-9.

Campbell, I.D., Cooke, R.M., Baron, M., Harvey, T.S. and Tappin, M.J. (1989) The solution structures of epidermal growth factor and transforming growth factor alpha. Prog. Growth. Factor. Res., 1, 13-22.

Carpenter, G. and Cohen, S. (1990) Epidermal growth factor. J. Biol. Chem., 265, 7709-12.

Chapdelaine, J.M. (1989) MTT Reduction - A Tetrazolium-Based Colorimetric Assay for Cell Survival and Proliferation. In MAXline Application Notes. (Molecular Devices).

Chayen, N.E. (1997) A novel technique to control the rate of vapour diffusion, giving larger protein crystals. J. Appl. Cryst., 30, 198-202.

Chirgadze, D.Y., Hepple, J.P., Zhou, H., Byrd, R.A., Blundell, T.L. and Gherardi, E. (1999) Crystal structure of the NK1 fragment of HGF/SF suggests a novel mode for growth factor dimerization and receptor binding. Nat. Struct. Biol., 6, 72-9.

Chobotova, K. and Heath, J.K. (2000) in press. .

Clackson, T., Ultsch, M.H., Wells, J.A. and de Vos, A.M. (1998) Structural and functional analysis of the 1:1 growth hormone:receptor complex reveals the molecular basis for receptor affinity. J. Mol. Biol., 277, 1111-28.

Clackson, T. and Wells, J.A. (1995) A hot spot of binding energy in a hormone-receptor interface. Science, 267, 383-6.

Clore, G.M. and Gronenborn, A.M. (1991) Comparison of the solution nuclear magnetic resonance and crystal structures of interleukin-8. Possible implications for the mechanism of receptor binding. J. Mol. Biol., 217, 611-20.

Cockett, M.I., Bebbington, C.R. and Yarranton, G.T. (1990) High level expression of tissue inhibitor of metalloproteinases in Chinese hamster ovary cells using glutamine synthetase gene amplification. Biotechnology NY, 8, 662-7.

Coligan, J.E., Dunn, B.M., Ploegh, H.L., Speicher, D.W. and Wingfield, P.T. (1995) Current protocols in protein science. John Wiley & Sons, Inc.

Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography. Acta. Cryst., D50, 760-763.

Cosman, D. (1993) The hematopoietin receptor superfamily. Cytokine, 5, 95-106.

Cosman, D., Lyman, S.D., Idzerda, R.L., Beckmann, M.P., Park, L.S., Goodwin, R.G. and March, C.J. (1990) A new cytokine receptor superfamily. Trends Biochem. Sci., 15, 265-70.

Cunningham, B.C., Bass, S., Fuh, G. and Wells, J.A. (1990) Zinc mediation of the binding of human growth hormone to the human prolactin receptor. Science, 250, 1709-12.

Cunningham, B.C., Jhurani, P., Ng, P. and Wells, J.A. (1989) Receptor and antibody epitopes in human growth hormone identified by homolog-scanning mutagenesis. Science, 243, 1330-6.

Cunningham, B.C., Mulkerrin, M.G. and Wells, J.A. (1991a) Dimerization of human growth hormone by zinc. Science, 253, 545-8.

Cunningham, B.C., Ultsch, M., De Vos, A.M., Mulkerrin, M.G., Clauser, K.R. and Wells, J.A. (1991b) Dimerization of the extracellular domain of the human growth hormone receptor by a single hormone molecule. Science, 254, 821-5.

Cunningham, B.C. and Wells, J.A. (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. Science, 244, 1081-5.

Cunningham, B.C. and Wells, J.A. (1993) Comparison of a structural and a functional epitope. J. Mol. Biol., 234, 554-63.

D'Anna, J.A., Tobey, R.A. and Gurley, L.R. (1980) Concentration-dependent effects of sodium butyrate in Chinese hamster cells: cell-cycle progression, inner-histone acetylation, histone H1 dephosphorylation, and induction of an H1-like protein. Biochemistry, 19, 2656-71.

Daopin, S., Piez, K.A., Ogawa, Y. and Davies, D.R. (1992) Crystal structure of transforming growth factor-beta 2: an unusual fold for the superfamily. Science, 257, 369-73.

Darnell, J.E., Jr., Kerr, I.M. and Stark, G.R. (1994) Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. Science, 264, 1415-21.

Davis, S., Aldrich, T.H., Stahl, N., Pan, L., Taga, T., Kishimoto, T., Ip, N.Y. and Yancopoulos, G.D. (1993) LIFR beta and gp130 as heterodimerizing signal transducers of the tripartite CNTF receptor. Science, 260, 1805-8.

de La Fortelle, E. and Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods in Enzymology, 276, 472-494.

De Serio, A., Graziani, R., Laufer, R., Ciliberto, G. and Paonessa, G. (1995) In vitro binding of ciliary neurotrophic factor to its receptors: evidence for the formation of an IL-6-type hexameric complex. J. Mol. Biol., 254, 795-800.

de Vos, A.M., Ultsch, M. and Kossiakoff, A.A. (1992) Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. Science, 255, 306-12.

Dealwis, C., Fernandez, E.J., Thompson, D.A., Simon, R.J., Siani, M.A. and Lolis, E. (1998) Crystal structure of chemically synthesized [N33A] stromal cell-derived factor 1alpha, a potent ligand for the HIV-1 "fusin" coreceptor. Proc. Natl. Acad. Sci. USA, 95, 6941-6.

Deller, M.C., Hudson, K.R., Bravo, J., Jones, E.Y. and Heath, J.K. (2000) Manuscript in preparation.

Deutscher, M.P. (1990) Guide to protein purification. Academic Press.

Di Marco, A., Gloaguen, I., Graziani, R., Paonessa, G., Saggio, I., Hudson, K.R. and Laufer, R. (1996) Identification of ciliary neurotrophic factor (CNTF) residues essential for leukemia inhibitory factor receptor binding and generation of CNTF receptor antagonists. Proc. Natl. Acad. Sci. USA, 93, 9247-52.

Diederichs, K., Boone, T. and Karplus, P.A. (1991) Novel fold and putative receptor binding site of granulocyte-macrophage colony-stimulating factor. Science, 254, 1779-82.

Drenth, J. (1994) Principles of protein X-ray crystallography. Springer-Verlag, New York.

Ealick, S.E., Cook, W.J., Vijay Kumar, S., Carson, M., Nagabhushan, T.L., Trotta, P.P. and Bugg, C.E. (1991) Three-dimensional structure of recombinant human interferon-gamma. Science, 252, 698-702.

Eck, M.J. and Sprang, S.R. (1989) The structure of tumor necrosis factor-alpha at 2.6 A resolution. Implications for receptor binding. J-Biol-Chem, 264, 17595-605.

Engh, R.A. and Huber, R. (1991) Accurate bond and angle parameters for X-ray protein-structure refinement. Acta. Cryst., A47, 392-400.

Esnouf, R. (1993) GROPAT, unpublished program. Oxford, U.K. .

Esnouf, R.M. (1997) An extensively modified version of Molscript that includes greatly enhanced colouring capabilities. J. Mol. Graphics, 15, 133-138.

Faham, S., Hileman, R.E., Fromm, J.R., Linhardt, R.J. and Rees, D.C. (1996) Heparin structure and interactions with basic fibroblast growth factor. Science, 271, 1116-20.

Fan, Q.R., Mosyak, L., Winter, C.C., Wagtmann, N., Long, E.O. and Wiley, D.C. (1997) Structure of the inhibitory receptor for human natural killer cells resembles haematopoietic receptors. Nature, 389, 96-100.

Feng, Y., Klein, B.K. and McWherter, C.A. (1996) Three-dimensional solution structure and backbone dynamics of a variant of human interleukin-3. J. Mol. Biol., 259, 524-41

Ferré-D' Amaré, A.R. and Burley, S.K. (1994) Use of dynamic light scattering to assess crystallizability of macromolecules and macromolecular assemblies. Structure, 2, 357-359.

Foa, P. (1991) Erythropoeitin: Basic and Clinical Aspects. In Spivak, J.L. (ed.) Acta Haematol. W. B. Saunders, Philadelphia, Vol. 86, p. 162.

Fontaine, V., Savino, R., Arcone, R., de Wit, L., Brakenhoff, J.P., Content, J. and Ciliberto, G. (1993) Involvement of the Arg179 in the active site of human IL-6. Eur-J-Biochem, 211, 749-55.

French, S. and Wilson, K. (1978) On the treatment of negative intensity observations. Acta. Cryst., A34, 517.

Fukunaga, R., Ishizaka Ikeda, E. and Nagata, S. (1993) Growth and differentiation signals mediated by different regions in the cytoplasmic domain of granulocyte colony-stimulating factor receptor. Cell, 74, 1079-87.

Furey, W. and Swaminathan, S. (1996) PHASES-95: A program package for the processing and analysis of diffraction data from macromolecules. In Carter, C. and Sweet, R. (eds.), Methods in Enzymology. Academic Press, Orlando, Fl. U.S.A.

Gearing, D.P. and Bruce, A.G. (1992) Oncostatin M binds the high-affinity leukemia inhibitory factor receptor. New. Biol., 4, 61-5.

Gearing, D.P., Thut, C.J., VandeBos, T., Gimpel, S.D., Delaney, P.B., King, J., Price, V., Cosman, D. and Beckmann, M.P. (1991) Leukemia inhibitory factor receptor is structurally related to the IL-6 signal transducer, gp130. EMBO. J., 10, 2839-48.

Gilman, A.G. (1987) G proteins: transducers of receptor-generated signals. Annu. Rev. Biochem., 56, 615-49.

Giovannini, M., Djabali, M., McElligott, D., Selleri, L. and Evans, G.A. (1993) Tandem linkage of genes coding for leukemia inhibitory factor (LIF) and oncostatin M (OSM) on human chromosome 22. Cytogenet. Cell. Genet., 64, 240-4.

Gordon, M.S. (1996) Thrombopoietic activity of recombinant human interleukin 11 in cancer patients receiving chemotherapy. Cancer Chemother. Pharmacol., 38, S96-8.

Gorman, C.M., Howard, B.H. and Reeves, R. (1983) Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate. Nucleic. Acids Res., 11, 7631-48.

Gouet, P., Courcelle, E., Stuart, D.I. and Metoz, F. (1999) ESPript: multiple sequence alignments in Postscript. Bioinformatics, 15, 305-308.

Gough, N.M. (1992) Molecular genetics of leukemia inhibitory factor (LIF) and its receptor. Growth Factors, 7, 175-9.

Graham, F.L., Smiley, J., Russell, W. and Nairn, R. (1977) Characterisation of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virology, 36, 59-72.

Grove, R.I., Eberhardt, C., Abid, S., Mazzucco, C., Liu, J., Kiener, P., Todaro, G. and Shoyab, M. (1993) Oncostatin M is a mitogen for rabbit vascular smooth muscle cells. Proc. Natl. Acad. Sci. USA, 90, 823-7.

Hage, T., Reinemer, P. and Sebald, W. (1998) Crystals of a 1:1 complex between human interleukin-4 and the extracellular domain of its receptor alpha chain. Eur. J. Biochem., 258, 831-6.

Hage, T., Sebald, W. and Reinemer, P. (1999) Crystal structure of the interleukin-4/receptor alpha chain complex reveals a mosaic binding interface. Cell, 97, 271-81.

Harlos, K. (1992) Micro-bridges for sitting-drop crystallizations. J. Appl. Cryst., 25, 536-538.

Harlos, K., Martin, D.M., O'Brien, D.P., Jones, E.Y., Stuart, D.I., Polikarpov, I., Miller, A., Tuddenham, E.G. and Boys, C.W. (1994) Crystal structure of the extracellular region of human tissue factor. Nature, 370, 662-6.

Heath, J.K. (1992) Cytokines. Can there be life without LIF?. Nature, 359, 17.

Heldin, C.H. (1995) Dimerization of cell surface receptors in signal transduction. Cell, 80, 213-23.

Hendrickson, W.A., Horton, J.R. and LeMaster, D.M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. EMBO. J., 9, 1665-72.

Hill, C.P., Osslund, T.D. and Eisenberg, D. (1993) The structure of granulocyte-colony-stimulating factor and its relationship to other growth factors. Proc. Natl. Acad. Sci. USA, 90, 5167-71.

Hilton, D.J. (1992) LIF: lots of interesting functions. Trends Biochem. Sci., 17, 72-6.

Hilton, D.J., Hilton, A.A., Raicevic, A., Rakar, S., Harrison Smith, M., Gough, N.M., Begley, C.G., Metcalf, D., Nicola, N.A. and Willson, T.A. (1994) Cloning of a murine IL-11 receptor alpha-chain; requirement for gp130 for high affinity binding and signal transduction. EMBO J., 13, 4765-75.

Hilton, D.J. and Nicola, N.A. (1992) Kinetic analyses of the binding of leukemia inhibitory factor to receptor on cells and membranes and in detergent solution. J. Biol. Chem., 267, 10238-47.

Hilton, D.J., Nicola, N.A. and Metcalf, D. (1988) Purification of a murine leukemia inhibitory factor from Krebs ascites cells. Anal. Biochem., 173, 359-67.

Hinds, M.G., Maurer, T., Zhang, J.G., Nicola, N.A. and Norton, R.S. (1998) Solution structure of leukemia inhibitory factor. J. Biol. Chem., 273, 13738-45.

Hiraoka, O., Anaguchi, H., Asakura, A. and Ota, Y. (1995) Requirement for the immunoglobulin-like domain of granulocyte-colony-stimulating factor-receptor in formation of a 2/1 receptor-ligand complex. J. Biol. Chem., 270, 25928-25934.

Hiraoka, O., Anaguchi, H. and Ota, Y. (1994) Evidence for the ligand-induced conversion from a dimer to a tetramer of the granulocyte-colony-stimulating factor-receptor. Febs Lett., 356, 255-260.

Ho, A.S., Liu, Y., Khan, T.A., Hsu, D.H., Bazan, J.F. and Moore, K.W. (1993) A receptor for interleukin 10 is related to interferon receptors. Proc. Natl. Acad. Sci. USA, 90, 11267-71.

Hoffman, R.C., Moy, F.J., Price, V., Richardson, J., Kaubisch, D., Frieden, E.A., Krakover, J.D., Castner, B.J., King, J., March, C.J. and Powers, R. (1996) Resonance assignments for Oncostatin M, a 24-kDa alpha-helical protein. J. Biomol. NMR, 7, 273-82.

Horan, T., Wen, J., Narhi, L., Parker, V., Garcia, A., Arakawa, T. and Philo, J. (1996) Dimerization of the extracellular domain of granulocyte-colony stimulating factor receptor by ligand binding: A monovalent ligand induces 2:2 complexes. Biochemistry, 35, 4886-4896.

Horan, T.P., Simonet, L., Jacobsen, R., Mann, M., Haniu, M., Wen, J., Arakawa, T., Kuwamoto, M. and Martin, F. (1998) Coexpression of G-CSF with an unglycosylated G-CSF receptor mutant results in secretion of a stable complex. Protein Expression and Purification, 14, 45-53.

Horn, D., Fitzpatrick, W.C., Gompper, P.T., Ochs, V., Bolton Hansen, M., Zarling, J., Malik, N., Todaro, G.J. and Linsley, P.S. (1990) Regulation of cell growth by recombinant oncostatin M. Growth Factors, 2, 157-65.

Horsten, U., Muller Newen, G., Gerhartz, C., Wollmer, A., Wijdenes, J., Heinrich, P.C. and Grotzinger, J. (1997) Molecular modeling-guided mutagenesis of the extracellular part of gp130 leads to the identification of contact sites in the interleukin-6 (IL-6).IL-6 receptor.gp130 complex. J. Biol. Chem., 272, 23748-57.

Howell, P.L. and Smith, G.D. (1992) Identification of heavy-atom derivatives by normal probability methods. J. Appl. Cryst., 25, 81.

Hubbard, S.J. and Thornton, J.M. (1993) 'NACCESS'. Computer program. Department of Biochemistry and Molecular Biology., University College, London. .

Hubbard, S.R. (1997) Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. EMBO J., 16, 5572-81.

Hudson, K.R., Vernallis, A.B. and Heath, J.K. (1996) Characterization of the receptor binding sites of human leukemia inhibitory factor and creation of antagonists. J. Biol. Chem., 271, 11971-8.

Hui, W., Bell, M. and Carroll, G. (1996) Oncostatin M (OSM) stimulates resorption and inhibits synthesis of proteoglycan in porcine articular cartilage explants. Cytokine, 8, 495-500.

Hutchinson, E.G. and Thornton, J.M. (1990) HERA-a program to draw schematic diagrams of protein secondary structures. Proteins, 8, 203-12.

Ibanez, C.F., Ebendal, T., Barbany, G., Murray-Rust, J., Blundell, T.L. and Persson, H. (1992) Disruption of the low affinity receptor-binding site in NGF allows neuronal survival and differentiation by binding to the trk gene product. Cell, 69, 329-41.

Ihle, J.N. and Kerr, I.M. (1995) Jaks and Stats in signaling by the cytokine receptor superfamily. Trends-Genet, 11, 69-74.

Ihle, J.N., Witthuhn, B.A., Quelle, F.W., Yamamoto, K., Thierfelder, W.E., Kreider, B. and Silvennoinen, O. (1994) Signaling by the cytokine receptor superfamily: JAKs and STATs. Trends. Biochem. Sci., 19, 222-7.

Inoue, M., Minami, M., Matsumoto, M., Kishimoto, T. and Akira, S. (1997) The amino acid residues immediately carboxyl-terminal to the tyrosine phosphorylation site contribute to interleukin 6-specific activation of signal transducer and activator of transcription 3. J. Biol. Chem., 272, 9550-5.

Inoue, M., Nakayama, C., Kikuchi, K., Kimura, T., Ishige, Y., Ito, A., Kanaoka, M. and Noguchi, H. (1995) D1 cap region involved in the receptor recognition and neural cell survival activity of human ciliary neurotrophic factor. Proc. Natl. Acad. Sci. USA, 92, 8579-83.

Jancarik, J. and Kim, S.H. (1991) Sparse Matrix Sampling: A screening method for crystallization of proteins. J. Appl. Cryst., 24, 409-411.

Johnson, D.L., Farrell, F.X., Barbone, F.P., McMahon, F.J., Tullai, J., Hoey, K., Livnah, O., Wrighton, N.C., Middleton, S.A., Loughney, D.A., Stura, E.A., Dower, W.J., Mulcahy, L.S., Wilson, I.A. and Jolliffe, L.K. (1998) Identification of a 13 amino acid peptide mimetic of erythropoietin and description of amino acids critical for the mimetic activity of EMP1. Biochemistry, 37, 3699-710.

Jones, E.Y., Stuart, D.I. and Walker, N.P. (1989) Structure of tumour necrosis factor. Nature, 338, 225-8.

Jones, T.A., Zou, J.-Y., Cowan, S.W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta. Cryst., A47, 110.

Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers, 22, 2577-637.

Kallestad, J.C., Shoyab, M. and Linsley, P.S. (1991) Disulfide bond assignment and identification of regions required for functional activity of oncostatin M. J. Biol. Chem., 266, 8940-5.

Karlson, G.B., Butters, T.D., Dwek, R.A. and Platt, F.M. (1993) Effects of the imino sugar N-butyldeoxynojirimycin on the N-glycosylation of recombinant gp120. J. Biol. Chem., 268, 570-576.

Karpusas, M., Hsu, Y.M., Wang, J.H., Thompson, J., Lederman, S., Chess, L. and Thomas, D. (1995) 2 A crystal structure of an extracellular fragment of human CD40 ligand. Structure, 3, 1031-9.

Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst., 24, 946-950.

Laemmli, U.K. (1970) Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4. Nature, 227, 680-685.

Laron, Z., Pertzelan, A. and Mannheimer, S. (1966) Genetic pituitary dwarfism with high serum concentation of growth hormone--a new inborn error of metabolism? Isr. J. Med. Sci., 2, 152-5.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures. J. Appl. Cryst., 26, 283-291.

Layton, J.E., Iaria, J., Smith, D.K. and Treutlein, H.R. (1997) Identification of a ligand-binding site on the granulocyte colony-stimulating factor receptor by molecular modeling and mutagenesis. J. Biol. Chem., 272, 29735-41.

Layton, J.E., Shimamoto, G., Osslund, T., Hammacher, A., Smith, D.K., Treutlein, H.R. and Boone, T. (1999) Interaction of granulocyte colony-stimulating factor (G-CSF) with its receptor. Evidence that Glu19 of G-CSF interacts with Arg288 of the receptor. J. Biol. Chem., 274, 17445-51.

Layton, M.J., Lock, P., Metcalf, D. and Nicola, N.A. (1994a) Cross-species receptor binding characteristics of human and mouse leukemia inhibitory factor suggest a complex binding interaction. J. Biol. Chem. , 269, 17048-55.

Layton, M.J., Owczarek, C.M., Metcalf, D., Clark, R.L., Smith, D.K., Treutlein, H.R. and Nicola, N.A. (1994b) Conversion of the biological specificity of murine to human leukemia inhibitory factor by replacing 6 amino acid residues. J. Biol. Chem., 269, 29891-6.

Le, J. and Vilcek, J. (1987) Tumor necrosis factor and interleukin 1: cytokines with multiple overlapping biological activities. Lab. Invest., 56, 234-48.

LeMaster, D.M. and Richards, F.M. (1982) Preparative-scale isolation of isotopically labeled amino acids. Anal. Biochem., 122, 238-47.

Leong, L.E.-C., Walker, P.A. and Porter, A.G. (1992) Efficient expression and purification of a protease from the common cold virus, human rhinovirus type 14. J. Cryst. Growth, 122, 246-252.

Li, C., Schwabe, J.W.R., Banayo, E. and Evans, R.M. (1997) Coexpression of nuclear receptor partners increases their solubility and biological activities. Proc. Natl. Acad. Sci. USA, 94, 2278-2283.

Li, M., Sendtner, M. and Smith, A. (1995) Essential function of LIF receptor in motor neurons. Nature, 378, 724-7.

Linsley, P.S., Bolton Hanson, M., Horn, D., Malik, N., Kallestad, J.C., Ochs, V., Zarling, J.M. and Shoyab, M. (1989) Identification and characterization of cellular receptors for the growth regulator, oncostatin M. J. Biol. Chem, 264, 4282-9.

Linsley, P.S., Kallestad, J., Ochs, V. and Neubauer, M. (1990) Cleavage of a hydrophilic C-terminal domain increases growth-inhibitory activity of oncostatin M. Mol. Cell Biol., 10, 1882-90.

Livnah, O., Johnson, D.L., Stura, E.A., Farrell, F.X., Barbone, F.P., You, Y., Liu, K.D., Goldsmith, M.A., He, W., Krause, C.D., Pestka, S., Jolliffe, L.K. and Wilson, I.A. (1998) An antagonist peptide-EPO receptor complex suggests that receptor dimerization is not sufficient for activation. Nat. Struct. Biol., 5, 993-1004.

Livnah, O., Stura, E.A., Johnson, D.L., Middleton, S.A., Mulcahy, L.S., Wrighton, N.C., Dower, W.J., Jolliffe, L.K. and Wilson, I.A. (1996) Functional mimicry of a protein hormone by a peptide agonist: the EPO receptor complex at 2.8 A. Science, 273, 464-71.

Livnah, O., Stura, E.A., Middleton, S.A., Johnson, D.L., Jolliffe, L.K. and Wilson, I.A. (1999) Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. Science, 283, 987-90.

Lovejoy, B., Cascio, D. and Eisenberg, D. (1993) Crystal structure of canine and bovine granulocyte-colony stimulating factor (G-CSF). J. Mol. Biol., 234, 640-53.

Lubkowski, J., Bujacz, G., Boque, L., Domaille, P.J., Handel, T.M. and Wlodawer, A. (1997) The structure of MCP-1 in two crystal forms provides a rare example of variable quaternary interactions. Nat. Struct. Biol., 4, 64-69.

Lucklow, V.A. and Summers, M.D. (1988) Trends in the development of baculovirus expression vectors. Biotechnology, 6, 47-55.

Lutticken, C., Kruttgen, A., Moller, C., Heinrich, P.C. and Rose John, S. (1991) Evidence for the importance of a positive charge and an alpha-helical structure of the C-terminus for biological activity of human IL-6. FEBS Lett., 282, 265-7.

Maenaka, K., Juji, T., Stuart, D.I. and Jones, E.Y. (1999) Crystal structure of the human p58 killer cell inhibitory receptor (KIR2DL3) specific for HLA-Cw3-related MHC class I. Structure, 7, 391-8.

Malik, N., Graves, D., Shoyab, M. and Purchio, A.F. (1992) Amplification and expression of heterologous oncostatin M in Chinese hamster ovary cells. DNA Cell. Biol., 11, 453-9.

Malik, N., Kallestad, J.C., Gunderson, N.L., Austin, S.D., Neubauer, M.G., Ochs, V., Marquardt, H., Zarling, J.M., Shoyab, M., Wei, C.M. and et al. (1989) Molecular cloning, sequence analysis, and functional expression of a novel growth regulator, oncostatin M. Mol. Cell Biol., 9, 2847-53.

Malkowski, M.G., Wu, J.Y., Lazar, J.B., Johnson, P.H. and Edwards, B.F. (1995) The crystal structure of recombinant human neutrophil-activating peptide-2 (M6L) at 1.9-A resolution. J. Biol. Chem., 270, 7077-87.

Maniatis, T., Sambrook, J. and Fritsch, E.F. (1989) Molecular Cloning: A Laboratory Manual. New York, Cold Spring Harbor Laboratory.

May, L.T., Sanathanam, U. and Sehgal, P.B. (1991) On the multimeric nature of human interleukin-6. J. Biol. Chem, 266, 9950-9955.

McDonald, N.Q., Lapatto, R., Murray-Rust, J., Gunning, J., Wlodawer, A. and Blundell, T.L. (1991) New protein fold revealed by a 2.3-A resolution crystal structure of nerve growth factor. Nature, 354, 411-4.

McDonald, N.Q., Panayotatos, N. and Hendrickson, W.A. (1995) Crystal structure of dimeric human ciliary neurotrophic factor determined by MAD phasing. EMBO J., 14, 2689-99.

McPherson, A. (1995) Increasing the size of microcrystals by fine sampling of pH limits. J. Appl. Cryst., 28, 362-365.

Merritt, E.A. and Murphy, M.E.P. (1994) Raster3D version 2.0. A program for photorealistic molecular graphics. Acta. Cryst., D50, 869-873.

Middleton, S.A., Barbone, F.P., Johnson, D.L., Thurmond, R.L., You, Y., McMahon, F.J., Jin, R., Livnah, O., Tullai, J., Farrell, F.X., Goldsmith, M.A., Wilson, I.A. and Jolliffe, L.K. (1999) Shared and unique determinants of the erythropoietin (EPO) receptor are important for binding EPO and EPO mimetic peptide. J. Biol. Chem., 274, 14163-9.

Milburn, M.V., Hassell, A.M., Lambert, M.H., Jordan, S.R., Proudfoot, A.E., Graber, P. and Wells, T.N. (1993) A novel dimer configuration revealed by the crystal structure at 2.4 A resolution of human interleukin-5. Nature, 363, 172-6.

Miles, S.A., Martinez Maza, O., Rezai, A., Magpantay, L., Kishimoto, T., Nakamura, S., Radka, S.F. and Linsley, P.S. (1992) Oncostatin M as a potent mitogen for AIDS-Kaposi's sarcoma-derived cells. Science, 255, 1432-4.

Miller, M.D. and Krangel, M.S. (1992) Biology and biochemistry of the chemokines: a family of chemotactic and inflammatory cytokines. Crit. Rev. Immunol., 12, 17-46.

Miller, R.T., Jones, D.T. and Thornton, J.M. (1996) Protein fold recognition by sequence threading: Tools and assessment techniques. FASEB Journal, 10, 171-178.

Mongkolsapaya, J., Grimes, J.M., Chen, N., Xu, X.N., Stuart, D.I., Jones, E.Y. and Screaton, G.R. (1999) Structure of the TRAIL-DR5 complex reveals mechanisms conferring specificity in apoptotic initiation. Nat. Struct. Biol., 6, 1048-1053.

Moreau, J.F., Donaldson, D.D., Bennett, F., Witek Giannotti, J., Clark, S.C. and Wong, G.G. (1988) Leukaemia inhibitory factor is identical to the myeloid growth factor human interleukin for DA cells. Nature, 336, 690-2.

Mosley, B., De Imus, C., Friend, D., Boiani, N., Thoma, B., Park, L.S. and Cosman, D. (1996) Dual oncostatin M (OSM) receptors. Cloning and characterization of an alternative signaling subunit conferring OSM-specific receptor activation. J. Biol. Chem., 271, 32635-43.

Moy, F.J., Li, Y.C., Rauenbuehler, P., Winkler, M.E., Scheraga, H.A. and Montelione, G.T. (1993) Solution structure of human type-alpha transforming growth factor determined by heteronuclear NMR spectroscopy and refined by energy minimization with restraints. Biochemistry, 32, 7334-53.

Muller, Y.A., Christinger, H.W., Keyt, B.A. and de Vos, A.M. (1997) The crystal structure of vascular endothelial growth factor (VEGF) refined to 1.93 A resolution: multiple copy flexibility and receptor binding. Structure, 5, 1325-38.

Muller, Y.A., Ultsch, M.H. and de Vos, A.M. (1996) The crystal structure of the extracellular domain of human tissue factor refined to 1.7 A resolution. J. Mol. Biol., 256, 144-59.

Murakami, M., Hibi, M., Nakagawa, N., Nakagawa, T., Yasukawa, K., Yamanishi, K., Taga, T. and Kishimoto, T. (1993) IL-6-induced homodimerization of gp130 and associated activation of a tyrosine kinase. Science, 260, 1808-10.

Murakami, M., Narazaki, M., Hibi, M., Yawata, H., Yasukawa, K., Hamaguchi, M., Taga, T. and Kishimoto, T. (1991) Critical cytoplasmic region of the interleukin 6 signal transducer gp130 is conserved in the cytokine receptor family. Proc. Natl. Acad. Sci. USA, 88, 11349-53.

Murakami Mori, K., Taga, T., Kishimoto, T. and Nakamura, S. (1995) AIDS-associated Kaposi's sarcoma (KS) cells express oncostatin M (OM)-specific receptor but not leukemia inhibitory factor/OM receptor or interleukin-6 receptor. Complete block of OM-induced KS cell growth and OM binding by anti-gp130 antibodies. J. Clin. Invest., 96, 1319-27.

Nair, B.C., DeVico, A.L., Nakamura, S., Copeland, T.D., Chen, Y., Patel, A., O'Neil, T., Oroszlan, S., Gallo, R.C. and Sarngadharan, M.G. (1992) Identification of a major growth factor for AIDS-Kaposi's sarcoma cells as oncostatin M. Science, 255, 1430-2.

Navaza, J. (1994) Amore - an Automated Package For Molecular Replacement. Acta. Cryst., 50, 157-163.

Neer, E.J. and Smith, T.F. (1996) G protein heterodimers: new structures propel new questions. Cell, 84, 175-8.

Nicola, N.A. (1994) An introduction to the cytokines. In Nicola, N.A. (ed.) Guide book to cytokines and their receptors. Oxford University Press, Oxford.

Oefner, C., D'Arcy, A., Winkler, F.K., Eggimann, B. and Hosang, M. (1992) Crystal structure of human platelet-derived growth factor BB. EMBO J., 11, 3921-6.

Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods in Enzymology, 276, 307-326.

Owczarek, C.M., Layton, M.J., Metcalf, D., Lock, P., Willson, T.A., Gough, N.M. and Nicola, N.A. (1993) Inter-species chimeras of leukaemia inhibitory factor define a major human receptor-binding determinant. EMBO J., 12, 3487-95.

Palacios, R. and Steinmetz, M. (1985) IL3-Dependent Mouse Clones that Express B-220 Surface Antigen, Contain Ig Genes in Germ-Line Configuration, and Generate B Lymphocytes In Vivo. Cell, 41, 727-734.

Panayotatos, N., Everdeen, D., Liten, A., Somogyi, R. and Acheson, A. (1994) Recombinant human CNTF receptor alpha: production, binding stoichiometry, and characterization of its activity as a diffusible factor. Biochemistry, 33, 5813-8.

Panayotatos, N., Radziejewska, E., Acheson, A., Somogyi, R., Thadani, A., Hendrickson, W.A. and McDonald, N.Q. (1995) Localization of functional receptor epitopes on the structure of ciliary neurotrophic factor indicates a conserved, function-related epitope topography among helical cytokines. J. Biol. Chem., 270, 14007-14.

Pandit, J., Bohm, A., Jancarik, J., Halenbeck, R., Koths, K. and Kim, S.H. (1992) Three-dimensional structure of dimeric human recombinant macrophage colony-stimulating factor. Science, 258, 1358-62.

Paonessa, G., Graziani, R., De Serio, A., Savino, R., Ciapponi, L., Lahm, A., Salvati, A.L., Toniatti, C. and Ciliberto, G. (1995) Two distinct and independent sites on IL-6 trigger gp 130 dimer formation and signalling. EMBO J., 14, 1942-51.

Patterson, P.H. (1992) The emerging neuropoietic cytokine family: first CDF/LIF, CNTF and IL-6; next ONC, MGF, GCSF? Curr. Opin. Neurobiol., 2, 94-7.

Paul, J. (1980) Cell and Tissue Culture. Churchill Livingstone, New York.

Petsko, G.A. (1975) Protein crystallography at sub-zero temperatures: Cryoprotective mother liquors for protein crystals. J. Mol. Biol., 96, 381-392.

Piquet Pellorce, C., Grey, L., Mereau, A. and Heath, J.K. (1994) Are LIF and related cytokines functionally equivalent? Exp. Cell. Res., 213, 340-7.

Powers, R., Garrett, D.S., March, C.J., Frieden, E.A., Gronenborn, A.M. and Clore, G.M. (1993) The high-resolution, three-dimensional solution structure of human interleukin-4

determined by multidimensional heteronuclear magnetic resonance spectroscopy. Biochemistry, 32, 6744-62.

Priestle, J.P., Schar, H.P. and Grutter, M.G. (1988) Crystal structure of the cytokine interleukin-1 beta. EMBO J., 7, 339-43.

Purvis, D.H. and Mabbutt, B.C. (1997) Solution dynamics and secondary structure of murine leukemia inhibitory factor: a four-helix cytokine with a rigid CD loop. Biochemistry, 36, 10146-54.

Radhakrishnan, R., Walter, L.J., Subramaniam, P.S., Johnson, H.M. and Walter, M.R. (1999) Crystal structure of ovine interferon-tau at 2.1 A resolution. J. Mol. Biol., 286, 151-62.

Read, R.J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. Acta. Cryst.., A42, 140.

Reed, C., Fu, Z.Q., Wu, J., Xue, Y.N., Harrison, R.W., Chen, M.J. and Weber, I.T. (1997) Crystal structure of TNF-alpha mutant R31D with greater affinity for receptor R1 compared with R2. Protein Eng., 10, 1101-7.

Reeves, R. and Cserjesi, P. (1979) Sodium butyrate induces new gene expression in Friend erythroleukemic cells. J. Biol. Chem., 254, 4283-90.

Reidhaar-Olson, J.F., De Souza-Hart, J.A. and Selick, H.E. (1996) Identification of residues critical to the activity of human granulocyte colony-stimulating factor. Biochemistry, 35, 9034-41.

Remy, I., Wilson, I.A. and Michnick, S.W. (1999) Erythropoietin receptor activation by a ligand-induced conformation change. Science, 283, 990-3.

Rhodes, G. (1993) Crystallography made crystal clear. Academic Press, New York.

Richards, C.D., Brown, T.J., Shoyab, M., Baumann, H. and Gauldie, J. (1992) Recombinant oncostatin M stimulates the production of acute phase proteins in HepG2 cells and rat primary hepatocytes in vitro. J. Immunol., 148, 1731-6.

Richardson. (1981) Adv. Prot. Chem, 34, 167-330.

Robinson, R.C., Grey, L.M., Staunton, D., Vankelecom, H., Vernallis, A.B., Moreau, J.F., Stuart, D.I., Heath, J.K. and Jones, E.Y. (1994) The crystal structure and biological function of leukemia inhibitory factor: implications for receptor binding. Cell, 77, 1101-16.

Robinson, R.C., Heath, J.K. and Jones, E.Y. (2000) Manuscript in preparation.

Robledo, O., Fourcin, M., Chevalier, S., Guillet, C., Auguste, P., Pouplard Barthelaix, A., Pennica, D. and Gascan, H. (1997) Signaling of the cardiotrophin-1 receptor. Evidence for a third receptor component. J. Biol. Chem., 272, 4855-63.

Rose, T.M., Lagrou, M.J., Fransson, I., Werelius, B., Delattre, O., Thomas, G., de Jong, P.J., Todaro, G.J. and Dumanski, J.P. (1993) The genes for oncostatin M (OSM) and leukemia inhibitory factor (LIF) are tightly linked on human chromosome 22. Genomics, 17, 136-40.

Rose, T.M., Weiford, D.M., Gunderson, N.L. and Bruce, A.G. (1994) Oncostatin M (OSM) inhibits the differentiation of pluripotent embryonic stem cells in vitro. Cytokine, 6, 48-54.

Rozakis Adcock, M. and Kelly, P.A. (1991) Mutational analysis of the ligand-binding domain of the prolactin receptor. J. Biol. Chem., 266, 16472-7.

Sakabe, N. (1991) X-ray diffraction data collection system for modern protein crystallography with a Weissenberg camera and imaging plate using synchrotron radiation. Nuclear Inst. Meth. Phys. Res., A303, 448-463.

Samudzi, C.T., Burton, L.E. and Rubin, J.R. (1991) Crystal structure of recombinant rabbit interferon-gamma at 2.7-A resolution. J. Biol. Chem., 266, 21791-7.

Sanchez, R. and Sali, A. (1997) Evaluation of comparative protein structure modeling by MODELLER-3. Proteins-Structure Function and Genetics, 50-58.

Savino, R., Lahm, A., Giorgio, M., Cabibbo, A., Tramontano, A. and Ciliberto, G. (1993) Saturation mutagenesis of the human interleukin 6 receptor-binding site: implications for its three-dimensional structure. Proc. Natl. Acad. Sci. USA, 90, 4067-71.

Scheufler, C., Sebald, W. and Hulsmeyer, M. (1999) Crystal structure of human bone morphogenetic protein-2 at 2.7 A resolution. J. Mol. Biol., 287, 103-15.

Schneider, I. (1972) Cell lines derived from late embryonic stages of Drosophila melanogaster. J. Embryol. Exp. Morphol., 27, 353-65.

Schreuder, H., Tardif, C., Trump-Kallmeyer, S., Soffientini, A., Sarubbi, E., Akeson, A., Bowlin, T., Yanofsky, S. and Barrett, R.W. (1997) A new cytokine-receptor binding mode revealed by the crystal structure of the IL-1 receptor with an antagonist. Nature, 386, 194-200.

Scott, J. (1996) New chapter for the fat controller. Nature, 376, 113-114.

Senda, T., Saitoh, S. and Mitsui, Y. (1995) Refined crystal structure of recombinant murine interferon-beta at 2.15 A resolution. J. Mol. Biol., 253, 187-207.

Simmons, D.C. (1992) Cloning cell surface molecules by transient expression in mammalian cells. Cellular interactions in development. IRL Press.

Simpson, R.J., Hammacher, A., Smith, D.K., Matthews, J.M. and Ward, L.D. (1997a) Interleukin-6: structure-function relationships. Protein Sci., 6, 929-55.

Simpson, R.J., Hammcher, A., Smith, D.K., Matthews, J.M. and Ward, L.D. (1997b) Interleukin-6: Structure-function relationships. Protein Sci., 6, 929-955.

Sims, J.E., March, C.J., Cosman, D., Widmer, M.B., MacDonald, H.R., McMahan, C.J., Grubin, C.E., Wignall, J.M., Jackson, J.L., Call, S.M. and et al. (1988) cDNA expression cloning of the IL-1 receptor, a member of the immunoglobulin superfamily. Science, 241, 585-9.

Skelton, N.J., Aspiras, F., Ogez, J. and Schall, T.J. (1995) Proton NMR assignments and solution conformation of RANTES, a chemokine of the C-C type. Biochemistry, 34, 5329-42.

Skelton, N.J., Quan, C., Reilly, D. and Lowman, H. (1999) Structure of a CXC chemokine-receptor fragment in complex with interleukin-8. Structure, 7, 157-68.

Smith, D.B. and Johnson, K.S. (1988) Single step purification of polypeptides expressed in Eschericia coli as fusions with glutathione S-transferase. Gene, 67, 31-40.

Smith, L.J., Redfield, C., Smith, R.A., Dobson, C.M., Clore, G.M., Gronenborn, A.M., Walter, M.R., Naganbushan, T.L. and Wlodawer, A. (1994) Comparison of four independently determined structures of human recombinant interleukin-4. Nat. Struct. Biol., 1, 301-10.

Sogabe, S., Stuart, F., Henke, C., Bridges, A., Williams, G., Birch, A., Winkler, F.K. and Robinson, J.A. (1997) Neutralizing epitopes on the extracellular interferon gamma receptor (IFNgammaR) alpha-chain characterized by homolog scanning mutagenesis and X-ray crystal structure of the A6 fab-IFNgammaR1-108 complex. J. Mol. Biol., 273, 882-97.

Somers, W., Stahl, M. and Seehra, J.S. (1997) 1.9 A crystal structure of interleukin 6: implications for a novel mode of receptor dimerization and signaling. EMBO J., 16, 989-97.

Somers, W., Ultsch, M., De Vos, A.M. and Kossiakoff, A.A. (1994) The X-ray structure of a growth hormone-prolactin receptor complex. Nature, 372, 478-81.

Sporeno, E., Paonessa, G., Salvati, A.L., Graziani, R., Delmastro, P., Ciliberto, G. and Toniatti, C. (1994) Oncostatin M binds directly to gp130 and behaves as interleukin-6 antagonist on a cell line expressing gp130 but lacking functional oncostatin M receptors. J. Biol. Chem., 269, 10991-5.

Stanley, P. and Carver, J.P. (1977) Lectin receptors and lectin resistance in chinese hamster ovary cells. Adv. Exp. Med. Biol., 84, 265-84.

Staunton, D., Hudson, K.R. and Heath, J.K. (1998) The interactions of the cytokine-binding homology region and immunoglobulin-like domains of gp130 with oncostatin M: implications for receptor complex formation. Protein Eng., 11, 1093-102.

Stewart, C.L., Kaspar, P., Brunet, L.J., Bhatt, H., Gadi, I., Kontgen, F. and Abbondanzo, S.J. (1992) Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature, 359, 76-9.

Stuart, D.I., Levine, M., Muirhead, H. and Stammers, D.K. (1979) Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 A. J. Mol. Biol., 134, 109-42.

Stura, E.A. and Wilson, I.A. (1992) Seeding techniques. Crystallization of nucleic acids and proteins: A practical approach. Oxford University Press, pp. 99-125.

Syed, R.S., Reid, S.W., Li, C., Cheetham, J.C., Aoki, K.H., Liu, B., Zhan, H., Osslund, T.D., Chirino, A.J., Zhang, J., Finer-Moore, J., Elliott, S., Sitney, K., Katz, B.A., Matthews, D.J., Wendoloski, J.J., Egrie, J. and Stroud, R.M. (1998) Efficiency of signalling through cytokine receptors depends critically on receptor orientation. Nature, 395, 511-6.

Taga, T. and Kishimoto, T. (1992) Cytokine receptors and signal transduction. FASEB J., 6, 3387-96.

Taga, T. and Kishimoto, T. (1995) Signaling mechanisms through cytokine receptors that share signal transducing receptor components. Curr. Opin. Immunol., 7, 17-23.

Tarentino, A.L., Quinones, G., Trumble, A., Changchien, L.M., Duceman, B., Maley, F. and Plummer, T.H., Jr. (1990) Molecular cloning and amino acid sequence of peptide-N4-(N-acetyl-beta-D-glucosaminyl)asparagine amidase from flavobacterium meningosepticum. J. Biol. Chem., 265, 6961-6.

Ten Eyck, L.F. (1973) Crystallographic fast Fourier transforms. Acta. Cryst., A29, 183.

Teng, T.-Y. (1990) Mounting of crystals for macromolecular crystallography in a free-standing thin film. J. Appl. Cryst., 23, 387-391.

Terwilliger, T.C. and Berendzen, J. (1999) Automated MAD and MIR structure solution. Acta. Cryst., 55, 849-861.

Thoma, B., Bird, T.A., Friend, D.J., Gearing, D.P. and Dower, S.K. (1994) Oncostatin M and leukemia inhibitory factor trigger overlapping and different signals through partially shared receptor complexes. J. Biol. Chem., 269, 6215-22.

Ullrich, A. and Schlessinger, J. (1990) Signal transduction by receptors with tyrosine kinase activity. Cell, 61, 203-12.

Ultsch, M., Lokker, N.A., Godowski, P.J. and de Vos, A.M. (1998) Crystal structure of the NK1 fragment of human hepatocyte growth factor at 2.0 A resolution. Structure, 6, 1383-93.

Vigers, G.P., Anderson, L.J., Caffes, P. and Brandhuber, B.J. (1997) Crystal structure of the type-I interleukin-1 receptor complexed with interleukin-1beta. Nature, 386, 190-4.

Walker, P.A., Leong, L.E., Ng, P.W., Tan, S.H., Waller, S., Murphy, D. and Porter, A.G. (1994) Efficient and rapid affinity purification of proteins using recombinant fusion proteases. Biotechnology NY, 12, 601-5.

Walter, M.R., Cook, W.J., Ealick, S.E., Nagabhushan, T.L., Trotta, P.P. and Bugg, C.E. (1992a) Three-dimensional structure of recombinant human granulocyte-macrophage colony-stimulating factor. J. Mol. Biol., 224, 1075-85.

Walter, M.R., Cook, W.J., Zhao, B.G., Cameron, R.P., Jr., Ealick, S.E., Walter, R.L., Jr., Reichert, P., Nagabhushan, T.L., Trotta, P.P. and Bugg, C.E. (1992b) Crystal structure of recombinant human interleukin-4. J. Biol. Chem., 267, 20371-6.

Ward, L.D., Hammacher, A., Howlett, G.J., Matthews, J.M., Fabri, L., Moritz, R.L., Nice, E.C., Weinstock, J. and Simpson, R.J. (1996) Influence of interleukin-6 (IL-6) dimerization on formation of the high affinity hexameric IL-6.receptor complex. J. Biol. Chem., 271, 20138-44.

Ward, L.D., Howlett, G.J., Discolo, G., Yasukawa, K., Hammacher, A., Moritz, R.L. and Simpson, R.J. (1994) High affinity interleukin-6 receptor is a hexameric complex consisting of two molecules each of interleukin-6, interleukin-6 receptor, and gp-130. J. Biol. Chem., 269, 23286-9.

Ward, L.D., Howlett, G.J., Hammacher, A., Weinstock, J., Yasukawa, K., Simpson, R.J. and Winzor, D.J. (1995) Use of a biosensor with surface plasmon resonance detection for the determination of binding constants: measurement of interleukin-6 binding to the soluble interleukin-6 receptor. Biochemistry, 34, 2901-7.

Waterman, M.S., Smith, T.F. and Katcher, H.L. (1984) Algorithms for Restriction Map Comparison. Nucleic Acids Res., 12, 237-242.

Watowich, S.S., Yoshimura, A., Longmore, G.D., Hilton, D.J., Yoshimura, Y. and Lodish, H.F. (1992) Homodimerization and constitutive activation of the erythropoietin receptor. Proc. Natl. Acad. Sci. USA, 89, 2140-4.

Wiesmann, C., Fuh, G., Christinger, H.W., Eigenbrot, C., Wells, J.A. and de Vos, A.M. (1997) Crystal structure at 1.7 A resolution of VEGF in complex with domain 2 of the Flt-1 receptor. Cell, 91, 695-704.

Wilken, J., Hoover, D., Thompson, D.A., Barlow, P.N., McSparron, H., Picard, L., Wlodawer, A., Lubkowski, J. and Kent, S.B. (1999) Total chemical synthesis and high-resolution crystal structure of the potent anti-HIV protein AOP-RANTES. Chem. Biol., 6, 43-51.

Wilkie, A.O., Morriss Kay, G.M., Jones, E.Y. and Heath, J.K. (1995) Functions of fibroblast growth factors and their receptors. Curr. Biol., 5, 500-7.

Wlodawer, A., Pavlovsky, A. and Gustchina, A. (1992) Crystal structure of human recombinant interleukin-4 at 2.25 A resolution. FEBS Lett., 309, 59-64.

Wong, G.G., Witek, J.S., Temple, P.A., Wilkens, K.M., Leary, A.C., Luxenberg, D.P., Jones, S.S., Brown, E.L., Kay, R.M., Orr, E.C. and et al. (1985) Human GM-CSF: molecular cloning of the complementary DNA and purification of the natural and recombinant proteins. Science, 228, 810-5.

Woolfson, M.M., Yao, J.X. and Fan, H.F. (1997) New techniques for applying anomalous-scattering and isomorphous-replacement data incorporated in ANOMIR - a general application package. Acta. Cryst., 53, 673-681.

Wrighton, N.C., Farrell, F.X., Chang, R., Kashyap, A.K., Barbone, F.P., Mulcahy, L.S., Johnson, D.L., Barrett, R.W., Jolliffe, L.K. and Dower, W.J. (1996) Small peptides as potent mimetics of the protein hormone erythropoietin. Science, 273, 458-64.

Wu, H., Lustbader, J.W., Liu, Y., Canfield, R.E. and Hendrickson, W.A. (1994) Structure of human chorionic gonadotropin at 2.6 A resolution from MAD analysis of the selenomethionyl protein. Structure, 2, 545-58.

Xu, G.Y., Yu, H.A., Hong, J., Stahl, M., McDonagh, T., Kay, L.E. and Cumming, D.A. (1997) Solution structure of recombinant human interleukin-6. J. Mol. Biol., 268, 468-81.

Yamasaki, K., Naito, S., Anaguchi, H., Ohkubo, T. and Ota, Y. (1997) Solution structure of an extracellular domain containing the WSxWS motif of the granulocyte colony-stimulating factor receptor and its interaction with ligand. Nat. Struct. Biol., 4, 498-504.

Yawata, H., Yasukawa, K., Natsuka, S., Murakami, M., Yamasaki, K., Hibi, M., Taga, T. and Kishimoto, T. (1993) Structure-function analysis of human IL-6 receptor: dissociation of amino acid residues required for IL-6-binding and for IL-6 signal transduction through gp130. EMBO J., 12, 1705-12.

Zarling, J.M., Shoyab, M., Marquardt, H., Hanson, M.B., Lioubin, M.N. and Todaro, G.J. (1986) Oncostatin M: a growth regulator produced by differentiated histiocytic lymphoma cells. Proc. Natl. Acad. Sci. USA, 83, 9739-43.

Zdanov, A., Schalk Hihi, C., Gustchina, A., Tsang, M., Weatherbee, J. and Wlodawer, A. (1995) Crystal structure of interleukin-10 reveals the functional dimer with an unexpected topological similarity to interferon gamma. Structure, 3, 591-601.

Zhang, F., Basinski, M.B., Beals, J.M., Briggs, S.L., Churgay, L.M., Clawson, D.K., DiMarchi, R.D., Furman, T.C., Hale, J.E., Hsiung, H.M., Schoner, B.E., Smith, D.P., Zhang, X.Y., Wery, J.P. and Schevitz, R.W. (1997a) Crystal structure of the obese protein leptin-E100. Nature, 387, 206-9.

Zhang, J.G., Farley, A., Nicholson, S.E., Willson, T.A., Zugaro, L.M., Simpson, R.J., Moritz, R.L., Cary, D., Richardson, R., Hausmann, G., Kile, B.J., Kent, S.B., Alexander, W.S., Metcalf, D., Hilton, D.J., Nicola, N.A. and Baca, M. (1999) The conserved SOCS box motif in suppressors of cytokine signaling binds to elongins B and C and may couple bound proteins to proteasomal degradation. Proc. Natl. Acad. Sci. USA, 96, 2071-6.

Zhang, J.G., Owczarek, C.M., Ward, L.D., Howlett, G.J., Fabri, L.J., Roberts, B.A. and Nicola, N.A. (1997b) Evidence for the formation of a heterotrimeric complex of leukaemia inhibitory factor with its receptor subunits in solution. Biochem. J., 325, 693-700.