Synthesis of neo-glycosylated human erythropoietin and investigation into interaction with the erythropoietin receptor

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy (PhD) in the Faculty of Engineering and Physical Sciences

Dominique Richardson

2013

School of Chemistry
Manchester Institute of Biotechnology
Contents

ABSTRACT 8

DECLARATION and COPYRIGHT STATEMENT 9

Abbreviations 10

ACKNOWLEDGMENTS 19

1 Introduction 21
  1.1 Therapeutic proteins ........................................... 21
  1.2 The role glycosylation in therapeutic proteins ............... 22
    1.2.1 Biosynthesis of glycoproteins ............................ 22
    1.2.2 The role of glycosylation in physiology and disease .... 25
    1.2.3 Glycosylation as a tool in protein-based therapeutics ... 27
  1.3 Additional modification strategies for improving therapeutic protein efficiencies ........................................ 29
  1.4 Expression platforms for pharmaceutical proteins ............ 31
  1.5 Chemical and enzymatic synthesis of glycoproteins .......... 34
    1.5.1 Native Chemical Ligation (NCL) and Expressed Protein Ligation (EPL) ........................................ 35
    1.5.2 Cysteine free native chemical ligation .................... 36
    1.5.3 Modification of peptide backbones ........................ 37
    1.5.4 Summary ...................................................... 39
  1.6 Erythropoietin .................................................. 40
    1.6.1 Erythropoietin: erythropoiesis and tissue protection .... 40
    1.6.2 Structure of erythropoietin ............................... 40
## Contents

1.6.3 Erythropoiesis: Regulation and mode of action .................................. 41  
1.6.4 Erythropoietin mediated tissue protection: Regulation and mode of action ................................................................. 46  
1.6.5 Effect of glycosylation on EPO function and activity ......................... 48  
1.6.6 The effects of modifications on the therapeutic use of Erythropoietin ............................................................... 49  
1.6.7 Methods of EPO detection and analysis ........................................... 57  
1.6.8 Summary ......................................................................................... 59  
1.7 Aims and Objectives ........................................................................... 61

2 Sub-cloning and expression of WT EPO and EPO cysteine variants 63
  2.1 Sub-cloning of WT EPO and EPO cysteine variant genes ...................... 63  
    2.1.1 Sub-cloning of WT EPO gene into pET16b vector ....................... 63  
    2.1.2 Introducing cysteine mutations at key EPO-(EPOR)\textsubscript{2} binding sites ........................................................................... 65  
  2.2 Expression of WT EPO and EPO cysteine variants .............................. 70  
    2.2.1 Optimisation of WT EPO gene expression ................................. 70  
    2.2.2 Expression and purification of EPO cysteine variants ................. 71  
  2.3 Oxidative refolding of WT EPO and EPO cysteine variants .................. 76  
    2.3.1 Pulse refolding of EPO .............................................................. 76  
    2.3.2 Circular dichroism .................................................................... 77  
  2.4 Summary ........................................................................................... 77

3 Glycosylation of the EPO cysteine variants 80
  3.1 Introduction ....................................................................................... 80  
  3.2 The use of Liquid chromatography–mass spectrometry (LCMS) to monitor the glycosylation of EPO ....................................... 82  
    3.2.1 Optimising the LCMS conditions .......................................... 83  
  3.3 Reaction of EPO cysteine variants with α-mannosyl iodoacetamide .... 85  
    3.3.1 Optimisation of the glycosylation reaction ............................... 87  
    3.3.2 Improving the selectivity of the EPO glycosylation reaction ....... 92  
  3.4 Reaction of double cysteine EPO variants with α-mannosyl iodoacetamide ................................................................. 96
4 Purification of the EPO glycosylation reaction

4.1 Introduction ....................................................... 99
4.2 Purification of the EPO glycosylation reaction using thiol selective column chromatography .......................................................... 102
  4.2.1 Binding of Green Fluorescent Protein to the Activated thiol sepharose .......................................................... 102
  4.2.2 Binding of EPO variants to the activated thiol sepharose ...... 105
4.3 Purification of the EPO glycosylation reaction using hydrophobic interaction chromatography ................................. 108
4.4 Summary ............................................................... 110

5 Development of an assay to monitor EPO-(EPOR)\(_2\) binding interactions

5.1 Protein interaction detection methods ...................................... 112
5.2 Detecting the EPO-(EPOR)\(_2\) interaction .................................. 116
5.3 Optimisation of the EPO-(EPOR)\(_2\) binding assay ......................... 121
5.4 Optimisation of surfactant concentration .................................... 123
5.5 Optimisation of blocking agent ............................................. 124
5.6 Summary ............................................................... 125

6 The effect of site specific glycan addition to EPO on EPO-(EPOR)\(_2\) binding

6.1 Introduction .............................................................. 127
6.2 Screening EPO variants by ELISA ........................................ 129
  6.2.1 Generating a dose–response curve for the binding of WT EPO to (EPOR)\(_2\) ......................................................... 129
  6.2.2 Screening EPO variants ............................................. 130
6.3 Reaction of WT EPO with α-mannosyl iodoacetamine and analysis by ELISA ......................................................... 136
6.4 Analysis of mannosylated K45C+K97C EPO variant on the ELISA ... 137
  6.4.1 Initial analysis of mannosylated EPO by ELISA ................. 138
  6.4.2 Procedure optimisation ............................................. 140
6.4.3 Final analysis of mannosylated EPO by ELISA .............. 142
6.5 Summary ................................................................. 146

7 Conclusions and outlook ............................................... 148

8 Experimental detail ....................................................... 153
  8.1 General experimental detail ........................................... 153
    8.1.1 Chemicals and Reagents ........................................... 153
    8.1.2 List of buffers .................................................... 153
    8.1.3 Oligonucleotide Primers ......................................... 154
    8.1.4 Gene Sequences ................................................ 155
  8.2 Experimental detail for chapter 2 .................................. 158
    8.2.1 Cloning of WT EPO gene into pET16b ............................ 158
    8.2.2 Site directed mutagenesis to produce EPO cysteine variants . 161
    8.2.3 Cloning of WT EPO gene into pPICZαA .......................... 162
    8.2.4 General procedure for expression of WT EPO and EPO cysteine variants ......................................................... 164
    8.2.5 General method for the purification and refolding of EPO variants ................................................................. 165
    8.2.6 General method for the refolding of EPO .......................... 166
  8.3 Experimental detail for chapter 3 .................................... 167
    8.3.1 Detecting EPO by LCMS ........................................... 167
    8.3.2 Chemical glycosylation of EPO variants with α-mannosyl iodoacetamide ................................................................. 168
  8.4 Experimental detail for chapter 4 .................................... 169
    8.4.1 General procedure for purification using Activated Thiol Sepharose 169
    8.4.2 General procedure for purification using hydrophobic interaction chromatography (HIC) ......................................................... 169
  8.5 Experimental detail for chapter 5 and chapter 6 ................... 171
    8.5.1 General EPO-(EPOR)_2 binding assay procedure .................. 171
    8.5.2 Preparation of non-modified EPO samples for screening by ELSIA ................................................................. 172
8.5.3 Preparation of glycosylated EPO samples for analysis by ELISA ........................................ 173
8.5.4 Second generation protocol: preparation of glycosylated EPO
samples for analysis by ELISA ........................................ 173

References .................................................................. 175
Abstract

Synthesis of neo-glycosylated human erythropoietin and investigation into the effect this modification has on the interaction of EPO with its erythropoietic receptor

Dominique Richardson, PhD. The University of Manchester, 2013

Erythropoietin (EPO) is the glycoprotein hormone responsible for regulating the production of erythrocytes. EPO also plays a cytoprotective role in a variety of tissues, including the brain and heart, by preventing apoptosis of healthy cells after ischemic injury (tissue protection). Recombinant human EPO (rhEPO) used clinically to treat anaemia is unsuitable for administration as a tissue protective agent due to the adverse side effects associated with overstimulation of erythropoiesis. As such, there have been a number of attempts to develop a tissue protective EPO derivative by eliminating erythropoietic function. Reported here is the synthesis of neo-glycosylated human erythropoietin and investigation into the effect this modification has on the interaction of EPO with its erythropoietic receptor (EPOR)₂.

In order to develop a non-erythropoietic EPO derivative, cysteine residues were introduced into the EPO sequence to act as chemical modification sites. EPO variants were then glycosylated at the cysteine residues using a one-step synthesis with α-mannosyl iodoacetamide before analysis by ELISA. Residues located within the two binding sites of EPO were targeted for mutation, and a total of 13 EPO variants were generated and expressed from E. coli. Upon expression and purification, EPO cysteine variants were glycosylated with α-mannosyl iodoacetamide. Reaction of single EPO variants yielded a mixture of unmodified and mono-glycosylated EPO species. Reaction of the double cysteine EPO variants yielded a mixture of three species; the unmodified EPO, the mono-glycosylated EPO and the di-glycosylated EPO. To determine the effect of this glycosylation on the EPO-(EPOR)₂ interactions, an ELISA based assay was developed. Initially, all EPO variants were screened by ELISA for WT-like binding to the (EPOR)₂. Of the 13 variants screened only the K45C+K97C EPO variant was used in the proof-of-concept study. Comparison of the WT EPO before and after reaction with the α-mannosyl iodoacetamide on the ELISA showed no inhibition of the EPO-(EPOR)₂ interactions while comparison of the K45C+K97C EPO before and after reaction showed a statistically significant difference in the EPO-(EPOR)₂ binding. This outcome indicates that it is possible to inhibit the EPO-(EPOR)₂ binding by introducing non-natural glycosylation sites into the EPO sequence.
DECLARATION and COPYRIGHT STATEMENT

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning. The following four notes on copyright and the ownership of intellectual property rights must be included as written below:

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
Abbreviations

**asialoEPO**  Desialylated rhEPO derivative

**βcR**  beta common receptor

**BSA**  Bovine serum albumin

**CEPO**  Carbamylated EPO; EPO with all eight lysine carbamylated

**CV**  Column volume

**DLS**  Dynamic light scattering

**dNTP’s**  Deoxyribonucleotide’s

**DTNB**  5,5’-dithiobis-(2-nitrobenzoic acid)

**ELISA**  Enzyme-linked immunosorbent assay

**EPO**  Erythropoietin

**EPOR**  Monomer of the erythropoietin receptor found on the surface of erythrocyte progenitor cells

**(EPOR)_2**  Pre-associated erythropoietin receptor dimer found on the surface of erythrocyte progenitor cells

**EPO_variants**  All EPO genes produced/ purchased for use in this thesis are labelled with the amino acid variant and position, e.g. K45C EPO corresponds to EPO with the lysine at position 45 replaced by a cysteine residue. See section 8.1.4

**ER**  Endoplasmic reticulum

**ESI**  Electrospray ionization

**GFP**  Green fluorescent protein
HIC  Hydrophobic Interaction Chromatography
His-tag  Polyhistidine-tag
HPLC  High-performance liquid chromatography
HRP  Horseradish peroxidase
IPTG  Isopropyl β-D-1-thiogalactopyranoside
JAK2  Janus kinase 2 kinases
LCMS  Liquid chromatography–mass spectrometry
MAPK  Rasmitogen-activated protein kinase
MS  Mass spectrometry
PBS  Phosphate buffered saline
PCR  Polymerase chain reaction
PEG  Polyethylene glycol
PI3K  Phosphatidylinositol 3-kinase
PTM  Posttranslational modification
rhEPO  Commercially produced recombinant human Erythropoietin used in the treatment of anaemia
SDS  Sodium dodecyl sulfate
SDS-PAGE  Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SPR  Surface plasmon resonance
STAT5  Signal Transducer and Activator of Transcription
TMB  3,3’,5,5’-tetramethylbenzidine
**TNF-α**  Tumor necrosis factor alpha

**TPR**  Erythropoietin receptor responsible for tissue protective activity

**WT_EPO**  The WT EPO sequence with an N-terminal His-tag (see section 8.1.4)
## List of Figures

1.1 Biosynthesis of $N$-linked glycoproteins ............................................. 23
1.2 Structures of glycan chains found on $N$-linked glycoproteins ..................... 24
1.3 Common $O$-linked (mucin type) glycan core structures found in mammals ........ 25
1.4 Glycan structures of Siayl-Lewis-X antigen and ABO blood groups .................. 26
1.5 Examples of PEGylating compounds .................................................... 30
1.6 Glycan chains observed when utilising different glycoprotein expression systems.
   Yeast expression systems yield high mannose type glycan chains. Human and
   animal expression systems yield very similar glycan structures; however animal
   expression systems yield both Neu5Ac and Neu5Gc moieties, while humans have
   lost the ability to synthesise Neu5Gc$^1$. .................................................. 33
1.7 Peptide coupling by native chemical ligation ........................................... 35
1.8 Cysteine free ligation methods ........................................................... 37
1.9 Lansbury Aspartylation ....................................................................... 38
1.10 Amino acid sequence of rhEPO ............................................................... 41
1.11 Crystal structure of EPO-(EPOR)$_2$ complex .......................................... 42
1.12 Stages of erythrocyte development ......................................................... 42
1.13 Mechanism of erythropoiesis ................................................................. 43
1.14 Signal regulation of erythropoiesis .......................................................... 44
1.15 Schematic of EPO bound to (EPOR)$_2$. ................................................... 45
1.16 Mechanism of EPOs tissue protective function ......................................... 46
1.17 Proposed structure of the EPO tissue protective receptor ......................... 47
1.18 The crystal structure of EPO with $\alpha$-helix B highlighted ......................... 56
1.19 Methods of EPO detection and analysis .................................................. 59
1.20 Modification of EPO to selectively block EPO-(EPOR)$_2$ interactions while pre-
   serving tissue protective function ........................................................... 61
1.21 Stepwise approach to EPO-(EPOR)\textsubscript{2} inhibition ........................................ 62
2.1 The WT EPO vector construct with EPO gene flanked by NdeI and BglII downstream of the His-tag. ........................................................................................................ 64
2.2 Interacting residues of EPO and (EPOR)\textsubscript{2} in Site 1 ........................................ 66
2.3 Interacting residues of EPO (green) and (EPOR)\textsubscript{2} in Site 2 ................................ 67
2.4 Rounds of mutagenesis to yield single EPO cysteine variants and double EPO cysteine variants ............................................................ 69
2.5 SDS PAGE gel showing time points for the expression of WT EPO ..................... 71
2.6 SDS Gels of the expression of the A: K45C, B: K52C, C: K97C, D: K45C+K52C and E: K45C+K97C EPO gene variants .............................................. 72
2.7 SDS PAGE of Ni-NTA purification of EPO cysteine variants ............................ 73
2.8 SDS PAGE gel of A: truncated expression of the gene optimised WT EPO from GeneArt and B: Complete expression of WT EPO gene from GeneArt .......... 74
2.10 CD spectra ............................................................................................................ 77
2.11 Amino acid sequence of EPO with the His-tag and the positions of the cysteine mutations highlighted ................................................................. 78
3.1 Representation of the reaction of Top: The WT EPO with \(\alpha\)-mannosyl iodoacetamide and Below: Reaction of an EPO cysteine variant with \(\alpha\)-mannosyl iodoacetamide ..................................................................................... 81
3.2 Reaction of Ellman’s reagent with thiol containing compounds .......................... 82
3.3 Multiple-charge ion electrospray ionisation mass spectrum of the WT EPO before reduction with DTT ............................................................... 83
3.4 Mass spectrum of the WT EPO after reduction with DTT .................................. 84
3.5 Mass spectra of a typical glycosylation reaction .................................................. 86
3.6 Mass spectra following the reaction of K97C EPO with \(\alpha\)-mannosyl iodoacetamide under varying conditions after 1 hour ................................. 88
3.7 Mass spectra following the reaction of K97C EPO with \(\alpha\)-mannosyl iodoacetamide under varying conditions after 5 hours ................................. 89
3.8 Mass spectra following the reaction of WT EPO with \(\alpha\)-mannosyl iodoacetamide under varying conditions ......................................................... 91
List of Figures

3.9 Mass spectra following the reaction of K97C EPO with \( \alpha \)-mannosyl iodoacetamide at pH 7.4 ........................................ 93

3.10 Mass spectra of EPO cysteine variants (1 mg ml\(^{-1}\)) after reaction with \( \alpha \)-mannosyl iodoacetamide (1 mM). Spectra show the unmodified protein samples (M) and the mono-glycosylated protein (M+[Man]). ........................................ 95

3.11 Mass spectrum of the reaction of K45C+K97C EPO with \( \alpha \)-mannosyl iodoacetamide ........................................ 97

4.1 Reaction of cysteine EPO variants with \( \alpha \)-mannosyl iodoacetamide ......................... 100

4.2 Purification of EPO variant glycosylation reaction using thiol selective chromatography and lectin affinity chromatography ........................................ 101

4.3 Selective binding of thiol containing proteins. ........................................ 102

4.4 SDS PAGE gel of WT GFP and I229C GFP fractions from thiol sepharose column purification ........................................ 103

4.5 SDS PAGE gel of fractions from the WT and I229C GFP bound to the thiol sepharose column under denaturing conditions ........................................ 105

4.6 SDS PAGE gel of fractions from the WT and K52C EPO bound to the thiol sepharose column under denaturing conditions ........................................ 107

4.7 SDS PAGE gel of fractions from the WT and K52C EPO bound to the thiol sepharose column under denaturing conditions ........................................ 108

4.8 SDS PAGE gel of fractions from the HIC column containing protein ........................................ 109

5.1 Schematic of different ELISA set ups ........................................ 114

5.2 Co-immunoprecipitation method ........................................ 115

5.3 Schematic of EPO-(EPOR)\(_2\) binding assay ........................................ 117

5.4 Absorbance at 652 nm monitored over time for the binding of biotinylated rhEPO to (EPOR)\(_2\) on high binding plates, ........................................ 118

5.5 Layout of the controls used to optimise the ELISA assay ........................................ 120

5.6 Dose-response curve of the WT EPO expressed from \( E. \) coli ........................................ 122

5.7 The effect of n-lauryl sarcosine on the signal and the background noise from the binding assay. ........................................ 124

5.8 Comparison of the efficiencies of different blocking agents on ELISA ........................................ 125

6.1 Comparing dose-response curves for an EPO variant and the mannosylated EPO variant ........................................ 128
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2</td>
<td>Flow diagram outlining the selection process applied to all EPO cysteine variants</td>
</tr>
<tr>
<td>6.3</td>
<td>Dose-response curves of EPO variants discarded from further analysis after initial screening on the ELISA</td>
</tr>
<tr>
<td>6.4</td>
<td>Dose-response curve of S104C EPO as the WT EPO on the ELISA</td>
</tr>
<tr>
<td>6.5</td>
<td>Dose-response curves of EPO variants exhibiting similar binding to the (EPOR)$_2$ as the WT EPO on the ELISA</td>
</tr>
<tr>
<td>6.6</td>
<td>Dose-response curves of the final three EPO variants</td>
</tr>
<tr>
<td>6.7</td>
<td>Dose-response curve for the WT EPO and reacted WT EPO</td>
</tr>
<tr>
<td>6.8</td>
<td>Crystal structure of the WT EPO bound to the homodimer EPO receptor. Binding Site 1 and Site 2 are highlighted to demonstrate the location of the two cysteine residues in the K45C+K97C EPO variant.</td>
</tr>
<tr>
<td>6.9</td>
<td>Dose-response curves of K45C+K97C EPO on ELISA before and after reaction with α-mannosyl iodoacetamide</td>
</tr>
<tr>
<td>6.10</td>
<td>Optimised protocol for the analysis of EPO samples by ELISA</td>
</tr>
<tr>
<td>6.11</td>
<td>Dose-response curve for the WT EPO using the new protocol</td>
</tr>
<tr>
<td>6.12</td>
<td>Procedure for the expression, purification, reaction and analysis of EPO variants.</td>
</tr>
<tr>
<td>6.13</td>
<td>Final dose-response curve for the WT EPO and K45C+K97C EPO after reaction with α-mannosyl iodoacetamide</td>
</tr>
<tr>
<td>7.1</td>
<td>Methods to orientate (EPOR)$_2$ 'upwards' on plate surface.</td>
</tr>
</tbody>
</table>
List of Tables

1.1 ESA’s available within the US and the EU ................................. 49
1.1 ESA’s available within the US and the EU ................................. 50
1.1 ESA’s available within the US and the EU ................................. 51
1.2 List of published EPO mutations and effects on specific bioactivity . 53
1.2 List of published EPO mutations and effects on specific bioactivity . 54

2.1 List of interactions between EPO and (EPOR)$_2$ in Site 1 and Site 2 at < 3.5Å distance. .......................................................... 66
2.2 List of residues in the EPO sequence targeted for mutation to cysteine residues . 67
2.3 List of cysteine mutations successfully introduced into the EPO gene sequence . 70
2.4 List of EPO cysteine variants successfully expressed in E. coli. .............. 79

3.1 Observed and expected masses of all EPO cysteine variants after LCMS analysis 85

5.1 Elements of the experiments used to optimise the ELISA assay ............. 119

6.1 Components of reactions run in parallel to determine the effect of glycosylation on EPO-(EPOR)$_2$ interactions. ................................. 139
I would like to thank Professor Sabine Flitsch for the opportunity to work within her research group and for all of the help and guidance over the last four years. The Flitsch group is a wonderful place to work and I am lucky to have had the opportunity to be here. Thanks to Oxyrane for their support of the project, and thanks to Professor Stanley Roberts for his time and enthusiasm during our meetings. Thanks to Andrew Martin for his friendship (and for making all of the compounds in this Thesis).

To the Turner-Flitsch group; thanks for making the MIB a brilliant place to work! You are all lovely people who are always willing to lend a hand; Mark and Mirja, thanks for reading my thesis! I have enjoyed my time in the group and met so many amazing people! Martin, thanks for introducing me to German Breakfast, and thanks to Bas and Paul for Friday night drinks.

Special thanks to Anthony ‘BigTony’ Green, Sarah ‘Trixie’ Lovelock, Shahed ‘Chad’ Hussain and Cesar ‘Brad’ Iglesias for being brilliant friends!

All my thanks to Mat for always being supportive, generally amazing and for putting up with me. Finally, thanks to my family; Mum, Dad, and Sister! You are always so supportive (annoying) and encourage me to do my best. You have all helped make me into the person I am today (even if you think I’m useless) and I love you all so much.
For my Family and especially my Big Sister Kelly, who I will always
look up to (even though she is smaller 😊)

\textbf{Figure 0:} Team Richardson on an adventure
1 Introduction

1.1 Therapeutic proteins

Protein-based therapeutic agents account for an increasing percentage of the biopharmaceutical market; worldwide sales of protein-based therapeutics grossed US$90-108 billion in 2010 and are predicted to reach US$140 billion by 2017\textsuperscript{2}. Protein-based therapeutic agents encompass a diverse range of biomolecules including antibodies, large glycoproteins and small therapeutic peptides. Around 50% of the revenue was generated by sales of monoclonal antibodies, with the remaining sales attributed to varying cytokines, hormones, colony stimulating factors, erythropoietins, insulins and more\textsuperscript{3}. In first generation biopharmaceuticals many of these protein-based drugs exhibited suboptimal efficiencies, half-lives, and \textit{in vivo} and \textit{in vitro} stabilities\textsuperscript{4,5}.

As such, several strategies have been developed to improve certain characteristics of current therapeutics by engineering their physiochemical and pharmacological properties. Modification of therapeutic proteins including glycosylation, PEGylation and mutation, as well as alternative delivery methods including polymeric nanoparticulates, micelle formation and biodegradable polymers have proven to be effective methods for optimising the pharmacological efficiencies and activities of a second generation of therapeutic proteins. Of these modifications, glycosylation (and to a lesser extent PEGylation) is the most common method of peptide modification\textsuperscript{6,7}. 

21
1.2 The role glycosylation in therapeutic proteins

Glycosylation refers to the modification of protein surface with carbohydrate moieties (glycans). These glycans influence the activity and stability of glycoproteins and are responsible for mediating a variety of biological processes.\textsuperscript{4,8,9}

Introducing / increasing glycosylation on therapeutic proteins has been shown to improve pharmacological properties including: protein stability, half-life (including a decrease in clearance and proteolytic degradation), \textit{in vivo} and \textit{in vitro} stability and activity.\textsuperscript{10,21} As such, there are a number of established and emerging methods for homogenous glycoprotein and glycopeptide synthesis, some of which will be discussed here.

1.2.1 Biosynthesis of glycoproteins

\textbf{Biosynthesis of }$N$\textbf{-linked glycosylation}

$N$-linked glycosylation is the co- and post-translational modification of proteins which occurs in the endoplasmic reticulum (ER). $N$-linked glycosylation occurs on asparagine (Asn) side chains where the Asn is in the sequence Asn-Xxx-Ser/Thr. Xxx can be any amino acid except proline, however some amino acids are preferred over others. The diversity and specificity found in glycosyltransferases results in highly heterogeneous $N$-linked glycosylation. However, all eukaryotic $N$-linked glycans share a conserved core structure; Man$_3$GlcNAc$_2$(Figure 1.1).

Assembly of this core $N$-linked glycan structure occurs as follows:

\textbf{Stage 1: Assembly of a precursor glycan structure:} A dolichol phosphate linked oligosaccharide precursor is sequentially synthesised from nucleotide sugar donors by various glycosyltransferases. This precursor is ‘flipped’ across the endoplasmic reticulum (ER) membrane to the luminal side where a further 7 sugar moieties are attached. This lipid linked 14-mer oligosaccharide is used as a glycan donor.
1.2 The role glycosylation in therapeutic proteins

Figure 1.1: Biosynthesis of N-linked glycoproteins. The dolichol phosphate linked 7-mer glycan precursor is synthesised, ‘flipped’ across the endoplasmic reticulum (ER) membrane to the lumen where a further 7 sugar moieties are attached. Oligosaccharyltransferase (OST) catalyses the transfer of the 14-mer glycan donor to an acceptor polypeptide. Figure adapted from\textsuperscript{12}

Stage 2: Attachment of the donor glycan to a polypeptide backbone: In the lumen, oligosaccharyltransferase (OST) catalyses the transfer of the 14-mer glycan donor to an acceptor polypeptide chain containing the sequence Asn-Xxx-Thr. Note that not all Asn residues within this sequence are glycosylated\textsuperscript{13}.

Stage 3: Glycan trimming: Trimming involves removal of the three glucose (Glc) residues by Glycosidase I (trims terminal α1-2 Glc) and II (trims inner α1-3Glc). This trimming signals the migration of the nascent glycoprotein from the ER

23
Stage 4: Maturation in the Golgi: Here the protein is exposed to mannosidases that remove some or all of the mannose residues. Glycosyltransferases then rebuild the glycan structures and it is at this stage that the glycan heterogeneity is introduced. The resulting glycoproteins can be broadly categorised into two types; complex and high mannose type glycans (Figure 1.2).

Biosynthesis of O-linked glycosylation

O-linked glycosylation is a post-translational modification occurring in the Golgi apparatus and is initiated by the covalent attachment of a monosaccharide to a serine or threonine residue. The most common type of O-linked glycan contains an initial N-acetyl galactosaminyl (GalNAc) residue (mucin type). Other O-linked glycans include GlcNAc, galactose, mannose and fucose as the initial sugar residue.

O-linked glycan structures have no single conserved core structure but a few shorter core sequences consisting of two or three sugars. Common cores structures for O-linked glycan are shown in Figure 1.3.
1.2 The role glycosylation in therapeutic proteins

Figure 1.3: Common O-linked (mucin type) glycan core structures found in mammals

1.2.2 The role of glycosylation in physiology and disease

Exhaustive study into the role of glycosylation has revealed the vast range of functions that are mediated by the glycan chains on proteins and coating cell surfaces.

Glycans have been implicated in a number of roles encompassing cell adhesion and protein-cell interactions; bulky carbohydrates can inhibit antibody binding and protect against proteolysis\(^\text{15}\). The structure and function of glycoproteins are influenced by their glycan chains; correct glycosylation is essential for the folding of some glycoproteins and often results in increased solubility as the glycan chains shield hydrophobic regions of the peptide backbone. Glycans on the surfaces on cells and proteins are also significant factors in a number of recognition processes including the role of glycans in disease, infection, and immunity\(^\text{16}\). Some of the more prominent and recent examples of the importance of glycosylation are discussed here.

Glycans in cell-cell interactions and recognition

Glycan structures on the surface of cells, on glycoproteins, glycolipids and on glycoconjugates play key roles in protein-cell, cell-cell and glycan-glycan recognition. In many cases, the terminal structure of a glycan chain on a cell or glycoprotein can mediate the binding of antibodies or lectins to glycoproteins and cells.

One prominent example is the role which glycans play in the fertilisation process. It has been reported that the terminal sialyl-Lewis x structure (Figure 1.4) expressed on the glycoprotein membrane of an oocyte is essential for binding of the spermatozoa to initiate fertilisation\(^\text{17}\).
Similarly, ABO blood groups are dependent upon the terminal sugar on the cell surface antigen of erythrocytes\textsuperscript{18}. Figure 1.4 shows the three blood group types, A, B and O. These each differ by only one terminal sugar moiety; group A has a terminal GalNAc while group B has a terminal Gal residue and absence of either of these antigens is responsible for the O blood group.

**Figure 1.4:** Glycan structures of A: Terminal Siayl-Lewis-X antigen and B: Terminal antigens found on the surface of red blood cells responsible for the ABO blood groups.

**Role of glycans in disease**

Carbohydrate motifs can be used as a quality control for protein folding in the ER, clearance of degraded glycoproteins from circulation and as an indicator for certain disorders and conditions\textsuperscript{15}. Absence or modification of specific glycan motifs has been linked to a variety of congenital disorders\textsuperscript{19,20}. Mutations in the enzymes required for the synthesis of N-linked glycans can result in the development of congenital disorders such as a variety of CDG’s (Congenital Disorders of Glycosylation) and muscular dystrophy. One example is the glycosylation of $\alpha$-dystroglycan. $\alpha$-dystroglycan is a mucin like transmembrane glycoprotein which serves as a receptor in muscles. Hyperglycosylation of the mucin domain results in inhibited $\alpha$-dystroglycan binding. This hyperglycosylation is seen in people with varying congenital muscular dystrophy disorders where patients are afflicted with progressive muscular weakening, muscle spasms and a limited range of movement\textsuperscript{21,22}.
1.2 The role glycosylation in therapeutic proteins

In addition to congenital disorders, the under- and over-expression of glycoproteins and the modification and absence of glycan motifs has been implicated in diabetes, autoimmune disorders, and tumour growth and progression in different types of cancer.\(^{23-25}\)

The Sialyl-Lewis-X antigen mentioned above is overexpressed in some cancers, most notably in breast cancer.\(^{26;27}\) As such, the use of glycans as cancer biomarkers in glycan and lectin arrays has become a prominent feature in cancer research.\(^{28-30}\)

**Role of glycans in bacterial and viral infection**

Glycans play a key role in disease and infection including roles in the immune response, antibody recognition and lectin binding. Almost all cells, bacteria and some viruses have surface glycans which mediate cell-cell recognition and lectin (carbohydrate binding protein receptors) binding.

Infection by many viruses and bacteria require surface glycans and/or lectins to mediate binding to the host cells. Binding of the influenza virus to cell surfaces is mediated by the HA (hemagglutinin) protein which specifically binds to the sialic acid moieties coating the host’s cell surface before internalisation.\(^{31-35}\)

Glycans have also been shown to assist viruses in evading the body’s immune response. The most prominent example of this is the HIV virus which evades antibodies though the expression of a densely packed N-linked glycan shield.\(^{36-39}\)

1.2.3 Glycosylation as a tool in protein-based therapeutics

In cases where the body is unable to produce sufficient levels of a protein to mediate and initiate biological processes (e.g. anaemia caused by chronic kidney disease or diabetes) the corresponding therapeutic protein (EPO and insulin respectively) can be commercially synthesised and administered for treatment.\(^{40;41}\) In other conditions where this strategy is not relevant, artificial therapeutic glycoproteins, glycopeptides and antibodies are developed and optimised for treatment\(^{42-45}\) (e.g. Infliximab for treatment of autoimmune diseases such as Crohn’s). Many of these
therapeutic protein-based drugs require glycosylation for activity, while in others glycosylation is used to manipulate certain characteristics such as solubility and half-life\(^{1,4-11,42-43,46-51}\).

Glycosylation of proteins has been shown to regulate uptake, improve in vitro stability, in vivo lifetimes, and solubility. As glycosylation is a common posttranslational modification, introducing human-like glycosylation onto exogeneous proteins can aid in preventing unwanted immunogenic responses; insufficient sialylation can result in the clearance of glycoproteins through the asialoglycoprotein lectin or the mannose-binding lectin (MBL) produced in the liver\(^{52-55}\). There is also evidence that some mammalian cells express sialylated glycoproteins that can provoke an immune response (discussed in section 1.4)\(^1\).

A major advantage of glycosylation is the ability of glycans to improve a protein’s half-life by increasing its circulation time\(^{10,11,46-48,51}\). Proteins which are not glycosylated can be susceptible to a number of clearance and degradation mechanisms including hepatic uptake, immune clearance and proteolytic degradation.

Introducing additional glycosylation sites to a therapeutic protein is a method frequently employed to extend in vivo half-life resulting in the decrease in the frequency of drug administration\(^{46,49,56}\). One example of this is the introduction of additional glycosylation sites in to the Erythropoietin hormone to produce a therapeutic derivative with increased half-life; ARANESP\(^{57}\).

Erythropoietin (EPO) is the glycoprotein hormone responsible for the proliferation of red blood cells and is produced commercially for the treatment of anaemia (rhEPO). In 2001 the effects on biological activity of increasing the sialic acid content of rhEPO by introducing additional glycosylation sites (Asn-Xxx-Thr/Ser) were investigated. Naturally occurring EPO has three N-linked and one O-linked glycan structures giving between 9-14 terminal sialic acid moieties depending on the glycoform\(^{57}\). Egrie and Browne introduced two additional N-linked glycosylation sites and compared the pharmacokinetic properties of ten rhEPO isoforms containing from 4-22 terminal sialic acid residues. A linear correlation between sialic acid content and serum half-life and in vivo activity was shown as an increase in terminal sialic acid residues resulted in prolonged in vivo half-life. Receptor binding affinity
1.3 Additional modification strategies for improving therapeutic protein efficiencies

however exhibited an inverse relationship with sialic acid content. The 5\(N\)-linked rhEPO derivative has a ~3-fold longer serum half-life and increased potency compared to the naturally occurring 3\(N\)-linked rhEPO. This 5\(N\)-linked glycosylation analogue of rhEPO is now marketed as ARANESP\textsuperscript{®} or darbepoetin alfa and requires less frequent administration compared to the commercially available rhEPO\textsuperscript{57}.

ARANESP\textsuperscript{®} is an example of how glycosylation can also inhibit binding of a protein to a receptor. However, this effect can be used to modulate the effectiveness of therapeutic protein treatments; using glycosylation to reduce receptor binding affinities while simultaneously increasing the half-life can prevent peaks of activity which cause overstimulation of some pathways while decreasing the frequency of administration\textsuperscript{7;58;59}.

### 1.3 Additional modification strategies for improving therapeutic protein efficiencies

Additional modification strategies for tailoring the pharmacological properties of therapeutic drugs include PEGylation, carboxylation, hydroxylation, sulfation and acylation\textsuperscript{9}. The most commonly employed of these methods is PEGylation; the covalent attachment of polyethylene glycol (PEG) polymer chains to a protein. Figure 1.5 shows commonly used PEGylation compounds.

The first PEGylated protein-based therapeutic was approved by the FDA in 1990 (Adagen\textsuperscript{®}), however in these first generation biopharmaceuticals, attachment of the PEG moiety was non-specific yielding proteins with diverse pharmacological profiles\textsuperscript{60;61}. In the current, second generation PEGylated biopharmaceuticals, the PEG moiety is covalently attached at specific sites; lysine, cysteine and \(N\)-terminal modification is most common. The relatively low abundance of cysteine residues in peptide backbones makes it an ideal target for the covalent attachment of certain modifiers (e.g PEGs, glycans).

PEGylation is used to increase the size, MW and hydrophilicity of a therapeutic protein resulting in improved stability and pharmacokinetic properties. This alteration
in the size, shape and charge of a protein leads to prolonged circulation times due to clearance inhibition\textsuperscript{60–64}. PEGylation is also known to reduce immunogenicity by masking epitopes in protein sequences\textsuperscript{65}.

One example of how PEGylation can moderate the properties of therapeutic proteins is the PEGylated EPO derivative Mircera\textsuperscript{®} (methoxy polyethylene glycol-epoetin beta); a derivative of EPO developed for the erythropoiesis stimulating agent family (ESA) which is used to treat anaemia. PEGylation of EPO at the N-terminus results in improved circulation times while preserving receptor binding affinity and protein structure\textsuperscript{64}.

One of the more common effects of PEGylation is the loss of binding affinity of some proteins for their respective receptor due to the bulky nature of PEGs. However, when utilising PEGylation as a modification strategy, the decreased binding affinity or activity is usually offset by an increase in circulation time. As such, the successful PEGylation of therapeutic proteins requires finding a balance of pharmacokinetic (circulation time) and pharmacodynamic (affinity and activity) properties\textsuperscript{61}. Additionally, when considering PEGylation the degree of polydispersity (higher MW PEGs yield wider ranges of MW proteins) and the number of available attachment sites on the protein backbone must be taken into consideration\textsuperscript{60,66}.

Other methods of modification include the encapsulation of therapeutic drugs using
1.4 Expression platforms for pharmaceutical proteins

biodegradable microparticles in order to control release rates and activity profiles of therapeutics. Encapsulation methods include the use of PGLA (Poly (D,L-lactic-coglycolic acid - a polymer of lactic acid and glycolic acid), microspheres, liposomes, protein scaffolds and micelle formation\(^6,9\).

Summary

Glycosylation and PEGylation are common strategies employed to tune the pharmacological properties of protein based therapeutics\(^9,61,65\). As PEGylation is an artificial modifier, PEGylated proteins and peptides can be chemically synthesised. However, the chemical synthesis of glycoproteins is much more complex and time consuming; the established method for producing commercial glycoproteins is to express them in mammalian cell lines to yield human like glycosylation patterns. However, the limited control over this process had led to need for more effective methods of glycoprotein synthesis.

1.4 Expression platforms for pharmaceutical proteins

While some pharmaceutical proteins are still isolated from natural sources, the majority are produced via recombinant protein expression\(^67\). Organisms for the production of proteins include mammals, bacteria, yeasts, insects, plants and using cell lysate in cell free protein synthesis (CFPS). As of 2014, approximately 30% of all pharmaceutical protein production in done in Chinese hamster ovary (CHO) cells, another 30% percent is produced in \textit{E. coli}, and yeast derived proteins account for around 12% of the market (ref our paper). Choosing an expression system is protein dependent and some of the more common choices are discussed here.

\textit{Escherichia coli}

The use of \textit{E. coli} as an expression system has clear advantages; it is fast, cheap, has a relatively well studies genome, and can provide high yields. In addition, the
well-studied process of protein production has facilitated the engineering of some strains to produce proteins with desired characteristics\textsuperscript{68–70}.

However, drawbacks of expressing in \textit{E. coli} include the incorrect formation of multiple disulfide bonds, limited PMTs (e.g. glycosylation is not possible), and expression of incorrectly folded proteins that yield inclusion bodies. As discussed previously, glycosylation plays an important role in protein activity, solubility and stability, and as such pharmaceutical glycoproteins that are expressed in \textit{E. coli} can exhibit impaired or no \textit{in vivo} activity\textsuperscript{4;8;71}. In some cases, these limitations can be overcome by expression followed by chemical modifications such as PEGylation to improve pharmacokinetic properties, a methodology that has come to the forefront of protein engineering in recent years\textsuperscript{72}.

Despite these advances, the majority of pharmaceutical proteins which require glycosylation for activity are expressed from mammalian cells\textsuperscript{1}.

\section*{Yeast}

Yeast are eukaryotic fungal organisms and \textit{Saccharomyces cerevisiae}, \textit{Pichia pastoris}, and \textit{Hansenula polymorpha} are currently the only expression strains used in the production of commercially available biologics\textsuperscript{4;5;8;9;72}. Advantages to using yeast include; high yields, cost effective methodologies, ability to grow to high cell densities, expression of proteins with multiple disulfide bonds, correct protein folding (compared to \textit{E. coli}), and its ability to glycosylate proteins\textsuperscript{73}. As such, proteins that do not fold correctly in \textit{E. coli} are often produced in yeast. Examples of yeast expressed therapeutics include some insulins produced by NovoNordisk and a range of Hepatitis B surface antigens (both expressed from \textit{S. cerevisiae})\textsuperscript{74}.

The N-linked glycan structures of \textit{P. pastoris}, \textit{H. polymorpha} and \textit{S. cerevisiae} differ slightly, however expression of therapeutic agents in all three yields hypermannosylated glycoforms (see Figure.1.6)\textsuperscript{75}. The core structures of human and yeast N-linked glycans are identical (Man\textsubscript{3}GlcNAc\textsubscript{2}), however, upon translocation to the Golgi and mannosidase trimming, glycosyltransferases catalyse the addition of further mannose residues in yeast cells.
1.4 Expression platforms for pharmaceutical proteins

Figure 1.6: Glycan chains observed when utilising different glycoprotein expression systems. Yeast expression systems yield high mannose type glycan chains. Human and animal expression systems yield very similar glycan structures; however, animal expression systems yield both Neu5Ac and Neu5Gc moieties, while humans have lost the ability to synthesise Neu5Gc\(^1\). Due to this lack of sugar diversity, expressed hypermannosylated therapeutic proteins are considered highly immunogenic. Even when overlooking the immunogenic aspect, the presence of terminal mannose residues on these glycoproteins would result in rapid clearance from the body by lectins. There have been a number of attempts to ‘humanise’ the yeast expression pathway via genetic engineering of strains to yield more human-like glycosylation. There has already been success in the generation of human-like hybrid and complex N-glycan structures in \textit{P. pastoris} and optimization of these systems in ongoing\(^{76–82}\).

**Mammalian cells**

Proteins requiring more complex PTM’s, e.g., glycosylation, are expressed in mammalian cells as they yield glycoproteins with human-like glycan structures (see Figure 1.6). The structure of glycan chains is determined by both the expression organism; it is well known that there are significant differences in glycan structures...
between proteins expressed in human, mammalian and yeast cells, and by the cell line \(^{83}\).

One drawback of mammalian expression is that a number of glycoforms of the same glycoprotein are expressed. Additionally, while these heterogeneous glycoforms are considered to have low immunogenicity, there is evidence that glycoproteins expressed in some cell lines contain terminal Neu5Gc (\(N\)-glycolylneuraminic acid) rather than the human Neu5Ac (\(N\)-acetylneuraminic acid) (Figure 1.6). These non-human sialic acids moieties may affect immunogenicity as antibodies against the Neu5Gc have been identified\(^1\).

### 1.5 Chemical and enzymatic synthesis of glycoproteins

The chemical/ chemoenzymatic synthesis of glycoproteins can be broadly grouped into three steps; the synthesis of complex oligosaccharides, the synthesis/expression of peptide backbones, and the subsequent joining of these fragments\(^{84}\). The chemical synthesis of complex oligosaccharides encompasses a wide range of synthetic strategies and protecting group chemistries and as such is not discussed here. There are a number of recent reviews concerning chemical\(^{85-87}\) and chemoenzymatic\(^{88-90}\) strategies to oligosaccharide synthesis.

Common methods in the chemical synthesis of peptide and proteins include solid phase peptide synthesis (SPPS), native chemical ligation (NCL) and expressed protein ligation (EPL). SPPS allows the synthesis of natural peptides, the incorporation of unnatural amino acids, and peptide/protein backbone modification. A peptide is immobilized on small porous beads. The free \(N\)-terminal amine of a peptide attached to the solid-phase is coupled to a single \(N\)-protected amino acid unit. The resulting unit is then deprotected, revealing a new \(N\)-terminal amine to for further amino acid addition. Repeating this cycle of coupling and deprotection generates the desired peptide\(^{91}\). SPPS is limited by yields to peptides of around 50 amino acids and so can only be used for the synthesis of small peptide fragments. To
1.5 Chemical and enzymatic synthesis of glycoproteins

obtain longer peptide lengths, two smaller peptide chains can be coupled together using native chemical ligation.\(^{92}\)

1.5.1 Native Chemical Ligation (NCL) and Expressed Protein Ligation (EPL)

NCL pioneered by Kent et al is used to couple two small peptide fragments.\(^{92}\) The coupling of peptide fragments is based around the reversible exchange of thioesters with the free thiols of cysteine residues. A peptide containing an N-terminal cysteine residue is coupled to the C-terminal thioester on the desired peptide. This reversible step yields a thioester intermediate which undergoes an irreversible intramolecular S → N acyl transfer, forming a new amide bond (Figure 1.7).

There are a number of examples of the synthesis of complex glycoproteins using NCL. Danishefsky has recently reported, among others, the chemical synthesis of
Erythropoietin, and the α- and β- subunits of the Human Follicle-Stimulating hormone. Other examples include the synthesis of RNase C and Diacylglycerol kinase.

EPL is employed to obtain large peptide fragments that cannot be synthesised by SPPS or NCL. These fragments are expressed in bacteria followed by coupling to a synthetic peptide using NCL.

Drawbacks of NCL and EPL include the instability of thioesters and the low occurrence of cysteine residues in proteins. As such, a number of methods for the cysteine free native chemical ligation of peptides have been developed.

1.5.2 Cysteine free native chemical ligation

One of the most common cysteine free methods is auxiliary-assisted chemical ligation; a removable auxiliary is attached to the N-terminus of a peptide, mimicking the function of a cysteine to promote the initial coupling followed by Y → N acyl transfer as in NCL. This auxiliary can then be removed after the ligation to yield the native peptide chain (Figure. 1.8).

Other methods of cysteine free ligation include the traceless Staudinger ligation; coupling of an azide and phosphinothioester to form an amide and Sugar assisted ligation (SAL); coupling of a sulphydryl group at the C2 position of a sugar with a peptide thioester (Figure. 1.8). SAL holds the advantage of introducing a glycan into the polypeptide backbone which can be further modified for glycoprotein synthesis.

Alternative methods include alanine ligation, leucine ligation, valine ligation and desulfurization. The major drawback of chemical synthesis of glycoproteins is the comparatively low yields and larger production times compared to cell systems.
1.5 Chemical and enzymatic synthesis of glycoproteins

Figure 1.8: A: Auxiliary NCL, B: Traceless Staudinger ligation for the coupling of an azide (PEP$_2$) and phosphinothioester (PEP$_1$) to yield an amide bond. C: Sugar assisted ligation (SAL). A sulfhydryl containing sugar (PEP$_2$) linked to a peptide is coupled to a peptide thioester (PEP$_1$).

1.5.3 Modification of peptide backbones

Routes for the chemical posttranslational modification of peptide backbones are dependent upon the type of modification desired. To simulate glycosylation without mammalian expression, glycan chains are either chemically synthesised separately before covalent attachment at specific sites, or built up enzymatically.

One route to building these glycan structures is to introduce a sugar moiety into the backbone during SPPS. This sugar (usually a GlcNAc) can then be modified enzymatically. Common enzymes employed to build $N$-linked glycan chains are those from the Endo-$\beta$-$N$-acetylglicosaminidase family; Endo A can catalyse the trans-glycosylation of a sugar moiety from a GlcNAc residue to the desired glycopeptide$^{116,117}$. However, the use of Endo A is limited by the specificity of the
reaction and the expense and availability of the nucleotide sugar donors.

Alternatively, glycans can be introduced to the peptide backbone chemically. The Lansbury aspartylation is a convergent method for the synthesis of \( N \)-glycopeptides. The key reaction involves the coupling of a glycosylamine to an Asp-containing partially protected peptide (Figure 1.9)\(^\text{96;113;118}\).

While this method can be used to synthesize a variety of complex glycopeptides using mild deprotection conditions, the efficiency of the reaction is compromised by the poor solubility of the peptide fragments. Another complication is the competing side reaction where the peptide undergoes ring closure to provide an aspartimide side product.

**Non-natural modifications**

Non-natural modifications such as PEGylation and non-natural glycosylation are commonly attached to protein backbones using regio- and chemo-selective methods. PEG derivatives (such as PEG maleimides, PEG amines and PEG NHS esters) as well as derivatized glycans (such as glycosylamines) are usually coupled to lysines, \( N \)-termini or cysteine residues on peptide backbones. Examples include the PEGylation of Erythropoietin as discussed in section 1.3.
1.5 Chemical and enzymatic synthesis of glycoproteins

A method for the chemical glycosylation of proteins using iodo-glycosylamines was developed and later adapted to the semi-synthesis of homogenous Erythropoietin (Figure 1.9). EPO variants containing cysteine residues at the natural $N$-glycosylation sites were expressed in $E.\ coli$. These cysteine variants were selectively modified by reaction with iodo-glycosylamides to yield homogenous glycosylated EPO.

1.5.4 Summary

Posttranslational modification is utilised to tune the pharmacological properties of protein-based therapeutics. While most therapeutic proteins are expressed in mammalian cell lines, there are a number of alternative expression platforms in development including expression in yeast systems and NCL in conjunction with chemoenzymatic modification.

Glycosylation, and to a lesser extent PEGylation, is by far the most common strategy employed to alter the pharmokinetic properties of proteins. Glycosylation can be used to modulate half-life, immunogenicity, binding and $in\ vivo$ stability and activity.
Chapter 1 Introduction

1.6 Erythropoietin

1.6.1 Erythropoietin: erythropoiesis and tissue protection

Erythropoietin (EPO) is the glycoprotein hormone responsible for the proliferation and differentiation of red blood cell precursors (erythropoiesis). During erythropoiesis, EPO is expressed in the kidney and binds to the EPO receptor, (EPOR)$_2$, present on the surface of red blood cell (rbc) precursors.

Discovery of this EPO receptor in non-erythroid tissues (e.g. the brain, heart) led to the identification of a novel biological function of EPO. As a response to infection/metabolic stress, one of the body’s reactions is to kill off healthy cells by expressing pro-inflammatory cytokines to kill damaged cells/tissue. Locally expressed EPO has been shown to prevent the destruction of viable cells which surround a site of injury that can be damaged by these cytokines$^{121}$. This process is termed erythropoietin mediated tissue protection and has been observed in the heart, kidney, brain, nervous system and other tissues.

Recombinant human erythropoietin (rhEPO) is currently used for the treatment of anemia in association with chronic kidney disease (CKD) and/or chemotherapy treatment$^{122,123}$. However, there have been a number of attempts to exploit the tissue protective function to develop a therapeutic agent for the treatment of patients after ischemic injury (e.g. stroke)$^{124–128}$.

1.6.2 Structure of erythropoietin

Recombinant human erythropoietin (rhEPO) is a small (166 amino acids) heavily glycosylated protein. As Figure 1.10 shows, rhEPO has an $\alpha$-helical structure with four helices; $\alpha$A (residues 1-26), $\alpha$B (residues 47-83), $\alpha$C (residues 190-121), and $\alpha$D (residues 138-161), and two internal disulfide bonds (C7-C161 and C29-C33). Approximately 40% of the MW of rhEPO is a result of the three $N$-linked (Asn 24, 38 and 83) and one $O$-linked (Ser126) glycan chains which are highlighted in
1.6 Erythropoietin

Figure 1.10: Amino acid sequence of rhEPO showing the natural glycosylation sites, αA (green), αB (purple), αC (blue), and αD (grey). Amino acids forming loops are shown in yellow and the two internal disulfide bonds are shown.

Figure 1.10, giving rhEPO a molecular mass ranging from 30-34 kDa depending on the carbohydrate content. The α-helical structure of rhEPO can be clearly seen in Figure 1.11 which shows the crystal structure of non-glycosylated EPO solved in 1998 by Syed et al\textsuperscript{129}. The interaction of EPO with (EPOR)\textsubscript{2} is discussed in detail in the following section.

1.6.3 Erythropoiesis: Regulation and mode of action

Red blood cell production

EPO controls the production of rbc’s by promoting differentiation and proliferation of erythroid progenitor cells in the bone marrow. As shown in Figure 1.12, erythropoiesis is controlled by several cytokines and EPO acts in the later stages of erythrocyte development to inhibit apoptosis of the cell.
Figure 1.11: Crystal structure of EPO (green) bound to the homodimer receptor, (EPOR)$_2$ ref. αD, αA and the AB loop interaction with one EPOR monomer (yellow) in Site 1. In Site 2, the αC and αA helices interact with residues on the second EPOR monomer (purple). PDB reference 1EER$^{129}$.

The primary site of EPO production is in the kidney and the fetal liver$^{130}$. EPO expression occurs under control of an oxygen sensing, hypoxia inducible, factor dependent mechanism; When the kidney detects low levels of oxygenation in the blood

Figure 1.12: Stages of erythrocyte development. Various cytokines involved in erythrocyte development are mentioned, EPO acts in the CFU-E stage of development. SCF = Stem Cell Factor, GM-CSF = Granulocyte Macrophage Colony-Stimulating Factor, IL = Interleukin, EPO = Erythropoietin.
Figure 1.13: Mechanism of erythropoiesis. Upon detection of hypoxia, the kidney stimulates the production of EPO. EPO binds to a cell surface receptor on the erythroid progenitor cells in the red bone marrow preventing apoptosis. The resultant increase in rbc count causes an increase in the blood oxygenation level and the kidney downscales EPO expression. Figure adapted from http://www.tarleton.edu/Departments/anatomy/erythro.html

(hypoxia) it will stimulate the production of EPO. EPO is then released into circulation where it binds to the transmembrane EPO receptor, (EPOR)$_2$, present on the surface of the erythroid progenitor cells. Binding of EPO induces signal transduction to prevent cell apoptosis and the resultant increase in rbc count causes an increase in the blood oxygenation level. Once normal levels of oxygen are detected, EPO expression is down-regulated. This feedback signalling loop is shown in Figure 1.13.

Signal Regulation

Erythropoiesis is mediated by a transmembrane homodimeric receptor, (EPOR)$_2$, present on the surface of red blood cell precursors. The mechanism of activation of the intracellular pathways involved in erythropoiesis are shared by other members of the type I and type II cytokine receptor family.

The extracellular domains of (EPOR)$_2$ are held in a confirmation as to separate
Figure 1.14: Signal regulation of erythropoiesis. Binding of EPO to (EPOR)$_2$ results in a conformational change in the (EPOR)$_2$. This causes phosphorylation of the pre-associated tyrosine residues on (EPOR)$_2$ and activates the downstream signalling pathway to prevent apoptosis of erythrocyte precursor cells.

the intracellular domains and their pre-associated Janus kinase 2 (JAK2) kinases (see Figure 1.14). Binding of EPO to (EPOR)$_2$ causes a conformational change in (EPOR)$_2$ and brings together the intracellular domains resulting in the autophosphorylation of the JAK2 kinases. The activated JAK2s then phosphorylate the (EPOR)$_2$ tyrosine residues providing a docking site for signaling proteins such as Signal Transducer and Activator of Transcription (STAT5), Rasmitogen-activated protein kinase (MAPK) and phosphatidylinositol 3-kinase (PI3K). STAT5 binds to the phosphorylated (EPOR)$_2$, is phosphorylated, homo dimerizes and translocates to the nucleus to activate target genes for the up-regulation of anti-apoptotic proteins$^{131-134}$.

**Erythropoietin receptor binding**

The structure of non-glycosylated EPO expressed from *E. coli* bound to (EPOR)$_2$ is given in Figure 1.11. EPO binds to (EPOR)$_2$ via two binding sites located on opposite faces of the protein; Site 1 and Site 2. Site 1 (Kd approx 1 nM) has a higher
1.6 Erythropoietin

Figure 1.15: Schematic of EPO bound to (EPOR)$_2$.

affinity and has almost twice as many interactions with the receptor than Site 2 (Kd approx 1 $\mu$M). A sequential binding model for the EPO-(EPOR)$_2$ binding has been proposed, in which EPO binds first via the high affinity interaction (Site 1) with one (EPOR)$_2$ monomer and then through the low affinity interaction (Site 2) with the second (EPOR)$_2$ monomer.

A schematic of the relative binding sites of EPO in relation to the receptor is shown in figure X. Residues from helices $\alpha$A, $\alpha$B and $\alpha$D, and part of the AB loop contribute to the Site 1 EPO-(EPOR)$_2$ interactions. Site 2 is composed of residues from $\alpha$A and $\alpha$C. The interactions are dominated by the interaction of lysine and arginine residues on EPO and asparate and glutamate residues on the (EPOR)$_2$. A more in depth analysis of the interacting residues is given in section 2.1.2.
1.6.4 Erythropoietin mediated tissue protection: Regulation and mode of action

In addition to its erythropoietic function, EPO also plays a key role in the body’s innate immune response by acting as an antagonist of pro-inflammatory cytokines\(^{121,135}\). Upon infection, trauma or stress to tissue, the body produces proinflammatory cytokines (e.g. tumour necrosis factor-alpha, TNF-\(\alpha\)) to kill the damaged cells. However this response is self-amplifying, resulting in the death of healthy cells and the spread of inflammation (Figure 1.16). In response to this, EPO expression in the surrounding tissue is up-regulated to counteract the effects of these proinflammatory cytokines and prevent apoptosis via a tissue protective receptor (TPR)\(^{125–127,135,136}\).

This TPR is not highly expressed under normal conditions but is immediately up-regulated following injury/stress. Expression of TNF-\(\alpha\) upregulates the production of the TPR, but inhibits expression of EPO which is found predominantly in the surrounding tissue. The extent of the tissue damage is thus a result of a balance between the production of proinflammatory cytokines and EPO. It has been found that almost every tissue studied under conditions of hypoxia or stress express EPO to mediate tissue protection\(^{121,137}\).

![Mechanism of EPOs tissue protective function](image)

**Figure 1.16**: Mechanism of EPOs tissue protective function. After an initial injury, locally expressed EPO prevents the apoptosis of healthy tissue. Figure adapted from reference 100.
1.6 Erythropoietin

**Tissue protective receptor**

In contrast to erythropoiesis, the receptor mediating tissue protection has yet to be isolated and details of the signaling cascade downstream of EPO binding remain to be determined. Tissue protective signaling varies depending on the cell and type of injury, however most responses are initiated by phosphorylation of JAK2 (similar to erythropoiesis)\(^{125–127;135;136}\).

The predominant theory is that the TPR is a heterodimer of (EPOR)\(_2\) and a beta common receptor (\(\beta_c R\)). A number of groups have reported the importance of the \(\beta_c R\) in tissue protection, including a report by Coldewey et al. where the activation of \(\beta_c R\) was essential for renoprotection in mice treated with rhEPO after the development of sepsis\(^{135}\). The proposed structure of the TRP is shown in Figure 1.17; as \(\beta_c R\) exists as a preformed homodimer the exact stoichiometry of the TRP is uncertain.

While the structure of the TPR remains elusive, the binding epitope on EPO is known to consist of amino acids in the \(\alpha B\) helix (amino acids 58-82) (Figure 1.18). This EPO-TPR interaction lies in a distinct region isolated from the classic erythropoietic epitope which is an important factor when considering tissue protective variants of EPO (see section 1.6.6)\(^{126}\).

---

**Figure 1.17:** Proposed structure of the EPO tissue protective receptor postulated in the literature. All structures are a complex of \(\beta_c R\) and (EPOR)\(_2\), different only in receptor ratios.
1.6.5 Effect of glycosylation on EPO function and activity

EPO has three natural $N$-linked glycosylation sites at positions Asn24, Asn38, and Asn83 and one $O$-linked (mucin type) glycosylation site at Ser 126 which is not essential for biological activity\textsuperscript{138}. The carbohydrate chains on EPO are highly heterogeneous with over 54 glycoforms associated with one EPO sample\textsuperscript{139}. These glycan chains account for ~40% of the MW of EPO and as such it is essential that the biological role of these sugar chains is well understood.

The glycan chains present on EPO are essential for protein stability and solubility; loss of the glycan chains results in low stability and protein aggregation. Nahri \textit{et al} compared the conformation and stability of glycosylated rhEPO (expressed in CHO cells) and non-glycosylated EPO (expressed in \textit{E. coli})\textsuperscript{140}. Both rhEPO and non-glycosylated EPO were tested for their stability to guanidine hydrochloride induced denaturation, pH changes, and heat induced denaturation. The non-glycosylated EPO exhibited a much lower tolerance to all of the destabilising conditions. While the glycan chains are essential for protein stability, they are not required for in vitro binding to the (EPOR)$_2$. In fact, non-glycosylated EPO has a higher affinity for the (EPOR)$_2$ than the rhEPO\textsuperscript{141}.

The presence of the terminal sialic acid residues is essential for the \textit{in vivo} half-life of the protein\textsuperscript{142}. Removal of the terminal sialic acid residues to yield asialoEPO (desialylated EPO) results in a decrease in the \textit{in vivo} activity due to rapid clearance from circulation. The serum half-life is reduced from ~2 hours (rhEPO) to ~10 minutes (asialoEPO)\textsuperscript{143}. The introduction of additional glycosylation sites onto EPO to improve protein half-life has already been discussed in section 1.2.3.

To date there has been no reports on the effect of the natural glycosylation on erythropoietin mediated tissue protection.
1.6 Erythropoietin

1.6.6 The effects of modifications on the therapeutic use of Erythropoietin

Commercially available erythropoiesis stimulating agents and their modifications

As of 2014, there are over 30 erythropoietin based therapeutics (or erythropoiesis stimulating agents, ESA’s) approved within the EU and the US for the treatment of anemia in association with chronic kidney disease (CKD) and/or chemotherapy treatment (see www.fda.com and www.ema.eu). Most of these ESA’s are based upon the original rhEPO, epoetin alfa, and Table 1.1 lists some examples of these ESA’s separated by their International Nonproprietary Names (INN).

Table 1.1: ESA’s available within the US and the EU listing the active substance, production method, and modification (if any) to the EPO peptide backbone

<table>
<thead>
<tr>
<th>INN</th>
<th>Market Name</th>
<th>Manufacturer (authorisation year)</th>
<th>Expression system</th>
<th>Modification and effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epoetin alfa</td>
<td>Epogen/Procrit®</td>
<td>Amgen (1989)</td>
<td>CHO</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Eprex®</td>
<td>Janssen-Cilag Ltd (1999)</td>
<td>CHO</td>
<td></td>
</tr>
<tr>
<td>Epoetin beta</td>
<td>Neorecormon®</td>
<td>Roche (1997)</td>
<td>CHO</td>
<td></td>
</tr>
<tr>
<td>HX575</td>
<td>Abseamed® MEDICE</td>
<td>CHO Arzneimittel (2007)</td>
<td></td>
<td>Epoetin alfa biosimilar</td>
</tr>
<tr>
<td>Binocirt®</td>
<td>Sandoz (2007)</td>
<td>CHO</td>
<td></td>
<td>Epoetin alfa biosimilar</td>
</tr>
<tr>
<td>Epoetin alfa</td>
<td>HEXAL AG</td>
<td>CHO</td>
<td></td>
<td>Epoetin alfa biosimilar</td>
</tr>
<tr>
<td>HEXAL®</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biopin®</td>
<td>CT Arzneimittel</td>
<td>CHO</td>
<td></td>
<td>Epoetin beta biosimilar</td>
</tr>
<tr>
<td>Epoetin theta</td>
<td>HEXAL®</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

49
Table 1.1: ESA’s available within the US and the EU listing the active substance, production method, and modification (if any) to the EPO peptide backbone

<table>
<thead>
<tr>
<th>INN</th>
<th>Market Name</th>
<th>Manufacturer (authorisation year)</th>
<th>Expression system</th>
<th>Modification and effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eporatino®</td>
<td>Ratiopharm GmbH</td>
<td>CHO</td>
<td></td>
<td>Epoetin beta biosimilar</td>
</tr>
<tr>
<td>Ratioepo®*</td>
<td>Ratiopharm GmbH</td>
<td>CHO</td>
<td></td>
<td>Epoetin beta biosimilar</td>
</tr>
<tr>
<td>Epoetin zeta</td>
<td>Retacrit® SB309</td>
<td>CHO</td>
<td></td>
<td>Epoetin alfa biosimilar</td>
</tr>
<tr>
<td>Silapo®</td>
<td>Stada Arzneimittel</td>
<td>CHO</td>
<td></td>
<td>Epoetin alfa biosimilar</td>
</tr>
<tr>
<td>Epoetin delta</td>
<td>Dynepo®* Shire</td>
<td>HT-1080</td>
<td></td>
<td>Expression in HT-1080 results in glycan chains containing no terminal Neu5Gc moieties.</td>
</tr>
<tr>
<td>Darbepoietin alfa</td>
<td>Aranesp® Amgen</td>
<td>CHO</td>
<td></td>
<td>Two additional glycosylation sites at positions Ala 30 and Trp86. Aranesp has a lower affinity for the EPO receptor (Aranesp IC₅₀ of 1.05 compared to epoetin alfa IC₅₀ of 0.54 ng) but an increased serum half-life (Aranesp has a t₁/2 of 26 hours compared to 8.5 hours for epoetin alfa).</td>
</tr>
<tr>
<td>Methoxy polyethylene</td>
<td>Mircera® Hoffman-La-Roche</td>
<td>CHO 30kDa PEG at positions Lys52 or Lys46 (Mircera has a t₁/2 of 134 hours compared to 8.5 hours for epoetin alfa)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1.6 Erythropoietin

Table 1.1: ESA’s available within the US and the EU listing the active substance, production method, and modification (if any) to the EPO peptide backbone

<table>
<thead>
<tr>
<th>INN</th>
<th>Market</th>
<th>Manufacturer (authorisation year)</th>
<th>Expression system</th>
<th>Modification and effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peginesatide</td>
<td>Hematide®</td>
<td>Affymaxn (2012)</td>
<td>Synthetic peptide</td>
<td>Small peptide mimetic</td>
</tr>
</tbody>
</table>

Epogen®/ Procrit® was the first ESA to be approved and the active ingredient, epoetin alfa, is recombinant human erythropoietin (rhEPO) expressed in CHO cells. Epoetin beta (marketed as Neorecormon® by Roche), submitted under a different INN, is also rhEPO expressed in CHO cell lines; however epoetin beta differs from epoetin alfa in its glycan composition. Epoetin beta is known to contain isoforms with a lower degree of N-glycan sialylation and a greater proportion of more basic isoforms; however these differences do not impact the clinical effectiveness of epoetin beta.

Upon the expiration of the EU patent for epoetin alfa in 2004, a number of biosimilar products appeared on the market (Table 1.1). These include HX575 and epoetin zeta (epoetin alfa biosimilars), and epoetin theta (an epoetin beta biosimilar). These biosimilars have the same peptide sequence as their scaffolds and differ in glycan content and formulation only: studies of HX575 products revealed more high mannose structures than epoetin alfa, while studies of epoetin zeta revealed isoforms with less O-linked glycans and lower levels of Neu5Gc (the immunogenic non-human sialic acid moiety found in proteins expressed in CHO cells, discussed in section 1.4). Epoetin delta, an epoetin alfa biosimilar, is expressed in human fibrosarcoma cell line HT-1080 and this ESA is unique as its glycan chains show no terminal Neu5Gc residues. Despite the differences in glycan composition and drug formulation, these ESA’s biosimilars are considered interchangeable with the original products (epoetin alfa or beta) as they exhibit only small differences in a clinical setting.
Darbepoietin alfa (Aranesp®), methoxy polyethylene glycol-epoetin beta (Mircera®) and peginesatide (Hematide®) are the only approved ESA’s with significant modification to the original protein backbone and highlight the effectiveness of protein engineering in generating therapeutic agents. Aranesp®, discussed in detail in section 1.3, is the long acting EPO derivative where two additional N-linked glycosylation sites have been engineered into the peptide backbone. The resulting increase in sialic acid content gives Aranesp® a lower affinity for the EPO receptor (Aranesp® IC₅₀ of 1.05 compared to epoetin alfa IC₅₀ of 0.54 ng) but an increased serum half-life (Aranesp® has a t₁/₂ of 26 hours compared to 8.5 hours for epoetin alfa)⁵⁷. As a result Aranesp® can be administered less frequently than epoetin alfa products. In a similar manner, Mircera® is PEGylated epoetin beta which acts as a continuous erythropoietin receptor activator⁵². After expression in CHO cells, epoetin beta is chemically modified with a single 30kDa PEG moiety to generate Micera®. Again this modification results in a lower affinity for the EPO but an increased serum half-life (t₁/₂ of 134 hours compared to 8.5 hours for epoetin alfa)⁵³. Hematide® is a PEGylated 20 amino acid synthetic cyclic peptide dimer which was approved in 2012 for the treatment of anemia but withdrawn in 2013 for safety reasons⁵⁴.

**Current research into the modification of Erythropoietin to improve therapeutic efficiency**

Before the crystal structure of the EPO–(EPOR)₂ complex was solved by x-ray diffraction, mutations were focused on identifying structurally important domains on the EPO protein. As such, there are single point mutations for almost every amino acid reported in the literature, the majority of which are alanine substitutions⁶⁴,⁵⁵,⁵⁶,⁵⁷. Table 1.2 lists these mutations, and their effect on the specific bioactivity of EPO (specific bioactivity quoted is proliferation unless otherwise stated). As expected, comparison of this data to the EPO crystal structure reveals that mutation of amino acids that are located within the two EPO biding sites results in some or complete loss of bioactivity. Interestingly, mutation of the R103 residue to any amino acid results in complete loss of bioactivity, while the effect of other substitutions varies depending upon the new amino acid.
### Table 1.2: List of published EPO mutations and effects on specific bioactivity (UT7-EPO assay).

Table shows No highlight: WT like activity (within 2 fold), Underlined: < 60 percent active, Bold: < 20 percent active, and Red Bold: < 2 percent active

<table>
<thead>
<tr>
<th>Residue</th>
<th>Effect of mutation on bioactivity</th>
<th>Effect of mutation on bioactivity</th>
<th>Effect of mutation on bioactivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>C, C-20kDaPEG, C-30kDaPEG</td>
<td>Q59 A, N</td>
<td>R110 T, E</td>
</tr>
<tr>
<td>P3</td>
<td>C, C-20kDaPEG</td>
<td>V61 A</td>
<td>A114 S</td>
</tr>
<tr>
<td>R4</td>
<td>N</td>
<td>E62 T, A</td>
<td>Q115 T, C, C-20kDaPEG</td>
</tr>
<tr>
<td>I6</td>
<td>S</td>
<td>W64 F, A</td>
<td>K116 E</td>
</tr>
<tr>
<td>C7/C161</td>
<td>S/S</td>
<td>Q65 T, A</td>
<td>E117 A, C, C-20kDaPEG</td>
</tr>
<tr>
<td>D8</td>
<td>S</td>
<td>G66 A</td>
<td>P121 N</td>
</tr>
<tr>
<td>S9</td>
<td>A, N</td>
<td>L67 S</td>
<td>D123 A</td>
</tr>
<tr>
<td>R10</td>
<td>A, I</td>
<td>A68 S</td>
<td>A124 P</td>
</tr>
<tr>
<td>V11</td>
<td>S</td>
<td>L69 A</td>
<td>A125 P</td>
</tr>
<tr>
<td>L12</td>
<td>A</td>
<td>L70 A</td>
<td>S126 V, H, A, E, T, C, C-20kDaPEG</td>
</tr>
<tr>
<td>E13</td>
<td>A</td>
<td>S71 A</td>
<td>C-20kDaPEG</td>
</tr>
<tr>
<td>R14</td>
<td>A, L, Q, E</td>
<td>E72 A</td>
<td>A127 T</td>
</tr>
<tr>
<td>Y15</td>
<td>I, F</td>
<td>A73 S, G</td>
<td>A128 S, C, C-20kDaPEG</td>
</tr>
<tr>
<td>L16</td>
<td>A, S</td>
<td>V74 A</td>
<td>L130 T</td>
</tr>
<tr>
<td>L17</td>
<td>A, S</td>
<td>L75 S</td>
<td>R131 T</td>
</tr>
<tr>
<td>E18</td>
<td>S, A</td>
<td>R76 A</td>
<td>T132 N, A</td>
</tr>
<tr>
<td>A19</td>
<td>S</td>
<td>Q78 A, E, R</td>
<td>T134 N</td>
</tr>
<tr>
<td>K20</td>
<td>E, I, R, A</td>
<td>A79 S</td>
<td>D136 T, A</td>
</tr>
<tr>
<td>E21</td>
<td>I, A</td>
<td>L81 S</td>
<td>F138 V</td>
</tr>
<tr>
<td>E23</td>
<td>S</td>
<td>N83 C</td>
<td>R139 A</td>
</tr>
<tr>
<td>N24</td>
<td>C</td>
<td>S85 C</td>
<td>K140 T, A, R, M</td>
</tr>
<tr>
<td>T26</td>
<td>C, C-20kDaPEG</td>
<td>Q86 S</td>
<td>F142 I</td>
</tr>
<tr>
<td>T27</td>
<td>A</td>
<td>P87 V</td>
<td>R143 A, E</td>
</tr>
<tr>
<td>C29/C33</td>
<td>S/S, Y/Y</td>
<td>W88 F</td>
<td>V144 A</td>
</tr>
<tr>
<td>A30</td>
<td>S</td>
<td>E89 C, C-20kDaPEG, C-30kDaPEG</td>
<td>Y145 F, I</td>
</tr>
</tbody>
</table>
Table 1.2: List of published EPO mutations and effects on specific bioactivity (UT7-EPO assay). Table shows No highlight: WT like activity (within 2 fold), Underlined: < 60 percent active, Bold: < 20 percent active, and Red Bold: < 2 percent active

<table>
<thead>
<tr>
<th>Residue</th>
<th>Effect of mutation on bioactivity</th>
<th>Residue</th>
<th>Effect of mutation on bioactivity</th>
<th>Residue</th>
<th>Effect of mutation on bioactivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>E31</td>
<td>I</td>
<td>S146</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H32</td>
<td>I</td>
<td>N147</td>
<td>A, K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E37</td>
<td>S</td>
<td>Q92</td>
<td>T, A</td>
<td>F148</td>
<td>Y</td>
</tr>
<tr>
<td>N38</td>
<td>C</td>
<td>L93</td>
<td>A</td>
<td>L149</td>
<td></td>
</tr>
<tr>
<td>P42</td>
<td>N</td>
<td>H94</td>
<td>E</td>
<td>R150</td>
<td>A, Q, E</td>
</tr>
<tr>
<td>T40</td>
<td>C</td>
<td>V95</td>
<td>A</td>
<td>G151</td>
<td>A</td>
</tr>
<tr>
<td>D43</td>
<td>V</td>
<td>D96</td>
<td>A, R</td>
<td>K152</td>
<td>A</td>
</tr>
<tr>
<td>T44</td>
<td>I</td>
<td>K97</td>
<td>A, R, E</td>
<td>L153</td>
<td></td>
</tr>
<tr>
<td>K45</td>
<td>D, A, R, I, C, K-20kDaPEG</td>
<td>A98</td>
<td>S</td>
<td>S, A, R, C, C-20kDaPEG</td>
<td></td>
</tr>
<tr>
<td>V46</td>
<td>A</td>
<td>S100</td>
<td>R, E, T, A</td>
<td>L155</td>
<td>A, N</td>
</tr>
<tr>
<td>N47</td>
<td>A</td>
<td>G101</td>
<td>A, I</td>
<td>Y156</td>
<td>F, I, A</td>
</tr>
<tr>
<td>F48</td>
<td>I, S, G</td>
<td>L102</td>
<td>A</td>
<td>T157</td>
<td>A</td>
</tr>
<tr>
<td>Y49</td>
<td>S, F</td>
<td>R103</td>
<td>A, E, H, N, Q, K</td>
<td>G158</td>
<td>A</td>
</tr>
<tr>
<td>A50</td>
<td>S</td>
<td>S104</td>
<td>A, I</td>
<td>E159</td>
<td>A, S</td>
</tr>
<tr>
<td>W51</td>
<td>F, N, S</td>
<td>L105</td>
<td>A</td>
<td>R162</td>
<td>C, C-20kDaPEG</td>
</tr>
<tr>
<td>K52</td>
<td>Q, S, R, E</td>
<td>T106</td>
<td>A, I</td>
<td>C-30kDaPEG</td>
<td></td>
</tr>
<tr>
<td>R53</td>
<td>G</td>
<td>T107</td>
<td>A, L, S</td>
<td>T163</td>
<td>C, C-20kDaPEG</td>
</tr>
<tr>
<td>M54</td>
<td>L</td>
<td>L108</td>
<td>K, A</td>
<td>K24/K126 N²-levulinyl -SEP</td>
<td></td>
</tr>
<tr>
<td>E55</td>
<td>I</td>
<td>L109</td>
<td>A</td>
<td>All</td>
<td>Carbamylation</td>
</tr>
</tbody>
</table>

Most current research is focused on the design of modified erythropoietin analogs with preserved bioactivity but improved pharmacokinetic profiles. Methods for improving the pharmacokinetics of proteins are discussed in section 1.3, and include
approaches such as PEGylation and glycoengineering. In addition to point mutations, Table. 1.2 also lists the effect of conjugation and/or modification of amino acid residues on EPO bioactivity. The majority of modifications listed are PEGylations as this is a common approach to increasing protein half-life. Table. 1.2 lists 15 EPO cysteine variants conjugated to either 20- or 30-kDa PEG moieties. Of those tested, PEGylation at positions K45C-20kDa PEG, K154C-20kDa, R162C-30kDa PEG, E89C-30kDa PEG and A1C-30kDa PEG, showed a reduction (<60%) in EPO bioactivity\(^{64;156}\). The PEGylation of other EPO cysteine variants resulted in only small or no changes in specific bioactivity and at least a 2-fold increase in circulating half-life. These data, in combination with the data from the PEGylated EPO analog Mircera\(^ {®}\), show that it is possible to modify EPO to improve its effectiveness without drastically affecting EPOs proliferative activity and binding to (EPOR)\(^ {2;64;152;155;156}\). Since the discovery of EPO’s tissue protective function, there has been interest in developing EPO as a therapeutic agent for the prevention of injury and restoration of function in non-haematopoietic tissue\(^ {128}\). This requires the generation of an EPO variant that it tissue protective but has no erythropoietic activity.

**Derivatives of EPO that are tissue protective but non-hematopoietic**

The time delay (12-24 hours) observed in peak EPO expression after an initial injury provides a window for the administration of exogenous EPO to aid in the recovery of healthy cells. However, initiation of erythropoietic activity requires circulating EPO concentrations in the pM range while tissue protective function requires nM concentrations. Use of rhEPO would therefore overstimulate erythropoiesis, resulting in severe side effects such as tumor growth, increased risk of heart attack, thrombosis, and formation of blood clots\(^ {124;127;134}\).

This limitation can be overcome by developing derivatives of EPO that are tissue protective but non-hematopoietic. Brines et al have led the way in the development of such derivatives with Carbamylated EPO (CEPO). CEPO is an EPO analogue in which all eight lysine residues have been converted into homocitrulline by reaction with cyanate [Brines2008a]. All eight lysine residues are located within Sites 1
and Site 2 on EPO and as such, modification inhibits the EPO-(EPOR)$_2$ interaction whilst preserving the tissue protective epitope. CEPO was shown to be tissue protective but non hematopoietic in vitro and in vivo and is currently in clinical trials$^{163}$.

More recently Brines et al have been working on identifying the exact sequence responsible for the tissue protective function of EPO. It is known that the residues 58-82 in α-helix B are involved in the interactions and Brines et al have synthesized this short peptide sequence and found it to be non-hematopoietic yet neuroprotective in vivo and in vitro$^{126;128;164;165}$. Based upon this work, a second generation peptide was developed based upon the residues in α-helix B which are adjacent in the 3-dimensional fold of the protein (but not adjacent in the sequence)$^{164}$. This 11-mer peptide (QEQLERALNSS) was shown to be an effective tissue protective agent (Figure 1.18). Despite the short plasma half-life (~2 minutes), this peptide (ARA290) is currently in clinical development$^{103}$.

Berezin et al have reported two additional EPO mimetics which are neuroprotective; EPObis and EPOtris$^{166;167}$. These are short peptides based on the EPO sequence

Figure 1.18: The crystal structure of EPO with α-helix B highlighted (amino acids 58-85 above, yellow). The tissue protective peptide ARA290 (QEQLERALNSS) is bases upon the 3D adjacent structure of α-helix B.
1.6 Erythropoietin

of the AB loop in Site 1 (EPObis) and α-helix C in Site 2 (EPOtris). While EPObis is both haematopoietic and neuroprotective, EPOtris, a 20-mer peptide composed of residues 91-111 of EPO is neuroprotective yet non-hemopoietic. *In vitro* studies show that EPOtris binds to the EPOR receptor, but with a $10^3$-fold lower potency than rhEPO, and is not hemopoietic *in vivo*. It has also been shown that this peptide does not bind to the βc receptor, providing further evidence that the EPOR-βcR complex proposed by Brines et al is not responsible for mediating the tissue protective activity of EPO.

1.6.7 Methods of EPO detection and analysis

Due to its prominence in research, diagnostic settings, and its use as a doping agent in endurance sports, the ability to detect and measure the activity of EPO in a sensitive and reliable manner is important. The choice of EPO assay is dependent upon the sample type (e.g serum or purified rhEPO) and the function to be measured (e.g concentration or activity).

*In vivo* techniques

Classic *in vivo* assays are based on monitoring the stimulation of red blood cell formation after administration of exogenous EPO in rats or mice. These normocythemic mouse assays are performed in animals injected with EPO and measure red blood cell formation by monitoring radioactive iron incorporation (59Fe) in the spleen or blood cells.

While these assays can give a measure of activity, *in vivo* assays can be labor intensive and the interference of other factors (e.g other steroids and hormones) limits the sensitivity. As such *in vivo* assays are not used to measure EPO levels in serum or plasma. Unlike erythropoiesis, EPO mediated tissue protection can occur in a number of tissues (e.g neuroprotection, renoprotection, cytoportection). As such, the choice of *in vivo* and *in vitro* assays is dependent upon the function under investigation and can vary widely.
Chapter 1 Introduction

**In vitro techniques**

*In vitro* methods involve the stimulation of proliferation and/or differentiation in EPO responsive cells in short term cell cultures. Cell lines used include AS-E2, TF1, UT-7 and UT-7/EPO. Of these, the UT-7 and UT-7/EPO cell lines are the most commonly quoted measure of EPO bioactivity. UT-7 cells are a human leukemia cell line whose growth is dependent upon GM-CSF, IL3, or EPO. The proliferation of UT-7/EPO cells is strictly EPO dependent.

*In vitro* assays are less labor intensive and more sensitive than *in vivo* assays, however a disadvantage is that *in vitro* assays are unable to account for circulating half-life, and so measured *in vivo* and *in vitro* activities can differ drastically. This is the case for asialoEPO which shows improved *in vitro* activity but is not active *in vivo* due to a much shorter half-life.

**Immunoassays**

Immunoassays for measuring the concentration of EPO are carried out using either monoclonal or polyclonal anti-EPO antibodies. Commonly used immunoassays are the enzyme linked immunosorbant assay (ELISA) and radioimmunoassay (RIA). The use of an ELISA is often preferable to the RIA, as RIA’s require the generation of radio labelled $^{125}$I-EPO. Sandwich ELISAs are used to detect EPO in samples, and there are a number of commercially available kits specifically for this purpose. The general procedure for a sandwich ELISA is to immobilize an EPO antibody to a well surface to bind EPO, and use a second antibody-conjugate or a primary antibody-secondary antibody pair to detect the EPO (Figure. 1.19). While immunoassays are a very sensitive technique, they are antibody dependent and can only detect ligand binding, not activity. An in depth discussion of the different ELISA techniques, and their advantages and disadvantages assays are discussed in detail in chapter 5.

Another useful technique for monitoring the interaction of a ligand and analyte is surface plasmon resonance (SPR). SPR monitors binding events on a sensor chip and can be used to determine binding, specificity, affinity, kinetics and concentrations.
1.6 Erythropoietin

Figure 1.19: **Left:** Generic Sandwich ELISA for EPO detection. An EPO antibody is immobilized to a well surface to bind EPO. A enzyme-conjugated antibody against EPO is used for detection. **Right:** Experimental configuration used in SPR.

of a variety of analytes. Figure 1.19 shows the experimental configuration used in SPR; a glass slide is coated with thin gold film to create the sensor surface, which is then coated with a dextran (or other) matrix for the attachment of molecules. Once this ligand is bound, an analyte is injected in a continuous flow. Binding of this analyte to the ligand results in a proportional change in the refractive index at the surface of the chip. To date SPR has been used to determine binding affinities of several EPO variants including new EPO derivatives and an investigation into the receptor binding kinetics of partially and fully un-glycosylated EPO.

1.6.8 Summary

EPO is a glycoprotein hormone responsible for both erythropoiesis and tissue protection. The two functions are mediated by independent epitopes within the EPO structure; however the structure of the TPR is yet to be confirmed. The four glycan chains on EPO are not required for EPO receptor binding interactions, however they are essential for *in vivo* stability, and as such physiological activity.

Several routes have been taken to develop a derivative of EPO that is tissue protective but non-erythropoietic for the treatment of ischemic injuries. Presented here is
the investigation into the use of glycans to inhibit the EPO-(EPOR)$_2$ interactions while preserving tissue protective function.
1.7 Aims and Objectives

The objective of this investigation is to determine the effect of site specific modification of EPO on the EPO-(EPOR)2 interaction.

As discussed in section 1.6.3, EPO is responsible for mediating both the production of red blood cells (erythropoiesis) and plays a role in controlling inflammation after ischemic injury (tissue protection). In order to exploit EPO as a therapeutic drug for aiding in the recovery of patients after an ischemic injury, the erythropoietic function must be eliminated; section 1.6.6 discusses the relative affinities of the (EPOR)2 and the tissue protective receptor for EPO, and how administration of commercial rhEPO for ischemic injury treatment would overstimulate red blood cell production causing complications in the patient. Here, we report investigation into the use of carbohydrates to inhibit the EPO-(EPOR)2 interaction.

As the (EPOR)2 and tissue protective receptors bind to non-overlapping, exclusive epitopes on EPO, the binding of EPO to (EPOR)2 could theoretically be blocked by selective modification at the EPO-(EPOR)2 interface (Figure 1.20). Carbohydrate

![Diagram showing modification of EPO to selectively block EPO-(EPOR)2 interactions while preserving tissue protective function.]

**Figure 1.20:** Modification of EPO to selectively block EPO-(EPOR)2 interactions while preserving tissue protective function.

61
moieties were used as the modifier as they are biologically combatable and can aid in promoting the proteins half-life\textsuperscript{57}.

The strategy for the proposed site selectively glycosylation of EPO was as follows (Figure 1.21):

1. Introduction of cysteine residues to the EPO sequence to provide handles for chemical modification.

2. Expression of these EPO cysteine variants in \textit{E. coli}.

3. Modification of these EPO cysteine variant using carbohydrate moieties.

4. Development of an assay to determine EPO-(EPOR)\textsubscript{2} binding and the effect of EPO modification on EPO-(EPOR)\textsubscript{2} binding.

\textbf{Figure 1.21:} Stepwise approach to EPO-(EPOR)\textsubscript{2} inhibition


2 Sub-cloning and expression of WT EPO and EPO cysteine variants

For the investigation into the effects of EPO modification on the EPO-(EPOR)\textsubscript{2} interaction, EPO genes were expressed in \textit{E. coli}. The choice of expression system is an important one and is discussed in section 1.4; each platform has advantages and disadvantages. For the initial proof-of-concept investigation, single glycoforms will be required in order to characterize the EPO variants and EPO reactions as heterogeneous glycan motifs make characterization difficult, especially in mass spectrometry. As such, \textit{E. coli} was chosen as the expression system for its lack of protein glycosylation. While expression in \textit{E. coli} has been reported, the protein is highly unstable and incorrectly folded due to the lack of posttranslational modification\textsuperscript{119}, hence further steps after expression are required to yield folded, soluble protein. These steps are discussed in this section.

2.1 Sub-cloning of WT EPO and EPO cysteine variant genes

2.1.1 Sub-cloning of WT EPO gene into pET16b vector

The WT EPO gene was amplified from a cloning vector using the polymerase chain reaction (PCR). In the forward primer an \textit{NdeI} site (CATATG) was placed immediately 5’ to the GCC codon (Ala) of the mature protein. In the reverse primer a \textit{BglII} (AGATCT) site was placed 3’ to the TGA stop codon (section 8.2).
Chapter 2  Sub-cloning and expression of WT EPO and EPO cysteine variants

Figure 2.1: The WT EPO vector construct with EPO gene flanked by NdeI and BglII downstream of the His-tag.

This amplified EPO gene flanked by the NdeI and BglII restriction sites was digested and ligated into the pET16b vector which allows for IPTG induction and ampicillin selectivity.

pET16b was chosen as it contains an N-terminal His-tag which facilitates an easier purification of the target protein and contains an N-terminal Factor Xa cleavage site for the removal of the tag after purification if required. However, there have been no cases where a His-tag has been shown to inhibit EPO binding to (EPOR)$_2$ in vitro.

The final product of the reaction was the pET16b vector with the EPO gene inserted downstream of the His-tag (Figure. 2.1). This plasmid was labeled WT EPO.

Sequencing confirmed that the mature EPO gene had been inserted into the vector in the correct reading frame, downstream of the His-tag. The mature EPO gene
2.1 Sub-cloning of WT EPO and EPO cysteine variant genes

with the His-tag corresponds to a 189 amino acid protein with a molecular weight of 21.048 kDa.

2.1.2 Introducing cysteine mutations at key EPO-(EPOR)₂ binding sites

As outlined in the aims (section 1.7), the goal of this investigation is to determine the effect of site specific modification of EPO on the EPO-(EPOR)₂ interaction. Cysteine residues were chosen to be introduced for chemical modification as the WT EPO contains four cysteine residues; however these residues are occupied in two internal disulfide bonds which are essential for EPO’s secondary structure. As such, introduction of an additional cysteine into the EPO sequence allowed for a selective and specific target for chemical reaction.

Choosing mutation sites

As mentioned previously (section 1.6.6), modification of all eight lysine residues in the EPO sequence yields a tissue protective but non-erythropoietic EPO derivative. Of these eight lysine residues, the five residues located in either Site 1 or Site 2 were initially targeted for mutagenesis; Lys45 (Site 1), Lys52 (Site 1), Lys97 (Site 2), Lys140 (Site 1), and Lys154 (Site 2).

The crystal structure of EPO bound to the EPOR homodimer receptor was used to predict further optimal sites to introduce cysteine mutations (Figure 2.2). EPO is bound to the homodimer EPO receptor via two distinct binding sites, Site 1 and Site 2. Table 2.1 lists the key interactions between EPO and the (EPOR)₂ and highlights that the majority of the contacts are in Site 1.
Table 2.1: List of interactions between EPO and (EPOR) in Site 1 and Site 2 at < 3.5 Å distance.

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPO</td>
<td>EPOR</td>
</tr>
<tr>
<td>T44</td>
<td>F93</td>
</tr>
<tr>
<td>L45</td>
<td>E62</td>
</tr>
<tr>
<td>V46</td>
<td>S92</td>
</tr>
<tr>
<td>N47</td>
<td>T87, A88</td>
</tr>
<tr>
<td>L140</td>
<td>D61</td>
</tr>
<tr>
<td>R143</td>
<td>E60</td>
</tr>
<tr>
<td>N147</td>
<td>F93, H114</td>
</tr>
<tr>
<td>R150</td>
<td>S204, P203, E117</td>
</tr>
</tbody>
</table>

Figure 2.2 shows the EPO-(EPOR) interactions in Site 1 and Site 2. In Site 1, Lys45, Val46 and Asn47 on EPO interact with Glu62, Thr87, Ala88 and Ser92 on (EPOR)2. Thr44 also interacts with Phe93 on the (EPOR)2 monomer. Mutation of this Phe93 (a F93A variant) eliminates detectable binding of EPO173 to (EPOR)2. Site 2 interactions are dominated by the Asp8, Arg14, Lys97, Ser100, Arg103 and Ser104 on EPO (Figure 2.3). In fact, it has been shown that R103A is an inactive EPO variant174. Taking these interactions into account, in addition to the lysine

Figure 2.2: Interacting residues of EPO (green) and (EPOR)2 (yellow and *) in Site 1. A: Highlights the interactions of the AB loop on EPO with (EPOR)2. B: Interactions of the zD helices on EPO with (EPOR)2.
2.1 Sub-cloning of WT EPO and EPO cysteine variant genes

**Figure 2.3:** Interacting residues of EPO (green) and (EPOR)$_2$ (red and *) in Site 2. Arg103 on EPO dominates the EPO-(EPOR)$_2$ interactions

residues already targeted for mutation, amino acids which have close contacts and/or hydrogen bonds with (EPOR)$_2$ residues were targeted. These include Thr44, Lys45, Val46, Asn147 and Arg150 in site 1 and Arg103 and Ser104 in Site 2.

In total, 11 residues were targeted for mutations (Table 2.2).

**Table 2.2:** List of residues in the EPO sequence targeted for mutation to cysteine residues

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>T44</td>
<td>K97</td>
</tr>
<tr>
<td>V46</td>
<td>R103</td>
</tr>
<tr>
<td>K45</td>
<td>S104</td>
</tr>
<tr>
<td>K52</td>
<td></td>
</tr>
<tr>
<td>K140</td>
<td></td>
</tr>
<tr>
<td>N147</td>
<td></td>
</tr>
<tr>
<td>R150</td>
<td></td>
</tr>
<tr>
<td>K154</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 2  
Sub-cloning and expression of WT EPO and EPO cysteine variants

The lysine to cysteine mutations were introduced by PCR site-directed mutagenesis and synthetic genes carrying the remaining mutations were purchased from GeneArt in order to save time.

**Site-directed mutagenesis of EPO gene**

The lysine to cysteine mutations were introduced by site-directed mutagenesis which requires two synthetic primers with the desired mutation and a vector template containing the gene of interest.

The WT EPO that was previously (section. 2.1.1) cloned into pET16b was used as the template DNA and the primers for each mutation are listed in section 8.1.3.

The mutations were carried out according to the QuikChange Site-Directed Mutagenesis Kit (200518) protocol using the *Pfu Ultra* polymerase. Optimal conditions for successful mutations required an annealing temperature of > 70°C and are outlined in section 8.2.

Of the five mutations attempted, the K45C, K52C, K97C and K140C mutations were successful and the sequences are given in section 8.1.4. The K154C mutagenesis was unsuccessful presumably due to unsuitable primers or pcr conditions. As 4 of the 5 mutations were successfully introduced, no further attempts were made to mutate the K154 position.

To expand the number of possible EPO variants and to improve the likelihood of EPO modification resulting in EPO-(EPOR)$_2$ inhibition, the plasmids for these single EPO variants were subjected to a second round of site-directed mutagenesis to generate double cysteine EPO variants (Figure 2.4).

In this round, the plasmids coding for the EPO cysteine variants were used as the templates for mutagenesis with a different primer (e.g the K45C plasmid was used as a template with the K97C primers). From this second round of mutagenesis, the double cysteine EPO variants K45C+K97C EPO and K45C+K52C EPO were generated.
2.1 Sub-cloning of WT EPO and EPO cysteine variant genes

For the remaining mutations (Table 2.2), the genes encoding for the variants were purchased from GeneArt to save time. The sequences were optimised for expression in *E. coli* and were supplied in a cloning vector with Amp resistance. The genes were sub-cloned from the cloning vector using the *NdeI* and *BamHI* restriction enzymes and sub-cloned into pET16b using the same restriction enzymes (section 8.2). Sub-cloning was successful for the following mutants: T44C, R103C, S104C, N147C, R150C, V46C+R103C (double mutant), and R103C+R150C (double mutant) EPO.
Chapter 2 Sub-cloning and expression of WT EPO and EPO cysteine variants

Summary

In total, 14 EPO variants (including the WT EPO) were generated and cloned into pET16b for expression and characterisation. Table 2.3 lists the positions of the successful mutations.

Table 2.3: List of cysteine mutations successfully introduced into the EPO gene sequence and cloned into pET16b. Each double variant contains a cysteine residue in both Site 1 and Site 2 except for * where both mutations are in Site 1.

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
<th>Double cysteine variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>T44C</td>
<td>K97C</td>
<td>K45C+K97C</td>
</tr>
<tr>
<td>K45C</td>
<td>R103C</td>
<td>K45C+K52C*</td>
</tr>
<tr>
<td>K52C</td>
<td>S104C</td>
<td>V46C+R103C</td>
</tr>
<tr>
<td>K140C</td>
<td></td>
<td>R103C+R150C</td>
</tr>
<tr>
<td>N147C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R150C</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Six of the variants contain cysteine mutations in Site 1 (T44C, K52C, K45C, N147C, K140C and R150C), and three of the variants contain cysteine mutations in Site 2 (K97C, R103C and S104C EPO).

The remaining four EPO variants contain two cysteine mutations, three with a cysteine in both Site 1 and Site 2 (K45C+K97C, V46C+R103C and R103C+R150C EPO), and one with the two cysteine residues in Site 1 (K45C+K52C EPO).

2.2 Expression of WT EPO and EPO cysteine variants

2.2.1 Optimisation of WT EPO gene expression

The vectors containing the EPO genes described in section 2.1.1 were transformed into BL21(DE3) for expression.
2.2 Expression of WT EPO and EPO cysteine variants

Figure 2.5: SDS PAGE gel showing time points for the expression of WT EPO at 30°C (left) and 37°C (right). Below each SDS PAGE gel is the corresponding western blot.

Once the gene is expressed, the protein produced is oxidised (the disulfides are formed), insoluble and incorrectly folded. To obtain EPO in its correctly folded, active conformation, the protein was solubilised, purified, and refolded\textsuperscript{119}.

WT EPO plasmids were transformed into BL21(DE3) cells and grown in 6 ml LB medium overnight. 400 ml flasks were inoculated and induced with IPTG when the OD\textsubscript{600nm} = 0.6 (approximately 4 hours). The expression was carried out at 37°C, 30°C and 16°C. Optimal protein production of the WT EPO occurred when cultures were grown at 30°C for 5-7 hours. Higher temperatures resulted in a truncation of the protein (confirmed by western blot, Figure 2.5). The protein is present in the insoluble fraction.

### 2.2.2 Expression and purification of EPO cysteine variants

The expression and purification methods developed for the WT EPO were used for the 13 EPO cysteine variants discussed in section 2.2. For the EPO variants that had been produced by site-directed mutagenesis, these conditions yielded the EPO
variants of the correct size (confirmed by SDS PAGE and western blot analysis) for the K45C, K52C, K97C, K45C+K97C (double variant) and K45C+K52C (double variant) EPO (see Figure 2.6). Overexpression of the K140C EPO mutation was not seen in expressions carried out as low as 16°C (data not shown). This EPO variant was therefore eliminated from any further studies.

**Preparation of the cell-free extract and purification of WT EPO and EPO cysteine variants**

The His-tag on the N-terminus of the protein allowed for purification by metal chelating affinity chromatography. Briefly, cells were centrifuged and lysed by sonication and contaminating DNA was digested with DNaseI. The protein was solubilised in a denaturing buffer (Purification Binding buffer: 6 M guanidine hydrochloride, 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 5 mM imidazole) ready for Ni$^{2+}$ immobilised metal affinity chromatography purification (section 8.2.4). Initially, the
2.2 Expression of WT EPO and EPO cysteine variants

Figure 2.7: SDS PAGE of Ni-NTA purification of EPO cysteine variants. Lane 1: Ladder, Lane 2: W EPO, Lane 3: K45C EPO, Lane 4: K52C EPO, Lane 5: K97C EPO, Lane 6: K45C+K97C EPO, Lane 7: K45C+K52C EPO.

Solubilised cell-free extract was loaded onto a 5 ml His-Trap Ni$^{2+}$ affinity column for purification. The protein was eluted in a three step gradient and gave average yields of 8 mg l$^{-1}$ culture. However, it was found that purification by Ni-NTA resin yielded purer protein and so Ni-NTA was used to purify all protein samples. Figure 2.7 shows an SDS PAGE of these EPO variants after purification using Ni-NTA resin. For the Ni-NTA protocol, 1-2 ml of the nickel resin was incubated with the protein samples for 1 hour. This slurry was then applied to an empty column, washed, and the protein eluted with 1 M imidazole (section 8.2.4).

Expression of purchased, codon optimised EPO variants

Expression of the purchased EPO genes under these conditions yielded a truncated version of the EPO variants (Figure 2.8). All ten of the EPO genes purchased from GeneArt yielded this truncated product, including the gene optimised WT EPO.

Codon optimising the EPO gene sequence for expression in *E. coli* was done in an effort to increase protein production. However, as Figure 2.8 shows, codon optimisation resulted in expression of the truncated protein. To investigate the reason for this truncation and to optimise the expression conditions, the WT EPO (purchased from GeneArt and codon optimised) was used as a model. Initially, the temperature...
was lowered to 16°C, however this did not result in the expression of a full length protein. Next, 1% glucose was added to the LB medium to suppress the expression. Glucose affects expression levels by modulating the concentration of cyclic AMP (cAMP) in E. coli. Cyclic AMP concentration is inversely proportional to the concentration of glucose. When lactose (or IPTG) is present and glucose concentrations are low, cAMP complexes to a catabolite activator protein and facilitates transcription by enhancing the binding of RNA polymerase to the lac operon. When the concentration of glucose is increased, the levels of cAMP drop, essentially inhibiting transcription. As such, addition of glucose to the media can suppress basal levels of expression. This resulted in the non-truncated expression of the WT EPO (Figure 2.8).

As such, 1% glucose was added to the cultures of all codon optimised EPO cysteine variants purchased from GeneArt (section 8.2.4). Figure 2.9 shows the SDS PAGE gels of the expression of all codon optimised EPO variants. The preparation of the cell free extract and purification was done exactly as described previously (section 2.2) to yield purified protein solubilised in denaturing buffer.

![Figure 2.8: SDS PAGE gel of (A): truncated expression of the gene optimised WT EPO from GeneArt and (B): Complete expression of WT EPO gene from GeneArt](image-url)
2.2 Expression of WT EPO and EPO cysteine variants

Figure 2.9: SDS Gels of the expression of the A: T44C, B: R103C, C: S104C, D: N147C, E: R150C, F: V46C+R103C and G: R103C+R150C EPO gene variants
Chapter 2 Sub-cloning and expression of WT EPO and EPO cysteine variants

2.3 Oxidative refolding of WT EPO and EPO cysteine variants

With the correctly expressed and purified protein prepared, a method for the refolding of EPO was devised. As EPO was expressed in the insoluble fraction, a reduction/oxidation refolding method was required\textsuperscript{119}.

Initially, a previously reported on-the-column refolding was attempted\textsuperscript{175}. In this protocol, purified EPO samples were loaded on to a His-Trap Nickel column in denaturing conditions and reduced on the column with 10 mM DTT (in Purification Binding buffer, see section 8.1.2). The EPO was left to reduce for 3 hours before a refolding buffer (Refolding buffer containing CuSO\textsubscript{4}, see section 8.1.2) was used to oxidatively refold the protein and reform the internal disulfide bonds. This native, correctly folded protein could then be eluted from the column in Native Elution buffer (section 8.1.2). However, when this method was attempted it was found that the nickel column was not compatible with DTT, even at concentrations as low as 1 mM DTT. As such, a different approach was attempted; pulse refolding.

2.3.1 Pulse refolding of EPO

To refold the protein, purified EPO samples were reduced with 1 mM DTT for 3 hours at 37\textdegree C and then equilibrated to 4\textdegree C for pulse refolding. Pulse refolding involves the drop wise addition of the EPO samples into a native buffer (Refolding buffer, section 8.1.2) which contains CuSO\textsubscript{4} to aid in the oxidative refolding of EPO. It was found that the final concentration of EPO must be kept < 0.1 mg ml\textsuperscript{-1} to prevent aggregation.

Once the protein had been refolded, EPO samples were recovered by addition of Ni-NTA resin to bind the protein. After incubation for 1 hour, the slurry was applied to an empty PD-10 column and eluted in Native elution buffer (section 8.1.2).
2.4 Summary

In total, 11 amino acid positions (not including double mutations) were targeted for mutation to cysteine. Of these 11 targets, 4 were successfully introduced via

**Figure 2.10:** CD spectra of A: WT EPO (——) and BSA (——) and B: K45C EPO (-----), K97C EPO (———) and K45C+K97C (———) EPO. BSA was used as a standard as it has an \( \alpha \)-helical structure.

### 2.3.2 Circular dichroism

Circular dichroism (CD) was used to confirm the secondary structure of the protein. CD is used to determine the secondary structure of proteins as different structural features (\( \alpha \)-helix, \( \beta \)-sheet, coil) give different characteristic spectra. Unfolded proteins have a lack of regular structure and give fairly flat spectra.

The secondary structure of EPO consists primarily of four \( \alpha \)-helices with a few loops. Figure 2.10 shows the CD spectra of the WT EPO (see section 7 for CD spectra of EPO variants) compared to BSA (an \( \alpha \)-helical structured protein). These spectra are consistent with EPO’s secondary structure, confirming that the protein had been folded into an \( \alpha \)-helical structure. This data indicated that the protein is correctly folded, however, does not confirm the activity of the protein. To confirm activity, a UT7-EPO proliferation assay would need to be performed as discussed in section X. An activity assay of the EPO variants is not carried out in this thesis; however the importance of performing said assay is discussed in section 7.
site directed mutagenesis (K45C, K97C, K52C and K140C EPO), one (K154C) was never successfully introduced, and 6 (T44C, V46C, N147C, R150C, R103C, and S104C) were bought from GeneArt. In addition to these single point mutations, 4 double mutations were also generated (K45C+K52C, K45C+K97C, V46C+R103C, and R103C+R150C) giving a total of 14 (including the WT) EPO variants to be expressed.

All 14 EPO variants (including the WT EPO) were cloned into pET16b and expressed in BL21(DE3). Of these 14 variants, only 13 were successfully expressed as overexpression of the K140C mutation was never observed. All other variants were expressed, solubilised, purified and successfully refolded to give EPO cysteine variants that can be modified and analysed for their interaction with the EPO receptor, (EPOR)₂. Table. 2.4 lists the EPO variants that were successfully expressed and refolded and used in this thesis. Figure 2.11 shows the amino acid sequence of EPO with the His-tag and the positions of the cysteine mutations highlighted.
2.4 Summary

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
<th>Double cysteine variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>T44C</td>
<td>K97C</td>
<td>K45C+K97C</td>
</tr>
<tr>
<td>K45C</td>
<td>R103C</td>
<td>K45C+K52C</td>
</tr>
<tr>
<td>K52C</td>
<td>S104C</td>
<td>V46C+R103C</td>
</tr>
<tr>
<td>N147C</td>
<td></td>
<td>R103C+R150C</td>
</tr>
<tr>
<td>R150C</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Before analysing the interaction of these EPO cysteine variants with the (EPOR)$_2$, a method for the derivatisation of the free thiols was required.
3 Glycosylation of the EPO cysteine variants

3.1 Introduction

With the successful introduction of cysteine mutations into the EPO sequence, and the expression of these genes, the purified WT EPO and EPO cysteine variants were modified to yield ‘glycosylated EPO’. As discussed previously, a common method for selective modification of proteins is through the exploitation of thiol chemistry. The relatively low abundance of cysteine residues in protein sequences combined with their tendency to form internal disulfides means that cysteine residues are ideal candidates for modification and tagging\(^\text{176}\).

The modification of thiol groups is an ideal method for glycosylating EPO as the protein sequence contains no free thiols. In EPO, all four cysteine residues are involved in disulfide bonds and are not available for reaction. The introduction of cysteine residues into the EPO variants provided a selective and regiospecific site for chemical modification. The proportion, if any, of the EPO cysteine variants that are misfolded to expose natural cysteine was considered to be very small as the secondary structures of the EPO variants was determined by CD, and misfolded EPO is generally insoluble. Common methods used to modify thiols include alkylation by haloacetamides, oxidation of cysteine residues, disulfide formations, and conjugation with Michael acceptors (e.g. Maleimide)\(^\text{176}\). While each method has its advantages, oxidative chemistry (disulfide bond formation) was ruled out as many of the protocols for the preparation of EPO samples involve a number of oxidation/reduction
3.1 Introduction

Figure 3.1: Representation of the reaction of **Top**: The WT EPO with α-mannosyl iodoacetamide and **Below**: Reaction of an EPO cysteine variant with α-mannosyl iodoacetamide

steps. Reduction of the protein would cleave any new disulfide bonds formed between EPO and the glycosylating reagent.

The chosen method for the modification of EPO is based on previously reported work where the asparagine residues at the three natural N-linked glycosylation sites on EPO were mutated to cysteine. These EPO cysteine variants were then modified using glycosyl-iodoacetamides to yield a homogenous glycoprotein.

To modify the EPO variants, a monosaccharide moiety was chosen as they were quicker and easier to synthesise than the more complex carbohydrate structures found on glycoproteins. As this was an initial proof-of-concept investigation, the glycosylation does not need to exactly mimic that found on glycoproteins – the position and selectivity of modification, rather than the modification itself, is the key factor.

To determine the effect of site specific glycosylation of EPO on EPO-(EPOR)$_2$ binding, EPO cysteine variants were expressed, the proteins purified and modified with α-mannosyl iodoacetamide to yield a EPO glycoprotein (Figure 3.1).

To yield this EPO glycoprotein, several issues needed addressing:

1. A suitable analytical method for monitoring the glycosylation reaction needed
2. Optimised conditions for the reaction needed to be determined, and
3. If the reaction did not proceed to a 100% yield, a suitable purification method would be required.

3.2 The use of Liquid chromatography–mass spectrometry (LCMS) to monitor the glycosylation of EPO

One method for monitoring the reaction of cysteine residues is to use Ellman’s reagent (5,5’-dithiobis-(2-nitrobenzoic acid) or DTNB). Upon reaction with free thiols, the disulfide in Ellman’s reagent is cleaved to yield 2-nitro-5-thiobenzoic acid (TNB$^{2-}$) which can be monitored at 412 nm (Figure 3.2). Ellman’s reagent can be used to follow the progress of cysteine modification as the number of free thiols present in the sample will decrease with an increasing reaction yield. The signal at 412 nm will decrease as the reaction progresses. However, Ellman’s reagent monitors the depletion of the starting material and gives no information on the product. As such, any additional glycosylation that may occur as a result of non-specific reaction will not be detected.

Another direct and more accurate method to detect the glycosylation of EPO is to use mass spectrometry to determine any change in the mass of the protein as the reaction proceeds. LCMS has a number of advantages over some of the other

![Figure 3.2: Reaction of Ellman’s reagent with thiol containing compounds. DTNB reacts with sulfhydryl groups to yield the mixed disulfide and TNB$^{2-}$ which is measured at 412 nm.](image)
3.2 The use of Liquid chromatography–mass spectrometry (LCMS) to monitor the glycosylation of EPO

methods considered to monitor the glycosylation of EPO. A detectable tag (e.g. fluorescent labels) rather than a sugar moiety, could be used. However, the presence of a tag may interfere with any future assays to monitor EPO-(EPOR)$_2$ interactions.

To confirm the site-specific and selective glycosylation of EPO, LCMS is the ideal tool as it can be used to monitor both the starting material (EPO) and the product (glycosylated EPO).

### 3.2.1 Optimising the LCMS conditions

Protein species were detected by running samples on the LCMS on a C4 silica gel column (SUPELCOSIL™ LC-304) with an acetonitrile / water solvent system. Prior to loading on the column, all EPO samples were diluted into denaturing conditions (6 M guanidine hydrochloride, 500 mM imidazole, pH 7.8) to promote protein solubility and improve detection. The C4 column separated the protein from the high salt concentrations and denaturing agents required to solubilise EPO. Initial attempts to detect EPO on the LCMS gave broad, unresolved peaks (Figure 3.3) eluting at 90% acetonitrile. Reduction of the samples with DTT (10 mM, 30 minutes) prior to analysis on the LCMS gave a clear, resolvable signal as seen in Figure 3.3. Proteins are often reduced before performing an MS analysis, as addition of reducing agents breaks the internal disulfide bonds of the protein, helping the structure to unfold. This gives a sharper, clearer MS signal. Attempts to run samples in native

![Figure 3.3: Multiple-charge ion electrospray ionisation mass spectrum of the WT EPO](image)

Left: before reduction with DTT, Right: After reduction with DTT
buffer resulted in a poor MS signal, likely due to the presence of n-lauryl sarcosine (essential for protein stability) in the buffer.

Figure 3.4 shows the Electrospray ionisation (ESI) spectrum of WT EPO. ESI of large biological molecules, such as peptides and proteins, yields a multiple-charge ion spectrum of the protein. Unlike in small molecules where the M+1 peak is most commonly observed, proteins can obtain positive charges from protonation of their amino acid side chains and also from the ESI of the sample. This results in multiple highly charged peaks, with each peak differing by a charge of 1. The signal observed for the WT EPO (Figure.3.4) is a typical example of the ESI spectrum of a protein.

Deconvolution of the WT EPO spectrum gave a mass of 20913 ± 3.87Da. This value corresponds to the mass of the WT EPO missing the N-terminal methionine. Cleavage of N-terminal methionine residues by methionine aminopeptidase is a well documented phenomenon, favoured when the side chain at the secondary position is small

Masses were obtained for all expressed EPO cysteine variants and are given in Table.3.1. The masses of all of the EPO variants were as expected from the amino acid sequence within the error of the MS instrument. Since the MS is run under reducing conditions, presence of any oxidation at the newly introduced cysteine residues will not be observed. However the data in Table.3.1 shows that the protein backbones are as expected.
3.3 Reaction of EPO cysteine variants with α-mannosyl iodoacetamide

<table>
<thead>
<tr>
<th>Expected mass (Da)</th>
<th>Observed mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT EPO</td>
<td>20916.83</td>
</tr>
<tr>
<td>T44C EPO</td>
<td>20918.87</td>
</tr>
<tr>
<td>K45C EPO</td>
<td>20891.81</td>
</tr>
<tr>
<td>K52C EPO</td>
<td>20891.81</td>
</tr>
<tr>
<td>K97C EPO</td>
<td>20891.81</td>
</tr>
<tr>
<td>R103C EPO</td>
<td>20863.79</td>
</tr>
<tr>
<td>S104C EPO</td>
<td>20932.90</td>
</tr>
<tr>
<td>K140C EPO</td>
<td>20891.81</td>
</tr>
<tr>
<td>N147C EPO</td>
<td>20905.87</td>
</tr>
<tr>
<td>R150C EPO</td>
<td>20863.79</td>
</tr>
<tr>
<td>K45C+K52C EPO</td>
<td>20866.78</td>
</tr>
<tr>
<td>K45C+K97C EPO</td>
<td>20866.78</td>
</tr>
<tr>
<td>V46C+R103C EPO</td>
<td>20867.80</td>
</tr>
<tr>
<td>R103C+R150C EPO</td>
<td>20810.75</td>
</tr>
</tbody>
</table>

3.3 Reaction of EPO cysteine variants with α-mannosyl iodoacetamide

With a method for identifying EPO species using LCMS in place, reaction of the EPO cysteine variants with α-mannosyl iodoacetamide was carried out.

To obtain samples for the reaction, the EPO gene of interest was expressed, the protein purified and refolded as outlined in section 8.2. After the refolding process and addition of the Ni-NTA to carry out affinity chromatography, samples were eluted from the nickel affinity column in denaturing buffer. Analysis of samples on an SDS PAGE gel revealed only one protein species, indicating pure EPO samples. As such, any contaminating protein species is undetectable by SDS PAGE (hence low concentrations) and is unlikely to interfere with EPO reactions. Furthermore, in the unlikely event that any contaminating protein is present and reacts with the α-mannosyl iodoacetamide, it is will most likely not appear in the MS signal. The α-mannosyl iodoacetamide was added to the desired concentration and the reaction
mixtures were reduced and analysed by LCMS.

All reactions were carried out in the dark, at pH between 7.5-8.0 and in denaturing buffer (6 M guanidine hydrochloride, 50 mM Tris-HCl, 200 mM NaCl, 1 M imidazole). These conditions were required for the following reasons:

Iodoacetamide agents are light sensitive and prone to hydrolysis. As such, solutions were prepared immediately before use and all reactions were conducted in the absence of light.

Cysteine is more reactive above its pKa (~8.2) as the thiolate ion is a better nucleophile than the thiol. However, reactions at these alkaline pH’s will increase the likelihood of unspecific amine modification. As such, reactions were conducted between pH 7.5–8.0

Denaturing conditions were used as it was found that reactions carried out under native conditions gave poor MS signals which could not be deconvoluted. This is

![Figure 3.5: Mass spectra of a typical glycosylation reaction. A: K97C EPO (1 mg/ml) was glycosylated with 1 mM α-mannosyl iodoacetamide. Samples were taken after 1 hour and reduced with DTT. B: Mass spectrum shows multiple-charge peaks for both the K97C EPO (M = 20894 Da) and the mono-glycosylated K97C EPO (M+Man = 21158 Da).](image-url)
likely due to the presence of residual n-lauroylsarcosine in the buffer. Removal of the n-lauroylsarcosine from the buffer results in a clearer MS signal; however the protein is then prone to aggregation and tends to precipitate. Imidazole is present in all reaction buffers as it has been reported to improve the reaction specificity\textsuperscript{179}.

Initial reaction conditions were based upon previously reported work and the manufacturers (iodoacetamide suppliers) recommended parameters\textsuperscript{179}. Analysing the reaction of EPO with $\alpha$-mannosyl iodoacetamide using LCMS resulted in the inseparable (by HPLC) elution of both the starting material (EPO) and the product (mannosylated EPO). Figure 3.5 shows the spectrum of an EPO cysteine variant after reaction with the $\alpha$-mannosyl iodoacetamide in non-optimized conditions. While the LCMS is not quantitative, the two species are likely to give signal intensities corresponding to their rations within the reaction mixture, as they are almost identical (and hence behave the same in the MS chamber). Figure 3.5 shows that the reaction is not very efficient after 1 hour as an excess of the $\alpha$-mannosyl iodoacetamide was used and the reaction did not proceed to completion (a 100% yield will give only one EPO species by MS). This data highlighted the need for the optimization of the reaction conditions.

### 3.3.1 Optimisation of the glycosylation reaction

Almost all of the single EPO cysteine variants were screened and the conditions optimised for their reaction with $\alpha$-mannosyl iodoacetamide. It was found that the EPO variants all exhibited similar trends in terms of reactivity. The EPO cysteine variant K97C EPO is therefore used consistently throughout this chapter to illustrate the optimisation process.

The reaction of the double cysteine EPO variants is outlined in section 3.4.

For each EPO cysteine variant, the concentration of both the EPO and the $\alpha$-mannosyl iodoacetamide was varied. As a control for each reaction, the WT EPO was exposed to the same conditions as the EPO cysteine variants. The data for the WT EPO is shown in Figure 3.8. Reactions were considered specific when an
Figure 3.6: Mass spectra following the reaction of K97C EPO with α-mannosyl iodoacetamide under varying conditions. Samples were taken after 1 hour of reaction and analysed by LCMS. Spectra show the non-modified K97C EPO (M = 20894 Da), the mono-glycosylated K97C EPO (M+[Man] = 21158 Da) and the di-glycosylated K97C EPO (M+2[Man] = 21420 Da). Addition of the second mannose residue is the result of an unspecific reaction. Conditions: 6 M Guanidine hydrochloride, 50 mM Tris-HCl, 200 mM NaCl, 1 M Imidazole, pH 7.8.

addition was seen in the cysteine variants, but not in the WT EPO; addition of the α-mannosyl iodoacetamide should occur exclusively at the cysteine position, and so
3.3 Reaction of EPO cysteine variants with α-mannosyl iodoacetamide

<table>
<thead>
<tr>
<th>Protein concentration</th>
<th>0.5 mg/mL</th>
<th>1 mg/mL</th>
<th>2 mg/mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 mM</td>
<td>N.D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 mM</td>
<td>M</td>
<td>M+[Man]</td>
<td>M+2[Man]</td>
</tr>
<tr>
<td>1 mM</td>
<td>M+Man</td>
<td>M+2Man</td>
<td></td>
</tr>
<tr>
<td>0.5 mM</td>
<td>M</td>
<td>M+Man</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.7:** Mass spectra following the reaction of K97C EPO with α-mannosyl iodoacetamide under varying conditions. Samples were taken after 5 hours of reaction and analysed by LCMS. Spectra shown the non-modified K97C EPO (M = 20894 Da), the mono-glycosylated K97C EPO (M+[Man] = 21158 Da) and the di-glycosylated K97C EPO (M+2[Man] = 21420 Da). Addition of the second mannose residue is the result of an unspecific reaction.

the WT EPO should show no change in mass.

Figure 3.6 and Figure 3.7 show the reaction of the K97C EPO variant with the α-mannosyl iodoacetamide under varying conditions after 1 hour and 5 hours respectively. Comparison of the same reaction conditions after 1 and 5 hours show
that the reactions proceeded quickly and are usually completed within 5 hours – there is no change in the mass spectrum after this time (data not shown). As shown in Figure 3.6 and Figure 3.7, it became clear that the reaction of α-mannosyl iodoacetamide did not proceed to completion as the presence of the starting material (EPO) was always detected. Rather than compete mono-glycosylation, multiple (di-, tri- and multi-) additions of the α-mannosyl iodoacetamide to the EPO cysteine variants were observed. This additional glycosylation was seen when concentration of the α-mannosyl iodoacetamide was increased in both the EPO cysteine variants and the WT EPO (see Figure 3.8, reaction of WT EPO with 2 mM α-mannosyl iodoacetamide). Detection of the addition reaction in both the cysteine variants and the WT EPO indicate reaction at a non-cysteine residue. Common side products in addition reactions to proteins occur on surface lysine residues and at the N-terminus. The LCMS data does not allow for the clarification of the site of the additional glycosylation; however a tryptic digest would yield the information if necessary.

As seen in Figure 3.7, increasing the concentrations of the reagents (2 mg/ml EPO variant, 2 mM α-mannosyl iodoacetamide) in an effort to increase the reaction yield resulted in the multiple addition, rather than the complete reaction, of the α-mannosyl iodoacetamide to the EPO variants. Using a large excess of the α-mannosyl iodoacetamide (above 5 mM) resulted in a poor MS signal which could not be deconvoluted.

In an attempt to inhibit the additional glycosylation, the reagent concentrations were decreased. However, this resulted in slower reaction times and a lower reaction yield (estimated from the peak ratios in the MS).

**Reaction of the WT EPO**

Reaction conditions that resulted in the selective mono-glycosylation of the EPO cysteine variants were screened against the WT EPO. Figure 3.8 shows the mass spectrum WT EPO after reaction with the α-mannosyl iodoacetamide under the same conditions as the EPO cysteine variants. As expected, most conditions resulted in no addition of the α-mannosyl iodoacetamide to the WT EPO. However,
3.3 Reaction of EPO cysteine variants with α-mannosyl iodoacetamide

<table>
<thead>
<tr>
<th>α-mannosyl iodoacetamide concentration</th>
<th>Protein concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5 mM</td>
<td>N.D.</td>
</tr>
<tr>
<td>1 mM</td>
<td>M</td>
</tr>
<tr>
<td>2 mM</td>
<td>M</td>
</tr>
</tbody>
</table>

**Figure 3.8:** Mass spectra following the reaction of WT EPO with α-mannosyl iodoacetamide under varying conditions. Samples were taken after 5 hours of reaction and analysed by LCMS. Spectra shown the unmodified WT EPO (M) and the mono-glycosylated WT EPO (M+[Man]). Addition of the any mannose residue is the result of an unspecific reaction.

Increasing the α-mannosyl iodoacetamide concentration to 2 mM resulted in unspecific addition, confirming that an excess of glycosylating agent resulted in unspecific
addition (Figure 3.8).

### 3.3.2 Improving the selectivity of the EPO glycosylation reaction

Figure 3.6 and Figure 3.7 show that the specificity of the reaction is compromised when excess α-mannosyl iodoacetamide is used. Under these conditions, multiple additions, rather than complete reaction to the mono-glycosylated species, is observed on both the EPO cysteine variants and on the WT EPO.

The additional glycosylation is most likely the result of either; reaction of at the N-terminus of EPO with the α-mannosyl iodoacetamide. This has been previously reported. Or, unspecific reaction at the lysine residues.

The aim of this investigation is to determine if the site specific glycosylation of EPO will inhibit the EPO-(EPOR)₂ interaction. If there are unknown sites of glycosylation present on EPO, then it cannot be confidently said that any inhibition observed is due to a specific modification. As such, concentrations of reagents were limited to values where only monoglycosylation was observed. This included reaction at 0.5 mg/ml < [EPO] < 1 mg/ml and [α-mannosyl iodoacetamide] ~1 mM.

**The effect of pH on specificity of the glycosylation reaction**

The non specific addition of the α-mannosyl iodoacetamide to either the surface lysine residues or the N-terminus of the EPO protein is favoured in alkaline pH. To improve the selectivity of glycan addition, reactions were carried out at pH 7.4 (previously reactions were conducted at pH 7.8).

The protocol for the reaction remained the same; EPO samples were purified, refolded and eluted from a nickel affinity chromatography column in denaturing buffer (6 M Guanidine hydrochloride, 200 mM NaCl, 1 M imidazole, pH 7.4).

Figure 3.9 shows the mass spectra of the reaction of K97C EPO with α-mannosyl iodoacetamide after 5 hours at pH 7.4. Reactions carried out under these conditions
3.3 Reaction of EPO cysteine variants with α-mannosyl iodoacetamide

<table>
<thead>
<tr>
<th>α-mannosyl iodoacetamide concentration</th>
<th>Protein concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 mM</td>
<td>1 mg/mL</td>
</tr>
<tr>
<td>1 mM</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>M+Man</td>
</tr>
<tr>
<td>0.5 mM</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.9:** Mass spectra following the reaction of K97C EPO with α-mannosyl iodoacetamide at pH 7.4. Samples were taken after 5 hours (*16 hours, **36 hours) of reaction and analysed by LCMS. Spectra show the unmodified K97C EPO (M) the mono-glycosylated K97C EPO (M+[Man]) and the di-glycosylated K97C EPO (M+2[Man]). Addition of the second mannose residue is the result of an unspecific reaction.

...proceeded at a comparatively slower rate and with lower overall yields than reactions conducted at pH 7.8. When the concentrations of the reagents were increased at pH 7.4, no improvement in the specificity was observed; the reaction led to multiple, unspecific additions of the α-mannosyl iodoacetamide.

**Summary**

From the mass spectra analysis of different reaction conditions, optimal yields were gained when 1 mg/ml EPO cysteine variants (single cysteine variants) were glyco-
sylated using 1 mM α-mannosyl iodoacetamide for ~ 5 hours in the absence of light.

The yield for the glycosylation of EPO was estimated to be ~50% (this was determined qualitatively, not quantitatively, by comparison of mass peak intensity between starting material and product). Figure 3.10 shows the mass spectra of EPO cysteine variants after reaction with the α-mannosyl iodoacetamide under optimised conditions. The reactions yield similar mixtures of unmodified and monoglycosylated EPO species, independent of the cysteine position.
3.3 Reaction of EPO cysteine variants with α-mannosyl iodoacetamide

<table>
<thead>
<tr>
<th></th>
<th>1 hour</th>
<th>5 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTEPO</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td>T44CEPO</td>
<td><img src="image3" alt="Graph" /></td>
<td><img src="image4" alt="Graph" /></td>
</tr>
<tr>
<td>K45CEPO</td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
</tr>
<tr>
<td>KS2CEPO</td>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
</tr>
<tr>
<td>K97CEPO</td>
<td><img src="image9" alt="Graph" /></td>
<td><img src="image10" alt="Graph" /></td>
</tr>
<tr>
<td>N147CEPO</td>
<td><img src="image11" alt="Graph" /></td>
<td><img src="image12" alt="Graph" /></td>
</tr>
</tbody>
</table>

**Figure 3.10:** Mass spectra of EPO cysteine variants (1 mg/ml\(^{-1}\)) after reaction with α-mannosyl iodoacetamide (1 mM). Spectra show the unmodified protein samples (M) and the mono-glycosylated protein (M+\([\text{Man}]\)).
3.4 Reaction of double cysteine EPO variants with α-mannosyl iodoacetamide

The double cysteine EPO proteins contain two cysteine residues in the protein sequence. The expression and purification procedures were identical to the single cysteine EPO variants and are outlined in section 8.2.4.

Applying the conditions optimised for the reaction of the single EPO variants (1 mg ml\(^{-1}\) EPO, 1 mM α-mannosyl iodoacetamide, section 3.3) with α-mannosyl iodoacetamide to the reaction of the K45C+K97C EPO yielded a highly glycosylated product which could not be identified; analysis of the reaction by LCMS was not possible as defined peaks could not be identified. The reaction of the K45C+K97C EPO proceeded at an increased rate compared to the single cysteine variants.

Optimal yields for the double cysteine EPO variant was observed when the concentration of the α-mannosyl iodoacetamide was reduced to 0.5 mM and the reaction time limited to 30 minutes. The reaction of K45C+K97C EPO (Figure 3.11) yielded a mixture of three EPO species; the unmodified EPO, the mono-glycosylated EPO (addition to one cysteine) and the di-glycosylated EPO (addition to the second cysteine). The mass spectra following the reaction of K45C+K97C EPO with α-mannosyl iodoacetamide are shown in Figure 3.11.

Figure 3.11 shows that the glycosylation of K45C+K97C EPO proceeded in a stepwise manner with the addition of one mannosyl moiety after 10 minutes, followed by a second addition within 30 minutes. Reaction times longer than 30 minutes resulted in undefined spectra that could not be deconvoluted. As such, the reactions were quenched after 30 minutes yielding a mixture of the unmodified EPO, the mono-glycosylated EPO and the di-glycosylated EPO with ~1:1:1 ratio (Figure 3.11). The preference, if any, of the first mannosyl moiety addition is unknown. It is possible that one of the cysteine residues is more readily available for reaction and addition occurs at this position, followed by addition of the second mannose to the remaining cysteine residue.
3.5 Summary

Figure 3.11: A: Representation of the reaction of K45C+K97C EPO with α-mannosyl iodoacetamide. Reaction yields a mixture of three EPO species; the unmodified EPO, the mono-glycosylated EPO and the di-glycosylated EPO. B: Mass spectrum of the reaction of K45C+K97C EPO (1 mg/ml) with α-mannosyl iodoacetamide (0.5 mM) after Left: 10 minutes and Right: 30 minutes. Reaction yields a mixture of three EPO species; the non-modified EPO (M), the mono-glycosylated EPO (M+[Man]) and the di-glycosylated EPO (M+2[Man]).

3.5 Summary

Under optimised reaction conditions, the glycosylation of EPO cysteine variants yielded a mixture of unmodified EPO and the mono-glycosylated EPO. In the case of the double variants, three species were detected by mass spectrometry; the unmodified EPO, the mono-glycosylated EPO and the di-glycosylated EPO.

Attempts to increase the yield of mono-glycosylated EPO resulted in unspecific multiple addition of the α-mannosyl iodoacetamide. The most likely explanation for this that a percentage of the disulfide bonds are not free for reaction as that are already
involved in disulfide bonds. Any disulfide modification would not be observed by mass spectrometry (all observed masses of EPO samples correspond to the expected mass) as any disulfide would be cleaved when the samples are reduced prior to LCMS. The most likely source of a disulfide modification is an intermolecular disulfide bond between two EPO molecules. Other possibilities include the possibility of misfolded EPO; the introduced cysteine could be involved in an internal disulfide bond, while a natural cysteine is ‘free’ and possibly inaccessible for reaction, or the formation of disulfide bonds with a component of the cell during expression. The latter possibility is the least likely, as the EPO protein undergoes a reduction step during the refolding process, before the reaction.
4 Purification of the EPO glycosylation reaction

4.1 Introduction

As discussed in chapter 3, reaction of EPO cysteine variants with α-mannosyl iodoacetamide yielded a mixture of products. When glycosylating the single cysteine variants, the unmodified EPO and the mono-glycosylated EPO species were observed. Reaction of the double cysteine EPO variants yielded the unmodified EPO, the mono-glycosylated EPO and the di-glycosylated EPO species (Figure 4.1).

To obtain purified glycosylated EPO, the separation of the modified and unmodified EPO was attempted. There are a number of ways in which the differences between the modified and unmodified proteins can be exploited to obtain purified EPO species. Approaches considered included:

1. Exploitation of the free thiol on the non-modified EPO for thiol selective chromatography.

2. Selective binding of the mannosylated EPO on a lectin affinity chromatography column.

3. Separation of the two species based upon physical attributes, e.g. difference in hydrophobic character.

The thiol selective chromatography and lectin binding chromatography approaches share a common theme; both methods rely on the selective capture of one EPO
Figure 4.1: *Top:* Reaction of single cysteine EPO variant with α-mannosyl iodoacetamide to yield the unmodified and glycosylated EPO. *Below:* Reaction of double cysteine EPO variant to yield the unmodified, the mono-glycosylated EPO and the di-glycosylated EPO species.

species on a column while the other species passes freely through the medium (Figure 4.2).

In lectin affinity chromatography, a lectin specific for the sugar of interest (here, the α-mannose binding lectin Concanavalin A, ConA, could be used) is used to selectively bind the mannosylated EPO, while the unmodified EPO passes through the column. The mannosylated EPO can then be eluted from the column and used in further analysis.

Conversely, a thiol selective column uses disulfide chemistry to selectively bind proteins containing free thiol groups. When purifying the EPO reaction products on a thiol selective column, the cysteine residues on the non-modified EPO would undergo disulfide exchange with the column. This species would be retained on the column while the modified EPO would wash through the column.

This latter method, purification by thiol selective chromatography, holds an advantage over the use of lectin affinity chromatography as the species of interest (the
4.1 Introduction

Figure 4.2: Left: Purification of EPO variant glycosylation reaction using thiol selective chromatography. Any glycosylated EPO will elute during the wash steps while any unmodified EPO will undergo disulfide exchange with the thiol sepharose and retain on the column. The retained EPO can be recovered by reducing the newly formed disulfide. Right: Purification of EPO variant glycosylation reaction using lectin affinity chromatography. Any glycosylated EPO will bind to the lectin of choice while the unmodified EPO will elute in the wash step.

glycosylated EPO) will not bind to the column and can be isolated without any further steps. In lectin affinity chromatography, the glycosylated EPO is retained on the column and needs to be eluted. Therefore thiol selective chromatography was employed in initial attempts to separate the EPO species.
4.2 Purification of the EPO glycosylation reaction using thiol selective column chromatography

Before attempting to separate the glycosylation reaction products, optimal conditions for the binding of thiol containing proteins to the sepharose matrix were determined using a GFP model system.

4.2.1 Binding of Green Fluorescent Protein to the Activated thiol sepharose

GFP is frequently used as a reporter protein in expression and imaging studies as it is thermo- and pH stable, biologically compatible and its presence can be confirmed visually. For our investigation, two GFP variants were used; the WT GFP and a GFP cysteine variant, I229C GFP.

The WT GFP sequence contains a single free cysteine residue which is buried within the beta-barrel structure of GFP and as such is not available for modification. The I229C GFP contains an additional cysteine residue at the 229 position. As with

![Diagram of GFP binding](image)

**Figure 4.3:** Selective binding of thiol containing proteins.
4.2 Purification of the EPO glycosylation reaction using thiol selective column chromatography

the EPO cysteine variants, the WT protein should flow freely through the medium while the cysteine variant should bind to the thiol sepharose (Figure. 4.3)

These GFP variants were used as model systems for the binding of EPO to the activated thiol sepharose column.

GFP variants were supplied as 2 mg ml\(^{-1}\) solutions in PBS. The activated thiol sepharose was prepared and used according to the manufacturer’s instructions. The thiol sepharose freeze-dried powder was washed and swollen with distilled water, then equilibrated in PBS. Both the WT GFP and the I229C GFP were incubated with ~ 1 ml of thiol slurry per 1 mg of protein and left to bind for 2 hours at 4\(^{\circ}\)C. Incubation allows any free thiols in the samples to undergo disulfide exchange with the thiol sepharose matrix. The samples were then applied to empty columns and the flow through collected. The columns were washed with PBS and fractions analysed on an SDS PAGE gel. Any bound samples were eluted with PBS+20 mM DTT. It was found that any bound GFP eluted only after incubation with the DTT for >10 minutes (confirmed visually). Elution fractions were collected and analysed on an SDS PAGE gel.

Figure. 4.4 shows the results of binding both the WT GFP and the I229C GFP to the thiol sepharose matrix. For the WT GFP, the protein was eluted in the first 2 fractions of the wash step with no further elution after reduction of the column.

**Figure 4.4:** SDS PAGE gel of the wash (W\(^{1}\)-W\(^{4}\)) and reduction (E\(^{1}\)-E\(^{4}\)) fractions from the binding of WT GFP (panel A) and I229C GFP (panel B) to the thiol sepharose column (B=Blank well). Panel C shows the wash and elution fractions from the binding of glycosylated I229C GFP to the column.
This confirms that the cysteine residue in the WT GFP is not accessible for reaction or modification.

From the SDS PAGE gel (Figure. 4.4) it was shown that the I229C GFP binds selectively to the thiol column as protein was detected in the reducing fractions.

With the binding procedure optimised, the I229C GFP was glycosylated with α-mannosyl iodoacetamide. Upon incubation of the glycosylated GFP with the thiol sepharose the protein was eluted in the wash steps. The SDS PAGE gel in Figure. 4.4C shows only small amounts of GFP were eluted after incubation with the sepharose column. This shows that the majority of the GFP did not bind to the column due to successful blocking of the cysteine residues via the glycosylation reaction. The eluted material was presumed to be GFP that had not undergone the glycosylation reaction.

**Binding of GFP to activated thiol sepharose under denaturing conditions**

To determine the compatibility of the thiol sepharose with denaturing agents, non-modified WT and I229C GFP were incubated with the thiol sepharose resin in both 6M guanidine hydrochloride and 6M urea. The procedure was as above (section. 8.4.1) with the exception of the buffer. Fractions from the wash and reduction steps were collected, concentrated and analysed on an SDS PAGE gel.

As shown in Figure. 4.5, I229C GFP binds to the thiol column in the presence of 6 M urea. The SDS gel shows that the bound:unbound GFP ratio is ~50:50 (visual estimations by the intensity of the bands on SDS PAGE gel). From this it was clear that the binding of I229C GFP to the thiol sepharose is more efficient in a PBS buffer system than a urea system.

In contrast, running the column in 6 M guanidine hydrochloride resulted in no (or very little) binding of the I229C GFP (Figure. 4.5). The procedure also appears to have caused degradation/ cleavage of the GFP samples. Since guanidine hydrochloride is compatible with GFP, it was presumed to be a combination of the buffer and the thiol column that results in the cleavage observed.
4.2 Purification of the EPO glycosylation reaction using thiol selective column chromatography

Figure 4.5: SDS PAGE gel of the wash (W) and reduction (E) fractions of the WT and I229C GFP bound to the thiol sepharose column under denaturing conditions. Panel A: Binding in the presence of 6M urea, Panel B: Binding in the presence of 6M guanidine hydrochloride (C = GFP control).

From these results it can be concluded that PBS is the optimal buffer system to use when binding GFP to the thiol sepharose column. While 6M urea can be used as a denaturing alternative, 6M guanidine hydrochloride is not compatible with this purification method.

4.2.2 Binding of EPO variants to the activated thiol sepharose

The WT EPO and the K52C EPO were used in initial test for the binding of the EPO variants to the thiol sepharose column. The WT EPO contains no free cysteine residues and therefore flow freely through the medium. The K52C EPO variant has a single free cysteine residue and would therefore bind to the thiol sepharose.

As PBS was found to be the optimal buffer for the binding of GFP to the thiol sepharose, initial attempts to bind the EPO cysteine variants were done under native conditions. The WT and K52C EPO were purified and refolded as outlined in section 8.2.4 and the protein samples were eluted in Native Elution buffer (50mM Tris-HCl (pH 7.8), 200mM NaCl, 2% w/v n-lauryl sarcosine, 1M Imidazole).

Both the WT EPO and the K52C EPO were incubated with ~ 1ml of thiol slurry per 1mg of protein as above. These samples were then applied to empty columns
and washed with the Native Elution buffer and the fractions were collected and run on an SDS PAGE gel. An elution buffer (Native Elution buffer + 20 mM DTT) was applied to both columns to reduce any newly formed disulfide bonds and elute any bound material. These fractions were collected and analysed on an SDS PAGE gel to determine the presence of protein.

In the case of both the WT EPO and the K52C EPO, the protein was not eluted in either the wash or the reduction steps; both variants were retained on the column. This non-specific binding can be attributed to the presence of n-lauryl sarcosine in the buffer. N-lauryl sarcosine is an ionic surfactant that helps to solubilize proteins by preventing aggregation. The importance of n-lauryl sarcosine is discussed in section 5.4. N-lauryl sarcosine is known to interfere with various column chromatography procedures, therefore the use of native conditions for the purification of the glycosylation reaction was eliminated.

Binding of EPO variants to activated thiol sepharose under denaturing conditions

Once GFP had been used as a model system to determine the usefulness of the thiol column, attempts were made to purify the EPO samples. While guanidine hydrochloride was not compatible with the GFP system, both guanidine hydrochloride and urea were used in conjunction with the EPO samples to determine the optimal buffer.

Figure 4.6 shows the result of incubating K52C EPO and the WT EPO with the thiol sepharose column. Binding was done in 6 M, 5 M and 4 M guanidine hydrochloride to test for optimal conditions. Rather than selectively bind to the matrix, both protein samples retain loosely on the column and elute out over a series of washes. These results indicated that the column is not compatible with high concentrations of guanidine hydrochloride.

As discussed in section 4.2.1, the binding of GFP to the thiol column was successful in 6 M urea. Both the WT and K52C EPO were incubated with varying urea
4.2 Purification of the EPO glycosylation reaction using thiol selective column chromatography

Figure 4.6: SDS PAGE gel of the wash (W1-W4) and reduction (E1-E4) fractions of the WT and K52C EPO bound to the thiol sepharose column under denaturing conditions. Binding in the presence of; Panel A: 6 M guanidine hydrochloride, Panel B: 5 M guanidine hydrochloride, Panel C 4 M guanidine hydrochloride.

concentrations to determine the optimal binding conditions. Using 8 M urea resulted in elution of both the WT and K52C EPO from the column during the wash steps (Figure. 4.7). Decreasing the concentration to 6 M, 5 M and 4 M urea to promote the selective binding of the K52C EPO over the WT EPO resulted in no selective binding to the column (Figure. 4.7).

To ensure that these results were not specific to the K52C EPO variant, identical experiments were carried out using the double cysteine EPO variant, K45C+K97C EPO. The results confirmed that thiol sepharose chromatography was not a suitable method for purifying the EPO glycosylation reaction.
4.3 Purification of the EPO glycosylation reaction using hydrophobic interaction chromatography

Hydrophobic interaction chromatography (HIC) separates proteins based upon the differences in their surface hydrophobicity. The reversible binding of proteins to the HIC medium can be manipulated by varying the salt concentration in the buffer.

High ionic strength buffers promote the hydrophobic interactions between protein and column, resulting in the retention of the protein on the column. Using low salt buffers weakens these interactions and the protein elutes from the column. Hydrophobic proteins are eluted at these lower salt concentrations.

The separation of the unmodified and glycosylated EPO using HIC was attempted. As HIC is not compatible with denaturing agents, EPO samples were loaded and purified in a native buffer: 50 mM Tris-HCl, 500 mM imidazole, 2% n-lauryl sarcosine, 1 M (NH₄)₂SO₄ (or 3 M NaCl).

Unmodified K52C was used to optimise the binding conditions of EPO variants to the hydrophobic column.

Initially, ammonium sulfate was used to vary the ionic strength of the buffer. The column was equilibrated with the loading buffer; 50 mM Tris-HCl, 500 mM imidazole,
4.3 Purification of the EPO glycosylation reaction using hydrophobic interaction chromatography

2% n-lauryl sarcosine, 1 M (NH₄)₂SO₄, the K52C EPO was loaded to the column and a wash gradient of 1 M to 0 M (NH₄)₂SO₄ was run. Fractions were collected and analysed on an SDS PAGE gel to confirm the presence of protein.

It was found that K52C EPO eluted between 100 mM and 0 M (NH₄)₂SO₄. A small percentage of the protein was retained on the column and was only eluted under denaturing conditions. As such, the binding of the EPO to the column under these conditions was deemed too strong and a salt with a lower ionic strength was chosen; NaCl.

Repeating the chromatography procedure with 3 M NaCl resulted in the elution of K52C EPO at 1.5> [NaCl] >1.0 M (Figure 4.8).

**Separating unmodified and glycosylated EPO using HIC**

The separation of the products from the EPO glycosylation reaction was attempted using the conditions described in section 4.3. The K52C EPO was prepared and glycosylated with α-mannosyl iodoacetamide as outlined in section 8.3. The reaction products were exchanged into the loading buffer; 50 mM Tris-HCl, 500 mM imidazole, 2% n-lauryl sarcosine, 3 M NaCl and loaded onto the HIC column. A gradient of 3 M to 0 M NaCl was used to separate and elute the reaction products.

![Figure 4.8](image_url): SDS PAGE gel of fractions from the HIC column containing protein. Panel A: Elution of K52C EPO from HIC column, Panel B: Elution of K52C EPO from HIC column after reaction with α-mannosyl iodoacetamide.
Figure 4.8 shows the SDS gel of the fraction collected after HIC. No separation of the glycosylated and non-glycosylated K52C EPO was observed; both species eluted simultaneously.

No separation of the two EPO species was observed in further studies (data not shown) and so the use of HIC to purify the EPO glycosylation reaction was abandoned. No further techniques were attempted for the purification of glycosylated and non-glycosylated EPO species. This decision was influenced by the data already collected – the presence of stabilizing detergents is known to affect the binding of proteins to chromatography columns and so other chromatographic techniques were ruled out e.g. Lectin binding chromatography.

4.4 Summary

The use of thiol selective chromatography for the separation of thiol and non thiol containing GFP was optimised. The I229C GFP variant was selectively bound to the thiol column, with optimal binding occurring under native conditions.

However, the application of these conditions to an EPO system was unsuccessful. In native conditions, both the WT EPO and thiol containing EPO variants non-specifically retained on the column, likely due to the presence of detergent in the buffer.

Similarly, when incubating the EPO variants on the thiol column under denaturing conditions no binding was detected. Decreasing the denaturing agent concentrations resulted in non-specific retention of both the WT and cysteine EPO variants.

Purification of the glycosylation reaction using HIC was also unsuccessful as the two EPO species could not be separated based upon their hydrophobicity alone. The purification of the glycosylated and unmodified EPO variants was made difficult by the unstable nature of the protein.

It became apparent that the reaction products could not be separated using these techniques and so it was decided that the EPO samples would be used directly after
the glycosylation reaction without further modification. As a result, any techniques used to observe the binding and/or proliferation of the glycosylated EPO will be done using these unpurified EPO samples. As such, any effect on the binding of EPO to (EPOR)\textsubscript{2} caused by the glycosylation will be more difficult to monitor due to interference from the non-glycosylated EPO. Additionally, while variations in the glycosylation reactions between samples is very small (i.e. the extent of glycosylation of K45C varies little every time the reaction is done) without a quantitative measure, only experiments carried out at the same time can be compared. Experiments using sample glycosylated separately cannot be compared.
5 Development of an assay to monitor EPO-(EPOR)$_2$ binding interactions

In order to determine the effect of EPO modification on the EPO-(EPOR)$_2$ interaction, a method to monitor the binding of EPO to (EPOR)$_2$ was devised. The method used for detection of protein interactions depends on the protein, availability of proteins and antibodies, and the parameter to be measured. Discussed below are the methods for the detection of protein-protein interactions, and the development and optimization of an assay to monitor EPO-(EPOR)$_2$ interactions.

5.1 Protein interaction detection methods

Free solution techniques

Commonly used methods for the detection of protein interaction in solution include fluorescence resonance energy transfer (FRET) and isothermal titration calorimetry (ITC). FRET can be used to measure protein-protein interactions in vitro and in vivo. In this technique, a donor and acceptor fluorophore are tagged to the protein pair of interest. FRET is observed by excitation at the donor wavelength and fluorescence is measured at the donor and acceptor emission wavelengths. Upon close association of the two protein partners (1-10nm), fluorescence is seen predominantly...
5.1 Protein interaction detection methods

at the acceptor emission wavelength. A disadvantage of this method is the need for labeling of the two proteins of interest.

ITC is a label free method used to monitor the reversible binding of two biomolecules. In ITC, the heat changes associated with the binding of a ligand and a protein are monitored as small aliquots of ligand are added to the protein. As the molar ration of ligands to protein increases, the associated heat change begins to decrease, until there are more ligands than receptor. Upon reaching this saturation point, no more heat changes are observed. ITC is used to determine binding affinity, reaction stoichiometry, and enthalpy. The major disadvantage to ITC is the relatively larger sample volumes and concentrations required for the experiments. Protein cells are usually 1-2mL in capacity, significantly higher than techniques carried out in 96 well formats.

Surface bound techniques

Surface bound methods for the detection of protein interactions are SPR (discussed in section 1.6) and ELISA’s. ELISA’s are possibly the most common analytical tool employed to detect the presence of an antigen or substrate using antibodies. In a standard ELISA an antigen is attached to the surface of a 96 well plate.

In a standard ELISA an antigen is attached to the surface of a 96 well plate. An antibody which is specific to this antigen is then applied to the wells where it binds to an epitope on the antigen. This antibody is either directly linked to an enzyme (such as horseradish peroxidase, HRP), or contains an epitope for a secondary enzyme-linked antibody to bind. Finally, a substrate for the enzyme (such as 3,3’,5,5’-tetramethybenzidine, TMB, for HRP) is added to the wells to allow detection of the antigen. Figure 5.1 outlines the three variations on the classic ELISA; a direct ELISA (as above), an indirect ELISA (where a secondary antibody is used for detection) and a Sandwich ELISA (where one antibody is initially bound to a surface to capture the antigen before detection with a different primary antibody).

Another type of ELISA, a competitive ELISA is shown in Figure 5.1B. In a competitive ELISA, a plate is pre-coated with the antigen of interest. An unlabeled primary
Chapter 5 Development of an assay to monitor EPO-(EPOR)$_2$ binding interactions

Figure 5.1: A. Schematic of different ELISA set ups. **Direct ELISA:** An antigen is bound to a plate and detected using a primary antibody conjugated to a reporter enzyme (e.g. HRP) which turns over a substrate to give a detectable signal. **Indirect ELISA:** Similar to a direct ELISA with the exception that the reporter enzyme is conjugated to a secondary antibody which binds to the primary antibody. **Sandwich ELISA:** A capture antibody is bound to the plate and non-specific binding sites are blocked. A sample containing the antigen is then applied to the plate and binds to the capture antibody. A primary enzyme linked antibody binds specifically to the antigen and is detected using an appropriate substrate. **B. Competitive ELISA.** Protein-antigen complex is added to a pre-coated surface. Any unbound antibody in the sample interacts with the surface bound antigen. After washing, a secondary antibody conjugated to a reporter enzyme is used to detect the presence of the primary antibody.

antibody is incubated separately with a sample containing the same antigen and this complex is then added to the pre-coated plate. After incubation, unbound antibody is removed by washing. A secondary enzyme linked-antibody is added to bind the primary antibody and addition of substrate produces a signal. In this system, the more antigen in the samples, the less antibody available to bind to the antigen in the well, hence a lower signal. Competitive ELISAs are highly specific as they require two antibodies, and so are suitable for the analysis of complex mixtures. In some ELISA kits, a labeled antigen rather than a labelled antibody is used. In this
case, the labelled and unlabeled antigens compete for the antibody and so the more antigen in your sample, the weaker the signal. Competitive ELSIA can have direct, indirect or sandwich type protocols.

The disadvantage to most ELISA protocols is the high background signals, leading to low signal: noise ratios. These issues and methods of optimization of ELISA assays are discussed in this section and in section 7.

**Immunoprecipitation**

Immunoprecipitation (IP) is used for antigen purification and detection on a small scale. IP facilitates the identification of active proteins, can determine post-translational modifications on proteins, and can be used to study protein-protein and protein-nucleic acid interactions. In IP, target antigens are precipitated using an antibody which is pre-immobilized onto an insoluble support e.g agarose or sepharose. This support is incubated with samples containing the antigen (e.g. a cell lysate) before purification and elution of the target complex for analysis.

In Co-immunoprecipitation (Co-IP), two proteins are expressed and mixed in free solution to form protein-protein complexes (Figure 5.2). Antibodies against protein one are then added to react with the protein-protein complex. Modified sepharose is then added to capture the antibody-protein-protein complex. Samples are washed

![Figure 5.2: Co-immunoprecipitation method. Two interacting proteins, protein 1 and protein 2, are pre-incubated to form a protein-protein complex. Antibody linked sepharose is added to bind protein 1 and non-binding material is washed away. Protein-protein complex are released from the sepharose and separated on a SDS-PAGE gel. Protein 2 is detected via the corresponding antibody.](image-url)
and uncoupled from the sepharose. Eluted fractions are run on a SDS PAGE gel and detected via western blot using antibodies against protein two. Protein two can only be detected if associated with protein one. IP and Co-IP are useful tools as they allow interaction in free solution and can be used to detect low concentration of antigen in cell lysates. However, IP and Co-IP give only limited protein-protein binding information.

5.2 Detecting the EPO-(EPOR)\textsubscript{2} interaction

As EPO is commonly used as a performance enhancing drug in endurance sports, there are a number of kits available commercially for the detection of EPO in the blood. These kits usually consist of an assay plate pre-coated with an antibody that has been raised against EPO. Samples can be bound to this antibody and then detected with a secondary antibody raised against a different epitope on EPO. However, these kits cannot be utilised in this instance as the epitopes on EPO to which the antibodies are raised against are usually unknown.

Fortunately, the (EPOR)\textsubscript{2} receptor is commercially available and can be used to directly monitor the interaction of EPO with (EPOR)\textsubscript{2}. The (EPOR)\textsubscript{2} receptor was purchased from R& D systems (product number 307-ER-050/CF) and is supplied carrier free (no BSA is present) and has a >90% purity by SDS PAGE. The (EPOR)\textsubscript{2} is supplied purified and as a lyophilized powder with no additives. It is expressed from a mouse myeloma cell line, NS0-derived, meaning it will have mammalian like glycosylation motifs. All (EPOR)\textsubscript{2} used in these studies is from the same supplier, and reconstituted in PBS and stored as an aliquot at -80°C as per the manufacturers guidelines. Using this commercial (EPOR)\textsubscript{2}, an enzyme-linked immunosorbent assay (ELISA) based assay was developed to monitor the EPO-(EPOR)\textsubscript{2} interaction.

To detect the binding of EPO to the EPO receptor, (EPOR)\textsubscript{2}, a sandwich type ELISA was used and is shown in Figure 5.3. In the initial step, the commercially available (EPOR)\textsubscript{2} is bound to the plate, then any non-specific binding was blocked using an appropriate agent. Common blocking agents include BSA, milk powder,
5.2 Detecting the EPO-(EPOR)$_2$ interaction

**Figure 5.3:** Top: Schematic of EPO-(EPOR)$_2$ binding assay. A) Commercially available (EPOR)$_2$ is bound to a high binding 96 well plate. BSA is then used to block any unspecific binding to the well surface. B) Addition of EPO samples to the wells. C) The His-tagged EPO protein binds the Anti-His-HRP conjugate which is detected using the substrate 3,3',5,5'-tetramethylbenzidine (TMB) (Step D). **Below:** Mechanism of TMB oxidation by HRP.
casein and commercially available blocking solutions. After blocking, the EPO was incubated in the well and binds to the (EPOR)$_2$. Anti-histidine antibody raised against a hexa-histidine epitope conjugated to horse radish peroxidise (Anti-His HRP) was used to detect the histag present on the $N$-terminus of the EPO protein. A wash step, usually a solution of PBS + 0.05% Tween 20 is done after each assay stage to wash off any unbound material. This wash is especially important in the final stage before the addition of a substrate as any unbound Anti-His HRP will contribute to the signal. The substrate 3,3',5,5'-tetramethylbenzidine (TMB) was used for detection. TMB is a common substrate used in ELISA’s and colour development is the product of two successive one-electron transfer steps to give a diimine. The first oxidation of TMB results in formation of a radical cation and this species, in a complex with the TMB, results in a blue product with $\lambda_{\text{max}}$ at 652 nm. A second oxidation step results in the formation of the yellow diimine with $\lambda_{\text{max}}$ at 450 nm (Figure 5.3).

To test the validity of the assay, commercially available rhEPO (the natural, gly-

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.4.png}
\caption{Absorbance at 652 nm monitored over time for the binding of biotinylated rhEPO to (EPOR)$_2$ (-----) on high binding plates, the binding of biotinylated rHuEPO bound to (EPOR)$_2$ (.......) on untreated (not high binding) plates and binding of biotinylated rHuEPO bound to the plate without any (EPOR)$_2$ (......).}
\end{figure}
5.2 Detecting the EPO-(EPOR)_2 interaction

cosylated form of EPO used as a clinical (therapeutic) was used as a control. The
glycans on rhEPO has been biotinylated for detection with Streptavidin-HRP. The
assay for the biotinylated rhEPO was the same as described in Figure. 5.3, with the
exception that the biotinylated rhEPO was detected with Streptavidin conjugated
HRP instead of Anti-His HRP as the rhEPO has no His-tag. For the initial studies
high concentrations of both (EPOR)_2 (1 μg ml\(^{-1}\)) and rhEPO (10 μg ml\(^{-1}\)) were used.
However, the signal from the assay was very weak, leading to an unacceptably low
signal: noise ratio.

In an attempt to improve the signal, high binding 96 well plates were acquired
(Costar 9018). These plates are engineered for binding of medium sized proteins.
These plates resulted in a much improved signal over the original plates (Figure. 5.4).

Alongside the EPO-(EPOR)_2 assay, a series of control experiments were carried
out to check for any non-specific binding/ cross recognition between each of the
assay components. For instance, does the Anti-His HRP bind to any region on the
(EPOR)_2? Figure. 5.5 shows the set up of each of these experiments and they are
outlined in Table. 5.1.

Table 5.1: Elements of the experiments used to optimise the ELISA assay. (i), (ii), (iii) and
(iv) correspond to the figures in Figure. 5.5

<table>
<thead>
<tr>
<th>Elements</th>
<th>(EPOR)_2</th>
<th>BSA</th>
<th>EPO</th>
<th>Anti-His HRP</th>
<th>TMB</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i): Signal</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(ii): Background: No (EPOR)_2</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(iii): Background: No EPO</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(iv): Background: No HRP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>

As shown in Figure. 5.5(ii), no cross reactivity was observed between the Anti-His
HRP and the (EPOR)_2 (Table. 5.1(iii): Background: No EPO). In fact it was
found that the Anti-His HRP antibody did not bind at all to the well surface even
when applied directly to the plate (no blocking step, data not shown). While the
TMB is light sensitive, no oxidation was observed in the absence of the Anti-His
HRP (Table. 5.1(iv): Background: No HRP, data not shown).
Chapter 5: Development of an assay to monitor EPO-(EPOR)$_2$ binding interactions

Figure 5.5: A: Layout of the controls used to optimise the ELISA assay. (i): Binding of EPO to (EPOR)$_2$, (ii): Unspecific binding of EPO in absence of (EPOR)$_2$, (iii): Binding of Strep-HRP in the absence of EPO, and (iv): conversion of TMB in the absence of HRP. B: Absorbance at 652 nm monitored over time for the binding of biotinylated rhEPO to (EPOR)$_2$ (---), biotinylated rhEPO bound to the plate without any (EPOR)$_2$ (-----) and binding of Anti-His HRP in the absence of biotinylated rHuEPO (-----).

As there was no cross reactivity / recognition between the Anti-His HRP antibody or the TMB and the other assay components, (iii): **Background: No EPO** and (iv): **Background: No HRP** (Table.5.1) did not need to be carried out for
Optimal conditions for each of the assay components were determined using the WT EPO from E. coli as a model system. These optimisations can be split in to the different layers of the assay.

In the first layer, titrating the (EPOR)\textsubscript{2} concentration from 10\(\mu\)gml\textsuperscript{-1} to 0.01\(\mu\)gml\textsuperscript{-1} gives a working concentration of 0.1\(\mu\)gml\textsuperscript{-1}. Below this concentration, data becomes inconsistent. Any increase in the (EPOR)\textsubscript{2} concentration does not result in an increase in the signal. In the second layer, BSA (2% in PBS) was chosen as the blocking solution as it is the standard first choice in most ELISA protocols. Further optimisation of this blocking agent is discussed in section 5.5.

In the third layer, titration of WT EPO samples from 40 \(\mu\)M to 60 \(\mu\)M gives a working concentration of below 1 \(\mu\)M. Above this concentration, the signal becomes saturated as Figure 5.6 shows the response curve reaching a plateau in the signal intensity.
For the Anti-His HRP, the recommended working dilution of 1/2000 for use in immunoassays was used.

For the final step, TMB is added the well plate for detection. Addition of 50 μL of TMB (Pierce #34024) to each well results in a detectable signal. There are two ways to interpret the signal from an ELISA;

(i) End point assay – after a period of time, the colour development is quenched using acid. After acid addition, the assay is usually incubated for ~30 minutes, then a final, end point absorbance reading is taken. In this case the addition of the acid will produce the diimine which is detected at 450 nm.

(ii) Kinetic ELISA – the initial rate of reaction where the colorimetric development is linear, usually quoted as dOD/min, is calculated and plotted. In this case the initial colour development is the formation of the one electron transfer species with λmax at 652 nm. The main advantage to using the kinetic ELISA over an endpoint ELISA is that endpoint ELISA’s rely on a single data point while kinetic ELISA’s
5.4 Optimisation of surfactant concentration

As mentioned previously, the background signal for the EPO samples expressed in *E. coli* are more significant than the biotinylated rhEPO as the non-glycosylated EPO is prone to aggregation. In initial studies, the protein would more often than not aggregate resulting in a high background signal, sometimes higher than the signal from the EPO-(EPOR)$_2$ interaction. Upon investigation it became clear that the major factor contributing to the intensity of the background signal was the amount of n-lauryl sarcosine present in the buffers. N-lauryl sarcosine is anionic surfactant that is commonly used to stabilise proteins through the formation of micelles.

The concentration of surfactant above which micelle formation is favourable is defined as the critical micelle concentration (CMC). The literature reported CMC for n-lauryl sarcosine is \(~0.5\%\) w/v. Initially, \(0.5\%\) w/v n-lauryl sarcosine was used in all native buffers to stabilise the EPO protein. However, it later became apparent that \(0.5\%\) n-lauryl sarcosine was insufficient to stabilise EPO for use on the assay. Figure 5.7 shows that increasing the percentage of surfactant in the samples reduces the background signal. Increasing the n-lauryl sarcosine concentration from \(0.5\%\) to \(1\%\) resulted in a substantial reduction in the background intensity without interfering with the EPO-(EPOR)$_2$ signal. The use of \(5\%\) n-lauryl sarcosine gave the minimal background signal observed. However use of such a high surfactant concentration inhibited the protein – protein interaction resulting in a reduction in EPO-(EPOR)$_2$ signal intensity. Additionally, using \(5\%\) w/v surfactant has the disadvantage of being incompatible with many of the techniques employed in analysis of the protein (i.e. LCMS and column chromatography). As such, \(2\%\) n-lauryl sarcosine was chosen as the optimal surfactant concentration and is used in all native buffers.
Chapter Development of an assay to monitor EPO-(EPOR)$_2$ binding interactions

5.5 Optimisation of blocking agent

In an effort to further improve the signal: noise ratio, the blocking agent used to inhibit unspecific binding was optimised. As mentioned previously, common blocking agents include BSA and milk powder. In addition to these, there are also a number of commercially available blocking agents. To determine the optimal agent to use to block unspecific EPO binding, the blocking efficiencies of BSA (2% in PBS), milk powder (0.5% in PBS) and a commercially available blocking agent (SUPERBLOC, Pierce 37580) were compared. To determine the optimal blocking agent, no (EPOR)$_2$ was incubated on the ELISA; the blocking agents were added directly to the well before addition of EPO samples. Figure 5.8 shows the background signals detected when the ELISA plates were blocked with these three blocking agents. The commercially available SUPERBLOC blocking agent was not effective in blocking unspecific interactions.

**Figure 5.7:** The effect of n-lauryl sarcosine on the signal (EPO-(EPOR)$_2$ interaction, white bars) and the background noise (non-specific binding, black bars) from the binding assay. The percentage of n-lauryl sarcosine in the EPO samples (0.2μM EPO in 50 mM Tris-HCl, 200 mM NaCl, 1 M imidazole, and X% n-lauryl sarcosine) was increased from 0.5% to 5%. In samples with low concentrations of the surfactant (0.5% w/v), the background swamped any signal from the EPO-(EPOR)$_2$ interaction rendering it impossible to detect. Surfactant concentrations > 1% w/v dramatically reduced the noise, resulting in an acceptable signal: noise ratio.
Figure 5.8: Comparison of the efficiencies of the blocking agents SUPERBLOC (black bars), BSA (2% in PBS, white bars) and milk powder (0.5% in PBS, grey bars) at varying EPO concentrations.

EPO binding. The BSA and milk powder were much more effective, with the milk powder showing optimal blocking efficiency. However, as milk powder is known to exhibit lot-to-lot variability in its blocking efficiency, BSA was chosen as the blocking agent for this assay.

5.6 Summary

An ELISA based assay was developed for the analysis of the EPO-(EPOR)$_2$ interaction. The presence of >1% w/v n-lauryl sarcosine in all native buffers in preparation of the EPO samples is essential for protein stability and assay reproducibility. The optimised protocol was as follows:

Step 1: Commercially available (EPOR)$_2$ (100μL, 0.1μg ml$^{-1}$ in PBS) was incubated
Chapter 5
Development of an assay to monitor EPO-(EPOR)$_2$ binding interactions

on a high binding 96 well plate at 20°C for >5 hours.

Step 2: Wells were incubated with 250 µL Blocking buffer (2% BSA in PBS) for 1 hour to block any unspecific binding.

Step 3: Serial dilutions of EPO (60 nM-1 µM in Elution buffer: 50 mM Tris-HCl, 200 mM NaCl, 1 M imidazole, 2% w/v n-lauryl sarcosine, 50 µL) were incubated in the wells for 1 hour.

Step 4: A 1/2000 dilution of Anti-his HRP (in blocking buffer, 50 µL) was added to the wells and left for 30 minutes.

Step 5: TMB (50 µL) was added to each well and the absorbance at 652 nm monitored over 30 minutes. The response, $dOD/min$, was determined from the linear portion of the slope and plotted against the concentration.

Between each step, a PBS+0.05% Tween20 wash was carried out to remove unbound material.

Utilising this protocol, it was then possible to monitor a change in the binding of EPO to the (EPOR)$_2$. This assay is limited to detecting the binding of EPO to (EPOR)$_2$ only, and gives no indication of proliferative activity. The consequence of this, and an in-depth analysis of the advantaged and pitfalls of this assay are discussed in section 7. To determine the effect of glycans on EPO-(EPOR)$_2$ binding, mannosylated EPO variants need to be compared to the same non-reacted EPO variant on this assay.
6 The effect of site specific glycan addition to EPO on EPO-(EPOR)$_2$ binding

6.1 Introduction

With the development of the ELISA assay described in chapter 5 the binding of the mannosylated EPO and the WT EPO to the (EPOR)$_2$ were compared.

The protocol for the glycosylation of EPO variants yields a mixture of glycosylated and non-reacted EPO which could not be separated. As such, dose–response curves were generated by analysing these reaction mixtures directly on the ELISA. To determine the effect of glycan addition on EPO-(EPOR)$_2$ interactions, a series of comparisons were drawn.

Firstly, the dose-response curve for the WT EPO and the EPO variants were compared. This comparison determined the effect, if any, of introducing a cysteine residue into the EPO variants. Introducing cysteine residues to EPO may;

(i) Affect EPO-(EPOR)$_2$ interactions. Since the cysteine residues lie in key binding sites they may interrupt key interactions between the EPO and the (EPOR)$_2$ causing inhibition of binding.

(ii) Affect protein stability. The introduction of cysteine residues is known to affect protein stability and can result in aggregation of the protein and
The effect of site specific glycan addition to EPO on EPO-(EPOR)\textsubscript{2} binding loss of EPO binding affinity. Any variants exhibiting an effect described in (i) and (ii) will exhibit reduced (EPOR)\textsubscript{2} binding affinities compared to the WT EPO (Figure 6.1).

(iii) Have little or no consequence on the overall stability and binding of the EPO protein. These EPO variants will exhibit similar concentration – response curves to the WT protein.

Any EPO variants with similar binding affinities for the (EPOR)\textsubscript{2} as the WT protein were chosen as candidates for glycosylation; any inhibition of the EPO-(EPOR)\textsubscript{2} binding after reaction of these EPO variants with the \(\alpha\)-mannosyl iodoacetamide will be due solely to the presence of said glycan, and not some other factor (e.g. blocking of key interactions or loss of protein stability). These results are discussed in section 6.2.

Once the suitable EPO variants had been determined, the dose–response curves of

\textbf{Figure 6.1:} A: Comparing dose-response curves for an EPO variant and the mannosylated EPO variant. If the presence of a glycan inhibits receptor binding it will be reflected in the dose-response curve for the mannosylated EPO variant. B: EPO variants will be screened on the ELISA assay and suitable candidates chosen. Any variants showing inhibition of binding in this stage of the process will be discarded.
6.2 Screening EPO variants by ELISA

glycosylated and non-glycosylated variants needed to be compared. This comparison ensures that any inhibition of binding is due to the presence of the glycan and not due to loss of protein stability caused in the preparation of the samples. To achieve this, two simultaneous reactions were performed; one reaction was be performed with the addition of the α-mannosyl iodoacetamide reagent, the other was performed in its absence. It was expected that glycosylated EPO samples would exhibit dose-response curves indicative of a lower (EPOR)$_2$ binding affinity compared to the unreacted EPO samples (schematic shown in Figure. 6.1). These results are discussed in section 6.4.3.

In the final comparison, WT EPO was glycosylated in parallel with the EPO variant to serve as a control. Exposing the WT protein to the same conditions as the EPO variant will confirm that the glycosylation and the effects of the glycosylation are specific to the cysteine variants.

6.2 Screening EPO variants by ELISA

As the aim of this investigation is to determine the effect of modification of EPO on the EPO-(EPOR)$_2$ interaction, any changes in the binding as a result of the mutation alone needed to be eliminated. Each EPO variant was screened on the ELISA to determine EPO-(EPOR)$_2$ binding. EPO variants exhibiting dose-response curves similar to the WT EPO were used in further studies while any EPO variants showing reduced EPO-(EPOR)$_2$ binding compared to the WT protein were eliminated from further analysis (Figure. 6.1).

6.2.1 Generating a dose–response curve for the binding of WT EPO to (EPOR)$_2$

A dose–response curve of the WT protein was generated for use as a standard. To generate this standard curve, the WT EPO gene was expressed, the protein purified by metal affinity chromatography, followed by an oxidative refolding step
before serial dilution for the ELISA (section 8.5). The concentration-dependent curve (Figure 6.5) for the WT EPO shows saturation of the signal above 1μM EPO.

### 6.2.2 Screening EPO variants

All 12 EPO variants were screened on the ELISA and selected by their ability to bind to (EPOR)$_2$. A criterion for the selection process is to exhibit a dose–response curve similar to the WT EPO.

The process outlined in Figure 6.2 was used to determine the suitability of each variant. Initially, each variant was tested on the ELISA only once and any EPO

![Flow diagram](image.png)

**Figure 6.2:** Flow diagram outlining the selection process applied to all EPO cysteine variants. The stage of elimination for each EPO variant is shown in red.
6.2 Screening EPO variants by ELISA

Figure 6.3: Dose-response curves of EPO variants discarded from further analysis after initial screening on the ELISA. All variants are plotted against the WT EPO-(EPOR)_2 binding (error = S.E.M., n=3). The WT EPO background signal (■), EPO variant-(EPOR)_2 binding (○) and EPO variant background ( ● ) are all shown. **Panel A**: T44C EPO, **B**: K52C EPO, **C**: R103C EPO, and **D**: V46C+R103C EPO (for all variants error = S.D.).

variant showing an inhibited affinity to (EPOR)_2 was discarded. The results from this initial screening are shown in Figure 6.3 and Figure 6.5.

Figure 6.3 gives the dose-response curves for the EPO variants T44C, K52C, R103C, and V46C+R103C EPO. As these curves show, there is an apparent reduction in binding affinity and so these variants were deemed unsuitable. The rationale for eliminating the EPO variants exhibiting lower binding affinities is twofold. The loss of binding affinity can be due to the introduction of a cysteine residue at a key binding site. If this is the case then the variants will need to be eliminated for the reasons given above. Alternatively, the reduction in binding affinity can be explained
The effect of site specific glycan addition to EPO on EPO-(EPOR)$_2$ binding

Figure 6.4: Dose-response curve of S104C EPO. EPO variant-(EPOR)$_2$ binding (---○--) and EPO variant background (••••). (n=3 (technical repeats), error = S.D.)

by a loss of stability in the protein caused by the mutation itself. It is known that introducing cysteine residues can cause loss of stability, and since there is no way to determine which of these factors contributed to the reduction in binding without further investigation, these EPO variants were discarded.

The variants S104C, N147C, and R103C+R150C EPO generated signals with a very high background signals. Figure 6.4 shows the binding curve for the S104C EPO variant, where the background signal is in fact more intense than the binding curve. From the optimisation of the ELISA assay it is known that the use of n-lauryl sarcosine reduces the background by preventing aggregation (section 5.4). It was thought that the cause of the high background signals observed for these EPO variants was due to aggregation of the protein. These variants were therefore discarded on the basis of protein instability.

Figure 6.5 gives the dose-response curves for the variants K45C, K97C, R150C, K54C+K97C and K45C+K52C EPO which were all deemed suitable candidates by our criteria. To validate the consistency of these results, the ELISA’s for these 5 variants were repeated.

Results from this second round of screening revealed R150C and K45C+K52C EPO to give highly inconsistent data between experiments and so were discarded (data not shown). The concentration – response curves for the K45C, K97C and K45C+K97C EPO variants were consistent and reproducible and so are shown in Figure 6.6.
6.2 Screening EPO variants by ELISA

From these 3 variants, only the K45C+K97C EPO variant was chosen as a model system to glycosylate and analyse. This decision was made as the K45C+K97C EPO has two cysteine residues, rather than one, and the probability of observing an effect on the EPO-(EPOR)$_2$ interaction is therefore greater. As such, from the original 12 EPO variants, only 1 was glycosylated and analysed on the ELISA.
Chapter 6 The effect of site specific glycan addition to EPO on EPO-(EPOR)\(_2\) binding

Figure 6.5: Dose-response curves of EPO variants exhibiting similar binding to the (EPOR)\(_2\) as the WT EPO on the ELISA. All variants are plotted against the WT EPO-(EPOR)\(_2\) binding (---). The WT EPO background signal (●●●), EPO variant-(EPOR)\(_2\) binding (○○○) and EPO variant background (◆◆◆) are all shown. Panel A: WT EPO (error = S.E.M., n=3 (biological repeats)), B: K45C EPO, C: R150C EPO, D: K97C EPO, E: K45C+K52C EPO and F: K45C+K97C EPO (for all variants n=3 (technical repeats), error = S.D.)
6.2 Screening EPO variants by ELISA

Figure 6.6: Dose-response curves of the final three EPO variants. All variants are plotted against the WT EPO-(EPOR)$_2$ binding (−○−). The WT EPO background signal (—●—), EPO variant-(EPOR)$_2$ binding (−○−) and EPO variant background (—●—) are all shown. 

**Panel A:** K45C EPO, **B:** K97C EPO, and **C:** K45C+K97C EPO. Error = S.E.M, n=3
6.3 Reaction of WT EPO with α-mannosyl iodoacetamine and analysis by ELISA

Initial LCMS analysis demonstrated that it is possible to site specifically alkylate the EPO cysteine variants with the α-mannosyl iodoacetamide. Reaction with the WT protein under the same conditions results in no glycan addition, and as such no inhibition of the EPO-(EPOR)$_2$ binding should be observed. To confirm this, dose–response curves for the WT EPO and the WT EPO treated with the α-mannosyl iodoacetamide were compared.

The protocol for the specific di-glycosylation of the K45C+K97C EPO double variant required reaction of 1 mM α-mannosyl iodoacetamide with 1mgml$^{-1}$ protein (in Denaturing Elution buffer section 8.1.2) for 30 minutes with the exclusion of light. To generate the WT EPO standard curve, two samples of WT EPO were prepared simultaneously under these conditions, with the exception that in one sample the α-mannosyl iodoacetamide was omitted. After the reaction, 10 mM DTT was used to quench the reaction and the samples were prepared for assay analysis.

As expected, characterisation of both reactions by LCMS before refolding showed no addition of the α-mannosyl iodoacetamide to WT EPO. The dose-response curve

![Figure 6.7: Dose-response curve for the WT EPO (-○-), WT EPO background signal (-●-), WT EPO treated with α-mannosyl iodoacetamide (-◇-) and the corresponding background signal (-◆-). Error = S.E.M, n=3.](image-url)
for the WT protein and the WT protein treated with the glycosylating agent are shown in Figure 6.7.

**Statistical analysis**

Two distinct statistical analyses were carried out to compare the two sets of data. In one statistical analysis, a paired T-test comparing two points was carried out to determine the significance of any difference between the two data sets; a paired T-test gave \( P > 0.05 \) for all concentration values, hence any differences are not significant.

Another statistical test was carried out to compare the two curves directly. This test was carried out using Graph Pad Prism and the \( \log(\text{agonist}) \) vs variable slope (four parameter) equation was used to analyse the data. Within this analysis, the programme determines if there are any differences between the two data sets by attempting to fit one curve to both sets of data (the null hypothesis). If there is a significant difference \( (P < 0.05) \), then the null hypothesis is rejected. This equation uses the top and bottom plateau, the \( EC_{50} \) and the Hill slope as criteria for fitting the data curve.

For the data in Figure 6.7, a comparison of the two data sets gave a \( P \) value of 0.6137, hence the null hypothesis was accepted – one curve can be fitted to all data sets. These analyses show that treating WT EPO with the \( \alpha \)-mannosyl iodoacetamide has no effect on its affinity for the EPO receptor, (EPOR)\(_2\).

### 6.4 Analysis of mannosylated K45C+K97C EPO variant on the ELISA

In the initial screening of the EPO variants (section 6.2.2), the double variant K45C+K97C EPO was chosen as a prime candidate for the proof of concept studies. As mentioned in section 6.2.2, this variant was chosen because it has a similar
Chapter 6 The effect of site specific glycan addition to EPO on EPO-(EPOR)\(_2\) binding

Figure 6.8: Crystal structure of the WT EPO bound to the homodimer EPO receptor. Binding Site 1 and Site 2 are highlighted to demonstrate the location of the two cysteine residues in the K45C+K97C EPO variant.

affinity for the (EPOR)\(_2\) as the WT EPO, possesses a relatively high stability and has the added advantage of containing two cysteine residues which are available for modification.

Each cysteine residue is located in a different binding site, the K45C mutation lies in Site 1 and the K97C mutation lies in Site 2 (Figure. 6.8).

To determine how the presence of a glycan on EPO will affect the EPO-(EPOR)\(_2\) interaction, the K45C+K97C EPO variant was exposed to the same reaction conditions as the WT protein and analysed on the ELISA.

6.4.1 Initial analysis of mannosylated EPO by ELISA

The K45C+K97C EPO variant was chosen as a model protein to study the effects of introducing chemical modification sites for glycosylation into the EPO sequence. To determine the effect of glycosylation on EPO-(EPOR)\(_2\) interaction, 4 reactions were set up in parallel:
Table 6.1: Components of reactions run in parallel to determine the effect of glycosylation on EPO-(EPOR)\textsubscript{2} interactions.

<table>
<thead>
<tr>
<th></th>
<th>EPO sample</th>
<th>α-mannosyl iodoacetamide (1 mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT EPO control</td>
<td>WT EPO (1mgml\textsuperscript{-1})</td>
<td>✗</td>
</tr>
<tr>
<td>WT EPO reaction</td>
<td>WT EPO (1mgml\textsuperscript{-1})</td>
<td>✓</td>
</tr>
<tr>
<td>K45C+K97C EPO control</td>
<td>K45C+K97C (1mgml\textsuperscript{-1})</td>
<td>✗</td>
</tr>
<tr>
<td>K45C+K97C EPO reaction</td>
<td>K45C+K97C (1mgml\textsuperscript{-1})</td>
<td>✓</td>
</tr>
</tbody>
</table>

The protein samples were prepared as described in section 6.4 and the reactions were carried out as described in section 6.3 and Figure 6.10;

Samples were left to react at room temperature for 30 minutes with the exclusion of light. The reactions were quenched by addition and incubation with 10 mM DTT, followed by oxidative refolding into native buffer and purification using Ni-NTA resin. The protein samples were eluted in native conditions and concentrated using a size exclusion centrifugal column to determine an accurate concentration measurement. After dilution to the desired concentration, all four samples are analysed by ELISA.

Comparing the WT EPO and the WT EPO + α-mannosyl iodoacetamide reactions acts as a control. Comparing the K45C+K97C EPO variant with the K45C+K97C EPO + α-mannosyl iodoacetamide reaction shows how the presence of a glycan on the EPO binding site affects the EPO-(EPOR)\textsubscript{2} interaction.

When generating the WT EPO control curves the WT protein gave reproducible results and Figure 6.7 shows that incubation of the WT protein with the α-mannosyl iodoacetamide had no effect on receptor binding. However, as Figure 6.9 shows, using this protocol the K45C+K97C EPO variant produced inconsistent data with high background signals.

In an attempt to reduce these high background signals, the protocol for the preparation of the EPO samples was optimised to prevent aggregation of the protein (the most likely source of the high background signals). A new protocol was devised with the aim of reducing handling of the protein. Figure 6.10A outlines the protocol used.
Figure 6.9: Dose-response curves of K45C+K97C EPO on ELISA before and after reaction with α-mannosyl iodoacetamide. Three EPO variants. Shown are the K45C+K97C EPO-(EPOR)$_2$ binding (---). The K45C+K97C EPO background signal (●), K45C+K97C EPO reacted with α-mannosyl iodoacetamide-(EPOR)$_2$ binding (−−−), K45C+K97C EPO reacted with α-mannosyl iodoacetamide background (●●●). Error = S.D., n=3

to prepare samples for analysis on the ELISA. The final centrifugal concentration was identified as a key step for elimination; it was thought that this step contributes significantly to the destabilisation of the protein as the samples are in native buffer. In an optimised protocol this step, among others, was eliminated in an effort to improve assay reproducibility.

### 6.4.2 Procedure optimisation

In an attempt to improve the reproducibility of the assay, the reaction and ELISA analysis protocols were optimised. The main objective was to bypass the second reduction and refolding step, and the second centrifugal concentration step.

In the optimised protocol, upon reaction of the EPO with the α-mannosyl iodoacetamide, samples were directly diluted in to the native buffer to a final concentration of 1 μM, before serial dilution to the desired concentrations and analysis on the ELISA. Figure 6.10B outlines the steps of the optimised protocol.

In this optimised protocol, the α-mannose iodoacetamide was not quenched, simply diluted. It is reasonable to assume that under these diluted conditions (50X dilution)
the reaction is unlikely to continue. Even if the glycosylation reaction continues at a reduced rate, the effect of not removing the \(\alpha\)-mannosyl iodoacetamide would be seen in the WT EPO control.

Using the optimised protocol, a new concentration – response curve of the WT EPO was generated (Figure 6.11A). The response \((d\text{OD/min})\) observed in the new WT EPO curve increased compared to samples analysed under the previous protocol.

**Figure 6.10:** A: Initial protocol for the analysis of EPO samples on the ELISA after reaction with \(\alpha\)-mannosyl iodoacetamide. EPO samples were glycosylated and addition of DTT (10 mM, 1 hour) quenched the reaction and reduced the protein. The protein was then diluted into native buffer to facilitate refold. Ni-NTA was used to extract the protein from the other reagents and the protein is eluted in native buffer. The eluant was concentrated for accurate determination of the protein concentration and then diluted to the desired concentration for analysis on the ELISA. B: Optimised protocol for the analysis of EPO samples on the ELISA. After reaction with \(\alpha\)-mannosyl iodoacetamide, EPO samples were glycosylated using the optimised conditions for the desired reaction. The protein was then directly diluted into native buffer to a final concentration of 1 \(\mu\text{M}\) for analysis on the ELISA.
Chapter 6 The effect of site specific glycan addition to EPO on EPO-(EPOR)$_2$ binding (Figure 6.11B). This increase in response was seen consistently and attributed to the optimised procedure.

With a new WT EPO control curve generated, the original hypothesis could be addressed: Can the site specific glycosylation of EPO be used to inhibit the EPO-(EPOR)$_2$ binding interactions?

6.4.3 Final analysis of mannosylated EPO by ELISA

The K45C+K97C EPO was glycosylated and analysed on the assay using the optimised protocol as described in Figure 6.10. From the initial expression to the final analysis by ELISA, Figure 6.12 outlines the procedure for preparing the glycosylated K45C+K97C EPO.

The protein is expressed, purified and refolded. The protein is then denatured for reaction with $\alpha$-mannosyl iodoacetamide. Before the reaction, the sample is split; half of the protein undergoes reaction with $\alpha$-mannosyl iodoacetamide, while the other half is used as a control. Both reactions are incubated for 30 minutes.

Figure 6.11: A Dose-response curve for the WT EPO using the new protocol (Signal—○—, background—■—. Error = S.E.M, n=2). B Comparison of the dose-response curves generated for the WT EPO using the new protocol (Signal—○—, background—■—. Error = S.E.M, n=2), and the old protocol (Signal—○—, background—■—. Error = S.E.M, n=3). The new protocol was optimised to maintain the stability of the protein sample for the duration of the analysis. This was reflected in the results as an apparent increase in the response (dOD/min).
6.4 Analysis of mannosylated K45C+K97C EPO variant on the ELISA

**Figure 6.12:** Procedure for the expression, purification, reaction and analysis of EPO variants.

(with or without α-mannosyl iodoacetamide) before quenching with DTT. These two samples are then directly diluted to the desired concentrations in buffer for analysis by ELISA. This procedure was carried out for the K45C+K97C EPO and the WT EPO, giving a total of four experiments.

**Results for WT EPO**

The data for these four experiments are shown in Figure 6.13. Figure 6.13A shows the dose-response curves for the two WT EPO samples, one reacted with α-mannosyl iodoacetamide, the other without. After the reaction, the sample incubated with α-mannosyl iodoacetamide was analysed by MS. This MS signal is shown in Figure 6.13B. As expected, after incubation with the glycosylating reagent, the MS shows no addition of mannose.
Chapter 6: The effect of site specific glycan addition to EPO on EPO-(EPOR)_2 binding

As there was no mannose addition, the expected result is that the two dose-response curves for the WT EPO samples will be identical, as the two EPO samples are identical. To determine if the dose-response curves for the two WT EPO samples shown in Figure 6.13A differ from each other, two distinct statistical analyses were carried out (see section 6.3). In one statistical analysis, paired T-test comparing points at the same EPO concentrations were carried out to determine the significance of any difference between the two data sets. A paired T-test gave P > 0.05 for all concentration values, indicating no differences between each set of data points.

Comparison of the two curves by attempting to fit one curve to both sets of data (see section 6.3 for a discussion of the analysis method), gave a P = 0.3627, meaning that one curve can be fitted to all data sets. Hence by both statistical methods, treating WT EPO with the α-mannosyl iodoacetamide has no effect on its affinity for the EPO receptor, (EPOR)_2.

Results for K45C+K97C EPO

In Figure 6.13D, the MS of the K45C+K97C EPO after reaction with α-mannosyl iodoacetamide shows 3 mass peaks. These peaks are the starting material, the mono-glycosylated and the di-glycosylated species. In the same manner as the WT EPO, the two dose-response curves, one from the K45C+K97C EPO and the other from this mixture of the three EPO species, were compared to determine any differences in affinity for the (EPOR)_2 after reaction with α-mannosyl iodoacetamide.

These two dose-response curves are shown in Figure 6.13C, and there is an apparent inhibition of binding of the K35C+K97C EPO to (EPOR)_2 after reaction with α-mannosyl iodoacetamide. To confirm this, the two statistical analysis methods were preformed to compare points within the curves and the two curves themselves.

Statistical analysis of the two K45C+K97C EPO variants showed a significant difference for responses (dOD/min) measured at 0.3 µM (p < 0.05, statistically significant), 0.2 µM (p < 0.01, statistically significant), 0.15 µM (p < 0.05, statistically significant) and 0.1 µM (p < 0.001, statistically highly significant) EPO. This indicated
that the reaction of the K45C+K97C EPO variant with α-mannosyl iodoacetamide results in the inhibition of the EPO-(EPOR)_2 binding interactions.

Comparison of the two curves gave P < 0.0001. This means that the null hypothesis
The effect of site specific glycan addition to EPO on EPO-(EPOR)$_2$ binding is rejected (that one curve can be used to fit both sets of data), hence there is a statistically significant difference between the two data sets.

These data sets indicate that the addition of a glycan moiety to key binding sites on the EPO protein result in the inhibition of binding to (EPOR)$_2$. As this difference is seen only in the K45C+K97C EPO and non of the other controls, it can be stated with reasonable certainty that the loss of binding of the K45C+K97C EPO to the (EPOR)$_2$ is a direct result of the addition of mannose to the protein.

However, these data only serve to determine binding of the EPO variants to (EPOR)$_2$. The proliferative activity of the EPO variants has yet to be tested, and is the logical next step in this proof-of-concept study. As outlined in section 1.6, proliferative activity of EPO samples is almost always carried out using the UT7-EPO assay. In this assay, survival of UT7 cells are dependent upon the presence of EPO. However, as these assays are carried out in complex systems over a number of days/weeks, it is unlikely that these less stable, non-glycosylated EPO variants would be suitable for this assay. As such, before proliferative assays can be preformed, it would be prudent to express more stable EPO variants, possibly from $P$. $Pastoris$. Options for testing the activity of the EPO variants are discussed in chapter 7.

In addition, whilst this data supports the proof-of-concept, the tissue protective function of these EPO variants needs to be tested. While time restraints prevented the investigation into the tissue protective activity of these EPO variant, chapter 7 outlines a plan of action for continuation of the work.

## 6.5 Summary

In the process of investigating if the addition of glycans to EPO can be used to inhibit EPO-(EPOR)$_2$ interactions, WT EPO and EPO cysteine variants have been expressed and purified. The cysteine variants were chemically modified and the binding of these glycoproteins to the (EPOR)$_2$ has been examined on an ELISA.

The reproducibility of the data from the ELISA’s preformed on the glycosylated EPO samples suffered due to the inherent instability of the EPO samples. As
mentioned previously, the glycosylation of rHuEPO is essential for in vivo stability. Without these glycans, the protein is unstable and prone to aggregation, limiting the degree to which the samples can be manipulated. Building the reaction and ELISA protocols around these restraints and minimising the handling of the protein to bare minimum was essential to generate reproducible data.

Comparison of the WT EPO before and after reaction with the α-mannosyl iodoacetamide on the ELISA showed no inhibition of the EPO-(EPOR)$_2$ interactions while comparison of the K45C+K97C EPO before and after reaction showed a statistically significant difference in the EPO-(EPOR)$_2$ binding. These results suggested that the EPO-(EPOR)$_2$ interaction can be disrupted by the site specific glycan addition to EPO.
7 Conclusions and outlook

Analysis of methods and results

In order to develop a non-erythropoietic EPO derivative, cysteine residues were introduced into the EPO sequence to act as chemical modification sites. These EPO cysteine variants were successfully expressed in *E. coli*, the proteins purified and a method for the refolding of EPO was established. EPO variants were then glycosylated at the cysteine residues and analysed by ELISA.

Initially, 11 amino acid residues were targeted for mutation based upon the crystal structure of the EPO-(EPOR)$_2$ complex and previously reported data$^{125-127,135}$. These residues were located within the two binding sites of EPO, and had multiple and/or close contacts with the (EPOR)$_2$. Of these single mutations, 9 of the genes were successfully cloned into pET16b for expression in *E. coli*. The remaining mutation positions (K154C and V46C) were unsuccessful, presumably due to unsuitable PCR or ligation conditions. In addition to the 9 single mutations, 4 EPO cysteine variants with double mutations (two cysteine residues) were successfully cloned, giving a total of 13 EPO cysteine variants in addition to the WT EPO. These were deemed a suitable number of candidates for this study and so no further mutations were attempted. Table 2.3 in chapter 2 lists these mutations. Of these 14 (including the WT) EPO variants 7 were generated in the lab from the same gene. The other 6 were purchased from GeneArt to save time, and the sequences of these genes were codon optimised for expression in *E. coli*. This was done in an attempt to improve the yield of protein from each expression; however this had little effect on the overall expression yields.
Conclusions and outlook

Upon the successful cloning of these EPO variants, the WT EPO was used to optimise the expression and purification conditions. Expression of WT EPO was optimised to growth at 30°C, with cells harvested 5 hours after induction with IPTG. Expression at higher temperature yielded a mixture of full length EPO and truncated EPO as shown by western blot (see Figure 2.5). These expression conditions were successfully applied to the 5 of the 6 EPO cysteine variants generated in house (the K140C EPO variant was not overexpressed). Expression of the purchased genes under these conditions yielded only truncated EPO products, even at low temperatures. The successful expression of these variants was done in the presence of 1% glucose (suppresses basal expression). From these data, it is recommended that any EPO variants generated in future studies not be codon optimised for expression in *E. coli*.

When EPO is expressed from *E. coli* it is present in inclusion bodies and must be purified in denaturing conditions. A protocol for this purification was built around the WT EPO and successfully applied to all expressed EPO cysteine variants.

Upon purification, EPO cysteine variants were glycosylated with α-mannosyl iodoacetamide in denaturing conditions and reactions were monitored using LCMS. Reaction of single EPO variants yielded a mixture of unmodified and mono-glycosylated EPO species. Reaction of the double cysteine EPO variants with the α-mannosyl iodoacetamide yielded a mixture of three species; the unmodified EPO, the mono-glycosylated EPO and the di-glycosylated EPO. Attempts to purify the glycosylated species from the unmodified EPO species were unsuccessful.

The data from these reactions suggest that a percentage of the cysteine residues in the EPO samples are unavailable / inaccessible for reaction. One possibility is that a proportion of the EPO variants are misfolded, yet soluble. In this case, the introduced cysteine residue could be taken up in a disulfide bond, leaving one of the original cysteine residues ‘free’. In this misfolded conformation, the ‘free’ cysteine may be located in an inaccessible position (e.g. buried). Another possibility is that a percentage of the cysteine residues are involved in intermolecular disulfide bonds between EPO molecules. To determine if this is the case, dynamic light scattering may be used to probe the particle size within the EPO samples.
Indepant of the cause of the low reaction yields (approx. 50%), the most effective method to improve the efficiency of the glycosylation reaction would be introduce natural glycosylation sites in place of the cysteine residues and proceed to an expression system which will glycosylate these new variants. While in no way simple, if any of the variants are successfully expressed, they will have full length glycan chains at the desired position. Options for expression systems include yeast, insect and mammalian cells. The possibilities of cell systems in discussed later.

To determine the effect of glycosylation on the EPO-(EPOR)$_2$ interactions, an ELISA based assay was developed using the WT EPO as a standard. In this ELISA the (EPOR)$_2$ protein was bound to the surface of a 96 well plate, the plate was then blocked with BSA (2% in PBS) before addition of the EPO sample for incubation and binding. After washing away unbound material, an Anti-His HRP antibody was added to bind to the His-tag present on the $N$-terminus of all EPO variants before addition of the TMB substrate.

During the development of this assay the signal: noise ratio was optimised by minimising the background signal. While successful, an oversight in the development of the assay was that the orientation of the (EPOR)$_2$ was never optimised. In the protocol described above, the (EPOR)$_2$ is bound to the surface of the well by adsorption. In this case, the orientation of the receptor is unknown and likely diverse (see Figure 7.1). If the (EPOR)$_2$ were orientated ‘upwards’, this could result in an increase in the signal intensity, and an improvement in the signal:noise ratio.

Figure 7.1 outlines two possible methods to orientate the (EPOR)$_2$ ‘upward’. One method is to covalently link the (EPOR)$_2$ at a known site (possibly a cysteine) to

![Current random orientation of (EPOR)$_2$ on plate surface](image1)

![Orientation of (EPOR)$_2$ on plate surface via a covalent attachment](image2)

![Orientation of (EPOR)$_2$ on plate surface via an antibody](image3)

**Figure 7.1:** Methods to orientate (EPOR)$_2$ ‘upwards’ on plate surface.
the surface of the plate (there are functionalised plates available specifically for these purposes (e.g. maleimide activated plates). Another possibility is to coat the plates with an antibody which has been raised against a specific epitope on the (EPOR)$\textsubscript{2}$ to capture the (EPOR)$\textsubscript{2}$ in a certain orientation. Before any further assays are carried out using this protocol, it is recommended that the (EPOR)$\textsubscript{2}$ orientation is optimised, and the resultant change in signal (if any) analysed.

Initially, all EPO variants were screened by ELISA for WT-like binding to the (EPOR)$\textsubscript{2}$. Any variants showing an apparent reduction in binding affinity before the glycosylation reaction were discarded. Of the 13 variants screened, the K45C, K97C and K45C+K97C EPO variants were the only suitable candidates. The other variants either exhibited inhibited binding to (EPOR)$\textsubscript{2}$ due to the presence of a cysteine residue within the binding site, or were not stable enough for analysis. In order to analyse more EPO variants, it would be prudent to express them in a system which will glycosylate the protein in order to improve the protein stability. Due to time restrictions, the K45C+K97C EPO only was chosen as the final candidate for further analysis.

**Conclusion and future prospects**

The K45C+K97C EPO and the WT EPO were simultaneously glycosylated and the products analysed on the ELISA. Comparison of the WT EPO before and after reaction with the $\alpha$-mannosyl iodoacetamide showed no statistically significant inhibition of the EPO-(EPOR)$\textsubscript{2}$ interactions.

Comparison of the K45C+K97C EPO before and after reaction showed a statistically significant difference in the EPO-(EPOR)$\textsubscript{2}$ binding. This outcome indicates that it is possible to inhibit the EPO-(EPOR)$\textsubscript{2}$ binding by introducing non-natural glycosylation sites in to the EPO sequence.

This project has laid the foundations for the development of a tissue protective but non-erythropoietic therapeutic agent. However, the lack of glycosylation present on the EPO samples as a result of expression from *E. coli* yields unstable proteins.
which are prone to aggregation. This intrinsic instability imposed limitations on the extent of modification and analysis that could be carried out.

To develop the initial concept further, expression in alternative systems is the logical next step. Expression in systems that will add glycan chains to the EPO variants will increase the stability of the protein thereby facilitating more accurate analysis. In addition, glycosylated EPO can be used in UT7-EPO cell assays to determine the activity of the variant.

Expression systems considered include mammalian and yeast systems. While mammalian cell cultures are costly and time sensitive, expression of EPO in *Pichia pastoris* has been reported and the glycosylation present acts to stabilise the protein\(^{156}\). As such, expression of these EPO variants in *Pichia pastoris* followed by modification is a distinct possibility for the continuation of this work. To that end, the WT EPO gene has been cloned into the pPICZ\(\alpha\)A vector for expression in *Pichia pastoris* (section 8.2.3).

Another option is the use of cell free expression systems to generate the EPO variants\(^{180-182}\). In a cell free system, genes can be added directly to *E. coli* cell lysate for protein expression. These cell free systems bypass the need for lysis thereby reducing handling steps and minimising purification times.

If a cell free system were considered, the use of *Leishmania tarentolae* lysate is recommended for these studies. Leishmania cell lysates have been shown to express proteins with mammalian like glycan chains\(^{183-186}\). Hence, this cell free system has the advantages of mammalian cell cultures (glycan chains to add stability) with none of the drawbacks (long culture times, high costs, etc.). Recommendation for further work on this project would be to introduce natural glycosylation sites in place of the cysteine residues for all EPO variants described here and expression of these variants in a *Leishmania* cell free system. These glycosylated variants could then be screened on the ELISA developed in this work and analysed by UT7-EPO cell assay for proliferative activity. If any variants are identified as non-erythropoietic EPO derivatives, a suitable assay for the detection of tissue protective activity can be developed.
8 Experimental detail

8.1 General experimental detail

8.1.1 Chemicals and Reagents

Unless stated otherwise, all chemicals were of analytical grade and used as received from the supplier. LCMS running solvents of an appropriate grade were obtained from Romil or Sigma-Aldrich.

Competent cells XL1-Blue, TOP10, BL21(DE3) and Origami 2(DE3) were obtained from Invitrogen. The pET16b and P. pastoris constructs were obtained from in-house sources. All restriction endonucleases were obtained from New England Biolabs and used according to the manufacturer’s instructions. Ampicillin stock solutions (100 mg ml\(^{-1}\)) and IPTG stock solutions (1 M) were stored at -20°C.

All (EPOR)\(_2\) binding assays were monitored in a Tecan infinite M200 plate reader.

The α-mannosyl iodoacetamide, α-mannosyl chloroacetamide and GFP variants were provided courtesy of Andrew Martin at the University of Manchester. The biotinylated rHuEPO was provided courtesy of Professor Paul Brenchley (Renal research labs, Manchester Royal Infirmary).

8.1.2 List of buffers

Lysis buffer: 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 5 mM imidazole, 1 mM PMSF
Purification Binding buffer: 6 M guanidine hydrochloride, 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 5 mM imidazole

Purification Elution buffer: 6 M guanidine hydrochloride, 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 1 M imidazole

Refolding buffer: 50 mM Tris-HCl (pH 7.8), 200 mM NaCl, 2% w/v N-lauryl sarcosine, 10 μM CuSO₄

Native Wash Buffer: 50 mM Tris-HCl (pH 7.8), 200 mM NaCl, 2% w/v N-lauryl sarcosine, 10 mM imidazole

Native Elution buffer: 50 mM Tris-HCl (pH 7.8), 200 mM NaCl, 2% w/v n-lauryl sarcosine, 1 M imidazole

Denaturing Wash buffer: 6 M guanidine hydrochloride, 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 5 mM imidazole

Denaturing Elution buffer: 6 M guanidine hydrochloride, 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 1 M imidazole

PMSF stock solution: 100 mM in isopropanol

8.1.3 Oligonucleotide Primers

Primers for cDNA amplification

Primer I FW: 5’-AAA AAA ACA TAT GGC CCC ACC ACG CCT CAT CTG TGA CAG CCG AGT.

Primer II: 5’-TTT TTT TTT TTT AGA TCT GAC GGG GTT CAG GAG TGG CGG GGG AGG.
8.1 General experimental detail

**Primers used site directed mutagenesis**

K45C: FW: 5’-CACTGTCCAGACACCCtgtGTGTTAATTCTATGCC,
RV: 5’-GGCATAGGAAAAAACacaGGTGCTCTGGACAGTG.

K52C: FW: 5’-GTGTTAATTTCTATGCCTGGtgtAGGATGGAGGTCGGGCAGC,
RV: 5’-GCTGCCCGACCTCCATCCTacaCCAGGCATAGAAATTAAC.

K97C: FW: 5’-CCCTGAgAGCTGACCATGTGAGATgtGCGTCAGTGGCCTTCGC,
RV: 5’-GCAGAAGGACCACCAGGCacaATCCACATGCAGCTGCAGGG.

K140C: FW: 5’-GCTGACACTTTCCGCtgtCTCTTTCCAGCTACTCC,
RV: 5’-GGAGTAGACTCGGAAGAGacaGCGGAAAGTGTCAGC.

K154C: FW: 5’-CCTCGGGGAAAGCTGTgtCTGTACACAGGGGAGGC,
RV: 5’-GCCTCCCCTGTGTACAGacaCAGCCTTTCCCGGAGG.

**8.1.4 Gene Sequences**

All numbering of the mutation positions (e.g. K45C) begins at the first A of the sequence, e.g. APPR......
<table>
<thead>
<tr>
<th>Mutation position</th>
<th>Original codon</th>
<th>New codon</th>
</tr>
</thead>
<tbody>
<tr>
<td>K45</td>
<td>AAA</td>
<td>TGT</td>
</tr>
<tr>
<td>K52</td>
<td>AAA</td>
<td>TGT</td>
</tr>
<tr>
<td>K97</td>
<td>AAA</td>
<td>TGT</td>
</tr>
<tr>
<td>K140</td>
<td>AAA</td>
<td>TGT</td>
</tr>
<tr>
<td>K45</td>
<td>AAA</td>
<td>TGT</td>
</tr>
</tbody>
</table>

Sequence for the WT EPO gene in pET16b and a list of the cysteine mutations.
Sequence for the optimised WT EPO gene in pET16b and a list of the cysteine mutations

<table>
<thead>
<tr>
<th>Mutation position</th>
<th>Original codon</th>
<th>New codon</th>
</tr>
</thead>
<tbody>
<tr>
<td>T44C</td>
<td>ACC</td>
<td>TGT</td>
</tr>
<tr>
<td>R103C</td>
<td>CGC</td>
<td>TGT</td>
</tr>
<tr>
<td>S104C</td>
<td>AGC</td>
<td>TGT</td>
</tr>
<tr>
<td>N147C</td>
<td>AAT</td>
<td>TGT</td>
</tr>
<tr>
<td>R150C</td>
<td>CGG</td>
<td>TGT</td>
</tr>
<tr>
<td>V46C</td>
<td>GTT</td>
<td>TGT</td>
</tr>
</tbody>
</table>
8.2 Experimental detail for chapter 2

8.2.1 Cloning of WT EPO gene into pET16b

**PCR amplification of EPO cDNA**

For the amplification of the WT EPO gene, forward and reverse primers were designed. The WT EPO template, the primers, NTP’s and additional required additives were mixed gently before the addition of the Taq polymerase. The following amounts were used:

- 40 ng cDNA template
- 20 pmol primer I (see section 8.1.3)
- 20 pmol primer II (see section 8.1.3)
- 1 Unit Taq DNA polymerase
- 2 µl dNTP’s (100 mM),
- 2 µl 10x Taq reaction buffer,
- 1 µl MgCl₂(25 mM)

make to 20 µL with H₂O.

The following conditions were used in a thermocycler:

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denaturation</td>
<td>94°C</td>
<td>1 min</td>
</tr>
<tr>
<td>Annealing</td>
<td>55°C, 60°C, 65°C</td>
<td>1 min</td>
</tr>
<tr>
<td>extension</td>
<td>72°C</td>
<td>6.5 min</td>
</tr>
<tr>
<td>repeat 18 cycles</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Products were resolved on an agarose gel and labelled pDR1.
8.2 Experimental detail for chapter 2

**Agarose gel electrophoresis**

Agarose gels for DNA visualisation were poured at 1.2% (w/v) agarose in 1x TAE buffer. SYBR Safe™ was added (10μl to 100 ml agarose). The solution was poured into a tray and left to set. DNA samples were mixed with 5 equivalent volumes of DNA loading buffer.

Electrophoresis was performed at 110 V for 20 minutes.

**DNA digestion with restriction enzymes**

All restriction enzymes were obtained from New England Biolabs. Gene and pET16b vector were digested according to manufacturer’s instructions. After amplification, DNA was purified from the agarose gel according to manufacturers instructions (QIAquick Gel Extraction Kit, 28704). Upon purification, digestion of the WT EPO gene was carried out using the appropriate restriction enzymes. Correspondingly, the desired vector (in this case pET16b) was also digested with the appropriate restriction enzymes. The conditions for both restrictions are as follows:

**Digestion of WT EPO gene for cloning into pET16b vector:**

- 9 μg EPO gene
- 40 Units Bgl II
- 40 Units Nde I
- 4.2 μl Buffer 2 (from NEB)

make to 20 μL with H₂O.

Incubate at 37°C for 12-16 hours

The digested DNA was analysed by agarose gel electrophoresis and isolated by gel extraction according to the QIAGen Gel Extraction kit (28704).

**Digestion of pET16b:**
0.3 μg pET16b vector

10 Units Bam HI

20 Units Nde I,

2 μl Buffer 2 (from NEB)

make to 20 μL with H2O.

Incubate at 37°C for 12-16 hours

**Dephosphorylation of plasmid DNA**

The digested pET16b vector was incubated at 37°C for 4 hours with 10 Units of calf intestinal alkaline phosphate (CIP) to prevent re-ligation.

**Ligation of pET16b and the EPO insert**

Ligation achieved by the overnight incubation at 14°C under the following conditions:

- 3:1 molar ratio of DNA: plasmid
- 2.4 μl T4 ligase
- 2 μl ligase buffer

make to 20 μL with H2O.

After ligation, reactions were transformed into TOP10 *E. coli* cells and plated onto LB+Amp agar plates. Colonies were picked, grown and sequenced.
8.2 Experimental detail for chapter 2

8.2.2 Site directed mutagenesis to produce EPO cysteine variants

To obtain EPO cysteine variants, primers for site directed mutagenesis at certain sites within the EPO sequence were designed. The mutations and transformations were carried out according to the QuikChange Site-Directed Mutagenesis Kit (200518) protocol. The Primers for each cysteine variant are given in section 8.1.3. For each variant, the WT EPO gene was used as a template. This template, in addition to the appropriate primers, dNTP’s and additional required additives were mixed gently before the addition of the Pfu Ultra polymerase. The following amounts were used:

PCR mutagenesis reactions were set up as follows:

- 70 ng WT EPO DNA template
- 125 pmol of relevant FW primer (see section 8.1.3)
- 125 pmol of relevant RV primer (see section 8.1.3)
- 5% DMSO
- 1 µl dNTP’s (100 mM)
- 5 µl 10x Pfu polymerase buffer

make to 50 µL with H₂O, then

- 2.5 U Pfu Ultra polymerase added to reaction mixture at final step.

The following conditions were used in a thermocycler:

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denaturation</td>
<td>94°C</td>
<td>1 min</td>
</tr>
<tr>
<td>Annealing</td>
<td>65°C</td>
<td>1 min</td>
</tr>
<tr>
<td>extension</td>
<td>68°C</td>
<td>6.5 min</td>
</tr>
</tbody>
</table>

repeat 18 cycles
Products were transformed into XL-1 Blue competent cells according to manufacturer’s instructions. Plasmids were extracted and sent for sequencing.

8.2.3 Cloning of WT EPO gene into pPICZαA

The WT EPO gene was cloned out of pET16b using the primers:

FW primer: 5’-AC CCG GAA TTC ATG GGC CAT CAT CAT CAT C-3’

RV primer: 5’-CC AGC TCT AGA TCA TCT GTC CCC TGT CCT -3’

PCR of WT EPO gene from pET16b and introduction of restriction sites:

The following conditions were used for the simultaneous amplification of the WT EPO from pET16b and the introduction of the desired restriction sites for cloning into pPICZαA.

- 30 ng EPO in pET16b construct
- 20 pmol FW primer (above, 10μM)
- 20 pmol RV primer (above, 10μM)
- 0.5 μL dNTP’s (25 mM each),
- 5 μL 10x PfU Ultra II buffer,
- 1 μl PfU Ultra II polymerase

make to 50 μL with H₂O.

The PCR cycle is as in section 8.2.2. The digested DNA was analysed by agarose gel electrophoresis and isolated by gel extraction according to the QIagen Gel Extraction kit (28704). The extracted and purified DNA was then digested with the appropriate restriction enzymes to confirm the successful introduction of the WT EPO gene into the pPICZαA vector:
Digestion of the EPO gene and pPICZαA vector

- 30 μl DNA (1.8 μg EPO or 10 μg vector)
- 10 Units *EcoRI*
- 20 Units *XbaI*
- 5 μl 10x NEB buffer 4
- 0.5 μl BSA (1 μg/ml)

make to 50 μL with H2O.

Incubate at 37°C for 1 hour

The digested EPO gene was incubated at 37°C for 1 hour with 10 Units of calf intestinal alkaline phosphate (CIP) to prevent re-ligation. CIP was heat inactivated for 20 minutes at 60°C.

Ligation of pPICZαA and the EPO insert.

Ligation was achieved by incubation for 10 minutes at room temperature as per manufacturer’s instructions.

- 10 μl digested EPO insert (60 ng/μl)
- 2 μl digested pPICZαA (10 ng/μl)
- 2 μl T4 ligase buffer

make to 20 μL with H2O.

- 1 μl T4 ligase

After ligation, reactions were transformed into TOP10 F’ *E. coli* cells and plated onto Low salt LB+Zeocin (25 μg/ml) agar plates. Colonies were picked, grown and sequenced.
8.2.4 General procedure for expression of WT EPO and EPO cysteine variants

Expression of EPO gene in BL21(DE3) cells

The plasmid containing the EPO gene of interest was transformed by heat shock into BL21(DE3) *E. coli* cells according to manufacturer’s instructions.

Cultures were grown in 400 ml of LB+Amp in 2 l flasks (37°C, 250 rpm) until the \( \text{OD}_{600nm} = 0.6 \). Expression was induced by addition of IPTG and incubated for 4-24 hours at 37°C, 30°C and 16°C.

Hourly samples were taken to monitor the EPO gene expression. These samples were spun down and loaded onto a gel for SDS PAGE and Western Blot analysis.

Cells were harvested by centrifugation (10 minutes, 8000 rpm, 4°C) and stored at -20°C prior to purification.

Expression of EPO gene in OrigamiTM 2(DE3) cells

The plasmid containing the EPO gene of interest was transformed by heat shock into Origami 2(DE3) cells according to manufacturer’s instructions. Expression of the EPO gene in Origami 2(DE3) cells was achieved in an identical manner to expression in BL21(DE3).

Expression of commercially obtained EPO cysteine variants

Plasmids containing the commercially purchased EPO genes were transformed into BL21(DE3) cells as per manufacturer’s instructions.

Cultures were grown in 400 ml of LB+Amp + 1% glucose in 2 l flasks (37°C, 250 rpm) until the \( \text{OD}_{600} = 0.6 \). Expression was induced by addition of IPTG and incubated for 4-6 hours at 30°C, 20°C and 16°C. Cells were harvested after 6 hours by centrifugation (10 minutes, 8000 rpm, 4°C) and stored at -20°C prior to purification.
8.2 Experimental detail for chapter 2

8.2.5 General method for the purification and refolding of EPO variants

Preparation of cell free extract

Cell pellets were thawed on ice and re-suspended in Lysis buffer (section 8.1.2). Cells were lysed on ice by sonication (10 cycles, 30 seconds on, 30 seconds off). After centrifugation (15 minutes, 18000 rpm, 4°C), pellets were re-suspended in Lysis buffer. DNaseI and MgCl₂ were added to a final concentration of 10 mM and 2 U ml⁻¹ respectively. The suspension was incubated at 20°C for 20 minutes before sonication (10 cycles, 30 seconds on, 30 seconds off), centrifugation (15 minutes, 18000 rpm, 4°C) and re-suspended in Purification Binding buffer (section 8.1.2). To completely solubilise the protein, the suspension was stirred on ice for 1 hour and any remaining insoluble material was collected by centrifugation (15 minutes, 18000 rpm, 4°C). The supernatant was sterilised by filtration (Minisart 0.45 μm).

HisTrap nickel column chromatography

Purification of EPO was done on a HisTrap (HisTrap HP, 5 ml, GE Life Sciences, 17-5248-01) nickel column in denaturing conditions. Filtered EPO samples were loaded onto the column which was integrated into the automated AKTA purification system (AKTA Pure, GE Life Sciences).

A linear gradient from 100% Purification Binding buffer (section 8.1.2) to 100% Purification Elution buffer (section 8.1.2) was used to selectively elute the target protein. All fractions were analysed using SDS-PAGE and western blot techniques.

Preparing the column: Columns were stripped with 2 CV of 0.05 M EDTA, washed with 5 CV of distilled water followed by nickel loading with 2 CV (column volumes) of 0.1 M NiSO₄. The column was washed with 5 CV distilled water and equilibrated with 5 CV of Purification Binding buffer (section 8.1.2).
Ni-NTA agarose gravity-flow chromatography

The solubilised EPO was purified under denaturing conditions using Ni-NTA agarose (QIAGEN, 30210) gravity flow chromatography. Incubation of Ni-NTA agarose (~1 ml Ni-NTA agarose per ~3 mg total protein) with cell free extract for 1 hour at 4°C ensured binding of the histidine-tagged EPO to the nickel resin. This slurry was then applied to an empty PD-10 column (GE Life Sciences, 17-0435-01) and the flow-through was discarded. The self-packed column was washed with 5 CV Purification Binding buffer (section 8.1.2). EPO was selectively eluted with 2 CV of Purification Elution buffer (section 8.1.2).

8.2.6 General method for the refolding of EPO

After the purification of EPO, the protein was reduced with 1-5 mM DTT and incubated for 3 hours at 30°C. Samples were equilibrated to room temperature before being added drop-wise to Refolding buffer (section 8.1.2) where the protein was oxidatively refolded. This method is referred to as pulse refolding. The final concentration of EPO in Refolding buffer needs to be < 0.1 mg ml$^{-1}$ to ensure the protein does not aggregate. The refolded protein was then incubated with Ni-NTA agarose for 1 hour at 4°C. As above, this slurry was applied to an empty PD-10 column and the flow-through was discarded. EPO was selectively eluted with 2 CV of Native Elution buffer (section 8.1.2). The secondary structure of EPO was confirmed using Circular Dichroism (CD).

Circular Dichroism

Refolded EPO in Native Elution buffer was concentrated and buffer exchanged into PBS using Vivaspin centrifugal concentrators (Vivaproductions, VS0602) until all imidazole was removed and the OD$_{280\ nm} = 1$. BSA was used as a reference protein (OD$_{280\ nm} = 1$). A cuvette with a path length of 1 mm was used.
8.3 Experimental detail for chapter 3

8.3.1 Detecting EPO by LCMS

Mass spectrometry of all protein samples was carried out on an Agilent 1100 series HPLC and LC/ MSD SL (G1956B) ESI mass spectrometer. Data were analysed using the LC/MSD Chemstation (Rev. B.01.03 SR1) and accompanying deconvolution software.

All EPO samples (final concentration before injection 0.25-1 mgml\(^{-1}\), 5-10 µl injection) were suspended in a Denaturing Wash buffer (section 8.1.2).

Samples were run in acetonitrile / water (+0.05% formic acid) on a C4 silica gel column (SUPELCOSIL™ LC-304) at a flow rate of 0.5 mlmin\(^{-1}\). Protein was eluted at > 95% acetonitrile.

**HPLC conditions:**

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>% acetonitrile</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>10</td>
</tr>
<tr>
<td>3.00</td>
<td>10</td>
</tr>
<tr>
<td>5.00</td>
<td>30</td>
</tr>
<tr>
<td>12.00</td>
<td>90</td>
</tr>
<tr>
<td>18.00</td>
<td>90</td>
</tr>
<tr>
<td>18.50</td>
<td>10</td>
</tr>
<tr>
<td>20.00</td>
<td>10</td>
</tr>
</tbody>
</table>

Column thermostat: 40°C

**MS running conditions:**

Positive mode. Mass range: 600-1700 Da. 10 minute lag before injection.
**MS Spray Chamber conditions:**

- Drying Gas Flow (l/min): 12.0
- Nebuliser pressure (psi): 50
- Drying gas temperature (°C): 350
- Capillary voltage (V): Positive 5000, Negative 5000

**8.3.2 Chemical glycosylation of EPO variants with α-mannosyl iodoacetamide**

EPO variants were eluted from Ni-NTA affinity columns in Denaturing Elution buffer (section 8.1.2) and the concentrations were determined by measuring the absorbance at 280 nm. α-mannosyl iodoacetamide was stored in 25 mM aliquots and thawed when required.

Reactions were typically carried out on a 50-200 μl scale at the concentrations discussed in chapter 3. All reactions were carried out at room temperature with shaking unless other stated. Reactions were left shaking for 1-24 hours in the absence of light.

To monitor the reactions, 5-10 μl aliquots were taken and reduced with 10 mM DTT before analysis using LCMS. Reaction of EPO variants with α-mannosyl chloroacetamides were carried out using an identical methodology.
8.4 Experimental detail for chapter 4

8.4.1 General procedure for purification using Activated Thiol Sepharose

EPO samples were refolded and eluted from Ni-NTA affinity column in the appropriate buffer (native or denaturing depending on experiment).

GFP was provided in PBS.

Purification on thiol sepharose medium was done using Activated Thiol Sepharose 4B from GE Life Sciences (17-0640-01). Activated Thiol Sepharose 4B is supplied freeze-dried and distilled water was used for swelling and washing the medium. For ~1 mg of powder, ~4 ml of slurry was formed and washed with 200 ml of distilled water.

This slurry was then equilibrated with binding buffer. The thiol sepharose slurry was added to the protein samples (1 ml of slurry per 1 mg protein) and samples were left to bind for 2 hours with shaking. The slurry with the protein bound was applied to an empty column, and washed with 4 CV of the appropriate buffer (denaturing or native buffer for EPO and PBS for GFP). The slurry was then incubated with the buffer containing 20 mM DTT for 10 minutes. The slurry was then washed with 4 CV of buffer containing 20 mM DTT.

Fractions were collected and analysed by SDS PAGE Gel to determine presence of protein.

8.4.2 General procedure for purification using hydrophobic interaction chromatography (HIC)

Purification using HIC was done on a Phenyl-6-Fast Flow column from GE Life Sciences (17-0973-03).
EPO samples were eluted from Ni-NTA affinity column in Native Elution buffer (section 8.1.2) and buffer exchanged into the column loading buffer: 50 mM Tris-HCl, 500 mM imidazole, 2% n-lauryl sarcosine, 1 M (NH$_4$)$_2$SO$_4$ (or 3 M NaCl).

The HIC column was equilibrated with 3 CV of loading buffer. EPO samples were loaded and the column was washed with 5 CV of loading buffer to wash out unbound sample. The column was then washed with 1 CV of each of the following buffers:

All buffers consist of 50 mM Tris-HCl, 500 mM imidazole + n M salt (X):

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Salt (X)</th>
<th>Salt (X)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X = (NH$_2$)$_4$SO$_4$</td>
<td>X = NaCl</td>
</tr>
<tr>
<td>1</td>
<td>1.0 M</td>
<td>3.0 M</td>
</tr>
<tr>
<td>2</td>
<td>800 mM</td>
<td>2.5 M</td>
</tr>
<tr>
<td>3</td>
<td>600 mM</td>
<td>2.0 M</td>
</tr>
<tr>
<td>4</td>
<td>400 mM</td>
<td>1.5 M</td>
</tr>
<tr>
<td>5</td>
<td>200 mM</td>
<td>1.0 M</td>
</tr>
<tr>
<td>6</td>
<td>100 mM</td>
<td>0.5 M</td>
</tr>
<tr>
<td>7</td>
<td>0 M</td>
<td>0 M</td>
</tr>
</tbody>
</table>

Fractions were collected, concentrated and run on SDS PAGE gel to determine presence of the protein.
8.5 Experimental detail for chapter 5 and chapter 6

8.5.1 General EPO-(EPOR)$_2$ binding assay procedure

Commercially available Recombinant Human Erythropoietin Receptor, (EPOreceptor, CF. R&D Systems, 307-ER-050/CF), suspended in PBS was stored at -80°C at 0.5 mg/ml$^{-1}$. Dilution of 10 mg/ml$^{-1}$ were stored at -20°C and thawed when required.

The assay to determine the binding of EPO to the (EPOR)$_2$ was carried out as follows:

(EPOR)$_2$ (0.1 mg/ml$^{-1}$, 100 µl) was incubated in high binding 96 well plates (Costar®, 9018) for 5 hours at room temperature.

After incubation, wells were washed with 3xPBS+0.05% Tween 20.

Incubation with BSA (2% w/v in PBS, 250 µl) for 1 hour blocked any non-specific binding to the well surface.

Wells were washed with 3xPBS+0.05% Tween 20.

EPO samples (50 µl, varying concentrations prepared) were added to the wells and incubated for 1 hour at 20°C.

Wells were washed with 3xPBS+0.05% Tween 20 to remove any unbound protein.

In the final step, Anti-His HRP antibody (1:2000 dilution in PBS, 50 µl, from Sigma, A7058)) was incubated in the wells for 30 minutes.

Wells were washed with 4xPBS+0.05% Tween 20 to remove excess antibody.

Plates were immediately placed in the plate reader (equilibrated to 20°C set up to dispense 50 µl of TMB substrate to each well. The absorbance was monitored at 652 nm.

All steps were performed at room temperature with shaking.
Analysis of the ELISA data

The rate of product formation, $d\text{OD}/\text{min}$, was determined by the linear portion of the absorbance at 652 nm. Concentrations of the EPO samples were varied to generate a concentration–response curve for each EPO variant.

Statistical analysis

Data were expressed as mean ± S.D. Comparisons between groups were made using the paired t-test and significance was defined at a $p$ value less than 0.05 ($p < 0.05$).

8.5.2 Preparation of non-modified EPO samples for screening by ELISA

This section refers to section 6.2 where all EPO variants were screened for their binding to the (EPOR)$_2$. The EPO gene of interest was expressed and the protein purified as outlined in section 8.2.4. Once purified, soluble protein (suspended in Denaturing Wash buffer) was reduced using 2-5 mM DTT. Samples were left to reduce for up to 3 hours at 30°C.

Once reduced, proteins were equilibrated to room temperature and pulse refolded into Refolding buffer (section 8.1.2) to a final concentration < 0.1 mg ml$^{-1}$. Ni-NTA agarose was added to the solution and left rotating at 4°C for 1 hour. The Ni-NTA bound protein solution was applied to a column, and washed with Native Wash buffer (section 8.1.2). The protein was eluted from the column in Native Elution buffer (section 8.1.2).

Centrifugal size exclusion spin columns were used to concentrate the samples before the exact concentrations were determined by measuring the absorbance at 280 nm. After determining the concentration, samples were diluted to 1 μM and serially diluted into Native Elution buffer for immediate analysis on the (EPOR)$_2$ binding assay.
8.5 Experimental detail for chapter 5 and chapter 6

8.5.3 Preparation of glycosylated EPO samples for analysis by ELISA

EPO samples were prepared as outlined above (section 8.5.2) with the exception that in the final step, EPO samples were eluted in Denaturing Elution buffer (section 8.1.2) to facilitate the glycosylation reaction. Reaction of the protein samples with the appropriate sugars was carried out in the absence of light at the desired concentrations as outlined in section 8.3.

After completion of the reaction, samples were reduced with 20 mM DTT and incubated at 30°C for 1 hour. 5 µl aliquots were taken for LCMS analysis.

Protein samples were pulse refolded to yield the native, correctly folded, mixture of unmodified and glycosylated EPO samples. Samples were extracted from the diluted solution by the addition of Ni-NTA agarose. 1 ml of Ni-NTA agarose was added to the solution and left at 4°C for 1 hour before the suspension was applied to a column. The protein was eluted from the column in Native Elution buffer (section 8.1.2).

Centrifugal size exclusion spin columns were used to concentrate the samples before the exact concentrations were determined by measuring the absorbance at 280 nm. After determining the concentration, samples were diluted to 1 µM and serially diluted for the assay.

The preparation of the EPO samples were carried out to coincide with the timing of the assay so that the samples could immediately be applied to the wells.

8.5.4 Second generation protocol: preparation of glycosylated EPO samples for analysis by ELISA

In the optimised procedure discussed in section 6.4.2, the methodology for EPO samples preparation is as above in section 8.5.3, with some minor alterations. The
methods differ in that, after the glycosylation reaction, the EPO samples were directly diluted into Native Elution buffer to a final concentration of 1 μM. This sample was then serially diluted and immediately analysed on the EPO-(EPOR)₂ binding assay.
References


176
References


[72] Richardson, Dominique; Casteleijn, M. European pharmaceutical review.


References


184
References


References


