"In vitro studies of STAT3 dimerisation and DNA binding"

By

Edwin Nkansah

A thesis submitted in partial fulfilment of the
requirement for the degree of
Doctor of Philosophy
University of London
2011

Department of Pharmaceutical and Biological Chemistry
29-39 Brunswick Square
London
WC1N 1AX

This thesis describes a study carried out from October 2007 to June 2011 at the School of Pharmacy, University of London under the supervision of Dr. Andy Wilderspin. I certify that the work is original and I have written all the text herein and have clearly indicated with suitable citations any part of the work that has already appeared in publications.

Signature: Date:
ABSTRACT

STAT3 (Signal Transducer and Activator of Transcription 3) is a transcription factor that homodimerises and binds to the response element of a number of genes that promote tumourigenesis when over-expressed. Tumourigenesis, in many cases, correlates with an aberrant signalling cascade that leads to unregulated activation of STAT3. Activation is widely understood to be a result of phosphorylation at tyrosine 705 which leads to dimerisation and DNA binding.

The dimerisation and DNA binding of STAT3βtc in an isolated cell-free system was investigated in order to understand the conditions which modulate these interactions, and to facilitate the discovery of small molecules capable of interfering with these interactions. A GFP-STAT3βtc construct was designed in which the N-terminal 126 amino acids of STAT3 were replaced with spectral variants of the GFP (green fluorescent protein) to act as a reporter. Phosphorylated and unphosphorylated GFP-STAT3βtc was purified from BL21 (DE3) TKB1 and Rosetta strains of E. coli respectively. The isoforms were structurally characterized by size exclusion chromatography as dimers when phosphorylated (~ 250 kDa) and as monomers when unphosphorylated (~ 100 kDa).

Using both conventional ELISA and fluorescence-based microtiter plate assays, it was demonstrated that both phosphorylated and unphosphorylated GFP-STAT3βtc can bind to immobilised dsDNA (M67; a modified c-fos sis inducible enhancer) or to an immobilised surrogate receptor (a phosphotyrosyl peptide derived from the interleukin-6 receptor subunit gp130), respectively. A DNA binding assay (PEMSA) was developed for the STAT3 transcription factor based on detection of a protein mobility shift rather than a dsDNA mobility shift, and which utilizes fluorescently tagged STATs. Using PEMSA, DNA binding was demonstrated for both phosphorylated and unphosphorylated GFP-STAT3βtc isoforms. Disruption of a pre-formed unphosphorylated GFP-uSTAT3βtc/dsM67 DNA complex was achieved with a phosphotyrosyl peptide at a DB50 of 30 µM. In contrast, 1000 µM peptide appeared to have no effect on phosphorylated GFP-pSTAT3βtc bound to the dsM67 DNA.

FRET was used to directly study dimerisation using the eCFP-uSTAT3βtc/eYFP-uSTAT3βtc pair in vitro. Low pH/low salt buffer as well as co-factors including dsM67 DNA promoted and stabilised STAT3/STAT3 assembly. FRET was also observed between eCFP-uSTAT3βtc and a 5'-FAM labelled phosphotyrosyl peptide which could also form the basis of a high throughput screening assay.
CONTENTS

ABSTRACT 2
CONTENTS 3
LIST OF FIGURES 15
LIST OF TABLES 19
ACKNOWLEDGEMENTS 20
ABBREVIATIONS 21
REFERENCES 268

1 GENERAL INTRODUCTION .................................................................................................................. 23

1.1 Cancer ........................................................................................................................................ 23

1.1.1 Protein-Protein Interaction ........................................................................................................ 24

1.1.2 Signal Transducers and Activators of Transcription (STAT) ......................................................... 24

1.1.2.1 STAT family and evolution .................................................................................................. 26

1.1.2.2 Isoforms of STAT .............................................................................................................. 26

1.1.2.3 Alternative splicing ............................................................................................................ 28

1.1.3 Activation of STAT transcription factors ..................................................................................... 29

1.1.3.1 JAK/STAT Signalling ......................................................................................................... 29

1.1.3.2 Dimerization of STAT protein ......................................................................................... 31

1.1.3.3 Nucleocytoplasmic transport of STATs ............................................................................ 31

1.1.3.4 STATs DNA binding sites ................................................................................................. 32

1.1.4 Crystal structures of STATs ....................................................................................................... 32

1.1.5 Inactive STAT dimers .............................................................................................................. 34

1.1.6 Role of STATs in mediating growth, differentiation and apoptosis ........................................... 34

1.1.7 Signal Transducer and Activator of Transcription-3 (STAT3) .................................................... 35

1.1.7.1 STAT3 target genes ......................................................................................................... 37

1.1.8 STAT3 and oncogenesis ............................................................................................................ 38

1.1.9 Regulating STAT3 signalling ..................................................................................................... 39

1.1.9.1 Positive regulation of STAT3 signalling ............................................................................. 39

1.1.9.1.1 Tyrosine phosphorylation ............................................................................................. 39

1.1.9.1.2 Serine phosphorylation ............................................................................................... 39

3
2.1.6 Antibodies................................................................................................................. 60
2.1.7 Equipment................................................................................................................ 60

2.2 METHODS ....................................................................................................................... 61

2.2.1 Molecular biology methods .......................................................................................... 61
  2.2.1.1 Escherichia coli (E.coli) stock preparation and transformation ...... 61
    2.2.1.1.1 Preparation of chemically competent E.coli cells with antibiotic marker 61
    2.2.1.1.2 Preparation of chemically competent E. coli cells with no antibiotic marker 62
    2.2.1.1.3 E.coli competent cell transformation .............................................................. 62
    2.2.1.1.4 Screening E.coli colonies .............................................................................. 62
    2.2.1.1.5 Glycerol stock of transformed E.coli cells .................................................. 63
    2.2.1.1.6 Mini preparation of plasmid DNA ............................................................... 63
    2.2.1.1.7 Maxi preparation of plasmid DNA ............................................................... 64
  2.2.1.2 Molecular cloning ............................................................................................ 65
    2.2.1.2.1 Restriction endonuclease digestion ................................................................. 65
    2.2.1.2.2 Analysis of DNA using flat bed agarose gel electrophoresis ....................... 66
    2.2.1.2.3 Flat bed agarose gel purification of restricted DNA ..................................... 67
    2.2.1.2.4 Precipitation of plasmid DNA using ethanol ............................................... 67
    2.2.1.2.5 Annealing oligonucleotide DNA primers using water-bath ...................... 67
    2.2.1.2.6 Designing PCR oligonucleotide primers ...................................................... 68
    2.2.1.2.7 General PCR-amplification reaction ............................................................. 69
    2.2.1.2.8 Purification of PCR products ....................................................................... 70
    2.2.1.2.9 Dephosphorylation of the 5’ end of the cloning vector ............................... 71
    2.2.1.2.10 Ligation reaction ......................................................................................... 71
    2.2.1.2.11 Sequencing analysis of selected clones ....................................................... 72
  2.2.1.3 Recombinant protein ..................................................................................... 72
    2.2.1.3.1 Small scale expression of recombinant protein in E.coli ....................... 72
    2.2.1.3.2 Large scale fermentation of phosphorylated (p) recombinant proteins in BL21 (DE3) TKB1 E.coli cells ................................................................. 72
    2.2.1.3.3 Large scale fermentation of unphosphorylated (u) recombinant protein in E.coli 74
  2.2.2 Extraction of recombinant protein from E.coli cells................................. 74
    2.2.2.1 E.coli cell lysis ............................................................................................... 74
    2.2.2.1.1 Mechanical disruption of E.coli cell wall .................................................. 75
2.2.2.1.2 Chemical disruption of E.coli cell membrane ................................................. 75

2.2.3 Purification processes................................................................................................................................. 76

2.2.3.1 Purification of soluble target protein........................................................................................................... 76

2.2.3.1.1 Ammonium sulphate precipitation.............................................................................................................. 77

2.2.3.1.2 Ion exchange chromatography .................................................................................................................. 78

2.2.3.1.3 Size exclusion chromatography (gel filtration)............................................................................................ 79

2.2.3.1.4 Affinity purification of polyhistidine tag recombinant proteins.......................................................... 80

2.2.4 Protein quantification techniques .................................................................................................................. 81

2.2.4.1 Protein quantification using the Bradford assay (Bio-Rad reagent). 81

2.2.4.2 Protein quantification by intrinsic GFP absorbance.................................................................................. 81

2.2.5 Protein analytical techniques ...................................................................................................................... 82

2.2.5.1 Electrophoresis........................................................................................................................................... 82

2.2.5.1.1 Sodium Dodecyl Sulphate−Polyacrylamide Gel Electrophoresis (SDS-PAGE) 82

2.2.5.1.2 Preparation of resolving gel..................................................................................................................... 83

2.2.5.1.3 Preparation of stacking gel .................................................................................................................... 83

2.2.5.1.4 Preparation of protein sample for SDS-PAGE .......................................................................................... 83

2.2.5.1.5 Electrophoresis of denatured protein samples ........................................................................................ 84

2.2.5.2 Electrophoretic transfer of denatured protein onto nitrocellulose membrane 84

2.2.5.2.1 Preparation for blotting .......................................................................................................................... 84

2.2.5.2.2 Development of nitrocellulose membrane.................................................................................................. 85

2.2.5.3 Non-denaturing polyacrylamide gel electrophoresis (Native-PAGE) 86

2.2.5.3.1 Triton X-100 polyacrylamide gel (Triton native gel).................................................................................. 86

2.2.5.3.2 Preparation and equilibration of Triton X-100 resolving gel.................................................................... 86

2.2.5.3.3 Preparation of protein sample for native-PAGE ....................................................................................... 87

2.2.5.3.4 Electrophoresis of non-denatured protein samples ................................................................................... 87

3 RECOMBINANT DNA ..................................................................................................................................... 88

3.1 Creating fluorescent expression constructs .................................................................................................... 88

3.1.1 Fluorescent STAT3βtc recombinant DNA expression................................................................................ 88

3.1.1.1 Developing GFP-STAT3βtc expression constructs (pET-32a (+) GFP-STAT3βtc ......................................................... 88

3.1.1.1.1 PCR-amplification of GFP cDNA variants .............................................................................................. 88
3.1.1.2 Preparation of cloning vector and insert DNA by restriction digest ........ 90
3.1.1.3 Ligation reaction ...................................................................................... 91
3.1.1.4 Transformation of JM109 cells with ligated products .................................. 92
3.1.1.5 Purification of potential clone .................................................................... 92
3.1.1.6 Selection of potential clones by restriction analysis ................................... 92
3.1.1.7 Screening of potential clones by PCR ...................................................... 93
3.1.1.8 Amplification of suspected clones (Maxiprep, Qiagen) .................................. 93
3.1.1.9 Restriction mapping of prospective clones .................................................. 94
3.1.1.10 Nucleotide sequence analysis of selected clones ......................................... 95
3.1.2 Cloning pET-28c (+) -GFP expression constructs ........................................ 95
3.1.2.1 PCR-amplification of eCFP, eGFP and eYFP ........................................... 96
3.1.2.2 Preparation of cloning vector and insert DNA by restriction digest . ......... 96
3.1.2.3 Ligation reaction ...................................................................................... 96
3.1.2.4 Transformation of XL1-Blue cells with ligated products ....................... 97
3.1.2.5 Screening of potential clones by restriction analysis ............................ 98
3.1.2.6 Restriction mapping of selected clones .................................................... 98
3.1.3 Creating pET-28c (+) -STAT3ptc-GFP expression constructs ..................... 99
3.1.3.1 Ligating STAT3ptc into pET-28c (+) GFP cloning vector ....................... 99
3.1.3.2 PCR amplification of STAT3ptc insert gene .......................................... 100
3.1.3.3 Preparation of cloning vector and insert DNA by restriction digest 100
3.1.3.4 Ligation reaction .................................................................................. 101
3.1.3.5 Selection of potential clones by restriction analysis ............................ 102
3.1.3.6 Restriction mapping of selected clones .................................................. 102
3.1.4 Creating pET-28c (+)-His-STAT3ptc expression vectors with C-tagged GFP 103
3.1.4.1 PCR amplification of STAT3ptc insert gene .......................................... 104
3.1.4.2 Preparation of cloning vectors and insert DNA by restriction digest ......... 104
3.1.4.3 Ligation reaction .................................................................................. 105
3.1.4.4 Selection of potential clones by restriction analysis ............................ 106
3.1.4.5 Screening potential clones by PCR ...................................................... 107
3.1.4.6 Restriction mapping of selected clones .................................................. 107
3.1.5 Creating pET-28c(+)-His–eCFP-TEV-eYFP construct

3.1.5.1 PCR-amplification eYFP cDNA

3.1.5.2 Preparation of cloning vector and insert DNA by restriction digest

3.1.5.3 Ligation reaction

3.1.5.4 Selection of potential clones by restriction analysis

3.1.5.5 Restriction mapping of selected clones

3.2 RESULTS

3.2.1 Creating fluorescent STAT3βtc expression constructs

3.2.1.1 pET-32a (+)-GFP-STAT3βtc expression constructs

3.2.2 pET-28c (+)-STAT3βtc-GFP expression constructs

3.2.2.1 pET-28c (+)-GFP expression constructs

3.2.2.2 pET-28c (+)-STAT3βtc-GFP expression constructs

3.2.3 pET-28c (+)-His-STAT3βtc-GFP expression constructs

3.2.4 pET-28c (+)-His-eCFP-TEV-eYFP expression constructs

3.2.5 Expression of recombinant, fluorescent chimaeric proteins in E.coli cells

3.2.5.1 Expression of recombinant, fluorescent STAT3βtc in E.coli cells

3.2.5.2 Expression of the spectral variants of GFP in E.coli cells

3.2.6 Lysis of E.coli cells for the purification of the soluble recombinant, fluorescent chimaeric proteins

3.2.6.1 Extraction of the recombinant GFP-STAT3βtc chimaeric protein using sonication

3.2.6.2 Extraction of the His-eCFP-TEV-eYFP chimaeric protein

3.2.7 Western Bolt analysis of the extracted recombinant, GFP-STAT3βtc chimaeric protein

3.2.8 Purification of N-terminal GFP-STAT3βtc chimaeric protein (GFP-STAT3βtc)

3.2.8.1 Ammonium sulphate precipitation

3.2.8.2 Purification of fluorescent STAT3βtc present in 40 % fraction using anion exchange chromatography

3.2.8.3 Ni²⁺ Sepharose™ affinity purification of recombinant His-eCFP-TEV-eYFP fusion protein
3.2.9 Spectroscopic properties of the fluorescent chimaeric proteins

3.2.9.1 Spectroscopic properties of GFP-STAT3βtc chimaeric proteins

3.2.9.1.1 Excitation and emission spectra obtained from enhanced (Cyan, Green and Yellow) GFP-STAT3βtc chimaeric proteins

3.2.9.1.2 Spectral characteristic of His-eCFP-TEV-eYFP

3.3 DISCUSSION

3.3.1 Construction, expression and extraction of fluorescent STAT3βtc expression constructs

3.3.2 Purification of fluorescent STAT3βtc

3.3.3 GFP-STAT3βtc retains spectral properties

4 STRUCTURAL CHARACTERIZATION

4.1 Characterization of phosphorylated and unphosphorylated GFP-STAT3βtc isoform

4.1.1 Native Triton X-100 Polyacrylamide Gel Electrophoresis (Native PAGE)

4.1.1.1 Solubilizing effect of Triton X-100 on GFP-STAT3βtc isoforms

4.1.1.2 Reducing GFP-STAT3βtc isoforms at varied temperatures

4.1.1.3 Disruptive effect of freeze/thaw treatment on GFP-STAT3βtc

4.1.1.2 Native Triton X-100 Polyacrylamide Gel Electrophoresis (Native PAGE)/Ferguson Plots

4.1.1.3 Size exclusion chromatography (SEC)

4.2 RESULTS

4.2.1 Analysis of purified GFP-STAT3βtc by native Triton X-100 PAGE

4.2.2 Characterization of GFP-uSTAT3βtc and GFP-pSTAT3βtc by Native-PAGE (TX)
4.2.3 Characterization of purified recombinant phosphorylated and unphosphorylated fluorescent STAT3βtc by size exclusion chromatography .... 160

4.3 DISCUSSION ................................................................................................................165

4.3.1 GFP-uSTAT3βtc multimerizes into aggregates via inter-chain disulphide bridging in response to increasing temperature ........................................................... 165

4.3.2 Characterizing phosphorylated and unphosphorylated GFP-STAT3βtc using Native-PAGE (TX)/Ferguson plot ................................................................. 167

4.3.3 Characterizing phosphorylated and unphosphorylated GFP-STAT3βtc using Size Exclusion Chromatography (SEC) .............................................................. 168

5 BINDING SPECIFICITY OF GFP-STAT3βtc ................................................................171

5.1 Enzyme-Linked Immuno-Sorbent Assays (ELISA) .............................................. 171

5.1.1 Preparation of Streptavidin coated plates ...................................................... 172

5.1.2 Evaluation of streptavidin coating ................................................................. 172

5.1.3 Immobilization of biotin-dsM67 DNA on streptavidin coated plate............. 173

5.1.3.1 Microtiter plate assay to demonstrate the capture of phosphorylated GFP-STAT3βtc (GFP-pSTAT3βtc) by immobilized dsM67 DNA ......................... 173

5.1.4 Immobilization of biotin-phosphotyrosyl peptide on streptavidin coated plate 174

5.1.4.1 Microtiter plate assay to demonstrate the capture of unphosphorylated GFP-STAT3βtc (GFP-uSTAT3βtc) by immobilized phosphotyrosyl peptide ................................................................. 175

5.1.5 Detecting captured GFP-STAT3βtc ............................................................... 176

5.1.5.1 Detection by intrinsic GFP fluorescence .............................................. 176

5.1.5.2 Detection by conventional ELISA colorimetry ...................................... 176

5.1.6 Inhibition of GFP-STAT3βtc/probe interaction ............................................. 177

5.1.6.1 Inhibition of GFP-pSTAT3βtc/dsM67 DNA complex with either dsM67 DNA (unlabeled) or phosphotyrosyl peptide (unlabeled) ........................................ 177

5.1.6.2 Inhibition of GFP-uSTAT3βtc/phosphotyrosyl peptide complex with either dsM67 DNA (unlabeled) or phosphotyrosyl peptide (unlabeled) .......... 178

5.2 RESULTS ....................................................................................................................... 179

5.2.1 Enzyme-Linked Immuno-Sorbent Assay (ELISA) formats ......................... 179

5.2.1.1 Evaluation of streptavidin coating .......................................................... 179
5.2.1.2 Analysis of the biological functionality of GFP-STAT3βtc isoforms

5.2.1.2.1 Detection of GFP-uSTAT3βtc bound to the immobilized biotin-phosphotyrosyl peptide .............................................................. 181

5.2.1.2.2 Detection of GFP-pSTAT3βtc bound to the immobilized biotin-DNA oligonucleotide probe ................................................................. 184

5.2.1.3 Disruption of GFP-uSTAT3βtc/phosphotyrosyl peptide or GFP-uSTAT3βtc/dsM67 DNA complexes using immobilized biotin-phosphotyrosyl peptide in vitro ................................................................................ 186

5.2.1.4 Disruption of GFP-pSTAT3βtc/phosphotyrosyl peptide or GFP-pSTAT3βtc/dsM67 DNA complexes using immobilized biotin-dsM67 DNA in vitro 188

5.3 DISCUSSION ................................................................................................................ 190

5.3.1 Streptavidin coated plates ................................................................................. 191

5.3.2 Immobilized capture probes .............................................................................. 192

5.3.3 Specific interaction of the recombinant GFP-uSTAT3βtc with the immobilized phosphotyrosyl peptide ................................................................. 192

5.3.4 Specific interaction of the recombinant GFP-pSTAT3βtc with the immobilized dsM67 oligonucleotide DNA ................................................................................ 193

5.3.5 H-pYLPQTV-NH2 and dsM67 DNA significantly inhibits GFP-uSTAT3βtc binding to the pY motif of the immobilized biotin- H-pYLPQTV-NH2 peptide sequence .................................................................................................. 195

5.3.6 H-pYLPQTV-NH2 and dsM67 DNA disrupts GFP-pSTAT3βtc binding to the STAT3 recognition sequence contained within the immobilized biotinylated dsM67 oligonucleotide DNA .............................................................................. 195

6 GFP-STAT3βtc/dsDNA COMPLEX IN GELS ............................................................... 197

6.1 Electrophoretic Mobility Shift Assay (EMSA) ....................................................... 197

6.1.1 EMSA in polyacrylamide gel .......................................................................... 198

6.1.1.1 Preparation of Tris-Boric-EDTA (TBE) polyacrylamide gel .................. 198

6.1.2 EMSA in agarose gels .................................................................................... 198

6.1.2.1 Preparation of Tris-Boric-EDTA (TBE) agarose gel ............................. 199

6.1.3 Preparation of binding reactions ..................................................................... 199

6.1.3.1 Preparation of unbound GFP-STAT3βtc control ...................... 199

6.1.3.2 Preparation of GFP-STAT3βtc/dsM67 DNA complex reaction ...... 199
6.1.3.2.1 Determination of the molar ratio for GFP-STAT3\(\beta\)tc/dsM67 complex formation

6.1.3.2.2 Analysis of the time needed for GFP-STAT3\(\beta\)tc/dsM67 DNA complex formation

6.1.3.3 Preparation of the complex between GFP-STAT3\(\beta\)tc and the phosphotyrosyl peptide

6.1.3.4 Inhibition or disruption of GFP-STAT3\(\beta\)tc/dsM67 DNA complex with a phosphotyrosyl peptide inhibitors

6.1.3.4.1 Disruption of pre-formed GFP-STAT3\(\beta\)tc/dsM67 DNA complex using the different forms of phosphotyrosyl peptide inhibitors

6.1.4 Electrophoresis of protein, protein/DNA and protein/peptide complexes

6.1.4.1 Electrophoretic transfer of binding reactions

6.1.4.2 Preparation for western blotting (Nitrocellulose membrane)

6.1.4.3 Preparation for southern blotting (nylon membrane)

6.1.4.4 Development of positively charged nylon membrane

6.1.4.5 Detecting biotin-labeled dsM67 oligonucleotide DNA

6.2 RESULTS

6.2.1 Electrophoretic Mobility Shift Assay (EMSA)

6.2.1.1 Demonstration of the differential DNA binding activities between the fluorescent STAT3\(\beta\)tc isoforms and the dsM67 DNA

6.2.1.2 Disrupting GFP-STAT3\(\beta\)tc/dsM67 DNA complex with a phosphotyrosyl peptide

6.2.1.3 Investigating the order of addition of the components for inhibition of the GFP-STAT3\(\beta\)tc/ dsM67 DNA complex formation

6.2.1.4 Investigating the inhibitory activity of the phosphotyrosyl peptide inhibitors in agarose gels

6.2.1.5 Comparative inhibition of the DNA-binding activity of GFP-uSTAT3\(\beta\)tc using either peptide or non-peptide small molecule inhibitors

6.3 Discussion

6.3.1 Protein Electrophoretic Mobility Shift Assay (PEMSA)

6.3.2 Differential DNA binding activity of the phosphorylated and the unphosphorylated recombinant, GFP-STAT3\(\beta\)tc isoforms
6.3.3 Disruption of the DNA binding activity of the purified, recombinant GFP-STAT3βtc by phosphotyrosyl peptide ................................................................. 220

7 FRET-BASED ANALYSIS OF GFP-STAT3βtc/GFP-STAT3βtc INTERACTIONS

7.1 Fluorescent Resonance Energy Transfer (FRET) ........................................... 226

7.1.1 Absorption and fluorescent measurement ............................................. 228

7.1.2 Preparation for FRET ........................................................................... 228

7.1.2.1 Preparing eCFP-uSTAT3βtc (donor)/eYFP-uSTAT3βtc (acceptor) FRET pair assay mixture ................................................................. 228

7.1.2.1.1 Stabilizing effect of dsM67 DNA in eCFP-uSTAT3βtc (donor)/eYFP-uSTAT3βtc (acceptor) interactions measured by FRET ............................... 229

7.1.2.1.2 Stabilizing effect of divalent cation (Mg²⁺) in eCFP-uSTAT3βtc (donor)/eYFP-uSTAT3βtc (acceptor) interactions measured by FRET .............................. 229

7.1.2.1.3 Stabilizing effect of glycerol in eCFP-uSTAT3βtc (donor)/eYFP-uSTAT3βtc (acceptor) interactions measured by FRET ................................................... 230

7.1.2.1.4 Ionic strength and eCFP-uSTAT3βtc (donor)/eYFP-uSTAT3βtc (acceptor) interactions measured by FRET ........................................................... 230

7.1.2.2 Preparing control assay for measuring FRET by donor fluorophore quenching (FRET efficiency) ................................................................. 231

7.1.2.3 Preparing control assay for measuring FRET by sensitized acceptor emission upon donor excitation (FRET ratio) ................................................... 232

7.1.2.4 Preparing control assay for non-specific interactions between donor (eCFP) and acceptor (eYFP) fluorophores .................................................. 232

7.1.2.5 Preparing control assay to demonstrate FRET using eCFP-TEV-eYFP chimaeric protein ................................................................. 233

7.1.3 FRET measurement ............................................................................. 233

7.2 RESULTS ............................................................................................... 235

7.2.1 Fluorescent Resonance Energy Transfer (FRET) .................................. 235

7.2.2 Demonstration of FRET ....................................................................... 236

7.2.3 Steady-state FRET assays between eCFP-uSTAT3βtc (donor) and eYFP-uSTAT3βtc (acceptor) protein species ......................................................... 239

7.2.4 Investigating selected ingredients of the FRET buffer that may improve the efficiency of energy transfer between FRET pair .................................. 242
7.2.5 Investigating the co-factors that influence FRET between eCFP-uSTAT3βtc and eYFP-uSTAT3βtc FRET pair species .............................................. 246

7.2.6 Disruption of FRET signals using STAT3-SH2 specific phosphotyrosyl peptide 249

7.3 DISCUSSION ................................................................................................................ 253

7.3.1 Protein-Protein Interaction between two GFP-uSTAT3βtc molecules is demonstrated by FRET ................................................................. 254

7.3.2 The GFP fusion construct demonstrates FRET ............................................. 255

7.3.3 FRET between eCFP-uSTAT3βtc and eYFP-uSTAT3βtc .......................... 256

7.3.4 FRET is stabilized in the presence of certain ingredients ........................... 257

7.3.5 Co-factor stabilizes FRET pair assembly ........................................................ 259

8 GENERAL DISCUSSION ......................................................................................... 263

8.1 FUTURE WORK ................................................................................................... 267
List of Figures

Figure 1.1: Schematic representation of the human STAT protein and their functional domains ..............................................................................................................................................27
Figure 1.2: Model of STAT nucleocytoplasmic shuttling ........................................................30
Figure 1.3: Model of the crystal structures STAT3β homodimer bound to DNA .......... 33
Figure 1.4: Schematic representation of the GFP-STAT3βtc construct used in this study ..............................................................................................................................................46
Figure 1.5: GFP structural features including approximate dimensions and mutations present in spectral variants ...........................................................................................................49
Figure 1.6: Excitation and emission spectra of the common spectral variants of Aequorea GFP ..................................................................................................................................................51
Figure 2.1: Summary of the procedure used to generate clones in the study ....................66
Figure 2.2: Summary of the oligonucleotide dsDNA sequence generated using the annealing protocol described in section 2.2.1.2.5 ................................................................ 68
Figure 2.3: Representative PCR product detailing the endonuclease restriction sites introduced ......................................................................................................................................70
Figure 3.1: Summary of the cloning procedure used to develop the fluorescent expression constructs, pET-32a (+)-eCFP-STAT3βtc, pET-32a (+)-eGFP-STAT3βtc and pET-32a (+)-eGFP-STAT3βtc ..............................................................................................................................................89
Figure 3.2: Summary of the cloning procedure used to develop the expression constructs, pET-28c (+)-eCFP, pET-28c (+)-eGFP and pET-28c (+)-eYFP .......................................................... 95
Figure 3.3: Summary of the cloning procedure used to develop the expression constructs; pET-28c (+)-STAT3βtc-eCFP, pET-28c (+)-STAT3βtc-eGFP and pET-28c (+)-STAT3βtc-eYFP ..............................................................................................................................................100
Figure 3.4: Summary of the cloning procedure for developing expression construct pET-28c (+)-GFP cloning vector to generate pET-28c(+)-His-STAT3βtc-eCFP, pET-28c(+)-His-STAT3βtc-eGFP and pET-28c (+)-His-STAT3βtc-eYFP expression constructs ..............................................................................................................................................104
Figure 3.5: Summary of the cloning procedure used to generate pET-28c (+) His-eCFP-TEV-eYFP expression construct .......................................................... 108
Figure 3.6: Plasmid DNA map of GFP- STAT3βtc ........................................................................... 112
Figure 3.7: Schematic representation of the GFP-STAT3βtc expression constructs. 112
Figure 3.8: Endonuclease restriction mapping of the pET-32a (+)-GFP-STAT3βtc expression constructs ......................................................................................................................113
Figure 3.9: Schematic representation of STAT3βtc-GFP expression constructs .... 114
Figure 3.10: Endonuclease restriction mapping of the pET-28c (+)-GFP expression constructs ..............................................................................................................................................115
Figure 3.11: Endonuclease restriction mapping of the pET-28c (+)-STAT3\(\beta\)tc-GFP expression constructs ................................................................. 116
Figure 3.12: Schematic representation of the His-STAT3\(\beta\)tc-GFP expression constructs ........................................................................................................... 118
Figure 3.13: Endonuclease restriction mapping of the pET-28c (+)-His-STAT3\(\beta\)tc-GFP expression constructs ................................................................. 119
Figure 3.14: Schematic representation of His-eCFP-TEV-eYFP constructs .... 120
Figure 3.15: Endonuclease restriction mapping of the pET-28c (+) His-eCFP-TEV-eYFP expression constructs ........................................................................... 121
Figure 3.16: Expression of the recombinant, STAT3\(\beta\)tc-GFP expression constructs in E.coli cell. ........................................................................................................ 123
Figure 3.17: Expression of the recombinant, His-STAT3\(\beta\)tc-GFP expression constructs in E.coli cells ......................................................................................................... 124
Figure 3.18: Photographic representation of recombinant, uSTAT3\(\beta\)tc and GFP-uSTAT3\(\beta\)tc expressed in E.coli cells ........................................................................ 125
Figure 3.19: Expression of the recombinant, GFP-STAT3\(\beta\)tc expression constructs in E.coli cell ................................................................. 126
Figure 3.20: Extraction of the recombinant GFP-uSTAT3\(\beta\)tc and GFP-pSTAT3\(\beta\)tc using sonication ................................................................. 128
Figure 3.21: Photograph of illuminated E.coli cells expressing recombinant, His eCFP-TEV-eYFP Chimaeric proteins ........................................................................ 129
Figure 3.22: Extraction of soluble recombinant, His-eCFP-TEV-eYFP chimaeric protein from E.coli cell by sonication ................................................................. 129
Figure 3.23: Western blot analysis of the GFP-STAT3\(\beta\)tc, ........................................... 130
Figure 3.24: Ammonium sulphate precipitate fractions analysed for recombinant, unphosphorylated (A) and phosphorylated (B) GFP-STAT3\(\beta\)tc chimaeric proteins. .. 132
Figure 3.25: Demonstration of fluorescent STAT3\(\beta\)tc content present in the fractions generated upon ammonium sulphate precipitation ........................................ 133
Figure 3.26: Purification of recombinant enhanced (cyan, green and yellow) unphosphorylated STAT3\(\beta\)tc chimaeric protein using anion exchange chromatography. ................................................................................................. 134
Figure 3.27: Purification of recombinant enhanced (cyan, green and yellow) phosphorylated STAT3\(\beta\)tc chimaeric protein using anion exchange chromatography. ................................................................................................. 135
Figure 3.28: Purification of recombinant His-eCFP-TEV-eYFP .................................. 136
Figure 3.29: Normalized absorption (A; dashed lines) and emission spectra (B; solid lines) for GFP-STAT3\(\beta\)tc chimaeric protein ........................................................................ 138
Figure 3.30: Normalized fluorescence emission spectra of His-eCFP-TEV-eYFP fusion protein

Figure 4.1: Schematic representation of the native Triton X-100 polyacrylamide gel.

Figure 4.2: Disruption and solubilisation of fluorescent STAT3βtc.

Figure 4.3: Migration of the purified fluorescent STAT3βtc isoforms with standard protein markers on native gels with different acrylamide content.

Figure 4.4: The relative migrations of the purified GFP-uSTAT3βtc, GFP-pSTAT3βtc and the standard protein markers in the series of native Triton X-100 polyacrylamide gels.

Figure 4.5: Evaluation of the molecular weights of GFP-uSTAT3βtc and GFP-pSTAT3βtc using the Ferguson plot.

Figure 4.6: Size exclusion chromatograms of purified, recombinant, GFP-pSTAT3βtc and GFP-uSTAT3βtc using Superose™ 6 column.

Figure 4.7: Chromatographic calibration curve for the various standard protein markers and GFP-pSTAT3βtc and GFP-uSTAT3βtc generated with Superose™ 6 column.

Figure 4.8: Elution profiles of the GFP-STAT3βtc isoforms analyzed by SDS-PAGE and western blot.

Figure 5.1: Schematic representation of the DNA binding assay.

Figure 5.2: Schematic representation of the peptide binding assay.

Figure 5.3: Maxisorp plate format for plate validation.

Figure 5.4: Validation of streptavidin coating using biotin-HRP.

Figure 5.5: Detection of bound GFP-uSTAT3βtc to immobilized biotin-phosphotyrosyl peptide surrogate by fluorescence.

Figure 5.6: Detection of bound GFP-uSTAT3βtc to immobilized biotin-phosphotyrosyl peptide surrogate by conventional ELISA.

Figure 5.7: Detection of bound GFP-pSTAT3βtc to immobilized biotin-dsM67 oligonucleotide probe by fluorescence.

Figure 5.8: Detection of bound GFP-pSTAT3βtc to immobilized biotin-dsM67 oligonucleotide DNA probe by conventional ELISA.

Figure 5.9: Percentage inhibition of biotin-phosphotyrosyl/GFP-uSTAT3βtc interaction.

Figure 5.10: Percentage inhibition of biotin-dsM67 DNA - GFP-pSTAT3βtc interaction as a function of inhibitor concentration.

Figure 6.1: Differential DNA binding activities of the of the purified fluorescent STAT3βtc isoforms.

Figure 6.2: Purified, recombinant, GFP-STAT3βtc isoforms bind strongly to biotinylated dsM67 DNA.

Figure 6.3: Disruption of GFP-STAT3βtc/dsM67 DNA-binding by a peptide.
Figure 6.4: Peptide probe, H-pY-LPQTV-NH₂, disrupts DNA binding activity of GFP-uSTAT3βtc. .................................................................................................................................................. 212

Figure 6.5: Demonstration of the disruptive effect of the H-pY-LPQTV-NH₂ peptide on the pre-formed GFP-STAT3βtc/dsM67 DNA complex ........................................................................................................ 214

Figure 6.6: Demonstration of the disruptive effect of the Ac-pY-LPQ-NH₂ peptide on the preformed GFP-STAT3βtc/dsM67 DNA complex ........................................................................................................... 215

Figure 6.7: Demonstration of the disruptive effect of the H-pY-LKTKF1-NH₂ peptide on the preformed GFP-STAT3βtc/cfsM67 DNA complex ........................................................................................................ 216

Figure 6.8: Comparative inhibitory of GFP-STAT3βtc/dsM67 DNA complex using a peptide- or a non-peptide-based inhibitors probes ........................................................................................................ 217

Figure 7.1: Jablonski diagram demonstrating electronic transition in FRET ................ 226

Figure 7.2: Schematic representation of the basic principles of FRET .................... 227

Figure 7.3: Disruption of FRET between covalently linked eCFP and eYFP GFP mutants by the recombinant TEV protease ........................................................................................................ 237

Figure 7.4: Demonstrating proteolysis of the His-eCFP-TEV-eYFP GFP fusion protein with the TEV protease .................................................................................................................. 238

Figure 7.5: Typical steady-state FRET efficiency assay ............................................. 240

Figure 7.6: Control curves for the steady-state FRET assay ....................................... 242

Figure 7.7: Demonstration of the effect of NaCl on the efficiency of energy transfer . 243

Figure 7.8: Demonstrate the attractive interactive effect of 5% glycerol in eCFP-uSTAT3βtc/eYFP-uSTAT3βtc protein interaction ................................................................. 244

Figure 7.9: Demonstration of the effect of glycerol on the efficiency of energy transfer between FRET pairs ........................................................................................................ 245

Figure 7.10: Demonstrate the stabilizing effect of Mg²⁺ ion on GFP-uSTAT3βtc protein-protein interaction and therefore FRET efficiency ........................................ 247

Figure 7.11: Demonstrating the effect of the high affinity dsM67 oligonucleotide DNA on GFP-uSTAT3βtc protein interactions and therefore the efficiency of energy transfer 248

Figure 7.12: Demonstrating FRET and the disruption of FRET signals ................. 251

Figure 7.13: Quantitative Steady-state FRET efficiencies between interacting eCFP-uSTAT3βtc and eYFP-uSTAT3βtc FRET pairs ............................................................... 252
List of Tables

Table 1.1: Cytokine specific JAK/STAT activation (data adapted from Schindler and Strehlow, 2000) ...............................................................................................................................36
Table 1.2: Mutation yielding spectral variants of green fluorescence protein............. 48
Table 2.1: Oligonucleotide primers (forward and reverse) used to develop expression constructs..........................................................................................................................................55
Table 2.2: Phosphotyrosyl peptide used in the functional assays............................... 57
Table 2.3: Characteristic properties of the expression vectors used in the study ...... 58
Table 2.4: Summary of the general parameters used in PCR-amplification................. 69
Table 4.1: Mean distance travelled by the standard protein markers and the purified, GFP-STAT3\textsuperscript{ptc} ..............................................................................................................................158
Table 4.2: Relative migrations of the standard protein markers and the GFP-STAT3\textsuperscript{ptc} isoforms expressed as 100[log (R_m * 100)] ........................................................................ 158
Table 4.3: Relative elution volumes (V_e) and their corresponding gel-phase distribution coefficients...................................................................................................................... 161

Edwin Nkansah
ACKNOWLEDGEMENTS

A significant number of pleasant individuals have contributed so much to make my time at SoP, and this thesis a success. I will begin by expressing my sincere gratitude to the Government of Ghana for appreciating the importance of this research project to the emerging research and development sector in Ghana and for funding the Ph.D. programme. I am also very grateful for the enormous material and developmental support I have received from the Cancer Research UK Protein-Protein Interaction Drug Discovery Research Group led by Prof. D.E. Thurston at The Department of Pharmaceutical and Biological Chemistry.

I am also deeply indebted to my supervisor Dr. A.F. Wilderspin for all the support and guidance over the period of my Ph.D. training. My journey with Andy has been special and fulfilling. Thank you.

Once again, I would like to express my sincere gratitude to Prof. D.E. Thurston for the kindness, support and encouragement over the period of my training.

I would especially like to thank Prof. F.A. Stephenson for her encouragement, teaching and support over the last 3½ years.

I also wish to thank Dr. G. Zinzalla and Dr. C. James for their scientific contribution to the project. Special thanks go to Mrs Asamoah and Auntie Grace for their support and guidance during my stay in the United Kingdom.

I am also extremely grateful to Dr. Emma Sharp and Mrs Samantha Essex for their enormous support in proofreading the thesis. Further, I would like to thank my friends and my colleagues including Jonathan Palmer, Chris Chamberlain, Hiroko Kachi, Lein, Sampson, Christina, Sibylle, Paul, Andreia, Tom, Rowena, Ashna, Isa and Ines. I would also like to extend my gratitude to Mr. Derek Marley, Dr. D. West, Dr. S. Cousins, Dr. M. Rahman and Dr. R. Smyth to their support and friendship. Also, many thanks go to the many project students I worked with in Lab 429.

On a special note, I owe my family including; my siblings as well as Ms Susana Owusu Appiah, Christina Asabere, Mr. E.Y. Nkansah, Dr. K.K. Sarpong for their patience, support and love over the period of my training. I would not have made it without your prayer. God Bless you all.
Edwin Nkansah

ABBREVIATIONS

\( \alpha \)                        alpha
\( \text{aa} \)                    amino acid
APS                           ammonium persulphate
bFGF                          basic fibroblast growth factor
BSA                          bovine serum albumin
\( \beta \)                    beta
CBP                         creb-binding factor
CCD                          coil-coiled domain
\( \text{DB}_{50} \)    DNA binding inhibition by 50%
DBD                         DNA binding domain
DNA                          deoxyribonucleic acid
DTT                         dithiothreitol
EDTA                       ethylenediamine tetraacetic acid
EGTA                      [ethylene bis(oxyethylenenitrilo)] tetraacetic acid
EGF                       epidermal growth factor
EGFP                      enhance green fluorescent protein
ELISA                   enzyme linked immunosorbent assay
EMSA                     electrophoretic mobility shift assay
FB                         fibrinogen
FRET                      fluorescence resonance energy transfer
\( \text{g} \)                   gram
GAS                       gamma activated site
GFP                        green fluorescent protein
GM-CSF                  granulocyte/macrophage-colony-stimulating factor
HEPES                 (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
\( \text{h} \)               hour
HPX                       hemopexin
HP                         Haptoglobin
HIF                      hypoxia-inducible factors
hSIE                    high affinity c-fos s/s-inducible element
IFN                interferon
IFNAR                   interferon alpha receptor chain
IL                        interleukin
IPTG                   isopropyl \( \beta \)-D-1-thiogalactopyranoside
IRF                  interferon regulatory factor
ISGF                  interferon-stimulated gene factor
ISRE                  interferon - stimulated response element
JAK                       janus kinase
kDa                       kilo Daltons
MAP                    mitogen-activated protein
mg                       milligram
mL                       millilitre
\( \mu \text{L} \)      microlitre
MMP                       matrix metalloproteinase
MWCO                  molecular weight cut off
PAGE                polyacrylamide gel electrophoresis
\( \text{pH} \)                  log [\( \text{H}^+ \)]
PEMSA              protein electrophoretic mobility shift assay
PCR                       polymerase chain reaction
PDGF                    platelet-derived growth factor
PEI                      polyethylene imine
PIAS                   protein inhibitors of activated STATs
PMSF                phenylmethylsulphonyl fluoride
S                        seconds
SAA  serum amyloid A
SAP  serum amyloid P-component
SDS  sodium dodecylsulphate
SH2  src homology 2
SHP2  SH2-domain-containing tyrosine phosphatase
SOSC  suppressor of cytokine signalling
STAT  signal transducer and activator of transcription
TBE  tris borate EDTA
TAD  transcriptional activation domain
TEMED  tetramethylethylenediamine
TRPSI  trichorhinophalangeal syndrome I
VEGF  vascular endothelia growth factor
1 GENERAL INTRODUCTION

1.1 Cancer

Cancer development involves dynamic changes in the genome. The disease develops when one or more cells lose their ability to regulate cell division and this results in out-of-control cell proliferation and homeostasis. The foundation of cancer can be traced to the genetic material of the cell and is a result of the accumulation of unwanted mutations in one or more genes that encourages clonal selection of cells with aggressive phenotypes, as a function of time, leading to invasive malignancy (Bartek and Lukas, 2001). This is underlined by alterations in the morphology of the cancer cell which is driven by two well known oncogenic categories; proto-oncogenes (dominant gain of function) which function as accelerators to activate the cell cycle; and tumour suppressor genes (recessive loss of function) that serve as brakes to slow cell growth (Bishop and Weinberg, 1996). The activity of these is manifested in all forms of cancer, and subtypes of tumours found in a specific organ. The genotype of all cancer cells is a manifestation of cardinal alterations in cell physiology that converged to direct malignant growth. Typically, a cancer cell exhibits features that include; loss of normal signal to stop proliferation, loss of signals for differentiation, sustained cell division, tissue invasion and metastasis, evasion of programmed cell death (apoptosis) and sustained angiogenesis. Each of these alterations is shared in common by most, if not all types of human tumours and therefore provides tumour cells with the capabilities to adapt and successfully breach anticancer defence mechanisms (Hanahan and Weinberg, 2000).

A total of 17 intracellular and intercellular highly conserved signal transduction pathways have been discovered, many of which are activated during embryogenesis and fetogenesis, as the equilibrium between cell division and apoptosis is designed to suit the developing organism. Significantly, all oncogenic proteins participate in cellular functions that involve signal transduction from the extracellular ligands through the membrane receptors into the cytoplasm and then towards the nucleus, where transcription is initiated to express proteins that directly contribute to the oncogenic phenotype (Hanahan and Weinberg, 2000). Each signalling pathway is made up of signal receivers (receptors), signal transmitters (membrane and/ or cytoplasmic proteins including kinases and phosphatase) and up- or down regulators of gene expression (transcription factors) (Blume-Jensen and Hunter, 2001) all of which operate by the principle of Protein-Protein Interaction (PPI).
1.1.1 Protein-Protein Interaction

Complex and advanced biological systems operate through efficiently controlled mechanisms of communications between the individual entities involved. To sustain this network of communication; cell-cell interactions, secreted factors, and extracellular matrix elements are crucial. Cells can regulate their population and sustain proper organ function by reacting to extracellular signals with proliferation, differentiation and apoptosis. This inherent regulatory instinct is impaired or lost during multiple genetic modifications such as point mutations, insertions and deletions through to chromosomal translocations and amplifications of the tumour cell’s DNA (Strausberg et al., 2004; Beerenwinkel et al., 2007; Boland et al., 2005). Normally, cells will respond to signals from the environment and cell surface receptor. The signals are subsequently transmitted from the membrane to the nucleus by intracellular signalling molecules to induce responsive genes. This regulated signalling paradigm is not observed in cells with no functional regulatory component. In such instances, a sustained signal transduction may be observed even in situations where the activating signal may not be present. As a result, cells may proliferate uncontrollably, resist apoptosis, induce vasculature formation and metastasize to invade distant organs. Since the expression and roles of such regulatory intracellular signalling entities are tightly controlled, it allows for them to be treated as a desirable target for therapeutic intervention (Bild et al., 2006). STAT, a family of transcription factors have gained popularity as an important signalling point for therapeutic intervention.

1.1.2 Signal Transducers and Activators of Transcription (STAT)

Research into the regulations of genes downstream of type I interferon (IFN; an antiviral and growth suppressing cytokine) led to the discovery of the STAT family of transcription factors. STAT1 and STAT2 were discovered as part of a protein complex called Interferon Stimulated Gene Factor 3 (ISGF3), purified from IFNa-stimulated HeLa cell nuclear extract (Darnell, 1994; Lackmann et al., 1998). Extensive research following the discovery of STAT1 and 2 in the interferon system has led to the discovery and identification of seven other mammalian and non mammalian forms, shown to play major roles in cell growth and development. STAT homologues have also been identified in other species including invertebrates such as; Drosophila, D-STAT/STAT92E or Marella (Hou et al., 1996; Yan et al., 1996); Anopheles gambiae (African malaria mosquito), Ag-STAT (Barillas-Mury et al., 1999) and lower eukaryotes such as slime mold, Dictyostelium discoideum (slime mould), Dd-STATa (Kawata et al., 1997); Caenorhabditis elegans, (round worm) ce-STAT-a and ce-STAT-b (Liu et al., 1999); Danio rerio (zebrafish), STAT1 and STAT3 (Oates et al., 1999); Xenopus laevis
Edwin Nkansah

(frog), STAT1 and STAT3 (Pascal et al., 2001). In addition, STAT genes from Bos Taurus (cow) and Sus scrofa (pig) have also been identified.

In vertebrates, cytokines such as Interferons (IFN) and Interleukins (IL) play a pivotal role in regulating and coordinating intercellular communications including immune responses (Leaman et al., 1996). They direct pleiotropic cellular responses by interacting with specific transmembrane receptors which, in turn activate intracellular signal transduction pathways, leading to the induction of gene expression. A significant number of transcription factors including the STAT family of transcription factors transmit signals in response to cytokines (Ihle, 1996; Darnell et al., 1997; O’Shea et al., 1997). As previously stated, the family in mammals consists of seven members named STAT1, STAT2, STAT3, STAT4, STAT5a, STAT5b and STAT6, encoded by different genes. All seven members are organized on three different chromosomes; STAT1 and STAT4 on chromosome 2; STAT2 and STAT6 on chromosome 12; STAT3, STAT5a and STAT5b on chromosome 17 (Copeland et al., 1995; Heim et al., 2003).

The initial revelation of a group of related family of transcription factors, each stimulated by a different cytokine receptor, suggested that the discovered STAT transcription factors would also demonstrate traits identical to the predicted requirements of carriers of extracellular signalling information, i.e., the retention of the inherent specificity of cytokine-receptor interactions (Levy et al., 1990). Although in vitro studies suggested that some of the earlier assertions were not altogether true, in vivo gene-targeting approaches via genetically modified mice, suggested that the various members of the STAT family demonstrated a high degree of specificity in their individual signalling pathways. This was apparent when genetically modified mice, with “knockout” of one or more STAT genes, demonstrated a relatively clear and distinct phenotype, thus assigning almost all the members of the STAT family to a relatively distinct pathway, but not so with STAT3 (Takeda et al., 1996).

Though STATs were originally known to mediate signals induced by IFNa/β/γ, further studies have revealed their role in mediating many different ligand-induced responses initiated by cytokines, growth factors, G-protein and oncogenic tyrosine kinase signalling (Bromberg, 2001; Schindler, 2002). In addition to their diverse biological functions including; roles in cell proliferation, differentiation, apoptosis, inflammation and oncogenesis (Calo et al., 2003; Yu et al., 2004), other functional studies have implicated STATs, especially STAT3 in angiogenesis under physiological and pathological conditions (Valdermbri et al., 2002; Osugi et al., 2002; Nui et al., 2002; Wei et al., 2003). Furthermore, based on the identity of the STAT protein as well as the
enhancer sequence it binds to, some members of the STAT family of transcription factors may act as either suppressor of growth (Kaplan et al., 1998) or promoter of growth (Catlett-Falcone et al., 1999).

1.1.2.1 STAT family and evolution

Chromosomal distributions coupled with the discovery of STAT in some primitive eukaryotes suggested that the STAT family of transcription factors originated from a single primordial gene. As eukaryotes evolve, the increasing need for cell to cell communication grows, hence, the observed duplicate of the gene locus. To complement this evolutionary pattern, STAT genes are structurally conserved and encode STAT proteins ranging in molecular weight from 90 to 115 kDa, i.e., 700 to 850 amino acids long, and 20 to 50 % pair-wise amino acid or nucleic acid sequence identity.

1.1.2.2 Isoforms of STAT

Structural and functional analysis of STAT’s has identified conserved domains shared by all seven isoforms (Chen et al., 1998; Vinkemeier et al., 1998). Structural domains associated with either signal transduction or gene regulation is common in all the STAT proteins (Horvath et al., 2000). Figure 1.1 is a schematic representation of the various domains shared by all seven STAT isoforms.

By convention, the conserved domains begin with the amino-terminal domain (NTD; ~ amino acids 1 to 125). The domain is independently folded and can easily be cleaved from the full length protein by controlled proteolysis (Vinkemeier et al., 1996). The structure of this isolated domain has been determined for STAT4 (Vinkemeier et al., 1998). Studies have implicated the NTD in cooperative DNA binding to tandem Gamma Activation Sites (GAS) elements, receptor association and nuclear translocation (Horvath et al., 2000; Shuai, 2000; Kisseleva et al., 2002; Strehlow et al., 1998).
The adjacent coiled-coil domain (CCD; ~ amino acids 125 to 325), is a flexible polypeptide sequence folded into four long helices that protrudes about 80 Å from the core structure. It provides a large hydrophobic surface area which allows interactions with each other, as well as mediates specific interactions with other helical proteins including c-Jun and StIP (Horvath et al., 2000). Functional studies have also shown that the CCD plays a significant role in receptor binding, tyrosine phosphorylation and nuclear export (Begitt et al., 2000; Zhang et al., 2000). The DNA binding domain (DBD; ~ amino acids 325 to 475) is made up of a beta sheet motif connected by unstructured loops. The STAT dimer identifies and binds to the proximal segment of the GAS element with nanomolar avidity. Since the DNA binding domain may slightly alter its conformation before and after activation of the STAT proteins, it is acceptable to assume that the domain may have other hidden functions (McBride et al., 2000). The adjacent linker domain (LD; ~ amino acids 475 to 575) connects the DNA-binding domain with the SH2 domain and it is made up of five helices. Mutational analyses have shown that the linker domain of STAT1 protein is involved in transcriptional regulation (Yang et al., 1999). Also, it is known to modulate the critical conformation between the DNA binding domain and the dimerization domain. The SH2 or src-homology 2 domain (SH2; ~ amino acids 575 to 680) is the most highly conserved domain in STATs. The SH2 phosphotyrosyl binding domain in STATs makes it a unique transcription factor. It recognises and binds to a specific phosphotyrosyl motif (pY) to play a significant role in signalling. It operates by directing the recruitment of monomeric STATs to the intracytoplasmic portion of the cytokine receptor, activation of STAT (Gupta et al., 1996) and homo- or heterodimerization of the STATs (Gupta et al., 1996; Barahmand Pour et al., 1998). The crystal structure of the STAT dimer clearly describes the reciprocal interaction between the SH2 domain of one STAT monomer and the phosphorylated tyrosine residue of the other STAT monomer. In addition, it allows STATs to mediate multiple signalling pathways by coupling to a number of proteins involved in signal transduction (Ihle, 2001). Directly adjacent to the SH2 domain is the tyrosine activation domain (residue ~ 700). The carboxyl terminal encodes the transcriptional activation domain (TAD; ~ 700 to 850). It is the most poorly
conserved domain with a greater degree of diversity. As such, it is believed to be an innately unstructured region which becomes structured on binding to its target site. As a result, it is able to mediate specific transcriptional responses. Splice variants have been discovered within this domain in a number of STAT proteins (Schaefer et al., 1995; Caldenhoven et al., 1996; Shuai et al., 1993). Transcriptional activation, however, is regulated by the phosphorylation of serine 727 in both STAT1 and STAT3.

1.1.2.3 Alternative splicing

At the 3' end of the gene, transcripts of STATs 1, 3, 4, 5a and 5b result in the alternative splice variants referred to as the beta (β) isoforms with a truncated TAD domain. The β isoform of STAT1 lacks the 38 C-terminal amino acid residues of the α isoform (Schindler, et al., 1992) whilst the β isoform of STAT3 lacks the 55 C-terminal amino acids of the α isoform, but gains a unique 7 amino acid sequence (Schaefer et al., 1995). In the case of STAT4, the β isoform isolated from the human peripheral blood lymphocytes is 44 amino acids shorter than the α isoform and was generated as a result of the insertion of a 369 bp exon that introduced a stop codon down-stream of the inserted exon (Hoey et al., 2003). The STAT5α and 5β originated from partially completed transcript splicing that resulted in C-terminal truncation of 77 and 80 kDa proteins respectively (Wang et al., 1996). STAT6a, 6b and 6c have been isolated from human embryonic lung fibroblast and represent the three naturally occurring STAT6 variants (Patel, et al., 1998).

As a result of the alternative splicing at the 3' end, the β isoforms of the STAT family lack the phosphoseryl residue (∼ amino acid residue 723), but retain the crucial phosphotyrosyl residue (∼ amino acid residue 705). Though the α and the β isoforms are distinctly different in their activation, transcriptional activities and biological functions, they are both efficiently phosphorylated at the critical tyrosine (∼ 705), homo- or heterodimerize with each other and bind to their promoter sequence. Further, studies have shown that the β isoforms of STAT1 (Shuai et al., 1993), STAT3 (Caldenhoven et al., 1996), STAT5 (Wang et al., 1996) and STAT6c (Patel et al., 1998), following over-expression assumes a dominant-negative role under certain conditions. Furthermore, the β isoforms of STAT3, STAT4 and STAT5 exhibit prolonged activation upon tyrosine phosphorylation (Schaefer et al., 1997; Hoey et al., 2003; Wang et al., 1996). On the basis of its extensive dimer stability, the β isoform of STAT3 is thought to be constitutively tyrosine phosphorylated, hence constitutively active for its DNA binding and gene transcriptional activities (Schaefer et al., 1995; Schaefer et al., 1997). Studies on the α and β isoforms of STAT1, STAT3 and STAT4
have demonstrated they activate their own distinct set of genes (Muller et al., 1993; Maritano et al., 2004; Hoey et al., 2003).

1.1.3 Activation of STAT transcription factors

The STAT family of transcription factors are activated by a large number of cytokines as well as growth factors and some hormones (Bromberg et al., 2001; Schindler et al., 2002). Whereas STAT2, STAT4 and STAT6 are reported to be activated by a limited set of cytokines, STAT1, STAT3, STAT5a and STAT5b are reported to be activated by a mixture of distinct and/or overlapping ligands. STAT1 is activated by IFN-α, β and γ whilst STAT3 is induced by cytokines and growth factors including the IL-6 family members and epidermal growth factor (EGF) (Bromberg et al., 2001). STAT4 and STAT6 are induced by IL-2 and IL-4 respectively, whilst STAT5a and STAT5b are engaged in prolactin and growth hormone signalling respectively.

A number of signalling pathways including the modes of activation have been reported. They include; the classical Janus kinase (JAK)/STAT pathway in cytokine signalling, receptors with intrinsic kinase activities, non-receptor kinases, G-protein coupled receptors and over-expressed adapter proteins such as MEKK1-C and islet-1 Figure 1.2.

1.1.3.1 JAK/STAT Signalling

Perhaps the best researched pathway for STAT activation is through the Janus Kinase, popularly known as the JAK/STAT signalling pathway demonstrated in Figure 1.2. Tyrosine phosphorylation is critical step in the JAK/STAT signalling pathway (Ihle et al., 2001).

JAK is a component part of the intracytoplasmic region of the cytokine receptor including IL-6R and IFNγR. In mammals, JAK family is made up of four members; JAK1, JAK2, JAK3 and JAK4. They play a central role in signalling initiated by cytokines and growth factors (Levy et al., 2002), that leads to the activation STAT proteins. They are associated with the proline-rich motif referred to as the box1/box2 region (Ihle et al., 2001). Upon stimulation by the cognate ligand, the cytokine receptor undergoes a conformational transformation that re-orientates JAK to allow activation via trans-phosphorylation (Remey et al., 1999).
On the basis of the accepted model of STAT activation obtained through mechanistic studies (Shuai et al., 1994), receptor mediated STAT activation begins with the binding of a ligand including IL-6, IFN, EGF, PDGF, etc, to its cognate receptors. Hormones such as angiotensin-II have also been found to activate STAT proteins through the serpentine receptor (Darnell, 1997; Leonard et al., 1998; Heinrich et al., 1998; Williams, 1999). Once stimulated, the cytokine receptors oligomerize into an active conformation (Well, 1996) with the concomitant activation of the receptor-associated tyrosine kinase, Janus kinase (JAKs). The activated JAKs subsequently trans-phosphorylate the receptor at multiple tyrosine residues in the endodomain (Turkson, 2004), and the phosphotyrosyl motif (e.g. pYXXQ in the gp130 receptor for recruiting STAT3) directs the recruitment of latent cytoplasmic STATs via their SH2 domain (Lamb et al., 1998). Once “docked”, the STAT protein is phosphorylated at a specific tyrosine residue on their cytoplasmic tail by the activated JAKs. Following phosphorylation, the active STAT molecule forms either a homo- or heterodimer with another active member of the STAT family through reciprocal intermolecular phosphotyrosine–SH2 domain interaction. The STAT dimers dissociate from the receptor and subsequently diffuse through the cytoplasm (Lillemeier et al., 2001) into the nucleus via the nuclear pore complex after forming a complex with importin-α (Lillemeier et al., 2001). In the nucleus, homodimers including STAT1-STAT1 and STAT3-STAT3 bind to a specific member of the GAS family of enhancer (palindrome, TTTCCNGGAAA) whilst heterodimer such as STAT1-STAT2 form a complex with IRF-9 before binding to a member of the ISRE family of enhancers (direct repeat AGTTTN3TTTCC) (Darnell, 1997; Levy et al., 2002). Tyrosine phosphorylation supported by, serine...
phosphorylation and lysine acetylation mediates the transactivation potentials of STATs (Levy et al., 2002; Yuan et al., 2005).

Following activation of target genes, specific tyrosine phosphatases dephosphorylate the active STAT protein, for example, the phosphatase Tc45 inactivates STAT1 before nucleocytoplasmic shuttling can resume (Meyer et al., 2002, Vinkemeier et al., 2004; Haspel et al., 1996; Doye et al., 1997; ten Hoeve et al., 2002).

However, unlike the cytokine receptors, growth factor receptors including epidermal growth factor (EGF), platelet-derived growth factor (PDGF) and fibroblast growth factor receptors show intrinsic tyrosine kinase activities and therefore activate STATs directly or indirectly by JAKs as demonstrated by EGF signalling in human breast cancer cell lines (Gracia et al., 2001).

1.1.3.2 Dimerization of STAT protein

Upon phosphorylation, active STATs dimerize via a reciprocal interaction between the phosphotyrosyl motif of one STAT monomer and the SH2 domain of the partner dimer. This interaction may allow the STAT dimer to assume a conformation that permits appropriate DNA contact on the dyad symmetrical binding site. The phosphotyrosyl-SH2 domain interaction is a polar interaction mediated principally by Lys591, Arg609, Ser611 and Ser613. This set of amino acid residues are conserved in all SH2 domains of other proteins (Kuriyan et al., 1997). Furthermore, the sequence between residues 702 and 722 including Tyr 705 has strongly been implicated in dimer formation.

1.1.3.3 Nucleocytoplasmic transport of STATs

Nuclear importation is not dependent on cytokine stimulation, since both unphosphorylated STAT (uSTAT) and phosphorylated STAT (pSTAT) proteins are translocated into the nucleus via different pathways. The nuclear pore complex is responsible for regulating nucleocytoplasmic transport of the STAT proteins. Unphosphorylated STAT is translocated into the nuclear by direct association with the nuclear pore complex (nucleoporins). The carrier-free nuclear translocation requires neither energy nor transport factors, but follows a concentration gradient (Marg et al., 2004). In contrast, the nuclear import of pSTAT is energy dependent and the binding of pSTAT to an import factor called importin (importin α and β). The pSTAT-Importin interaction requires the parallel orientation of both the monomers within the STAT dimer (Mertens et al., 2006; Mao et al., 2005). In addition, pSTAT nuclear importation revolves around the asymmetric distribution of the small GTPase Ran at the two regions of the nuclear membrane. The RanGDP/RanGTP gradient across the nuclear membrane results in the asymmetric distribution of RanGTP, which is transported into the nucleus and back out of the nucleus through the nuclear pore complex.
envelope forms the driving force behind the active transport and facilitates the importation of pSTAT into the nucleus against a concentration gradient (McBride et al., 2000).

The nuclear import of pSTAT and export of uSTAT molecules via the export receptor chromosomal region maintenance 1 (CRM1) is dependent on the asymmetric distribution of the RanGTP. The high concentration of RanGTP in the nucleus promotes the disintegration of the import complex and facilitates the binding of CRM1 to uSTAT during nuclear export (McBride et al., 2000).

1.1.3.4 STATs DNA binding sites

The activated STAT dimers bind to a specific DNA sequence; mostly, 9-base-pair (bp) consensus sequence with nanomolar binding constant (Kumar et al., 1997). The ISGF-3 complex, i.e., IRF-9/p48, STAT1 and STAT2 generated downstream of type I IFN signalling recognises and binds to the ISRE enhancer family (a direct repeat, AGTTTN₃TTTCCC) to initiate transcription. In contrast, STAT homodimers mostly formed downstream of type II signalling bind to a member of the GAS family of enhancers which is a palindromic DNA sequence as follows; TTTCCNGGAAA. Biochemical studies involving all the STAT proteins except STAT2 have shown optimum binding affinity for DNA sequences with the TTCN₂-4GAA consensus sequence.

1.1.4 Crystal structures of STATs

The DNA bound active human STAT1 homodimer (amino acids 132-713) has been crystallized and described at 2.9 Å (Chen et al., 1998) whilst the DNA-bound murine STAT3β complex (amino acids 136-716) has also been described at 2.25 Å (Becker et al., 1998). Both STAT1 and STAT3β were prepared without the NTD and TAD fragments, and were co-crystallized with the high affinity 18-mer duplex DNA corresponding to the high affinity c-fos sis-inducible element (hSIE, M67). The DNA binds tightly to both STAT1 and STAT3 dimers with high affinity (Wagner et al., 1990). From the side profile, the crystal structure of both STATs were very similar, thus both resembled a nutcracker, Figure 1.3. The major domains including CCD, DBD, LD and SH2 demonstrated extensive and close inter-domain interactions to generate a contiguous hydrophobic core.
However, some regions common to both STAT1 and STAT3 were disordered and included; the loop connecting the helices α1 and α2 in the CCD, few residues in the N- and C-terminals and the LD. The LD is rich in Pro, Gly and hydrophilic residues that confers the flexibility that permits the phosphotyrosyl tail to bind to the SH2 domain of its dimer partner, and not its own SH2 domain.

On the other hand, the crystal structure of the unphosphorylated STAT1 (amino acids 1-683), STAT3 (amino acids 127-688) and STAT5a (amino acids 128-712) have been resolved independently at 3.0, 3.0 and 3.21 Å respectively (Mao et al., 2005; Ren et al., 2008; Neculai et al., 2005). STAT1 was co-crystallized with a five residue phosphotyrosyl peptide derived from the IFN-γ receptor. The crystal structures of STAT1 and STAT3 revealed that the core fragment (amino acids ~ 130 to ~680) formed a reciprocal dimer also referred to as the anti-parallel dimer, the formation of which involved the CCD and DBD. The observed dimers are considered relatively weak compared to the phosphorylated dimers. This observation correlates with biochemical studies on STAT3, where an interaction between the CCD with the SH2 domain and TAD was observed (Zhong et al., 2005). However, this dimer interface observed for STAT1 and STAT5a was absent in the crystal structure of unphosphorylated STAT3, correlating to studies in solution that suggested that the core fragment of STAT3 was principally monomeric (Ren et al., 2008). Also, as observed in STAT1, there were little conformational changes in the core fragment of STAT3 before and after tyrosine phosphorylation. The core fragment in STAT3 shows a contiguous hydrophobic core that is an integral structure unit, identical to what was observed in the crystal structures of STAT1 and STAT5a (Neculai, et al., 2005).
1.1.5 Inactive STAT dimers

The ability for STATs to undergo nuclear translocation and DNA-binding is attributed to dimer formation (Chen et al., 1998). Although numerous studies support this argument, recent studies have challenged this classical STAT signalling model. Several groups have reported the pre-association of inactive cytoplasmic STAT molecules. Co-immunoprecipitation studies have revealed the existence of STAT3-STAT3 and STAT3-STAT1 complexes in untreated cells, suggesting that STATs dimerize before activation into inactive dimers irrespective of tyrosine phosphorylation (Novak et al., 1998; Haan et al., 2000). Also, evidence exists to support the formation of STAT1 complexes including STAT1 dimers in the cytoplasm of cells even in the absence of cytokine stimulation and tyrosine phosphorylation (Mao et al., 2005; Ten Hoeve et al., 2002). Further, the movement of unphosphorylated STAT1 from the cytoplasm to the nucleus and vice versa is well documented (Meyer et al., 2002; Meyer et al., 2003). Li and colleagues have reported the existence of STAT3 dimers in vitro using biochemical analytical tools (Li et al., 2004). Oxidative stress has also been implicated in STAT3 complex formation. Ndubuisi and colleagues reported that a relatively insignificant proportion of STAT3 molecules existed as monomer whilst the majority exists as a high molecular mass complex (200-400 kDa and 1-2 mDa) in association with other protein molecules including scaffolding proteins (Ndubuisi et al., 1999).

1.1.6 Role of STATs in mediating growth, differentiation and apoptosis

Cell culture experiments, animal experiments and clinical observation of human cancer cells have shown that STATs play a crucial role in regulating cell-cycle progression and apoptosis. These studies have defined and categorised the fundamental roles of all the seven STAT proteins in cell growth and development. STAT1 is essential for suppressing growth and is also considered as a modulator for the tumour suppressing pathway (Bromberg et al., 2000). It is also involved in antiviral and antibacterial responses as well as apoptosis. Mice deficient in STAT2, STAT4 or STAT6, show significant limitations in their activities during IFN-γ signalling and T-cell development (Kaplan et al., 1996; Shimoda et al., 1996; Takeda et al., 1996; Thierfelder et al., 1996). STAT4 is implicated in Th1 (T helper 1) development whereas STAT6 is involved in Th2 (T helper 2) development. Studies have shown that STAT3 and STAT5 are active promoters of tumourigenesis (Bromberg et al., 2000). STAT5 is involved in prolactin and growth hormone signalling.
With this knowledge, the question of how these STAT proteins, specifically STAT3 regulate biological responses is reviewed in this study. The study will extensively address the role of STAT, specifically STAT3 in cell growth and development.

1.1.7 Signal Transducer and Activator of Transcription-3 (STAT3)

The activation of STAT3 by a plethora of cytokines (entire family of IL-6-type cytokines), growth factors, oncogenes, IFN and other stimuli suggested that the activities of STAT3 may be more generally deployed, unlike other members of the STAT family and that STAT3 may represent a primordial STAT protein (Takeda et al., 2000; Hirano et al., 2000). STAT3 also plays a critical regulatory role in numerous cytokine-mediated cellular processes including; acute-phase and immune response, differentiation, proliferation, and cell survival (Bromberg et al., 2001). Interestingly, unlike all the other members of the STAT gene family of transcription factor, ablation of STAT3 results in embryonic lethality (Takeda et al., 1997).

STAT3 was initially describe as a DNA-binding activity in IL-6 stimulated hepatocytes that is capable of binding to the enhancer element in the promoter of acute-phase genes, also referred to as the acute-phase response elements (Akira et al., 1994; Zhong et al., 1994). Structurally, as previously discussed, STAT3 is very similar to all the other STATs. They possess a conserved NTD implicated in dimer-dimer interaction, a DBD that is sequence specific to a palindromic IFN-γ-activated sequence (GAS), SH2 domain that participate in receptor recruitment and dimer formation and a TAD. STAT3α and STAT3β are the two isoforms of STAT3. They are the alternatively spliced variants of the same gene, differing only at their C-termini. The STAT3β was discovered as a c-Jun interacting protein. It is a shorter form of STAT3α, has C-terminal 55 amino acid residue of STAT3α isoform replaced with 7 amino acid residues which are not found in the STAT3α isoform. As stated in the previous section, overexpression of the STAT3β isoform is reported to act as dominant negative factor. In addition, the activation, transcriptional activities and biological functions of the STAT3β isoform is distinctly different from the α isoform. This was evident in human CD34⁺ bone marrow and HL60 cells, where Granulocyte-Colony Stimulating Factor (G-CSF) was only activated by the β isoform but not the α isoform (Chakraborty et al., 1996). Moreover, following tyrosine phosphorylation, extended activation of the β isoform was observed relative to the α isoform. Also, due to the extensive dimer stability of the β isoform relative to the α, they are reported to show constitutive tyrosine 705 phosphorylation, DNA-binding and gene transcriptional activities even in the absence of cytokine stimulations. Further, using genetically manipulated mice, it was demonstrated that the β isoform was capable of stimulating distinct STAT3 target
genes in response to IL-6 ligand, thus indicating that the \( \beta \) isoform may not be a dominant-negative factor of STAT3\( \alpha \) in vivo. The \( \alpha \) isoform has non-redundant functions including regulating cellular responses to IL-6 and modulation of IL-10 function in macrophages (Maritano et al., 2004).

The level of STAT3 expression is regulated by the trichorhinophalangeal syndrome 1 (TRPS1) gene. The gene encodes for a different member of the GATA family of transcription factors (GATA-binding protein recognize a target site conforming to the consensus WGATAR \((W = A \text{ or } T \text{ and } R = A \text{ or } G)\) that binds to and represses the expression of the STAT3 protein. However, in the absence of the TRPS1, elevated levels of STAT3 are observed leading to a decreased proliferation of chondrocytes and skeletal disorder (Suemoto et al., 2007). As in all STATs, STAT3 activation follows the generally accepted STAT activation paradigm describe above. Many ligands including cytokines, growth factors, hormones and oncogenes have been shown to activate STAT3, Table 1.1 (Akira et al., 1994).

Table 1.1: Cytokine specific JAK/STAT activation (data adapted from Schindler and Strehlow, 2000)

<table>
<thead>
<tr>
<th>Ligands</th>
<th>Jak Kinases</th>
<th>STATs</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFN family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFN-( \alpha )/( \beta )/( \omega )/Limitin</td>
<td>Tyk2, Jak1</td>
<td>STAT1, STAT2, (STAT3, STAT4, STAT5</td>
</tr>
<tr>
<td>IFN-( \gamma )</td>
<td>Jak1, Jak2</td>
<td>STAT1, (STAT5)</td>
</tr>
<tr>
<td>IL-10</td>
<td>Tyk2, Jak1</td>
<td>STAT3</td>
</tr>
<tr>
<td>IL-19</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>IL-20</td>
<td>?</td>
<td>STAT3</td>
</tr>
<tr>
<td>IL-22</td>
<td>?</td>
<td>STAT3, (STAT5)</td>
</tr>
<tr>
<td>gp 130 family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL-6</td>
<td></td>
<td>STAT3, STAT1</td>
</tr>
<tr>
<td>IL-11</td>
<td></td>
<td>STAT3, STAT1</td>
</tr>
<tr>
<td>OSM</td>
<td></td>
<td>STAT3, STAT1</td>
</tr>
<tr>
<td>LIF</td>
<td></td>
<td>STAT3, STAT1</td>
</tr>
<tr>
<td>CNTF</td>
<td></td>
<td>STAT3, STAT1</td>
</tr>
<tr>
<td>( yC ) family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL-2</td>
<td>Jak1, Jak3</td>
<td>STAT5, (STAT3)</td>
</tr>
<tr>
<td>IL-7</td>
<td>Jak1, Jak3</td>
<td>STAT5, (STAT3)</td>
</tr>
<tr>
<td>CTSLPY( \delta )</td>
<td>?</td>
<td>STAT5</td>
</tr>
<tr>
<td>IL-9</td>
<td>Jak1, Jak3</td>
<td>STAT5, STAT3</td>
</tr>
<tr>
<td>IL-15</td>
<td>Jak1, Jak3</td>
<td>STAT5, (STAT3)</td>
</tr>
<tr>
<td><strong>IL-3 family</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>IL-3</td>
<td>Jak2</td>
<td>STAT5</td>
</tr>
<tr>
<td>IL-5</td>
<td>Jak2</td>
<td>STAT5</td>
</tr>
<tr>
<td>GM-CSF</td>
<td>Jak2</td>
<td>STAT5</td>
</tr>
</tbody>
</table>

**Single chain family**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EPO</td>
<td>Jak2</td>
<td>STAT5</td>
</tr>
<tr>
<td>GH</td>
<td>Jak2</td>
<td>STAT5 (STAT3)</td>
</tr>
<tr>
<td>PRL</td>
<td>Jak2</td>
<td>STAT5</td>
</tr>
<tr>
<td>TPO</td>
<td>Jak2</td>
<td>STAT5</td>
</tr>
</tbody>
</table>

**Receptor tyrosine kinases**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EGF</td>
<td>(Jak1, Jak2)</td>
<td>STAT1, STAT3, STAT5</td>
</tr>
<tr>
<td>PDGF</td>
<td>(Jak1, Jak2)</td>
<td>STAT1, STAT3</td>
</tr>
<tr>
<td>CSF-1</td>
<td>(Tyk2, Jak1)</td>
<td>STAT1, STAT3, STAT5</td>
</tr>
<tr>
<td>HGF</td>
<td>?</td>
<td>STAT1, STAT3</td>
</tr>
</tbody>
</table>

**G-protein coupled receptor**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1</td>
<td>Jak2</td>
<td>STAT1, STAT2</td>
</tr>
</tbody>
</table>

* mda7 and AK 155 have not yet been functionally characterized.

* There are 12 IFNα bind to related but γc independent receptors.

### 1.1.7.1 STAT3 target genes

STAT3 binds to target genes that regulate vital phases of oncogenesis including proliferation, tumour survival, metastasis, invasion, angiogenesis potential, and evasion of host immune response (Bromberg et al., 2001). Apoptosis suppressing genes such as Mcl-1, Bcl-Xl and Survivin genes are target genes for STAT3 transcriptional activity (Aoki et al., 2003). The expression of c-myc and cyclin D proliferation promoting genes are regulated by STAT3. Moreover, aside from stimulating genes noted for proliferation and survival of tumour cells, STAT3 also acts as the transcriptional activator for gene products that affect the surrounding cells including; MMP, which breaks down extracellular matrix to aid migration and metastasis and VEGF, bFGF, HIF-1α, which are involved in angiogenesis. STAT3 may also regulate the expression of IL-6, IL-10, and TGF-β involved in immunosuppression. Furthermore, STAT3 regulates p53 in a manner that inhibits its ability to suppress tumourgenesis (Haura et al., 2005). STAT3 is also implicated in the suppressed secretion of pro-inflammatory cytokines and chemokines which enhance tumour immune system evasion (Yu et al., 2004).
1.1.8 STAT3 and oncogenesis

STAT3 was originally identified as an acute-phase response factor in the liver, activated by IL-6 (Bromberg et al., 2001; Kishimoto et al., 2005). It is ubiquitously expressed in most tissues. STAT3 is activated in tumour cells and actively promotes their proliferation survival and the production of immunosuppressive factors (Wang et al., 2004). Cells transformed with v-Src gene clearly revealed the contrast between cellular transformation and STAT3 activation (Cao et al., 1996). Fibroblast cells became transformed with a constitutively active STAT3 gene, whilst the dominant-negative STAT3β gene, reversed the transformed phenotype thus leading to the original suggestion that STAT3 is an oncogene (Bromberg et al., 1998; Bromberg et al., 1999).

Numerous mouse and human malignancies have identified constitutively active STAT3, including lymphomas, leukemias, mycoses fungoides, multiple myeloma, brain, breast, prostate, lung, and head and neck cancer (Haura et al., 2005; Grandis et al., 2000). Biochemical analyses including immunohistochemical measurements have led to the discovery of constitutively active STAT3 proteins in breast, prostate and lung tumour samples (Alvarez et al., 2006; Gao et al., 2007). Thus, aberrant STAT3 signalling is a requirement for growth and survival of tumour cells (Garcia et al., 2001; Catlett-Falcon et al., 1999; Grandis et al., 2000). Their constitutively active nature has been attributed to atypical activities of cytokine receptor, growth factor receptors, cytoplasmic tyrosine kinases and other upstream activators associated with STAT3 signalling.

STAT3 operates as an integral modulator of autonomous tumour cell properties as well as regulating the communication between tumour and stromal or immune cells (Yu et al., 2007). It was shown to activate the expression of some principal genes implicated in the regulation of cell cycle progression and apoptosis. T-cells with knocked-out STAT3 gene have been used to demonstrate the crucial role STAT3 plays in anti-apoptotic signalling in response to IL-6 as well as in inflammatory helper T-cell activated by IL-6 and IL-23 (Yang et al., 2007).

Further, STAT3 plays a major role in hair and skin development. Tissue specific deletion of STAT3 gene in keratinocytes resulted in mice developing defects such as hyperkeratosis, and scaling in skin. In breast tissue, however, the absence of STAT3 has resulted in delayed involution and a reduced rate of programmed cell death (Chapman et al., 1999). Furthermore, mice with hepatocytes deficient in STAT3 showed defects in their ability to induce acute-phase response genes (e.g. SAP, FB, HP, SAA and Hpx) upon activation by IL-6 (Alonzi et al., 2001). However, these genes
are induced by corticosteroids. The efficient maintenance of the pluripotent mouse embryonic stem cell also requires STAT3 activation (Aoki et al., 2003).

The role of STAT3 have been extensively investigated in cell culture systems. Cytokine mediated STAT3 activation resulted in a varied response in different cells, including acute-phase response in hepatoma cells, stimulation of proliferation in B lymphocytes, activation of terminal differentiation and growth arrest in monocytes (Heinrich et al., 1998) and the retention of pluripotency of the embryonic stem cells (Niwa et al., 1998; Raz et al., 1999). By revealing the different and distinct roles of STAT3 in the varied cell lines, it suggests that STAT3 is major signal transducer downstream of the gpl30-like receptors.

1.1.9 Regulating STAT3 signalling

STAT plays a crucial role in modulating cellular activities. Thus, several control mechanisms are involved in regulating STAT signalling at various levels of STAT activation. These regulators are generally refered to as positive or negative regulators of STAT.

1.1.9.1 Positive regulation of STAT3 signalling

1.1.9.1.1 Tyrosine phosphorylation

In addition to cytokines and growth factors (Schindler et al., 2000), STATs, including STAT3 are activated by hormones (e.g. insulin, and angiotensin II) and non receptor tyrosine kinase; v-Src, v-Abl). STAT proteins may also be activated by G-proteins (chemokines, TSH) (Leonard and O'Shea, 1998; Marrero et al., 1995).

1.1.9.1.2 Serine phosphorylation

The realisation that STAT proteins can be post-translationally modified through Ser-727 phosphorylation revealed the possibility of cross-talk between signalling cascades (Decker et al., 2000). This was shown through phosphorylation of the PMSP sequence motif in STAT1 and STAT3 (Wen et al., 1995). The suggestion that proline-directed serine kinases (e.g. MAP kinases) (Gonzalez et al., 1991) are responsible for the observed modification is strongly supported by the conserved nature of the PMSP motif at the carboxyl-termini of STAT1, STAT3, STAT4 and STAT5. Studies involving reporter genes have shown that serine phosphorylation promotes transcription activity in STAT1 and STAT3 (Wen et al., 1995; Decker et al., 2000). Moreover, physiological studies carried out in cells originally with no STAT1 but re-engineered to accommodate STAT1 (STAT1 Ser727) revealed that serine phosphorylation solely promoted the
potential of STAT1 to activate the expression of a limited number of target genes (Kovarik et al., 2001).

1.1.9.2 Negative regulation of STAT3 signalling

Many regulatory mechanisms purported to control STAT signalling have been identified at different phases including the receptor phase and the JAK phase of STAT signalling. These are individually detailed.

1.1.9.2.1 Receptors endocytosis

Upstream receptor endocytosis is used extensively to negatively regulate STAT signalling either directly when a ligand binds to its cognate receptor or as a result of receptor phosphorylation after ligand–receptor interaction. The mechanism is regulated by an amino acid motif positioned within the endodomain of the gp130 IL-6 receptor chain. Studies have shown that in some cases, soluble receptors are secreted to compete with the membrane receptors for ligands needed to initiate the STAT signalling pathway (Mitsuyama et al., 2007)

1.1.9.2.2 Target degradation

The (suppressor of cytokine signalling) (SOCS) proteins actively suppress STAT activation by either competing with STAT for the docking sites on the phosphorylated receptor or binding to activated receptor complexes via its SH2 domain thereby inhibiting JAKs activities. They may also target receptors and JAK for ubiquitination. SOCS genes are known to be direct target genes of STATs protein, and thus, form part of a negative feedback chain (Yoshimura et al., 2003).

1.1.9.2.3 Dephosphorylation

Phosphatases such as SHP1 and SHP2 have been shown to reverse the activated state of JAK and STAT5 respectively via dephosphorylation. STAT3 is dephosphorylated and subsequently inactivated by protein tyrosine phosphatase, receptor type, T (PTPRT) phosphatase (Zhang et al., 2007).

Other factors including Grim-19, inhibit STAT3 by preventing it from entering the nucleus to carry out it transcriptional activities. Also, Protein Inhibitor of Activated STAT (PIAS), can also bind to STAT3 and prevent it from binding to the target gene. It may also regulate STAT3 transcriptional activities via small ubiquitin-like modifier (SUMO) ligase activity (Lee et al., 2007).
1.1.9.3 Inhibition of STAT3 in tumour cells

Since STAT3 is a vital player in tumour formation, growth, survival and progression, studies have been done to engineer and develop strategies to inhibit STAT3 signalling. Classes of molecules including RNA interference (RNAi), antisense oligonucleotides, phosphotyrosyl peptides and small molecular weight inhibitor have been analyzed and found to regress tumour cell growth and also promote cell death (Deng et al., 2005; Deng et al., 2007).

1.1.10 STAT3, a target for cancer drug discovery

In considering STAT3 as a valid target for cancer drug discovery, the application of tools in molecular biology as well as the availability of relevant research material including dominant negative and activated mutant variants of STAT3, antisense oligonucleotides, cell culture, animal models, and patient samples provide a significant amount of data to support STAT3 as a target for cancer drug discovery. A study conducted by Bromberg and colleagues, showed that a constitutively active STAT3 mutant is sufficient to induce transformation of cells which subsequently form tumours in vivo (Bromberg et al., 2000).

When constitutive STAT3 signalling is interrupted through pharmacological or genetic intervention, the expression of anti-apoptotic Bcl family members including Bcl-xL, Bcl-2, or Mcl-1 in multiple myelomas (Catlett-Falcone et al., 1999), head and neck squamous cell carcinomas xenograft model (Grandis et al., 2000), mycosis fungoides (Nielsen et al., 1999) and large granular lymphocytes leukemia (Epling-Burnette et al., 2000) are inhibited whilst increasing the expression of pro-apoptotic Bax protein (Nielsen et al., 1999). Also, the use of gene therapy (i.e. a dominant negative STAT3 variant) to inhibit constitutive STAT3 activity in tumour models including murine B16 melanoma resulted in significant tumour regression (Niu et al., 1999). The study provided strong evidence to support the idea that STAT3 is a desirable target for therapeutic intervention. In addition, a number of STAT3 target genes including cyclin-D1, p21, c-myc, and the Bcl-family members have been identified as having crucial roles in modulating cell proliferation and survival (Bromberg et al., 1999b; Chin et al., 1996; Bowman et al., 2000). Furthermore, cells without cyclin-D or c-myc genes do not respond to STAT3 mediated transformation (Bowman et al., 2000a). However, although extensive research has generated a lot of data on the roles and activities of STAT3 in various contexts, there is still a great deal to be learned. The identification of STAT3 as central player in the pathogenesis of varied human cancers makes it a prime target for novel pharmacological intervention.
1.1.10.1 Active STAT3 dimers and its therapeutic relevance

As previously stated, dimerization of STAT3 via reciprocal phosphotyrosine-SH2 interaction in response to cytokine stimulation renders the STAT3 protein competent to perform its biological functions including nuclear translocation and DNA binding (Park et al., 2000; Chen et al., 1998). Thus, mutations to prevent dimerization including substituting the critical tyrosine 705 with phenylalanine resulted in a STAT protein unable to efficiently perform its functions such as dimer formation, nuclear translocation, DNA binding and inducing transcription (Bromberg et al., 1998). Hence, on the basis of Bromberg’s work, blocking dimer formation with drug-like small molecules will provide us with the working tool to target and inhibit STAT3 signalling and subsequent functional activities.

Turkson and colleagues used a STAT3 SH2 domain-binding phosphotyrosyl peptide, PpYLKT (where pY represent the phosphotyrosyl motif), derived from the core of the native C-terminal STAT3 SH2 domain binding sequence to disrupt activated STAT3 dimers (Turkson et al., 2001; Moa et al., 2005). Further, other groups have either employed a tripeptide derivative of the PpYLKT or have used a high affinity phosphotyrosyl peptide, GpLPGTV, derived from the gp130 receptor to target and disrupt STAT3 so as to serve as molecular tool for proof-of-concept testing and to provide the framework for the design of STAT3 inhibitors (Turkson et al., 2001; Schust et al., 2004; Ren, et al., 2003). Numerous studies have demonstrated the relevance of the phosphotyrosyl motif of the SH2 binding region in STATs and STAT3 in particular, to their dimerization and DNA-binding activities (Shuai et al., 1993; Shuai, et al., 1994; Schaefer et al., 1997; Sasse et al., 1997). The phosphotyrosine-based peptidomimetics targeted to the STAT3 SH2 domain destabilize active STAT3 dimers by disrupting the phosphotyrosyl-SH2 domain interaction and therefore promotes heterodimer complexes with STAT3 over STAT3/STAT3 homodimers (Turkson et al., 2000).

In addition, drug-like small molecules including ISS840 (Gunning et al., 2007), STA-21 (Song et al., 2004) and Stattic (Schust et al., 2006) have been designed based on peptidomimetic probes to target and disrupt STAT3 activation and dimerization.

1.1.10.2 Inactive STAT3 dimer and its therapeutic relevance

The controversies surrounding the status, structure, localization, regulation and functions of latent inactive STAT3 have generated tremendous interest from the drug discovery point of view. Unphosphorylated STAT3 (uSTAT3), like all STATs, was initially reported as monomeric (Shuai et al., 1994; Schgal et al., 2000). However, co-
Edwin Nkansah

immunoprecipitation studies have revealed otherwise, suggesting that STAT3 dimerizes prior to their activation (Stancato et al., 1996; Lackmann et al., 1998; Haan et al., 2000; Novak et al., 1998). This was confirmed by Ndubuisi and colleagues. However, they used gel chromatography to analyze cellular extracts and then reported that a relatively small amount of STAT3 in the cytosol existed as monomers whereas majority of the proteins existed as large poorly characterized high molecular weight complexes of STAT3 and other scaffolding proteins ranging from 200-400 kDa (STATOSOME I) and 1-2 mDa (STATOSOME II) (Ndubuisi et al., 1999). Thus, on the basis of Ndubuisi's studies, it was premature to conclude that the co-precipitated STAT3 molecules were a result of direct protein-protein interaction between STAT3 species since the likelihood of multiple STAT3 molecules integrating into a STATOSOME was also feasible.

To make matters clearer, Kretzschmar and colleagues adopted a non-invasive technique and also reported the existence of STAT3 dimers in the absence of tyrosine phosphorylation following visualization of STAT3/STAT3 protein interactions within living cells using FRET. Through BRET and FRET, Schröder and colleagues also demonstrated the existence of pre-associated STAT3 proteins in living mammalian cells. They also suggested that STAT3 molecules existed in the cytoplasm as dimers and multimers that go through conformational changes after activation (Schröder et al., 2003). Though the crystal structures of STAT1 and STAT5a suggest that they existed as dimers independent of phosphorylation, the crystal structure of the unphosphorylated STAT3 (uSTAT3) core fragment lacked the dimer interface observed in uSTAT1 and uSTAT5a, thus suggesting that unlike STAT1 and STAT5a, STAT3 was primarily monomeric (Vinkemeier et al. 1998; Mao et al., 2005; Neculai et al., 2005; Ren et al., 2008). However, solution studies on the full length STAT3α, suggested that the protein existed as a dimer in low-salt buffer and PBS (Ren et al., 2008).

Using biochemical analytical tools including native-PAGE/Ferguson plot, sucrose gradient analysis, size exclusion chromatography, cross-linking, etc, Braunstein and colleagues demonstrated the existence of a stable unphosphorylated STAT3 homodimer in vitro (Braunstein et al., 2003). On the contrary, Becker and colleagues also reported that a slightly shorter construct of the β isoform, STAT3βtc (amino acids 127-722) existed as a monomer by gel filtration whilst the phosphorylated version existed as a dimer (Becker et al., 1998). Further, equilibrium sedimentation studies on the same construct also reported significant elevation of the presence of monomeric STAT3βtc at concentrations ranging from 0.2 to 1 mg/mL, however, full length STAT3
remained dimeric (Braunstein et al., 2003). The structural differences may be due to the NTD of the full length STAT3 that may be engaged in dimerization in a homotypic manner thereby, generating stable dimers of the full length STAT3 (Ota et al., 2004).

Over-expression of STAT3 mutants including Y705F-STAT3 and STAT3-null in human mammary epithelium (hTERT-HME1) and mouse embryonic fibroblast (MEF) cell lines was reported to have activated the transcription of specific genes, an indication that uSTAT3 may promote transcription through pathways different from that of pSTAT3 (Yang et al., 2005). Recent reports have also suggested that latent uSTAT3 may also serve as a microtubule stabilizer by binding to stathmin (microtubule destabilizing protein) and hindering its activities. These observations, though limited, suggest that uSTAT3 may be actively engaged in cellular dynamics and functions.

Thus, the characterization of STAT3 dimers is critical to appreciate the molecular function of individual STAT3 monomers and to identify and understand the different dimerization pathways that may be desirable targets for drug development and diagnostics (Davidov et al., 2003). To identify and distinguish between dimers of phosphorylated STAT3 (pSTAT3) and unphosphorylated STAT3 (uSTAT3) proteins as well as high molecular mass complexes, a method capable of circumventing some of the limitations of co-precipitation is needed.

1.2 Cell-Free approach for the interrogation of Protein-Protein Interactions

Considering the importance of dimer formation in protein-protein interactions, it is understandable that a significant number of assays have been put together for their study (Fu et al., 2004; Coligan et al., 2005). Analytical applications such as affinity chromatography (Ranter et al., 1974) and co-immunoprecipitation (Horvitz et al., 1974) as well as popular genetic screens including the yeast 2-hybrid system and its derivatives, (Dye et al., 2005), and the newly developed biophysical approaches such as Surface Plasmon Resonance, have been used extensively to characterize protein-protein interactions.

Despite the relatively large amount of work carried out towards characterizing the structural status of STAT3, almost all the techniques used were invasive, i.e., co-immunoprecipitation and gel chromatography. Thus, quantitative in vitro analysis of the STAT3/STAT3 interaction with its antagonist would be crucial to our understanding of the conditions that modulate the interaction and to enable the discovery of small molecules that are capable of interfering with the interaction. As previously described,
any small molecule that binds to the STAT3 SH2 domain has the potential of being the principal tool for the treatment of tumours dependant on constitutively active STAT3 proteins.

Three different fluorescence-based *in vitro* assays were therefore developed to investigate STAT3-STAT3, STAT3-DNA and STAT3-peptide interactions and their antagonists. Existing biochemical analytical techniques were modified including: Enzyme-Linked Immuno-Sorbent Assay (ELISA) which was used to develop a quick and less laborious fluorescence based microtiter plate assays: Electrophoretic Mobility Shift Assay (EMSA) which was used to develop a fluorescence-based Protein Electrophoretic Mobility Shift Assay (PEMSA); Fluorescence Resonance Energy Transfer (FRET) which was developed to study STAT3-STAT3 interactions in solution. The assays are cell-free, homogenous and fluorescence-based and therefore required purified, recombinant, fluorescent STAT3 chimaeric proteins.

### 1.2.1 Enzyme-Linked Immuno-Sorbent Assay (ELISA)

Haan and colleagues were the first group to report the use of ELISA to study STAT3. They applied the assay to determine the functionality of the STAT3 SH2 domain, by using immobilized biotinylated phosphotyrosyl peptides derived from either the gp130 receptor or the core amino acid sequence encompassing Y°® of STAT3 as the capture probe. The immobilized probe was detected with the phosphotyrosine antibody (pY20). This meant that the relative decrease in absorbance with increasing STAT3 SH2 domain concentration compared with the control, *i.e.*, SH2-free sample, suggested that the inhibition of the antibody-phosphotyrosyl peptide interaction by the SH2 domain of STAT3 was an indication of STAT3 SH2/phosphotyrosyl peptide interaction (Haan *et al.*, 1999). Also, through ELISA, Dourlat and colleagues evaluated the DNA-binding activity of the active STAT3 (pSTAT3). They applied either a phosphotyrosyl or a non phosphotyrosyl peptide to disrupt active STAT3 dimers and therefore prevent the protein from binding to the immobilized oligonucleotide sequence containing the consensus (5'-TTCCGGAA-3') binding site for STAT3 (Dourlat *et al.*, 2009).

### 1.2.2 Electrophoretic Mobility Shift Assay (EMSA)

EMSA is the most sensitive means of detecting active STAT3 proteins. Turkson *et al.* and Ren *et al* applied EMSA to identify small molecules that target and bind to the STAT3 SH2 domain. Through EMSA, Turkson and colleagues demonstrated that the phosphotyrosyl peptide derived from core amino acid sequence encompassing Y°® of STAT3 could inhibit phosphorylated STAT3 dimerization and DNA binding (Turkson *et al.*, 2001). Ren and colleagues on the other hand demonstrated with EMSA the inhibitory
activity of the phosphotyrosyl peptides derived from the gp130 receptor (Ren et al., 2003). Other research groups have also used EMSA to screen drug-like small molecules for their ability to target the STAT3 SH2 domain, block dimerization and therefore DNA binding activity (Gunning et al., 2007; Song et al., 2005; Schust et al., 2008; Zhang et al., 2010).

1.2.3 Fluorescence Resonance Energy Transfer (FRET)

While a few groups have applied a RET-based technique (e.g. FRET and BRET) to study STAT3-STAT3 interactions in living cell, (Kretzschmar et al., 2004; Schröder et al., 2004), there is no published report on a homogeneous assay based on either steady-state or time-resolved STAT3/STAT3 FRET analysis. Though Seidel and colleagues reported on a cell free assay for the JAK/STAT pathway, it was not detailed enough for evaluation (Seidel et al., 2000).

1.3 Fluorescent STAT3 constructs

Investigating STAT3/STAT3 interactions and DNA binding in an isolated cell-free system and to understand the conditions necessary to mediate such interactions as well as to facilitate the discovery of small molecules capable of interfering with the interactions, GFP-STAT3ptc construct in which the NTD (amino acids residues 1 to 126) of STAT3 was replaced with spectral variants of the GFP (Green Fluorescent Protein) to act as a reporter was designed. Figure 1.4 (A) and (B).

Figure 1.4: Schematic representation of the GFP-STAT3ptc construct used in this study (A), overview of the STAT3ptc (Ai) and GFP-STAT3ptc (Aii) used. (B), GFP indicating the point of fusion of the two proteins is shown in green, the N-terminal coiled-coil domain (CCD) is shown in magenta, the DNA binding domain (DBD) is shown in red, and the linker domain (LD) in in gold, the SH2 domain (SH2) is in blue and the C-terminal domain (TAD) is in yellow. The double stranded DNA is shown in cyan. Structure adapted from PDB; STAT3p/1BG1 and GFP/1EMA.
1.3.1 STAT3βtc

As previously discussed, section 1.1.2.3, the STAT3β (amino acids residues 1 to 722) is a naturally occurring splice variant of the full length STAT3α and was originally isolated following screening of an eosinophil cDNA library (Caldenoven et al., 1996). The differential splicing in the STAT family of transcription factors is not uncommon, section 1.1.2.3. Like the STAT3 gene, the STAT1 gene encodes two different proteins, the STAT1α and the STAT1β both of which are a result of alternative splicing (Schindler et al., 1992). Due to the alternative splicing event in STAT3β, the CTD’s 55 amino acid residues of STAT3α is substituted for a unique 7 amino acid residues (Schaefer et al., 1995). Constructs lacking the unstructured CTD are less likely to aggregate and facilitate purification.

In vivo following IL-5 stimulation, STAT3β is activated through Y705 phosphorylation, dimerizes and translocates into the nucleus where it binds efficiently to the palindromic IL-6/IFN-gamma response element (pIRE) from the ICAM-1 promoter, but unlike the STAT3α isoform, it fails to initiate transcription in the IL-5 stimulated cell due to the absent CTD (Caldenhoven et al., 1996). However, in vitro and in vivo studies have reported a marked quantitative difference in the DNA binding activities of the phosphorylated version of the two isoforms of STAT3 (Park et al., 1996; Schaefer et al., 1997). The active STAT3β isoform has a significantly higher affinity for DNA than the STAT3α isoform, a quality conferred by the more stable dimer formed by the active STAT3β protein molecules (Park et al., 2000), thus explaining the observed dominant negative effect of the STAT3β isoform over the STAT3α isoform in transfected COS cells (Caldenhoven et al., 1996).

In recognising that the active STAT3β was a stable dimer with a significantly higher affinity for DNA, Becker and colleagues co-crystallized the N-terminally truncated fragment of the active murine STAT3β (amino acid residues 127 to 722) dimer in complex DNA (Chen et al., 1998).

1.3.2 Green Fluorescent Protein (GFP)

The Green Fluorescent Protein (GFP) reporter used in this study is an important reporter molecule isolated from the Pacific Northwest jellyfish Aequorea Victoria in 1961 (Shimomura et al., 1962). The sequence remained unknown until 1992 when it was cloned (Prasher et al., 1992).

GFP has become an invaluable self-assembling, biologically compatible reporter molecule in developmental and molecular biology. It was first applied to highlight
Edwin Nkansah

sensory neurons in nematodes (Chalfie et al., 1994). However, recent applications include gene expression studies, protein localization, intracellular dynamics of proteins and organelles and biological screening. Further, GFP's have particularly been used as a non-toxic reporter molecule within organisms.

A sizable number of spectral variants of the GFP protein have been engineered by modifying the wild-type GFP, Table 1.2, these emit in the blue (BFP), cyan (CFP) and yellow (YFP) regions (Heim et al., 1994; Ormø et al., 1996; Tsien et al., 1998). The modifications have led to the development of an excellent range of reporter molecules that are capable of providing superior monitoring far exceeding that provided by fluorescent dye reporters including cyanine's Cy2, Cy3 and Cy5, thus, greatly enhancing fluorescence-based applications in vitro and in vivo. Remington suggested basic principles that could result in the alteration of the emitting colour (Remington, 2006). Altering the immediate surroundings of the chromophore including the position of charged amino acid residues, hydrogen bonding network and hydrophobic interaction within the GFP complex resulted in blue or red spectral shifts by as much as 40 nm, evident in the absorption and emission maxima. Such covalent structural alterations and the degree of π-orbital conjugation of the chromophore generate the spectral shift necessary to distinguish the common fluorescent protein classes, i.e., CFP, GFP and YFP, Figure 1.5 (B).

**Table 1.2: Mutation yielding spectral variants of green fluorescence protein**

<table>
<thead>
<tr>
<th>GFP variants</th>
<th>Mutations relative to wild-type GFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>eBFP*</td>
<td>F64L, S65T, Y66H, Y145F</td>
</tr>
<tr>
<td>eCFP*</td>
<td>F64L, S65T, Y66W, N149I, M153T, V163A</td>
</tr>
<tr>
<td>mCFP*</td>
<td>F64L, S65T, Y66W, N149I, M153T, V163A, A206K</td>
</tr>
<tr>
<td>eGFP*</td>
<td>F64, S65T</td>
</tr>
<tr>
<td>Emerald*</td>
<td>F64, S65T, S72A, N149K, M153T, I167T</td>
</tr>
<tr>
<td>eYFP*</td>
<td>S65G, V68L, S72A, T203K</td>
</tr>
<tr>
<td>mYFP*</td>
<td>S65G, V68L, S72A, T203Y, A206K</td>
</tr>
</tbody>
</table>

*Some clones of Aequorea fluorescent protein contains additional mutations believe to be neutral, such as K26R, Q80R, H231L, etc variant.

*Many GFP variant contain V inserted after Met1 so that the mRNA should contain an ideal translational start sequence. We number such a V as 1a to preserve the wild-type numbering for the rest of the sequence.
1.3.2.1 Properties of GFP

GFPs are versatile reporters (Chalfie et al., 1994), belonging to a group of evolutionally unrelated species. They are made up of 238-amino acid residues which are predominantly hydrophobic. The crystal structure of GFP reveals a rigid β-barrel made up of 11 β-sheets surrounding the α-helical region containing the chromophore. Figure 1.5 (A). The chromophore is generated from the spontaneous cyclization and oxidation of the tripeptide sequence -S65 (or T65)-Y66-G67-. However, chromophore formation and fluorescence emission is dependent on the native protein fold (Ormö et al., 1996). It is efficiently enclosed and protected from the external environmental by the β-barrel. The emitted fluorescence is generated from this region.

In addition to its remarkable chemical and photochemical stability including resistance to mild denaturing and reducing agents, bases and high temperatures, GFPs have the added feature of re-naturation following exposure to harsh conditions. The fluorescence intensity, however, decreases sharply with lowering pH, i.e., less than pH 6.0, though, the original fluorescent intensity is retained at pH above 6.0 (Patterson et al., 1997). Further, Ghoeh and colleagues have reported that GFPs is highly resistant to proteolysis, but could be cleaved under specific conditions and subsequently re-ligated to regenerate the same amount of fluorescence (Magliery et al., 2006).

Figure 1.5: GFP structural features including approximate dimensions and mutations present in spectral variants.

(A), Aequorea GFP β-barrel structure showing the approximate dimensions and the chromophore of the Aequorea victoria jellyfish GFP derivatives. (B), map of the Aequorea victoria jellyfish GFP showing the β-sheets, the central helix and the α-helices bearing mutations superimposed on a topological layout of the peptide structure. Note that the mutations are colour-coded to represent the GFP spectral variant they represent; BFP’s (blue), CFP’s (cyan), GFP’s (green), YFP’s (yellow) and Sapphire (violet). Structure (A) was adapted from PDB/1EMA (Ormö et al., 1996) whilst structure (B) was adapted from Shaner et al., 2007.
1.3.2.2 Activation of GFP

When expressed endogenously, wild-type GFP is photo-activated by a calcium-binding photoprotein called aequorin (Johnson et al., 1978). Aequorin consists of apoaquorin (apoprotein), molecular oxygen and coelenterazine (Kojima et al., 1997). Upon mechanical or electrical stimulation, the level of calcium ions (Ca²⁺) increases within the photogenic organs of the jellyfish where GFP and aequorin coexist (Johnson et al., 1978). The calcium ion binds to aequorin and activates it via oxidation. In vitro, activated aequorin emit blue light (λmax 470 nm), however, in a live jellyfish a resonance energy transfer takes place from the activated aequorin (BFP) to the GFP chromophore. The result is an excited GFP molecule from which green light (λmax 508 nm) is emitted (Morise et al., 1974; Morin et al., 1971). The eGFP fluorophore can easily be excited with argon ion laser and detected with fluorescein (Fluorescein Isothiocyanate, FITC) filter sets respectively.

1.3.2.3 Modification and applications of GFP

To improve and enhance monitoring and detection capabilities of exogenous wild-type GFP, several spectral variants of GFP have been developed, Table 1.2. The enhanced Green Fluorescent Protein (eGFP) is one of the most popular variants of GFP. It was developed through site directed mutagenesis. Two major mutations in the chromophore region of the wild-type GFP, i.e., F64L and S65T, resulted in a red-shifted GFP (eGFP) with a spectral excitation peak at 489 nm. This together with other silent base changes has lead to improved translation, chromophore formation and folding in mammalian cells.

The advent of red-shifted GFP variants has encouraged further modifications within the chromophore regions of red-shifted GFPSs. These mutations have successfully resulted in colour shifts into the yellow and blue region of the visible spectrum, Figure 1.6 (A), (B) (Hebshi et al., 2007).
The creation of the Yellow Fluorescence Protein was (YFP) was reported in 1996 by Ormø and colleagues and emits a yellowish green fluorescence. Human codon optimization and other mutations in the chromophore region have led to the development of variants with higher fluorescence intensity in mammalian cells. Other GFP red-shifted variants including eGFP have been engineered by altering some amino acid residues in the chromophore region. Cyan Fluorescence Protein (GFP) was also engineered to address limitations including rapid photobleaching and unstable BFP chromophore (Heim et al., 1994; Heim et al., 1996). Mutation, Y66W, shifted the excitation peak out of the near-ultraviolet region and the spectral emission peak from blue to cyan. Unlike BFP, CFP exhibits enhanced brightness and chromophore stability and has been used in FRET experiment with YFPs and GFPs. CFP and YFP have been identified as the ideal FRET pair for labelling proteins to study the dynamics of protein trafficking in cell and protein-protein interactions. Unlike other luminescent or chromogenic reporter molecules with higher sensitivity, modified GFP's and their spectral variants possess several features that make them invaluable reporter molecules. GFP needs no cofactor or substrate to fluoresce, hence can be monitored noninvasively. Furthermore, the relative nontoxic nature of GFP allows expression in numerous cell lines (mammalian and bacteria) without disturbing cellular activities.

1.4 Aims of the thesis

As has been addressed in this chapter, active STAT3 transcription factor targets genes that include regulators of critical stages in tumour development such as proliferation and survival of tumour cells, invasion, migration, angiogenesis and evasion of the host immune system. Therefore, the aims of this thesis were:
To design, purify and structurally characterize fluorescent STAT3 chimaeric proteins containing spectral variants of GFP fused to the N-terminally truncated beta isoform of the full length STAT3.

To use fluorescence-based isolated cell-free systems to study dimerisation and DNA binding of STAT3βtc in order to understand the conditions that mediate these interactions.

To develop assays to enhance the discovery of small molecules capable of interfering with these interactions.
2 MATERIALS AND GENERAL METHODS

Materials including biological, chemicals and equipments as well as some general protein and nucleic acid methods are describe in this chapter. Specific methodologies are described in the relevant chapters.

2.1 MATERIALS

2.1.1 Molecular biology reagents

Falcon polypropylene conical bottom tube were purchased from BD Bioscience, 5, 10 and 25 mL serological pipettes as well as pasteur pipettes, universal 20, 200, and 1000 μL micro pipette tips and Gel-loading pipette tips were purchased from Greiner Bio-One. Restriction endonucleases including; EcoRI and SphI and T4 DNA ligase, *Pfu* DNA polymerase, SI nuclease and their buffers and PCR nucleotide mix (dATP, dCTP, dGTP and dTTP) were purchased from Promega (Southampton, UK). Restriction endonucleases including; BamHI, Ncol, Ndel and Nhel as well as Antarctic Phosphatase, 1 kb DNA marker and Pre-stained Broad range protein marker were purchase from New England Biolabs (Hertfordshire, UK). Molecular biology kits for plasmid Miniprep and Maxiprep, PCR cleanup and Gel extraction were purchased from QIAGEN (West Sussex, UK). Isopropyl-β-D-galactopyranoside (IPTG), ampicillin sodium salt and Dithiothreitol (DTT) were purchased from Melford Laboratories (Suffolk, UK). Luria Bertani Broth (LB Broth), Agar powder, Casamino acid, organic solvents (methanol, ethanol and propan-2-ol) were purchase from Fisher Scientific (Leicestershire, UK). Kanamycin, tetracycline, chloramphenicol, ethidium bromide, Igepal, Triton X100, agarose powder and 3-Indoleacrylic acid were purchased from Sigma-Aldrich (Poole, UK). D (+) glucose was purchased from WWR international Ltd (Dorset, UK). Bradford reagent was purchased from Bio-Rad (Hertfordshire, UK).

2.1.2 Other chemical reagents and materials

NuSep coomassie brilliant blue G-250 dye and Instant brilliant blue were purchased from Genron (Berkshire, UK). Superose™ 6 (10/300 GL), superpose™ 12, agarose chromatography column, HiTrap™ QFF anion exchange column and Gel filtration calibration kits (HMW) were purchased from GE Healthcare (Buckinghamshire, UK). Protogel (30% [w/v] acrylamide: 0.8% [w/v] bis-acrylamide (37:5:1) was purchased from National Diagnostics (Yorkshire, UK). Ammonium persulphate, \( N,N,N',N' \)-tetramethylethlenediamine (TEMED), 5-amino-2,3-dihydro-1,4-phthalazinedione (luminol), trans-4-hydroxycinnamic acid (p-coumaric acid), Tween-20®, boric acid, dimethyl sulphoxide (DMSO), ethylene glycol-bis(2-aminoethylether)-N,N,N',N'
tetraacetic acid (EGTA), ethylenediaminetetraacetic acid (EDTA), 3,3'-diaminobenzidine (DAB), SIGMAFAST™ Protease Inhibitor Tablets and phenylmethanesulphonyl fluoride (PMSF) were purchased from Sigma-Aldrich (Poole, UK). Ammonium acetate, sodium chloride, potassium chloride and sodium hydroxide were purchased from VWR International Ltd (Dorset, UK). Glycerol, glycine, ammonium sulphate and magnesium chloride were purchased from Analar NORMAPUR. TRIS Base [Tris (hydroxymethyl) aminomethane], 4-(2-Hydroxyethyl)piperazine-1-ethanesulphonic acid (HEPES) were purchased from Melford Laboratories (Suffolk, UK). Bovine serum albumin (BSA) was purchased from Fisher scientific (Leicestershire, UK). Phosphate buffer saline (PBS) Tablets was purchased from OXOID (Hampshire, UK). Hybond C and Hybond N+ nitrocellulose membrane were purchased from GE Healthcare (Buckinghamshire, UK). Millex®GP syringe driven filter units (0.22 and 0.45 µm) were purchased from Millipore (Cork, Ireland). Parafilm was purchased from Pechiney plastic packaging. Where appropriate, all other buffer and general reagent including organic solvent were Analar or molecular biology grade and were purchased from commercial sources.

2.1.3 Oligonucleotide primers and probes used in the study

2.1.3.1 Primers used for Polymerase Chain Reaction (PCR)

The oligonucleotide primers used to clone the expression construct pET-32a (+)-GFP-STAT3ptc were synthesized by Eurofins MWG Operon (Ebersberg, Germany). Further, primers used to generate the following expression constructs; pET-28a (+)-GFP, pET-28a (+)-STAT3ptc-GFP, pET-28a (+)-His-STAT3ptc-GFP and pET-28a (+)-eCFP-TEV-eYFP were synthesized by Invitrogen™ (Paisley, UK). Table 2.1 describes the primers used to generate the expression constructs.
### Table 2.1: Oligonucleotide primers (forward and reverse) used to develop expression constructs

<table>
<thead>
<tr>
<th>Expression Construct</th>
<th>Forward Primer (5’ - 3’)</th>
<th>Reverse Primer (5’ - 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET-32a (+)-eCFP-STAT3βtc</td>
<td>dAAAACATATGGTGAGCAAGGGCGAGG</td>
<td>dAAAACATATGGTGAGCAAGGGCGAGG</td>
</tr>
<tr>
<td>pET-32a (+)-eGFP-STAT3βtc</td>
<td>dAAAACATATGCTGAGCAAGGGCGAGG</td>
<td>dAAAACATATGCTGAGCAAGGGCGAGG</td>
</tr>
<tr>
<td>pET-32a (+)-eYFP-STAT3βtc</td>
<td>dAAAACATATGGTGAGCAAGGGCGAGG</td>
<td>dAAAACATATGCTGAGCAAGGGCGAGG</td>
</tr>
<tr>
<td>pET-28a (+)-STAT3βtc-eCFP</td>
<td>dAAAACATATGGGCCAGGCCAACC</td>
<td>dAAAACATATGGGCCAGGCCAACC</td>
</tr>
<tr>
<td>pET-28a (+)-STAT3βtc-eGFP</td>
<td>dAAAACATATGGGCCAGGCCAACC</td>
<td>dAAAACATATGGGCCAGGCCAACC</td>
</tr>
<tr>
<td>pET-28a (+)-STAT3βtc-eYFP</td>
<td>dAAAACATATGGGCCAGGCCAACC</td>
<td>dAAAACATATGGGCCAGGCCAACC</td>
</tr>
</tbody>
</table>

Note: Restriction endonuclease sites underlined.

In generating the N-terminus GFP tagged STAT3βtc construct, i.e., pET-32a (+)-eCFP-STAT3βtc, pET-32a (+)-eGFP-STAT3βtc and pET-32a (+)-eYFP-STAT3βtc we used forward and reverse primers containing just the Ndel restriction endonuclease site (5’-CATATG-3’). Further, the primers used to generate the pET-28a (+)-GFP expression constructs contained the Ndel and the Nhel (5’-GCTAGC-3’) restriction endonuclease sites in the forward and reverse primers respectively.

In generating the C-terminus GFP tagged STAT3βtc constructs, pET-28a (+) - STAT3βtc-GFP, the STAT3βtc cDNA was PCR-amplified with forward and reverse
primers containing the NcoI and NdeI restriction endonuclease sites respectively and then cloned into pET-28a (+) -GFP cloning vector. In cloning pET-28a (+) - His-STAT3βtc-GFP expression constructs, both the forward and reverse primers contained NdeI restriction endonuclease sites respectively.

In generating the GFP fusion construct, i.e., pET-28c (+) -eCFP-TEV-eYFP, the CDNA containing the eYFP gene was PCR-amplified with forward and reverse primers containing BamHI and NotI respectively. The insert eYFP gene was then cloned into the pET-28c (+) -eCFP-TEV cloning vector.

In generating pET-28c (+) -eYFP-HIF1α expression construct, forward and reverse primers containing BamHI and NotI respectively restriction endonuclease sites were used. The restriction endonuclease sites have additional bases at the 5’ ends to serve as platforms during binding of the restriction endonuclease to the recognition site. The restriction sites are underlined. Developed clones were sent to Eurofins MWG Operon for sequencing (Raynes Park, UK).

2.1.3.2 Oligonucleotide primers for assay

The biotin labeled and unlabeled high affinity c-fos sis-inducible element (hSIE or dsM67), used in the ELISA, PEMSA and FRET assays were synthesized by Invitrogen™ Ltd (Renfrew, UK).

Biotin...5’-TCT CTC TCT CTG CAT TTC CCG TAA ATC T- 3’ (Forward primer)
3’-C GTA AAG GGC ATT TAG AA-5’ (Reverse primer)

2.1.3.3 Phosphotyrosyl peptides used for activity assays

The phosphotyrosyl peptides used in the ELISA, PEMSA and FRET assays (see Table 2.2) were either purchased from Cambridge Research Biochemical’s Ltd (Bellingham, UK) or Peptide Synthetic Protein Research Ltd (Fareham, UK).
Table 2.2: Phosphotyrosyl peptide used in the functional assays

<table>
<thead>
<tr>
<th>Phosphotyrosyl peptide sequence</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-FAM-GpY-LPQTV-NH₂</td>
<td>Cambridge Research Biochemical's (UK)</td>
</tr>
<tr>
<td>Ac-pY-LPQ-NH₂</td>
<td>Peptide Synthetic protein research (UK)</td>
</tr>
</tbody>
</table>

Note: p; phosphate group, FAM; 5-Carboxyfluorescein, Ac; acetyl group.

2.1.4 DNA CONSTRUCTS

2.1.4.1 Bacterial expression plasmids

The pET-32a (+)-STAT3βtc expression vector containing the nucleotides sequence coding for the 127 – 722 amino acid residues of murine STAT3βtc (identical to the human STAT3 on the protein level) was a gift from Prof. C.W. Müller (EMBL Greoble, France). The enhance Green Fluorescent Protein plasmid DNA vector (peGFP-NI) was a generous gift from Prof. R. Harvey (Department of Pharmacology, The School of Pharmacy, University of London. UK). The enhance Cyan Fluorescent Protein and enhance Yellow Fluorescent protein DNA plasmids DNA vectors, i.e., peYFP-NI and peCFP-NI were also generous gifts from Prof. F.A. Stephenson (Department of Pharmaceutical and Biological Chemistry, The School of Pharmacy, University of London. UK).

The pET-28c (+) plasmid DNA vector was purchased from Novagen (Sutton, UK). The expression vectors pET-32a (+)-eGFP-STAT3βtc, pET-32a (+)-eCFP-STAT3βtc and pET-32a (+)-eYFP-STAT3βtc (where the GFP tag was fused to the N-terminus of STAT3βtc) as well as pET-28c (+)-STAT3βtc-eCFP, pET-28c (+)-STAT3βtc-eGFP and pET-28c (+)-STAT3βtc-eYFP (where GFP was fused to the C-terminus of STAT3βtc) were developed in this study. Also developed in this study included; pET-28c (+)-His-STAT3βtc-eCFP, pET-28c (+)-His-STAT3βtc-eGFP and pET-28c (+)-His-STAT3βtc-eYFP (where the polyhistidine tag is fused to the N-terminus of STAT3βtc). In addition, expression constructs including; pET-28c (+)-eCFP, pET-28c (+)-eGFP, pET-28c (+)-eYFP and pET-28c (+)-eCFP-TEV-eYFP were also cloned in this study. Furthermore, pET-28c (+) –eCFP-HIF1α, pET-28c (+) -eCFP-HIF1β and pET-28c (+)-eCFP-Tev-HIF1β were also developed for other related study. The vector maps of the original and cloned expression constructs are shown in the methods section. A summary detailing
the antibiotic marker of the developed expression constructs as well as their respective recombinant proteins are shown in Table 2.3.

**Table 2.3: Characteristic properties of the expression vectors used in the study**

<table>
<thead>
<tr>
<th>Expression Construct</th>
<th>Recombinant Protein</th>
<th>Antibiotic Marker</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET-32a (+) -STAT3βtc</td>
<td>STAT3βtc</td>
<td>Ampicillin (100 µg/mL)</td>
</tr>
<tr>
<td>pET-32a (+) -eCFP-STAT3βtc</td>
<td>eCFP-STAT3βtc</td>
<td>Ampicillin (100 µg/mL)</td>
</tr>
<tr>
<td>pET-32a (+) -eGFP-STAT3βtc</td>
<td>eGFP-STAT3βtc</td>
<td>Ampicillin (100 µg/mL)</td>
</tr>
<tr>
<td>pET-32a (+) -eYFP-STAT3βtc</td>
<td>eYFP-STAT3βtc</td>
<td>Ampicillin (100 µg/mL)</td>
</tr>
<tr>
<td>pET-28c (+) -eCFP</td>
<td>eCFP</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET-28c (+) -eGFP</td>
<td>eGFP</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET-28c (+) -eYFP</td>
<td>eYFP</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET-28c (+) -STAT3βtc-eCFP</td>
<td>STAT3βtc-eCFP</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET-28c (+) -STAT3βtc-eGFP</td>
<td>STAT3βtc-eGFP</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET-28c (+) -STAT3βtc-eYFP</td>
<td>STAT3βtc-eYFP</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET-28c (+) -His-STAT3βtc-eCFP</td>
<td>His-STAT3βtc-eCFP</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET-28c (+) -His-STAT3βtc-eGFP</td>
<td>His-STAT3βtc-eGFP</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET-28c (+) -His-STAT3βtc-eYFP</td>
<td>His-STAT3βtc-eYFP</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET28c (+) -eCFP-HIF1β</td>
<td>eCFP-HIF1β</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET32a (+) -eCFP-Tev-HIF1β</td>
<td>eCFP-Tev-HIF1β</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET28c (+) -eYFP-HIF1α</td>
<td>eYFP-HIF1α</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
<tr>
<td>pET28c (+) -eCFP-TEV-eYFP</td>
<td>eCFP-TEV-eYFP</td>
<td>Kanamycin (30 µg/mL)</td>
</tr>
</tbody>
</table>

### 2.1.5 Bacterial strains

#### 2.1.5.1 Escherichia coli cells (E.coli)

**XL1-Blue™ E.coli** competent cells – Genotype: *recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac [F' proAB lacF'ZΔM15 Tn10 (Tet')]. XL1-Blue *E.coli* cell was used for
molecular cloning and optimal plasmid DNA amplification. This strain was purchased from Stratagene (Agilent Technologies, UK).

**JM109™ (DE3) *E.coli* competent cells** — Genotype: e14' (McrA'), recA1, endA, gyrA96, thi-1, hsdR17, (r_k', m_k'), supE44, relA1, Δ(lac-proAB) [F' traD36 proAB lacIqZAM15]. JM109 (DE) competent cell was used for routine molecular cloning application and plasmid DNA amplification. The strain was purchased from Stratagene (Agilent Technologies, UK).

**DH5α™ *E.coli* competent cells** — Genotype: F^- ϕ80lacZΔ15 Δ(lacZYA-argF)U169 deoR recA1 endA1 hsdR17(r_k', m_k') phoA supE44 thi-1 gyrA96 relA1 λ^- . DH5α was employed for routine subcloning application and high-quality DNA plasmid preparations. The strain was purchased from Invitrogen Ltd (Paisley, UK).

**BL21 (DE3) TKB1™ *E.coli* competent cells** — Genotype: *E.coli* B F^- dcm, ompT, hsdS (r_b' m_b') gal λ (DE3) [pTK Tet']. TKB1 strain carries the gene for T7 polymerase and also contains a plasmid encoded, inducible tyrosine kinase gene (pTK). It was used for the controlled expression and phosphorylation of recombinant proteins placed downstream of the T7 polymerase binding site. This strain was purchased from Stratagene (Agilent Technologies, UK).

**BL21 (DE3) Rosetta™ competent cells** — Genotype: F- ompT, hsdSB (r_b' m_b) gal, dcm, pRARE2 (Cam^R). The Rosetta™ strain enhances the expression of genes that encode rare *E.coli* codons. It was used to overcome translational limitations by the codon usage by *E.coli*. It was used to express recombinant proteins with no post-translational modifications. This strain was purchased from Novagen (Merckbioscience, UK).

**BL21 (DE3) Origami™ 2 *E.coli* competent cells** — Genotype: Δ (ara–leu) 7697 ΔlacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsL[F' lac+ lacI q pro] (DE3) gor522::Tn10 trxB (StrR, TetR)4. The Origami 2 strain was employed to solubilise the pET-28c (+) STAT3βc-GFP expression constructs. Due to mutations in the thioredoxin reductase (trxB) and glutathione reductase (gor) genes, disulphide bond formation is greatly enhanced in the cytoplasm. The strain was purchased from Novagen® (Merckbiosciences, UK).
2.1.6 Antibodies

STAT3 K-15 rabbit polyclonal IgG antibody raised against peptide region (626-640) of STAT3 of mouse origin was purchased from Santa Cruz Biotechnology (Wiltshire, UK). Also, STAT3 (Phospho pY705) rabbit monoclonal IgG antibody raised against a synthetic phospho-peptide corresponding to the residues surrounding the tyrosine 705 region of human STAT3 was purchased from AbCam (Cambridge, UK). The GFP rabbit polyclonal IgG antibody was raised against a highly purified recombinant full length GFP protein construct expressed in E.coli. The antibody was purchased from AbCam (Cambridge, UK) and it is reactive to the various variants of the Aequorea Victoria GFP proteins, i.e., S65T-GFP, RS-GFP, CFP, YFP, RFP and eGFP. The streptavidin-HRP conjugate (peroxidase-conjugated streptavidin biotin binding protein) was purchased from Pierce® (Rockford, USA). The monoclonal anti-FLAG® antibody produced in mouse and raised against the amino-terminal FLAG epitopes was purchased from Sigma-Aldrich (Poole, UK). The phosphotyrosine mouse IgG monoclonal antibody (pY20) raised against phosphorylated tyrosine residues was purchased from BD Transduction laboratories (Oxford, U.K). Anti-polylhistidine mouse IgG monoclonal antibody raised against a synthetic polylhistidine tags was purchased from Sigma-Aldrich® (Dorset, UK). ECL™ Peroxidase labeled Anti-Rabbit IgG horseradish peroxidase–linked whole antibody and ECL™ Peroxidase labeled Anti-Mouse IgG horseradish peroxidase–linked whole antibody were purchased from GE Healthcare (Buckinghamshire, UK).

2.1.7 Equipment

The Biometra Personal cycler was programmed for PCR, restriction digestion and general incubation; New Brunswick Innova™ 4000 rotatory incubator, used for bacterial culturing at 37 °C; CE2021 spectrophotometer used to estimate protein and DNA concentration; Leec incubator, programmed at 37 °C for overnight incubation; Electrolab P300 fermenter, used for growing up to 10L cultures; Sigma EK10 centrifuge, used to harvest bacterial cell cultures; Soniprep 150 sonicator, used to lyse E.coli; LKB Brown 2011 macrovue transilluminator, used to analyse ethidium bromide stained DNA in an agarose gel; Platform Shaker STR6, for incubation with shaking at either 4 °C or at room temperature; Sartorius basic balance, for weighing; Mettler delta 340, pH measurement; Beckman Coulter centrifuge (Avanti J-E), for separating bacteria lysate into soluble and insoluble fractions at RCF greater than 20,000 xg (rotor used JA 25.5 rotor); Bibby B212 hot plate and stirrer, for heating or stirring; Eppendorf table top centrifuge 5415C, used extensively for 30 second spins, miniprep and protein purification with the aid of the Vivapure protein purification and concentration spin columns; Bio-Rad’s Sub-Cell™ GT system, used for agarose gel electrophoresis; Bio-
Rad's Mini-Protean™ Tetra system, used for SDS-PAGE and Native PAGE analysis; Bio-Rad's Mini-Protean II™ system, used for Western Blots Hoeffer multiple gel caster® (Amersham Bioscience), used for SDS PAGE and native PAGE analysis; AKTA purifier chromatographic system, used for gel filtration chromatography; PerkinElmer Multilabel plate reader (Wallac EnVision 2101), used to read 96 and 384-well assay plates at different absorption wave lengths; PerkinElmer Luminescence LS55 spectrometer; used for measuring fluorescence intensity.

2.2 METHODS

2.2.1 Molecular biology methods

All solutions were sterilized using 0.22 μm syringe driven filters or autoclaved at 121 °C for 25 min.

2.2.1.1 Escherichia coil (E.coli) stock preparation and transformation

2.2.1.1.1 Preparation of chemically competent E.coli cells with antibiotic marker

Using a sterile 1 μL loop, thawed E.coli were streaked from a glycerol stock onto a Luria Bertani Broth (LB) 2.5 % (w/v), agar 1.5 % (w/v) plate supplemented with the appropriate antibiotics. The plate was gently inverted and incubated overnight (about 12 – 16 h) at 37 °C. A seed culture was generated with a 5 mL LB broth inoculated with a single colony and incubated at 37 °C for 4 to 6 h at 150 rpm. 1 mL of the seed culture was used to inoculate 500 mL sterile Luria Bertani Broth (LB). The culture was incubated at 37 °C while shaking at 150 rpm until the optical density (OD)_{A600} = 0.6 was achieved. At an OD of 0.6, the culture was quickly transferred into a pre-chilled 1 L sterile centrifuge tube and allowed to cool on ice for 10 – 15 min. The cell culture was cleared by centrifugation at 3,600 rpm for 20 min at 4 °C using a Hermle Z513K table top centrifuge (rotor, 220.70/V05/V06). The supernatant was quickly discarded the pellet was re-suspended in 50 mL sterile ice-cold 100 mM CaCl₂. Subsequent steps were performed on ice. Once re-suspended, the 50 mL cell suspension was further diluted to 250 mL with ice-cold 100 mM CaCl₂ and incubated on ice for 20 min. The suspension was cleared again via centrifugation at 3,600 rpm for 25 min at 4 °C. The supernatant generated was discarded and the pellet was gently re-suspended with 5 mL ice-cold CaCl₂ storage solution (20 % (v/v) glycerol, 100 mM CaCl₂). The cell suspension was immediately aliquoted into 200 μL fractions, snap frozen in liquid nitrogen and stored at −80 °C until use. The competent cells were used within 8 months of preparation.
2.2.1.1.2 Preparation of chemically competent *E. coli* cells with no antibiotic marker

*E. coli* strains, such as JM109 (DE3), do not carry intrinsic antibiotic markers. Specially prepared plates referred to as minimal plates were used to generate seed colonies for cells. The minimal plates were prepared as follows: 10 mL stock solutions of glucose (20%), thiamine HCl (1.0 M) and MgSO<sub>4</sub> (1.0 M) were prepared and filter sterilized. In addition, 100 mM CaCl<sub>2</sub> and 10 times concentrated M9 salt (420 mM Na<sub>2</sub>HPO<sub>4</sub>, 220 mM KH<sub>2</sub>PO<sub>4</sub>, 180 mM NH<sub>4</sub>Cl and 85 mM NaCl) were also prepared and sterilized (autoclaved at 121 °C for 20 min). The solutions were filter sterilized separately to avoid caramelizing the glucose and destroying the thiamine. A sterile molten 1.5 % (w/v) agarose was supplemented with 0.2 % (v/v) glucose, 1.0 mM thiamine HCl, 1.0 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, and 1X M9 salt before pouring into plates. Ten plates were generated from the 200 mL media; unused plates were sealed with parafilm and stored at 4 °C until use. A single colony of JM109 (DE3) was isolated from a minimal plate and competent cells were prepared as before, except the LB broth contained no antibiotics.

2.2.1.1.3 *E. coli* competent cell transformation

Plasmid amplification and antibiotic resistance selection steps were carried out in *E. coli* cells. An aliquot of frozen competent *E. coli* cells were thawed on ice and approximately 0.50 µg of the plasmid DNA was added. The mixture was gently agitated to mix and subsequently incubated on ice for 45 min. The cell-plasmid mix was transferred into thermo cycler PCR machine simply programmed to perform the heat shock stage, i.e., 42 °C for 90 s. The cells were then incubated on ice for a further 2 min to allow them to recover. They were then diluted 1/10 by the adding 400 µL of freshly prepared LB broth. The cells were grown with agitation (150 rpm) for 45 min at 37 °C. Following incubation, 100 µL of the culture media was plated on antibiotics selective plates, i.e., plates supplemented with either 12.5 µg/mL tetracycline in 2.5% (w/v) LB and 1.5 % (w/v) agar or 20 µg/mL chloramphenicol in 2.5% (w/v) LB and 1.5% (w/v) or 30 µg/mL kanamycin in 2.5% (w/v) LB and 1.5 % (w/v) agar or 100 µg/mL ampicillin in 2.5% (w/v) LB and 1.5 % (w/v) agar plates etc depending upon the resistance conferred by the plasmid. The plates were inverted and incubated at 37 °C for 16 h. Bacterial colonies were selected for further studies or the plates were wrapped in parafilm and stored at 4 °C for future use.

2.2.1.1.4 Screening *E. coli* colonies

Bacteria colonies were randomly picked and cultured in fresh broth (5 mL) prior to miniprep DNA extraction (section 2.2.1.1.6) or were PCR-amplified with the appropriate primers to amplify the region of interest prior to agarose gel electrophoresis.
2.2.1.5 Glycerol stock of transformed *E.coli* cells

Freshly prepared LB broth (5 mL) supplemented with the appropriate antibiotics was inoculated with a single colony of a pre-transformed *E.coli* cells. The culture media was grown at 37 °C with agitation (150 rpm) until an OD$_{600}$ of 0.6 was reached. Glycerol 50% (v/v) stocks of the culture media was prepared by diluting 500 µL culture with 500 µL 100 % sterile glycerol. The mixture was thoroughly mixed by inversion and stored at -20 °C until use.

2.2.1.6 Mini preparation of plasmid DNA

Freshly prepared LB broth (5 mL) was transferred into a 30 mL universals tube and then inoculated with a single colony of a pre-transformed *E.coli* cells picked from a selective plate. The 5 mL culture was supplemented with the appropriate selective antibiotics, either 100 µg/mL ampicillin or 12.5 µg/mL tetracycline or 30 µg/mL kanamycin or 20 µg/mL chloramphenicol, as appropriate. The culture was grown at 37 °C with shaking at 150 rpm overnight, i.e., 16 h. From 5 mL cell culture generated overnight, 1.5 mL was transferred into sterile 1.5 mL microcentrifuge tube and then cleared by centrifugation at 7000 xg for 2 min at 4 °C. The resultant supernatant was discarded and the cell pellet was gently re-suspended in 200 µL of re-suspension buffer (50 mM Tris (hydroxymethyl) aminomethane-HCl (Tris-HCl), pH 8.0, 10 mM ethylenediaminetetraacetic acid (EDTA) 100 µg/mL RNase A). Lysis buffer, 200 µL (200 mM NaOH, 1 % (w/v) sodium dodecyl sulphate (SDS)) was added to the cell suspension. The mixture was mixed gently by inverting 6 times. Neutralization buffer (3 M Potassium acetate, pH 4.8), 200 µL was added and then the mixture was further homogenized gently by inversion (6x) until white precipitates developed. The tubes were incubated at -20 °C for 10 min and the precipitates were sedimented by centrifugation at 13,000 xg for 10 min at 4 °C. The resultant supernatant was carefully transferred into a fresh 1.5 mL tube. 400 µL of ice-cold 100 % isopropanol was added to the supernatant and subsequently mixed by inversion. The tubes were centrifuged at 13000 xg for 10 min at 4 °C. The supernatant was discarded and to the near transparent pellet, 200 µL of ice-cold 70 % ethanol (v/v) was added. The tube was inverted 3 times to wash the DNA pellet. The tubes were centrifuge again at 13,000 xg for 5 min at 4 °C. The generated supernatant was discarded and the plasmid DNA pellet was air-dried for 10 min prior to suspension with 50 µL sterilized ddH$_2$O. The sample was stored at -20 °C for future use.
2.2.1.1.7 Maxi preparation of plasmid DNA

Qiagen Plasmid Maxiprep kit was employed for large scale plasmid DNA extraction and purification. A single colony of a pre-transformed cells picked from an antibiotic selective plate was used to inoculate 5 mL of freshly prepared LB broth supplemented with the appropriate selective antibiotics, either 100 µg/mL ampicillin or 12.5 µg/mL tetracycline or 30 µg/mL kanamycin or 20 µg/mL chloramphenicol. The culture was grown at 37 °C for 16 h at 150 rpm. The 5 mL culture was used to inoculate 500 mL of freshly prepared and sterile LB broth supplemented with the appropriate selective antibiotics, i.e., either 100 µg/mL ampicillin or 12.5 µg/mL tetracycline or 30 µg/mL kanamycin or 20 µg/mL chloramphenicol. The culture was further incubated at 37 °C for 16 h with agitation (150 rpm).

The cell culture was cleared by centrifugation at 3,600 rpm for 20 min at 4 °C using a Hermle Z513K table top centrifuge (rotor, 220.70/V05/V06). Using the Qiagen Maxi prep protocol, the cell pellet was re-suspended in 10 mL of the re-suspension buffer (P1) by gentle pipetting. Lysis buffer (P2), 10 mL was added and mixed gently by inverting 6 times and then incubated for 4 min at room temperature. Ice-cold neutralization buffer (P3), 10 mL was added to terminate lysis. The mixture was thoroughly mixed by inverting the tube 6 times until a white precipitate was formed. The resultant mixture was cleared by centrifuge at 20,000 xg for 30 min at 4 °C using the rotor JA25.50 in a Beckman centrifuge.

The resultant supernatant was bound to the equilibrated (equilibration buffer; buffer QBT) Qiagen-tip and allowed to empty by gravity flow. The Qiagen-tip was washed with twice 30 mL buffer QC and then eluted with 15 mL buffer QF. The eluted fraction containing the plasmid DNA was subsequently precipitated with 10.5 mL of isopropanol and immediately centrifuged at 20,000 xg for 30 min at 4 °C. The resultant precipitate DNA pellet was washed twice with 70 % ethanol to remove precipitated salts as well as substitute isopropanol with the more volatile ethanol. This step improves the solubility of the precipitate DNA. The ethanol-DNA suspension was cleared by centrifuged as described above. The plasmid DNA pellet was air-dried for 10 min before re-dissolving in 500 µL of sterilized double distilled H₂O.

The concentration of the purified plasmid DNA was estimated by reading the, i.e., \( \text{OD}_{260} = 260 \text{ nm} \). An \( \text{OD}_{260} = 260 \text{ nm} \) of 1.0 is equivalent to a DNA concentration of 50 µg/mL. The plasmid DNA preparation was evaluated by agarose gel electrophoresis after restriction endonuclease digestion to ensure that the correct plasmid has been amplified and purified. The sample was stored at -20 °C.
2.2.1.2 Molecular cloning

2.2.1.2.1 Restriction endonuclease digestion

Restriction endonuclease digestion also referred to as a restriction digest was carried out with a restriction endonuclease enzyme. It was employed for diagnostic and cloning purposes. For the purposes of diagnostics, the restriction digestion was mostly carried out in 0.5 mL sterile tubes containing about 0.5 – 1.5 μg (~1.0 to 10 μL) plasmid DNA, appropriate concentration of the restriction endonuclease buffer, i.e., 1 x final concentration and 5 – 15 units of restriction endonuclease enzymes. The reaction was always diluted with ddH₂O to the appropriate volume before adding the enzyme. The restriction mixture was incubated at 37 °C for 1 h and terminated either by adding the appropriate volume of 6 x loading dye (8.3 % (w/v) glycerol, 0.016 % (w/v) bromophenol blue, 1 x TBE) for the purposes of electrophoresis or by thermal deactivation of enzyme at 65 °C for 20 min for the purposes of dephosphorylation. The fragmented DNA was analyzed by 0.8 % (w/v) flat bed agarose gel electrophoresis, section 2.2.1.2.3.

For the purposes of cloning as described in Figure 2.1, following from the protocol described above, approximately 0.5 μg of the purified PCR-amplified product was restricted. The restriction mixture was incubated at 37 °C for 1 h, terminated using 6x loading dye and then analyzed by agarose gel electrophoresis, section 2.2.1.2.2. The restricted PCR product was ligated into the appropriate sites in the cloning vector, section 2.2.1.2.10.
Figure 2.1: Summary of the procedure used to generate clones in the study

Oligonucleotide primers were used to generate insert DNA with the appropriate endonuclease restriction sites via PCR. The PCR product was purified using a Qiagen PCR purification kit, restricted using the appropriate endonuclease restriction enzyme and then purified again using agarose gel electrophoresis. Together with the restricted cloning vector, the restricted insert DNA was used in the ligation reaction.

2.2.1.2.2 Analysis of DNA using flat bed agarose gel electrophoresis

Flat bed agarose gel electrophoresis (0.8 % (w/v)) was used to evaluate the integrity of plasmid DNA or PCR products. The agarose gel electrophoresis was performed with either a Bio Rad Wide Mini Sub-Cell® GT (analysis of more than 8 samples) or a Mini Sub-Cell® GT (analysis of less than 8 samples) electrophoresis system. A 0.8 % (w/v) agarose solution was prepared from 1 x Tris-borate-EDTA buffer (TBE, 100 mM Tris-HCl, pH 8.4, 90 mM boric acid, 4.0 mM EDTA). The agarose suspension was boiled for 60 to 80 sec to dissolve. Upon cooling to approximately 40 °C, ethidium bromide was added to a final concentration of 0.5 μg/mL. The molten agarose was poured into a plastic tank sealed at both ends with an autoclave tape. A comb was inserted about 2 cm from the top of the gel. The molten agarose was allowed to solidify for 30 to 45 min. Upon solidifying, the comb and the autoclave tape was removed and the agarose gel was transferred into the buffer reservoir. The gel was then covered in approximately 300 mL of 1 x TBE buffer supplemented with 0.5 μg/mL of ethidium bromide. The plasmid DNA or PCR product to be analyzed was mostly prepared by the addition of 1 or 2 μL 6 x loading dye. The DNA samples were loaded (either 5.0 or 10.0 μL) and separated out at 80 volt for 1 h. The ethidium bromide stained DNA fragments were
imaged using LKB Brown 2011 Macrovue Transilluminator UV light source (λ = 312 nm) and a Kodak DC 120 digital camera.

2.2.1.2.3 Flat bed agarose gel purification of restricted DNA

The Qiagen gel extraction kit was used to extract and purify restricted plasmid DNA vectors or PCR products from agarose gels. Agarose gel electrophoresis was used to fractionate the restricted reaction mixture. Upon visualizing the agarose gel using blue light, the desired DNA band was excised using a clean and sharp scalpel. The excised band was transferred into a sterile 1.5 mL centrifuge tube and then weighed. Three volumes of buffer QG was added to 1 volume of the gel (100 mg = 100 µL). The mixture was incubated at 50 °C until the gel was completely dissolved. One volume of isopropanol was added to the reaction and subsequently mixed. The mixture was transferred into a QIAquick spin column and centrifuged for 1 min at 13,000 xg at room temperature. The flow-through was discarded and the bound DNA washed with 0.5 mL of buffer QG via centrifugation as stated above. The bound DNA was further washed with 0.75 mL buffer PE via centrifugation and the generated flow-through discarded. The QIAquick spin column was further centrifuged as stated above to remove traces of ethanol present in the buffer PE. The column was then transferred into a sterile 1.5 mL centrifuge tube and the bound DNA eluted with 30.0 µL ddH2O. The purified cloning vector or PCR product was immediately used to set-up a ligation reaction.

2.2.1.2.4 Precipitation of plasmid DNA using ethanol

Where necessary, ethanol precipitation was used to further purify and concentrate plasmid DNA for sequencing. Using a sterile 1.5 mL tube, 0.12 volumes of 3 M sodium acetate and 2.5 volumes of ice-cold ethanol were added to 1 volume of plasmid DNA. The mixture was incubated at -80 °C for 2 h. Following incubation, the mixture was thawed and subsequently centrifuged for 30 min at 13,000 xg and at 4 °C. The plasmid DNA pellet was washed with 200 µL of ice-cold 70 % (v/v) ethanol after carefully discarding the supernatant. The ethanol-plasmid DNA suspension was further centrifuged at 13,000 xg for 5 min at 4 °C. The generated supernatant was carefully discarded and the plasmid DNA pellet air-dried for approximately 10 min. The dried pellet was re-suspended in appropriate volume of double distilled H2O and either stored at -20 °C or sent to Eurofins MWG Operon for sequencing.

2.2.1.2.5 Annealing oligonucleotide DNA primers using water-bath

An annealing mixture was prepared by mixing equimolar amounts of complementary DNA in a sterile 1.5 mL tube (e.g. 20 µL forward and reverse primers of 1 mM each). The reaction mixture was supplemented with 100 mM NaCl, and then diluted to 100 µL.
with sterile ddH$_2$O. The lid of the 1.5 mL tube was wrapped in parafilm and then supported with a floater in a water-bath (4 L). The water-bath was heated until boiling (100 °C) for 5 min and subsequently allowed to cool slowly to room temperature. The tube was centrifuged for 30 sec and the oligonucleotide DNA stored at -20 °C for future use. Examples of the double stranded oligonucleotide DNA prepared in section 2.2.1.2.5 are show in Figure 2.2.

(i). Biotinylated dsM67 oligonucleotide sequence for functional assays

```
  5' Biotin-TCT CTC TCT GAT TCC CCG TAA ATC T - 3'
  3' -C GTA AAG GGC ATT TAG AA - 5'
```

(ii). Tobacco Etch Virus (TEV) oligonucleotide for ligation reaction

```
  5' CTAGC GAA AAC TTG TAT TTC CAG GGC G - 3'
  3' G CTT TTG AAC ATA TTG GTC CCG C GATCCG - 5'
```

Figure 2.2: Summary of the oligonucleotide dsDNA sequence generated using the annealing protocol described in section 2.2.1.2.6.

Equimolar amounts of complimentary DNA primers were used to generate the double stranded DNAs' described above. (i) Biotinylated modified c-fos sis inducible enhancer element (dsM67) containing dimeric STAT3 DNA binding site (site designated as blue), (ii) Oligonucleotide sequence containing the DNA sequences that code for the TEV protease recognition site (site designated as red). The dsDNA sequence contains Nhel and BamHI sites at 5'end 3'end respectively.

2.2.1.2.6 Designing PCR oligonucleotide primers

A general criterion was followed when designing the oligonucleotide DNA primers used for PCR amplification. These included the following;

- Oligonucleotide DNA primers must not be less than 18 base pairs.
- Wallace et al., 1979 rule was used to calculate the melting point of the designed primer. The melting point formula is a follows
  Melting point: \(4[G + C] + 2[A + T]\), and defined as:
  \[G + C\] = number of guanine bases plus number of cytosine bases;
  \[A + T\] = number of adenine bases plus number of thymine bases. The primers were designed so that the optimal melting points were restricted between the temperature ranges 52 °C to 60 °C.
❖ The melting temperature of the forward and reverse primers was kept within 10°C of each other.
❖ The guanine and cytosine (GC) content of the oligonucleotides were kept between 40 to 60 %.
❖ Palindromic sequences within individual primers were avoided.
❖ Oligonucleotide primers were designed to have extra bases at the 5' end to facilitate restriction endonuclease binding during restriction of PCR products.

Oligonucleotide DNA primers were always dissolved in sterile ddH2O prior to application. Primer were purchased from either by Eurofins MWG Operon (Raynes Park, UK) or Invitrogen™ Ltd (Renfrew, UK)

2.2.1.2.7 General PCR-amplification reaction

The reaction mixture for PCR-amplification was generally prepared in specialized 0.5 mL PCR tube. The 50 μL reaction mixture contained the following ingredients; 1x Pfu polymerase buffer (20.0 mM Tris-HCl pH 8.8 at 25 °C, 10.0 mM KCl, 10.0 mM (NH₄)₂SO₄, 2.0 mM MgSO₄, 0.10 mg/mL nuclease-free BSA and 1% Triton®X-100), 200 μM deoxynucleoside-5'-triphosphates mix [(dNTP) dATP, dCTP, dGTP, and dTTP], 0.50 μM forward primer, 0.5 μM reverse primer, 0.5 μg template DNA and 2-3 units Pfu polymerase enzyme.

A control reaction containing no template DNA was prepared in parallel. The control reaction was prepared to confirm that the right gene had been amplified. The programme for PCR-amplification involved 30 cycles and it was carried out using a Biometra thermo cycler. The parameters described in Table 2.4 were used for amplification. The diagram in Figure 2.3 is a representative PCR product.

Table 2.4: Summary of the general parameters used in PCR-amplification

<table>
<thead>
<tr>
<th>STAGE</th>
<th>CYCLE</th>
<th>TEMPERATURE (°C)</th>
<th>TIME (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>95</td>
<td>45</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>95</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72</td>
<td>15 sec/kb</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>72</td>
<td>600</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td></td>
<td>A</td>
</tr>
</tbody>
</table>
The resultant PCR product was purified using a Qiagen PCR clean-up kit. The integrity of the sample was analyzed by a flat bed agarose gel electrophoresis. The concentration of the PCR product was estimated using GeneQuantII RNA/DNA calculator (Pharmacia Biotech, Cambridge, UK).

![Diagram of PCR product and restriction sites](image)

**Figure 2.3: Representative PCR product detailing the endonuclease restriction sites introduced.**

Upon PCR-amplification, new restriction endonuclease site were introduced to the 5' and 3' ends of the amplified DNA sequence described as target protein coding region.

### 2.2.1.2.8 Purification of PCR products

The amplified DNA fragment (PCR product) was purified using a Qiagen QIAquick PCR purification kit. The kit was used to isolate the PCR product from the constituents of the PCR reaction mixture as well as to transfer the product into a preferred buffer. A 2 μL aliquot of the PCR product was taken prior to PCR product purification. The remaining PCR product was diluted 5 times (v/v) with Qiagen PB buffer. The resultant mixture was transferred into a QIAquick column and then centrifuged at 13,000 xg for 60 s at room temperature. The resultant flow-through was discarded and the bound DNA washed with 0.75 mL PE buffer through centrifugation at 13,000 xg for 60 s at room temperature. The resultant flow-through was subsequently discarded and the QIAquick column re-centrifuged (using parameters described above) to remove traces of ethanol that was present in the PE buffer. The QIAquick column was transferred into a sterile 1.5 mL centrifuge tube and the bound PCR product eluted with 35.0 μL ddH₂O by centrifugation at 15,000 xg for 60 s at room temperature. The resultant purified PCR product was either restricted and then ligated into a cloning vector or stored at -20 °C for a very short period (approximately 1 week).
2.2.1.2.9 Dephosphorylation of the 5' end of the cloning vector

The phosphate group at the 5' end of the cloning vector was removed with alkaline phosphatase to prevent re-circularization of the empty cloning vector. Dephosphorylation was performed after restricting the cloning vector, section 2.2.1.2.1. The reaction was prepared in a 0.5 mL sterile centrifuge tubes and involved the following ingredients; 1.0 to 5.0 µg of restricted cloning vector, 1x Antarctic phosphatase buffer (50 mM Bis-Tris-Propane, 1 mM MgCl₂, 0.1 mM ZnCl₂, pH 6.0 at 25 °C) and 5.0 units of Antarctic phosphatase enzyme. The reaction was incubated at 37°C for 1 h. The reaction was terminated by heat de-activating at 65°C for 20 min. The dephosphorylated cloning vector was purified using 0.8% (w/v) flat bed agarose gel electrophoresis, section 2.2.1.2.3, prior to ligation.

2.2.1.2.10 Ligation reaction

Restricted PCR products were directly ligated into a restricted or restricted/dephosphorylated cloning vector either non-directionally or directionally by the addition of T4 DNA ligase. A 10.0 µL ligation reaction contains the following; restricted or restricted/dephosphorylated cloning vector, restricted PCR product (insert), 1x T4 DNA ligase buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM (DTT) and 1 mM ATP) and 3.0 units T4 DNA ligase. The vector to insert ratio was estimated in nanograms (ng) and using the equation below. Unless stated otherwise, 200 ng cloning vector was used. In addition to vector: insert ratio 1:1, other reactions contained vector to insert ratios such as; 1:3 (one part of vector to three parts of insert) and 1:6 (one part of vector to six parts of insert). Two control reactions, similar to the ligation reaction above were prepared in parallel. One control reaction was prepared minus PCR products whilst the other was prepared minus the T4 DNA ligase. The control reactions were prepared to establish the efficiency of the restriction, dephosphorylation and ligation processes.

\[
\frac{\text{mass of vector (ng)} \times \text{size of insert (kb)}}{\text{size of vector (kb)}} \times \text{molar ratio of insert to vector} = \text{mass of insert (ng)}
\]

The ligation reaction was incubated at room temperature for 1 h and then at 4 °C for 16 h. The ligated mixture was used to chemically transform one of the following; XL1-Blue, JM109, DH5α competent cells, Section 2.2.1.1.3. As described in section 2.2.1.1.4., possible clonal colonies were taken through the following processes; "mini-culturing, 5 mL culture (section 2.2.1.1.6), Miniprep DNA extraction method (section 2.2.1.1.6), plasmid identification by restriction digest and agarose gel electrophoresis (sections 2.2.1.2.1 and 2.2.1.2.2) and Maxiprep DNA extraction of the positive clone (section
2.2.1.7) for future use and nucleotide DNA sequencing. The DNA sequence data were analyzed for mutations and to ensure that the insert gene were in-frame and in the right orientation.

2.2.1.2.11 Sequencing analysis of selected clones

Approximately 2.5 to 3.0 µg (~ 20.0 µL) aliquot of the purified plasmid DNA was transferred into a sterile 1.5 mL tube. Label supplied by Eurofins MWG Operon (DNA sequencing company) was attached to the tubes. The plasmid DNA sample was sequenced in both direction using the promoter (5’-dTAATACGACTCACTATAGG G -3’) and terminator primers (5’-dGCTAGTTATTGCTCAGCGG-3’) of the T7 promoter present in the recombinant plasmid DNA. The sequence data was provided in a file containing the nucleotide sequence derived from a false coloured trace of the sequentially eluted fluorescent nucleotides.

2.2.1.3 Recombinant protein

2.2.1.3.1 Small scale expression of recombinant protein in E.coli

To evaluate the cloning process, a test expression of the clone in recombinant E.coli cells was performed following transformation of either BL21 (DE3) TKB1 or BL21 (DE3) Rosetta competent cells. A single colony of pre-transformed E.coli cell was picked from a selective plate and used to inoculate 5 mL sterile LB broth supplemented with the appropriate antibiotics, either 100 µg/mL ampicillin or 30.0 µg/mL kanamycin or 12.5 µg/mL tetracycline, or 20.0 µg/mL chloramphenicol, where appropriate. The culture was allowed to grow at 37 °C with 200 rpm agitation until light scatter from the cells gave an apparent optical density of 0.6 (A₆₀₀ = 0.6). The culturing temperature was lowered to 21 °C and the protein induction initiated by supplementing the culture with isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1.0 mM. The culture was allowed to grow at the reduced temperature for 4 h. The culture, 1.5 mL was transfer into a sterile 1.5 mL centrifuge tube and centrifuged at 6000 xg for 2 min at room temperature. The supernatant generated was discarded and the pellet either stored at -20 °C or lysed with 0.5 mL extraction buffer (20.0 mM HEPES pH 8.0, 10.0 mM MgCl₂, 100 mM NaCl, 20.0 mM DTT, 1.0 mM EDTA and 0.5 mM PMSF) supplemented with 2 % sodium dodecyl sulphate (SDS). The lysate was analyzed by SDS-PAGE (section 2.2.5.1.1) followed by western blot (section 2.2.5.2).

2.2.1.3.2 Large scale fermentation of phosphorylated (p) recombinant proteins in BL21 (DE3) TKB1 E.coli cells

Expression of phosphorylated recombinant proteins, i.e., GFP-pSTAT3tc was performed in BL21 (DE3) TKB1 E.coli cells. Due to solubility issues, large scale
expression of the STAT3βtc-GFP constructs, i.e., constructs with the GFP tag fused to the C-terminus of STAT3βtc, was not performed.

A 2 or 10 L fermenter flask containing sterile LB broth was used. A single colony of a freshly transformed TKB1 was selected and used to inoculate 20 mL sterile LB culture broth supplemented with 100 μg/mL ampicillin and 12.5 μg/mL tetracycline. In parallel with culturing the seed culture at 37 °C for 6 h at 200 rpm, the fermenter flask containing the culture media was also maintained 37 °C in preparation inoculation. The 20 mL seed culture was subsequently used to inoculate the culture media also containing the appropriate antibiotics. The fermentation process was carried out at 37 °C with agitation at 200 rpm and filtered air aeration of 2 L/min until the mid-log phase, $A_{600} = 0.6$ to 0.7.

At about, $A_{600} = 0.4$ to 0.45, the temperature of the media was gradually dropped to 21 °C with the help of an ice jacket wrapped around the flask. Upon reaching $A_{600} = 0.6$ to 0.7, the pre-induction control sample was taken and the media was supplemented to 1 mM IPTG to induce the expression of the recombinant protein. The cells were allowed to grow for 14 to 16 h at 200 rpm agitation and 2 L/min aeration. A single drop of antifoam 204 (Sigma) was added to the culture to suppress foaming during fermentation.

After taking a post-induction sample and recording the optical density ($A_{600} \sim 1.8$), the cells were harvested by centrifugation at 3,600 xg for 20 min at 4 °C. The supernatant generated was discarded and the pellet kept on ice.

To generate the activated recombinant protein (tyrosine phosphorylated recombinant protein), firstly, cells were separated from media by centrifugation and then washed with freshly prepared kinase media (activation media) via centrifugation. The cells were washed to remove residual tryptophan. The tryptophan binds to, and activates the trp repressor which sits on the promoter/operator sequence and stops expression. The kinase media contains the following ingredients; 1 x M9, 1 mM MgSO4, 11 mM D (+) glucose, 0.1 % (w/v) casamino acid, 1.5 μM Thiamine-HCl, 53 μM 3β-indoleacrylic acid, 100 μg/mL ampicillin and 12.5 μg/mL tetracycline. Casamino acid lacks tryptophan, so in the absence of tryptophan the trp repressor fails to bind and expression of the gene is turned on. Also, 3β-indoleacrylic acid (3βIAA) competes with any residual tryptophan that would otherwise repress expression. The washed cell pellet was subsequently re-suspended with the kinase media, i.e. approximately 2.5 g cell/L kinase media. The resultant cell suspension was further incubated at 37 °C for 2.5 h and 2 L/mL aeration to generated activated recombinant protein. Following
activation, the cells were harvested by centrifugation at 3,600 xg for 25 min at 4 °C. The spent media was discarded and the resulting pellet was aliquoted into 2 g portions by centrifugation. The resulting cell pellets were stored at -20 °C for future use. Typically, 2 L fermentation would yield approximately 10 g of cells whilst a 10 L culture will yield about 45 g of packed cell.

2.2.1.3.3 Large scale fermentation of unphosphorylated (u) recombinant protein in E.coli

Unless stated otherwise, production of the unphosphorylated or inactive recombinant proteins were solely carried out in either BL21- (DE3) Rosetta or BL21- (DE3) CodonPlus E.coli cells. A single colony from a freshly transformed cell was used to inoculate 20 mL LB seed culture supplemented with the appropriate antibiotics; 30 μg/mL kanamycin, BL21 (DE3) Rosetta; 20 μg/mL chloramphenicol, BL21 (DE3) CodonPlus. In parallel with culturing the seed culture at 37 °C for 6 h at 200 rpm, the fermenter flask containing the culture media was also kept at 37 °C. The 20 mL seed culture was subsequently used to inoculate the culture media (2L or 10L) also containing the appropriate antibiotics. The fermentation was carried out at 37 °C with 200 rpm agitation and 2 L/min aeration until the mid-log phase, $A_{600} = 0.6$ to 0.7.

The culture media was then treated as described in section 2.2.1.3.2. Upon IPTG induction and subsequent harvesting, the cells were packed into several 2 g portions and then stored at -20 °C for future use. Typically, 2 L fermentation yielded about 10 g of cells whilst a 10 L culture yielded approximately 45 g of cells.

2.2.2 Extraction of recombinant protein from E.coli cells

2.2.2.1 E.coli cell lysis

The first phase to purify or characterize recombinant proteins is to disrupt the cell wall and release the protein. For in vivo recombinant protein expression, such as in E.coli, the total protein content of the cell including the target and "house-keeping" should be extracted from the cell before a purification strategy can be applied to purify the target protein. Two known techniques were applied to lyse the E.coli cells. The methods included mechanical disruption of the cell membrane by sonication and enzymatic digestion of the cell membrane through the use of lysozymes. These lysis methods did not interfer with downstream application including purification, immunoblot analysis and the functional analysis, e.g., ELISA, PEMSA and FRET, of the target recombinant proteins.
Large scale lysis was performed by sonication due to the considerably high yield of the target protein relative to other techniques. However, small scale lysis was performed with commercially available cell lysis solution that contains mild non-ionic detergents to disrupt cells and solubilise proteins without denaturation (CelLytic™ B Cell Lysis reagent, Sigma-Aldrich, UK).

### 2.2.2.1.1 Mechanical disruption of *E.coli* cell wall

Cells were lysed mechanically via sonication. During sonication, high intensity ultrasound was used to disintegrate the cell wall so as to release the soluble cytosolic recombinant protein. A pellet was thawed on ice and subsequently re-suspended in the extraction buffer; 20 mM HEPES-HCl, pH 8.0, 100 mM NaCl, 1.0 mM EDTA, 10 mM MgCl₂, 20 mM DTT, 5 % (w/v) glycerol and 1 tablet SIGMAFAST™ protease inhibitor (4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF), pepstatin, E-64, bestatin, sodium EDTA, aprotinin and leupeptin) in a 1 g cells per 10 mL buffer ratio. However, with polyhistidine tagged GFP recombinant proteins, the extraction buffer was prepared minus DTT and EDTA due to the adverse effect of the two ingredients on downstream affinity purification process, section 2.2.3.1.4. A specially made inhibitor cocktail was applied to the cell-buffer suspension in the ratio 50 μL inhibitor cocktail; 4-(2 aminoethyl) benzenesulfonyl fluoride (AEBSF), pepstain, E-64, bestatin and sodium EDTA, per 1 g cell suspension. The cell-buffer suspension was incubated on ice for 15 min and then sonicated for 5 min with amplitude of 13 micron in a 15 s pulse, 15 s interval rhythm. The cell suspension was kept on ice-water mixture during sonication. Following lysis by sonication, an aliquot of the lysate was taken and the remainder transferred into 40 mL centrifuge tube (Beckman centrifuge tube). The lysate was cleared by centrifugation at 27,000 xg for 45 min at 4 °C using the Beckman JA25.50 rotor in a Beckman coulter centrifuge (Avanti J-E). The resultant supernatant, *i.e.*, cell extract containing the target protein, was transferred into a fresh 50 mL falcon tube and an aliquot taken for analysis. The remainder of the cell extract was either supplemented to 10 % glycerol and stored at -20 °C or subjected to a series of purification steps that may include ammonium sulphate precipitation, anion exchange chromatography and size exclusion chromatography or affinity chromatography to isolate, concentrate and quantify the target protein.

### 2.2.2.1.2 Chemical disruption of *E.coli* cell membrane

For small scale lysis, *i.e.*, less than 0.2 g cells, the commercially available CelLytic™ B Cell Lysis reagent from Sigma-Aldrich was used. Cells were suspended in the lysis reagent buffer in the ratio, 1 g cells per 10 mL of lysis reagent. The cell-reagent suspension was subsequently supplemented with the protease inhibitor cocktails.
Edwin Nkansah

described in section 2.2.2.1. The mixture was then incubated at room for 15 min with shaking and then supplemented with DNase I and MgSO₄ (required for DNase I activity) to final concentrations of 0.1 mg/mL and 100 mM respectively. The mixture was further incubated at room temperature for 30 min. Following incubation, an aliquot of the lysate was taken and stored at -20 °C. The lysate was then cleared by centrifugation at 12,000 xg for 15 min at 4 °C. The supernatant containing the soluble cell extract including the target protein was transferred into a 30 mL universal tube on ice. The resultant cell extract was treated as described in section 2.2.3, based on the target protein. The pellet was re-suspended in lysis buffer and an aliquot taken for analytical purposes.

2.2.3 Purification processes

2.2.3.1 Purification of soluble target protein

Purification, functional characterization and structural evaluation of the recombinant protein are crucial and challenging prospects in proteomics. The use of expression host including *E.coli* (*E.coli* contains ~ 300-400 mg/mL "house-keeping" macromolecules (Zimmerman and Trach, 1991)) for generating target heterogeneous proteins means a purification strategy was needed to isolate the recombinant protein from the thousands of contaminating *E.coli* "house-keeping" proteins. However, no single purification technique can be used to achieve homogeneity of a recombinant protein. Thus, any purification strategy should involve isolation of the protein from the matrix that confines it, separate the protein from the non-protein fraction and isolate the target protein from all other proteins. In isolating target protein, the purification techniques should exploit variations in the physio-chemical properties of proteins. Separating the target protein from the complex cell lysate would require knowledge of the protein's size, pl, charge, shape, hydrophobicity and binding affinity. In utilizing charge and hydrophobicity, separation may be carried out via ammonium sulphate precipitation. Also, with the knowledge of the proteins isoelectric point (net charge); separation may be performed via ion exchange chromatography. Further, protein separation may be carried out on a size exclusion column via chromatography according to the size or molecular weight of the protein. Thus, a customized purification protocols has to be designed to meet the requirements of a particular recombinant protein.

Hence, on the basis of the knowledge above, we purified GFP-STAT3βtc from contaminant proteins and non-protein macromolecules through ammonium sulphate precipitation, ion exchange chromatography and size exclusion chromatography. All polyhistidine tagged recombinant proteins were purified by affinity chromatography.
2.2.3.1.1 Ammonium sulphate precipitation

This technique was employed to clarify and to concentrate target and contaminating proteins present in the soluble cell extract. The technique is based on the fact that at high salt concentrations, the natural tendencies of recombinant proteins not to aggregate is overcome, since the charges are neutralized. Thus, the stepwise addition of the \((\text{NH}_4)_2\text{SO}_4\) salt gradually removes layers of water present at the protein's hydrophilic amino acid surface thereby inducing protein aggregation and subsequent precipitate pellet following centrifugation. Since individual proteins would aggregate at a characteristic \((\text{NH}_4)_2\text{SO}_4\) concentration, the technique provides a simple and easy way of targeting a particular protein in a mixture. Protein precipitates normally refold in the presence of aqueous buffers. In most cases, proteins precipitate at 70 % ammonium sulphate.

Thus, the cell extract containing the recombinant, GFP-STAT3βtc was subjected to various bouts of precipitation to isolate the target protein. To identify the salt concentration at which the GFP-STAT3βtc predominately precipitates, the cell extract was transferred into a fresh 100 mL beaker submerged in an ice - water bath (0 °C). The extract was then treated to 10, 20, 30, 40, 45, 50, 60 and 80 % \((\text{NH}_4)_2\text{SO}_4\) saturation using \((\text{NH}_4)_2\text{SO}_4\) granule at 243 g per L cell extract. The slurry extract-\((\text{NH}_4)_2\text{SO}_4\) mixture was incubated at 0 °C for 45 min with stirring and then transferred into a fresh 40 mL Beckman centrifuge tube. The slurry was cleared by centrifugation at 27,000 xg for 45 min at 4°C using a JA25.50 rotor in a Beckman Coulter centrifuge (Avanti J-E). The clear supernatant was transferred into a fresh beaker and then further saturated to 20 % with ammonium sulphate granules. Whilst incubating in an ice-water mixture for 45 min with stirring, the precipitate protein pellet was re-suspended in a buffer containing the following ingredient; 20 mM HEPES, 100 mM NaCl, 10 mM MgCl₂ and 5 mM DTT. The suspension was stored at -20 °C for further analysis whilst the 20 % cell saturated extract-\((\text{NH}_4)_2\text{SO}_4\) slurry mixture was further cleared by centrifugation, similar to that described above. The generated supernatant was further saturated to 30 % with \((\text{NH}_4)_2\text{SO}_4\) granules. The process was repeated until the extract was saturated to 80 % with \((\text{NH}_4)_2\text{SO}_4\), by which GFP-STAT3βtc would have been precipitated. The protein precipitate was either stored at 4 °C for short term storage or was immediately subjected to further downstream purification steps.

The precipitates suspensions were analyzed by SDS-PAGE, section 2.2.5.1.1, and further evaluated by fluorescence using the PerkinElmer EnVision plate reader. Usually, GFP-STAT3βtc protein is precipitated at 40 % or 1.6 M \((\text{NH}_4)_2\text{SO}_4\) saturation, i.e., 4.68g \((\text{NH}_4)_2\text{SO}_4/20\) mL extract.
2.2.3.1.2 Ion exchange chromatography

By ion exchange chromatography, proteins and other molecules in solution are separated based on differences in their net charge. Negatively charged molecules are captured by positively charged solid supports and positively charged molecules bind to negatively charged solid supports. Sodium chloride (NaCl) is used to elute the bound proteins regardless of the solid support used, i.e., a cation or an anion support. In the anion application, as in the case of GFP-STAT3\(\beta\)tc, the counter ion is the chloride ion (Cl\(^-\)) because the Cl\(^-\) ion is exchanged for the bound target and contaminant proteins which are then eluted. Conversely, in the cation application, the sodium ion (Na\(^+\)) is the counter ion because the Na\(^+\) is exchanged for the target and other contaminant protein that was bound to the column. The strength of the ionic interaction between the target protein and the solid support is a function of the pi of the target protein and the pH of the bind or elution buffer.

Hence, on the basis of the knowledge discussed above, GFP-STAT3\(\beta\)tc was further purified by anion exchange chromatography. GE Healthcare’s HiTrap QFF™ anion exchange column containing the quaternary ammonium ion resin (R-CH2-N+-(CH3)3Cl (Q)) was used for the purification. The column was prepared by washing with 4 column volumes (20 mL) of filtered 20 mM NaCl solution and then equilibrating with buffer S (20 mM HEPES pH 8.0, 10 mM MgCl\(_2\), 20 mM NaCl, 5 mM DTT and a SIGMAFAST™ protease inhibitor tablet (1 tablet per 100 mL)). Due to the theoretical pi of the recombinant protein, i.e., 6.2, the 40 % saturated (NH\(_4\))\(_2\)SO\(_4\) precipitate pellet (judged to contain predominantly GFP-STAT3\(\beta\)tc), section 2.2.3.1.1 was suspended in the pH8 buffer S to 4% (w/v). An aliquot, usually 100 \(\mu\)L was taken and then the suspension was filtered with 0.22 \(\mu\)m syringe filter to remove insoluble debris. The filtrate was then incubated on ice for 1 h. Following incubation, the filtrate (10 mL) was transferred into a 20 mL syringe barrel connected to the anion exchange column by an adapter. The plunger was gently inserted and pressure gently applied to generate approximately 2.5 mL/min flow rate. The eluate collected was labeled as “flow through” (FT) (FT may contain unbound target protein and contaminant proteins) and stored on ice. The column was then washed with 2 column volumes of buffer E (20 mM HEPES pH 7.0, 10 mM MgCl\(_2\), 30 mM NaCl, 5 mM DTT and a SIGMAFAST™ protease inhibitor tablet (1 tablet per 100 mL)). The eluted fraction was labeled as wash (W) and stored on ice. The eluate may contain protein samples that were loosely bound to the column. The bound protein was subsequently eluted linearly with 2 column volumes of buffer E supplemented with increasing concentration of NaCl, i.e., 100, 150, 200, and 300, 400 and 1000 mM. The process was repeated for the remaining 30 mL precipitate suspension following regeneration of the anion exchange column with 2 column
volumes of buffer E supplemented to 1000 mM NaCl and 4 column volumes of buffer S. The collected eluates were pooled and stored on ice for further analysis, i.e., either subsequent purification steps or assays development or analytical studies. Usually, 100 μL aliquots were collected at each stage of the purification process for analysis. Finally, the column was flushed with 20 mM NaCl solution and 20 % ethanol respectively before storage at 4 °C.

2.2.3.1.3 Size exclusion chromatography (gel filtration)

Size exclusion chromatography (SEC), also called gel-filtration or gel-permeation chromatography, separate molecules according to their hydrodynamic volume or size, by which, the molecular weight of the protein molecule can also be elucidated, chapter 4. By using porous particles, molecules that are smaller than the pore size can enter the particle and therefore have a longer path and longer transit time than larger molecules that are too big to penetrate the pore and therefore elute in the interstitial or void volume of the column. Hence, high molecular weight molecules elute faster from the SEC column than lower molecular weight components.

SEC is a relative technique for determining the molecular weight of a molecule, but not an absolute one; hence by necessity, the SEC column must be calibrated with molecular standard of known molecular weight.

Based on the fractionation potential of SE column, the buffer E eluate containing the GFP-STAT3βtc (i.e., fluorescent fractions eluted at 100 and 150 mM NaCl) was applied to the column for further purification. The superose S-12 (12 10/300GL) Tricorn™ high performance SEC column pre-packed with composite of cross-linked agarose powder was used in this study. The column was connected to an ÄKTA Purifier chromatographic system stored in a refrigerator at 4 °C and supported by unicorn 4.0 I software for data collection and analysis.

The pooled 40 mL fluorescent fraction (judged to contain the GFP-STAT3βtc) eluted with 100 and 150 mM NaCl supplemented buffer E, section 2.2.3.1.2, was filtered with 300,000 MWCO filter. The resultant filtrate was concentrated to 10 mL using a 100,000 MWCO filter by centrifugation at 2500 xg at 4 °C. The concentrated fluorescent retentate was finally cleared with a 0.22 μm filter to remove any insoluble protein precipitates or particulates, e.g., fibers from the filter membranes.

In parallel, the Superose column (S6 S12) was washed with 2 column volumes (48 mL) of filtered and degassed ddH₂O at 0.2 mL/min to remove the 20 % ethanol used
storage. Unless stated otherwise, the column was equilibrated with 2 column volumes, 0.5 mL/min, filtered and degassed buffer G (20 mM HEPES, pH 7.0 10 mM MgCl₂, 150 mM NaCl, and 5 mM DTT).

Following equilibration, the concentrated GFP-STAT3ptc retentate was loaded, i.e., 50 to 100 µL for S6 or 100 to 500 µL for S12, into the injection loop (either 200 µL or 1000 µL for S6 or S12 respectively) and then injected onto the column with buffer G at a flow rate of 0.5 mL/min (S6) or 1.0 mL/min (S12). Eluates were collected at a flow rate of 0.5mL/min. In addition to absorbance at 280 nm, the fluorescent tag of GFP-STAT3ptc allowed for visual tracking of the recombinant protein as it migrates through the column. Eluates were either stored in 10 % glycerol at −20 °C for future use and were quantified indirectly using a Bio-Rad assay kit via bound coomassie blue absorbance or directly via the fluorophore tag with both using Beer-Lamberts law (where absorbance a = GFP extinction factor (e)* optical cell length (l) * concentration (c)).

2.2.3.1.4 Affinity purification of polyhistidine tag recombinant proteins

Affinity purification is a technique that allows separation of a recombinant protein of interest in solution (mobile phase) by the specific interaction of the protein to a particular ligand that has been immobilized to a stationary material (solid phase). The solid phase is a matrix material to which a bio-specific ligand is covalently attached. During affinity purification, cell extract containing the solute of interest is passed over the immobilized ligand resin in a buffer condition that facilitates binding. Upon binding of the target protein, unbound components of the extract are washed off with additional buffer. By using an altered buffer condition, the bound protein is eluted from the resin by disrupting ligand-protein interaction.

In our study, we employed GE Healthcare's HisTrap FF™ Crude column, pre-packed with Ni²⁺ Sepharose™ 6 solid phase to purify the hexahistidine tagged recombinant proteins construct; His-eCFP, His-eGFP, His-eYFP and His-eCFP-TEV-eYFP. Prior to purification, the column was washed with 6 column volumes of filtered and degassed ddH₂O with the aid of a 20 mL syringe. It was then equilibrated with 4 column volumes (20 mL) of the wash buffer or equilibration buffer (20 mM Na₂HPO₄, pH8.0, 20 mM imidazole and 500 mM NaCl). The EDTA and DTT free cell extract (10 mL) generated in either sections 2.2.2.1.1 or 2.2.2.1.2 was then passed through the column at a rate of 2.5 mL/min. The FT fraction containing the unbound contaminant protein samples was collected into a 30 mL universal for further analysis. The bound target protein was subsequently washed with 10 column volumes of the wash buffer. The eluates were collected into 50 mL falcon tubes and stored on ice for future analysis. The washed
column was then eluted linearly with 3 column volumes of the wash buffer supplemented with 100, 200, 300, 400 and 500 mM imidazole. Eluates were collected in 30 mL white cap universal and kept on ice. 1 mL fractions of the collected eluates were transferred into 1.5 mL sterile tubes and stored at -20 °C for analysis. The column was regenerated with 2 column volumes (10 mL) of wash buffer supplemented with 1 M imidazole. It was subsequently equilibrated with 4 column volumes (20 mL) of wash buffer. The column was either use to purify the remainder of the cell extract or prepared for storage by passing 20 % ethanol through. The column was then stored at 4 °C.

Following purification, the collected 100, 200, 300, 400, and 500 mM imidazole eluates were transferred into separate pre-boiled dialysis tube. The cut edges of the tubes were properly secured with 2 x knot and a clip to prevent the eluates from leaking into the dialysis buffer (20 mM HEPES, pH 7.0, 200 mM NaCl, 5 mM EDTA, 5 mM DTT and protease inhibitor cocktail 1/200 dilution). Each tube was properly labeled and then transferred into the 2 L dialysis buffer prepared in a 4 L beaker. Dialysis was carried out for 16 to 24 h whilst stirring at 4 °C.

After dialysis, the concentrations of the purified fluorescent protein fractions were measured. The pure protein fractions were supplement with 10 % glycerol and divided into aliquots before storage at -80 °C.

2.2.4 Protein quantification techniques

2.2.4.1 Protein quantification using the Bradford assay (Bio-Rad reagent)

To generate a standard curve, 10 mg/mL stock solution of Bovine Serum Albumin (BSA) was diluted to a final concentration of 5 mg/mL using 50 mM HEPES pH7.5. In a 96 well plate, dilutions of the new stock solution were prepared to 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8 and 2.0 mg/mL to a total volume of 50 μL. The test samples were diluted 1/10 and 1/100 with 50 mM HEPES pH 7.5. Coomassie brilliant blue G-250 protein assay reagent, 200 μL, was added to each well. The samples were incubated at room temperature for 30 min and then the absorbance at 595 nm measured using Perkin Elmer plate reader programmed to read “assay 595 nm”. Together with the concentrations of the protein standards, the calibration curve was constructed to determine the total protein concentrations of the test samples.

2.2.4.2 Protein quantification by intrinsic GFP absorbance

In addition to providing us with an idea of the functional state of its fusion partner, a direct quantitative correlation exists between the GFP fluorescence intensity and the
concentration of the fusion protein. Hence, for more sensitive applications including; FRET, PEMSA and fluorescence-based microtiter plate assays, the applied protein sample was quantified by direct GFP fluorescence, hence, quantification by Beer-Lambert law (Equation 2.1) was applied:

\[ A = \varepsilon I c \]

Equation 2.1: Beer-Lambert Law where \( A \) = absorbance, \( \varepsilon \) = extinction factor, \( I \) = optical cell length and \( c \) = concentration.

The absorbance was measured in a 1 cm cell using a spectrophotometer (Biochrom Libra S22, UK). Prior to measurement, the spectrophotometer was set to zero using buffer G at the appropriate absorption wave length, i.e., 435, 485 and 514 nm for eCFP, eGFP and eYFP fluorescent tags respectively.

Test samples were diluted 1/25 and 1/50 with elution buffer G and absorbance measured at their respective wavelengths. With knowledge of their extinction factors; 28750, 55000 and 83400 M\(^{-1}\)cm\(^{-1}\) for eCFP, eGFP and eYFP respectively (Karasawa et al. 2004; Heim et al., 1995; Tsien et al., 1998), actual concentration of the fusion protein was estimated using Beer-Lamberts equation. Approximately 1 \( \mu \)M eCFP, eGFP and eYFP tagged STAT3\( \beta \)tc gave a peak absorbance of 0.028, 0.055 and 0.084 respectively.

Alternatively, the intrinsic absorbance at 280 nm may be used where many proteins exhibit an absorbance of 1 optical density (OD) for 1 μg/\( \mu \)L sample. The exact extinction coefficient is dependent on the percentage content of the tyrosine, tryptophan and phenylalanine content of the target protein.

2.2.5 Protein analytical techniques

2.2.5.1 Electrophoresis

2.2.5.1.1 Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE was performed under reducing conditions using a Bio-Rad mini gel system. Unless otherwise stated, 10 % polyacrylamide gels were used for analysis. Prior to gel preparation, the glass plates, i.e., hinged spacer plates; 10 cm x 8.3 cm and the short outer glass plates; 10 cm x 7.3 cm, were rinsed with 70 % ethanol and then ddH\(_2\)O. The short outer glass plate was placed on top of the hinged spacer plate, and the pair
assembled in the casting frames and then mounted on the casting stands. For electrophoresis, the plates were assembled into the buffer dam and then transferred into the Bio-Rad Mini-Protean Tetra system's buffer reservoir.

2.2.5.1.2 Preparation of resolving gel

Unless stated otherwise, all resolving polyacrylamide gels contained 10 % (w/v) acrylamide. Gels were prepared as follows: 6 mL Tris-HCl pH8.8, (400 mM), 5.0 mL ProtoGel (37: 5: 1 acrylamide to bis-acrylamide stabilized solution), 0.12 mL 10% SDS (0.08 % (v/v)) and 0.012 mL N,N,N'N'-tetramethylethylene diamine (9.3 mg/mL) (TEMED) made up to 15 mL with ddH₂O. Polymerization was initiated by the addition of 0.12 mL 0.1 mg/mL ammonium persulphate solution (APS). The solution was poured gently into the space between the two glass plates to about 3 cm from the top of the plate short glass plate. To ensure a uniform interface between the stacking and resolving gels, about 1 mL ddH₂O was gently overlaid on top of the resolving gel as it set. The gel was allowed to polymerize for 1 h at room temperature.

2.2.5.1.3 Preparation of stacking gel

The stacking gel mixture contains 5 % (w/v) acrylamide. The mixture includes the following; 0.676 mL 1M Tris-HCl pH6.8 (400 mM), 0.867 mL ProtoGel (37: 5: 1 acrylamide to bis-acrylamide stabilized solution), 0.052 mL 10% SDS (0.08 %), 0.0052 mL N,N,N'N'-tetramethylethylene diamine (TEMED) made up to 5 mL with ddH₂O. Polymerization was initiated by adding 0.052 mL 0.1 mg/mL ammonium persulphate (APS).

The stacking gel was immediately added onto the resolving gel and the appropriate comb (10 wells; 5mm x 1mm or 15 wells; 3 mm x 1mm) positioned. The stacking gel was allowed to polymerize for a further 30 min.

Upon polymerization, the gel cast was released from the cast frame and stand and subsequently fitted to the buffer dam. The gel cassette was introduced into the buffer reservoir tank, and the tank was filled with electrode buffer (25 mM Tris-HCl, pH 8.2, 250 mM Glycine and 0.1 % (w/v) SDS) to the appropriate mark. Residual polyacrylamide in the wells was flushed with tank buffer.

2.2.5.1.4 Preparation of protein sample for SDS-PAGE

To each protein sample, i.e., either crude or pure, 5 µL 6x SDS-PAGE loading dye buffer (360 mM Tris-HCl, pH 6.8, 60 % (v/v) glycerol, 12 % (w/v) SDS and 0.06 % (w/v) bromophenol blue) and 1µL 1M DTT (33.3 mM) were made up to 30 µL with
ddH₂O and allowed to dry. The preparation was mixed thoroughly and then centrifuged at 13,000 xg for 15 s. It was then boiled for 10 min at 95 °C and further centrifuged at 13,000 xg for 15 s to drain the condensed vapour to the bottom of the tube as well as sediment particulates and insoluble materials.

2.2.5.1.5 Electrophoresis of denatured protein samples

Up to 15 μL (10 wells) or 12 μL (15 wells) of the treated protein sample (section 2.2.5.1.4) was loaded into the gel lanes using a long loading tip. One lane per gel had 10 μL pre-stained broad range protein standards. The New England BioLab's (NEB) pre-stained protein marker, contained the following pre-stained proteins (kDa): MBP-β-galactosidase; 175 kDa, MBP-paramyosin; 80 kDa, MBP-CBD; 58 kDa, CBD-Mxe Intein-2CBD; 46 kDa, CBD-Mxe Intein; 30 kDa, CBD-BmFKBP-13; 25 kDa, Lysozyme; 17 kDa and Aprotinin; 7 kDa. The gel was operated at 40 volts through the stacking gel and 70 to 100 volts through the resolving gel. Once the bromophenol blue dye front reached the bottom of the gel, the run was stopped and the gel stained with coomassie brilliant blue or transferred onto a nitrocellulose membrane Hybond-C Extra (GE Healthcare, Buckinghamshire).

2.2.5.2 Electrophoretic transfer of denatured protein onto nitrocellulose membrane

Upon successful separation of the SDD-protein complex via SDS-PAGE, the denatured protein sample was transferred onto a nitrocellulose membrane (Towbin et al., 1979). The transfer of protein from polyacrylamide gel was performed in a BioRad Mini Protean® 3 Western Trans-blot system. The system includes; electrode module, gel holder cassettes (10 cm x 11 cm), bio-ice cooling unit and buffer chamber and lid. These accessories are thoroughly cleaned before use.

2.2.5.2.1 Preparation for blotting

The following accessories were prepared prior to blotting; the bio-ice cooling unit was filled with ddH₂O and frozen at -20 °C until ready to use, 800 mL transfer buffer (31 mM Tris, pH8.5, 192 mM Gylcine, 20 % methanol diluted to 800 mL with ddH₂O), 4 x 3 MM paper and 1 x nitrocellulose membrane cut to the dimension of the gel. The gel, nitrocellulose membrane, 3MM paper and fiber pads were soaked in the transfer buffer for 7-10 min. The sandwich was prepared in a tray partially filled with the transfer buffer as follow;

- The cassette was positioned with the grey side down.
- One pre-wetted fiber pad was placed on the grey side.
- 2 x pre-wetted 3MM papers were placed on top of the fiber pad.
The equilibrated gel was placed on the 3MM paper.

2 x pre-wetted 3MM papers were placed on the gel.

A second pre-wetted fiber pad was placed on the 3MM paper.

A glass tube (usually glass bottle) was gently rolled over the sandwich to remove bubbles.

The cassette was firmly closed and locked without disturbing the gel and 3MM paper.

Following preparation of the sandwich-cassette, it was placed in the module and then the whole unit transferred into the tank. The Bio-Ice cooling unit was also added. The tank was completely filled with transfer buffer. The lid was positioned and the blot was run for 1 h at 150 mA and 100 V.

Upon completing transfer, the sandwich was disassembled and the membrane removed for development, section 2.2.5.2.1. The system was cleaned with a detergent and rinsed with ddH₂O.

2.2.5.2.2 Development of nitrocellulose membrane

The transferred protein was visualized by immune-staining using commercially available antibodies. Upon protein transfer, the membrane was placed in a petri dish containing the blocking medium (10 mg/mL BSA in Tris-Saline buffer (TS); 10 mM Tris, pH 7.0, 150 mM NaCl), 15 mL per membrane. The membrane was either blocked for 1 h at room temperature or overnight at 4 °C. The blocking medium was replaced with a fresh blocking medium supplemented with Igepal to 0.05 % (non-ionic detergent) and the primary antibody at a 1/5000 dilution. The medium was either incubated at room temperature for 1 h 30 min or overnight at 4 °C. After probing the membrane with the primary antibody, it was washed with TS buffer (15 mL) for 10 min. The buffer was discarded and the membrane was further washed twice with a TS-Igepal buffer (10 mM Tris, pH 7.0, 150 mM NaCl, 0.05 % Igepal) for 2x 10 min. A final washing step with TS buffer (15 mL) was carried out for 10 min. After discarding the wash buffer, the secondary antibody, i.e., either horseradish peroxidase (HRP)-linked anti-rabbit IgG antibodies or HRP-linked anti-mouse IgG antibodies, was added to the TS-Igepal buffer at a 1/5000 dilution and then used to treat the membrane for 1 h 30 min with agitation by gentle rocking. The membrane was then subjected to several washes as follows; TS buffer (10 min), 2x TS-Igepal (2x 10 min) and then TS buffer (10 min). Following the washing steps, the transferred protein was visualized either by chemiluminescence or by using Sigma's SIGMA FAST™ 3, 3'-Diaminobenzidine (DAB) tablets.
Edwin Nkansah

Prior to chemiluminescence, the membrane was washed briefly with 15 mL 1 x PBS. The frozen 0.8 M luminol solution (100 mM Tris-HCl, pH 8.5, and 0.8 M luminal; 10 mL aliquots) and 11 mg/mL hydroxycinnamic acid (p-coumaric acid; 100 µL aliquot) were thawed at room temperature. The 100 µL p-coumaric acid was added to the 10 mL luminol solution and mixed thoroughly. In addition, 7 µL H$_2$O$_2$ was also added to the luminal – coumaric acid mix and the mixture vortexed to ensure uniform mixture. After discarding the 1x PBS wash buffer, the membrane was exposed to the luminal-coumaric-H$_2$O$_2$ mix for 1 min. The mixture was then discarded and the membrane carefully blotting at edge using a paper towel to remove excess luminal-coumaric-H$_2$O$_2$ mix. The membrane was transferred into a plastic sleeve and then imaged using GeneGenome chemiluminescence Capture and Analysis System (Syngene, Cambridge, UK).

Using SIGMA FAST™ 3, 3'-Diaminobenzidine (DAB) tablets, the membrane was stained with 20 mL DAB-Urea-H$_2$O$_2$ (1 DAB tablet and 1 urea H$_2$O$_2$ tablet in 20 mL ddH$_2$O). DAB is a precipitating substrate used to detect peroxidase activities through the development of an intense brownish-black stain whose production is described in the reactions below;

\[
\text{Peroxidase} + 2\text{H}_2\text{O}_2 \rightarrow \text{O}_2 + 2\text{H}_2\text{O} \text{ (pH 7.6)}
\]

\[
\text{O}_2 + \text{DAB} \rightarrow \text{Insoluble, brownish-black precipitate}
\]

Upon observing the immunoreactive bands, the DAB medium was discarded and the membrane was washed 3 times with ddH$_2$O. The stained membrane was dried between 2x 3 MM papers and then stored in a new set of 3 MM papers. The data generated was documented by scanning the image using HP PSC 1410. In both cases, the membrane was preserved between 2x 3 MM papers.

2.2.5.3 Non-denaturing polyacrylamide gel electrophoresis (Native-PAGE)

2.2.5.3.1 Triton X-100 polyacrylamide gel (Triton native gel)

Native Triton X-100 gel was performed under non-reducing conditions using Bio-Rad Tetra gel system. Percentage acrylamide used for analysis included 5, 6, 7 and 8. The apparatus used for the analysis was prepares as described in section 2.2.5.1.1.

2.2.5.3.2 Preparation and equilibration of Triton X-100 resolving gel

The percentage of acrylamide used was dependant on the nature of the experiment. A 5 % Triton X-100 gel, a gel mix identical to that described in section 2.2.5.1.2 was
prepared. However, instead of 0.08 % SDS, the mixture was supplemented with 0.08 % Triton X-100. Also, the percentage acrylamide was adjusted to 5% by adding 2.5 mL (37: 5: 1 acrylamide to bis-acrylamide stabilized solution) ProtoGel to the mixture. Upon adding APS, the gel solution was poured gently into the space between the two glass plates until it was completely filled. The appropriate comb (10 wells; 5mm x 1mm or 15 wells; 3 mm x 1mm) was positioned since no stacking gel was required. The gel was allowed to polymerize for 2 h at room temperature. After polymerization, the plates were assembled as described in section 2.2.5.3.1. The gel cassette was transferred into the buffer reservoir and then filled with electrode buffer supplemented with 0.1 % (w/v) Triton X-100. The wells were thoroughly flushed with buffer and the gel equilibrated by pre-running for 2 h at 150 V in the cold room with the electrophoresis tank packed in ice.

2.2.5.3.3 Preparation of protein sample for native-PAGE

To each protein sample, 5 μL 6x Triton X-100 loading dye buffer (360 mM Tris-HCl, pH 8.8, 60 % (v/v) glycerol and 12% (w/v) Triton X-100) and 40 mM DTT was made up to 30 μL with ddH2O. The sample was gently mixed and then centrifuged at 13,000 xg for 15 s to drain the protein mix to the bottom of the tube as well as sediment particulates and insoluble materials. The sample was incubated at room temperature for 30 min and then further centrifuged at 13,000 xg for 15 s before prior to loading.

2.2.5.3.4 Electrophoresis of non-denatured protein samples

Up to 7.5 μL (10 wells) or 5 μL (15 wells) of the treated protein sample was loaded into the gel lanes using a long loading tip. One lane per gel contained 10 μL (~5 μg) BSA (66 kDa) mixed with 6x Triton X-100 loading dye buffer (360 mM Tris-HCl, pH8.8, 60 % (v/v) glycerol and 12 % (w/v) Triton X-100 and 0.06 % (w/v) bromophenol blue). The marker lane served a dual purpose, i.e., as the gel front and as the sole protein marker. Electrophoresis was performed at 100 volts until the bromophenol blue dye had migrated approximately 2/3 to 3/4 down the length of the gel. The run was stopped and the separated GFP tagged parent proteins were visualized following illumination of the gel with blue light. Images of the separation were capture with the Kodak DC120 206 M digital camera. After imaging and subsequent documentation, the gel was either stained with coomassie brilliant blue dye or the fractionated protein transferred onto a nitrocellulose membrane via western blotting, sections 2.2.5.2., 2.2.5.2.1 and 2.2.5.2.2.
3 RECOMBINANT DNA

Cloning involves seven main steps that include; choosing the cloning vector and the host organism (e.g. *E.coli* cell), preparation of the cloning vector (e.g., restriction digest), preparation of the cloning insert DNA (e.g., PCR, restriction digest etc.), creation of recombinant DNA (e.g., ligation of insert into vector), transforming the host organism with recombinant DNA, screening for host organism containing recombinant DNA (e.g., antibiotic resistance marker) and screening for the clones with the desired DNA (e.g., nucleotide sequencing).

3.1 Creating fluorescent expression constructs

3.1.1 Fluorescent STAT3βtc recombinant DNA expression

In order to investigate molecular interactions involving the fluorescent STAT3βtc chimaeric protein, fluorescent expression constructs were developed containing STAT3βtc fused to either N- or C-termini of the spectral variants of the GFP, i.e. enhanced Cyan, Green and Yellow fluorescent protein. The constructs were designed to verify the appropriate order of fusion that will not disturb the biological functionality of STAT3βtc *in vitro*.

3.1.1.1 Developing GFP-STAT3βtc expression constructs (pET-32a (+) GFP-STAT3βtc)

In generating the N-terminal GFP tagged STAT3βtc constructs, pET-32a (+)-GFP-STAT3βtc, the various cDNA's of the spectral variants of GFP were PCR-amplified and ligated into the cloning vector containing STAT3βtc, pET-32a (+)–STAT3βtc.

3.1.1.1.1 PCR-amplification of GFP cDNA variants

In generating the fluorescent expression constructs, pET-32a (+)-GFP-STAT3βtc, encoding the following GFP-STAT3βtc chimaeric proteins: enhanced Cyan Fluorescent Protein; eCFP-STAT3β, enhanced Green Fluorescent Protein; eGFP-STAT3β and enhanced Yellow Fluorescent Protein; eYFP-STAT3β, the cDNA's coding for all three spectral variants of GFP *i.e.* peCFP-N1, peGFP-N1 and peYFP-N1 respectively, were PCR-amplified to introduce Ndel restriction endonuclease sites at the 5' and 3' ends. The PCR product was subsequently restricted with Ndel restriction endonuclease, section 3.1.1.1.2. The restricted PCR product was cloned into the Ndel site contained in the pET-32a (+)-STAT3βtc cloning vector (Ndel site at the 5' end of STAT3βtc). Below is a schematic representation of the cloning procedure:
Primers were designed to introduce of Ndel sites at both the 5' and 3' ends of the GFP (-) insert DNA by PCR-amplification. After amplification, the GFP insert is purified, restricted with Ndel restriction endonuclease and further purified. The restricted GFP insert is then ligated into an acceptor pET-32a (+)-STAT3βtc cloning vector that has been restricted with Ndel. Following transformation, the cells are selected with the appropriate antibiotic (ampicillin) for a clone.

Here, cDNA coding for the spectral variants of GFP were PCR-amplified and ligated to the 5' end of STAT3βtc. Both forward and reverse oligonucleotide primers were designed to introduce Ndel restriction endonuclease site at the 5' and 3' ends of the GFP insert gene. Care was taken not to introduce a stop codon at the 3' end of the gene. The PCR products were generated using the following primers: forward primer; 5'-dAAAACATATGGTGAGCAAGGGCAGG-3' (Tm: 56 °C, GC content (%): 59) and reverse primer; 5'-dAAAACATATGCCTTGACTACGCTCTCCATG-3' (Tm: 55 °C and GC content (%): 58). As described in section 2.2.1.2.7., a 50 μL PCR reaction cocktail was prepared with approximately 0.5 μM of each primer and about 0.5 μg cDNA templates. A control reaction mixture containing all the ingredients of the PCR “cocktail” except, the cDNA template was also prepared in parallel. Following amplification, the PCR tube was centrifuged for 15 s to drain any condensed vapour to the bottom of the tube. The PCR-amplification was performed in a thermo cycler (Biometra, Cambridge, UK) using the condition described in Table 2.4. A similar protocol was used to amplify eCFP and eYFP cDNA's.
Edwin Nkansah

From the product; 2.0 μL aliquots was taken (aliquots used analyzed by agarose gel electrophoresis) and the remaining sample was purified as described in section 2.2.1.2.8. The bound DNA product were eluted with 35 μL ddH2O into a sterile 1.5 mL tube and then stored on ice in preparation for restriction digests, section 3.1.1.1.2.

3.1.1.1.2 Preparation of cloning vector and insert DNA by restriction digest

Firstly, Ndel restriction endonuclease and its corresponding restriction buffer (D) (Promega, UK) were used to selectively cleave the cloning vector, pET-32a (+)-STAT3βtc, at the Ndel recognition site. The Ndel restriction site is located at the 5’ end of STAT3βtc gene (cloning site for restricted PCR products). The cloning DNA vector was restricted in a 30 μL reaction prepared as follows; approximately 2.7 μg (6.0 μL) of cloning vector was diluted to 30 μL with 1x restriction buffer D (NEB) (6.0 mM Tris-HCl pH 7.9, 150 mM NaCl, 6 mM MgCl2, and 1.0 mM DTT) and approximately 20 units Ndel restriction endonuclease. The restriction mixture was gently mixed and then incubated at 37 °C for 2 h. The restriction endonuclease was subsequently de-activated by heating the mixture at 65 °C for 20 min. The restricted reaction was then centrifuged for 15 s at 13,000 xg. A 1.0 μL aliquot labelled "Restricted" was stored on ice. The aliquot was used to estimate the efficiency of the restriction process.

To avoid self-ligation or re-circularization, the restricted cloning vector was de-phosphorylated as describe in section 2.2.1.2.9. Five units of the phosphatase enzyme (Antarctic phosphatase) and its corresponding buffer (1x concentrated) were diluted to 30.0 μL with 25.0 μL of the restricted cloning vector. The mixture was further incubated at 37 °C for 1 h during which the 5’ phosphate groups at the "TA" sticky ends were removed by the phosphatase enzyme. The phosphatase was also at 65 °C for 20 min. A 1.0 μL aliquot labelled "R_D" (i.e. restricted and de-phosphorylated) was taken for analysis.

Secondly, the GFP PCR product was prepared for ligation with Ndel restriction endonuclease. Similarly, the PCR product was restricted in a 30 μL reaction containing the following ingredients; 25 μL PCR products (either eCFP or eGFP or eYFP) was diluted to 30 μL with 1x restriction endonuclease buffer D (6.0 mM Tris-HCl pH 7.9, 150 mM NaCl, 6 mM MgCl2, and 1.0 mM DTT) with approximately 20 units Ndel restriction endonuclease enzyme. The reaction was mixed gently and then incubated at 37 °C for 2 h. The enzyme was also subsequently de-activated at 65 °C for 20 min.

The restricted/dephosphorylated cloning vector together with the 3 different restricted GFP inserts were purified by agarose gel electrophoresis (0.8% w/v), see section
2.2.1.2.3. The concentrations of the purified cloning vector and that of the insert DNA, i.e., eCFP, eGFP and eYFP were measured as; 0.075, 0.072, 0.064 and 0.060 μg/μL respectively. A total of 29.0 μL of each DNA fragment was left after 5.0 μL was used to generate the concentrations of the ligation reactants.

3.1.1.1.3 Ligation reaction

The purified and restricted PCR product was sub-cloned into the restricted/dephosphorylated cloning vector by adding a T4 DNA ligase and its corresponding buffer. As previously described in section 2.2.1.2.10, “vector: insert” ratios 1:3 (one part of vector to three parts of insert) and 1:6 (one part of vector to six parts of insert) were applied. Firstly, a 1:1 vector to insert ratio (ng) was calculated based on the concentrations of the ligation reactants describe in section 2.2.1.2.10. Unless stated otherwise, the volume of the ligation mixture was made up to 10 μL.

Firstly, in creating eCFP-STAT3ptc, the ligation reaction mixture for ratio 1:3 was prepared as follows; 200 ng cloning vector and 59.7 ng eCFP insert PCR product were supplemented 1x T4 DNA ligase buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM (DTT) and 1 mM ATP) and 3.0 units T4 DNA ligase enzyme. For ratio 1:6; 200 ng cloning vector and 119.40 ng eCFP insert PCR product were supplemented 1x T4 DNA ligase buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM (DTT) and 1 mM ATP) and 3.0 units T4 DNA ligase enzyme.

Secondly, the construct eGFP-STAT3ptc was created through the ratio 1:3 as follows: 200 ng cloning vector and 60.5 ng eGFP insert PCR product were supplemented 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme. Further, ratio 1:6 was also prepared as follows; 200 ng cloning vector and 120 ng eGFP insert PCR product were supplemented 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme.

Thirdly, in creating expression construct eYFP-STAT3ptc, the ratio 1:3 was initially prepared as follows; 200 ng cloning vector and 59.7 ng eYFP insert PCR product were supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme. On the other hand, ratio 1:6 was prepared as follows; 200 ng cloning vector and 119.4 ng eGFP insert PCR product were supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme.

Finally, two control reactions similar to the reactions described above were prepared in parallel. The first control reaction was prepared minus the PCR products and thus contained 200 ng cloning vector supplemented with 1x T4 DNA ligase buffer and 3.0
units T4 DNA ligase enzyme whilst the second reaction contained no PCR insert product and T4 DNA ligase, but contained 200 ng cloning vector supplemented with 1x T4 DNA ligase buffer. The control reactions were prepared to establish the efficiency of the restriction, de-phosphorylation and ligation processes. Reactions were diluted to 10 μL with sterile ddH2O. The reactions were incubated at room temperature for 1 h and then at 4 °C for 16 h.

3.1.1.1.4 Transformation of JM109 cells with ligated products

The resultant ligated products were used to transform JM109 (DE3) competent cells, section 2.2.1.1.3. As described in section 2.2.1.1.3., 50 μL of JM109 cells were transformed with the ligated mixture. Following transformation, colonies were cultured on ampicillin (100 μg/mL) selective LB-Agar plates since the cloning vector was ampicillin resistance. A total of 24 LB-Agar plates were used to culture the ligated test and control samples.

In addition to the ligation controls, the efficiency of the transformation process was evaluated with controls prepared as follows; negative control, 50.0 μL JM109 cells were transformed with 1.0 μL sterile ddH2O. In the case of the positive control, 50.0 μL JM109 cells were transformed with approximately 0.045 μg unrestricted or circular cloning vector, pET-32a (+) STAT3βtc. Colonies were documented and plates wrapped in parafilm and stored at 4 °C until the screening. The efficiency of transformation was evaluated by comparing the number of E.coli colonies on the experimental plates to that on the control plates. Colonies on the positive control suggested a successful transformation.

3.1.1.1.5 Purification of potential clone

Colonies cultured in section 3.1.1.1.5 were mini cultured (5 mL) in broth supplemented with 100 μg/mL ampicillin. The cultures were subsequently purified as described in section 2.2.1.1.6. The purified plasmid DNA samples representing the expression constructs; pET-32a (+)-eCFP-STAT3βtc, pET-32a (+)-eGFP-STAT3βtc and pET-32a (+)-eYFP-STAT3βtc were dissolved in 50 μL sterile ddH2O and stored at -20 °C for future analysis.

3.1.1.1.6 Selection of potential clones by restriction analysis

The plasmid DNA samples were linearized with EcoRI restriction endonuclease in a 10 μL reaction volume described as follows; 8.0 μL purified plasmid DNA was supplemented with 1x restriction endonuclease buffer H (1.0 mM Tris-HCl, pH 7.5, 1.0
mM MgCl₂, 40.0 mM NaCl, 0.01 mM EDTA, 0.1 mM DTT, 0.05 mM BSA and 5% glycerol) and 10 units EcoRI restriction endonuclease enzyme. The control restriction reaction containing about 0.5 µg of the original cloning vector was prepared in parallel. The reactions were restricted (section 2.2.1.2.1) and then analyzed by agarose gel 0.8% (w/v) electrophoresis, section 2.2.1.2.2. The gel was illuminated with UV light to select the samples that may be carrying the insert gene. Selected samples were further screened by PCR (see section 3.1.1.1.7).

3.1.1.1.7 Screening of potential clones by PCR

Since the GFP insert was ligated into the cloning vector non-directionally (clones generated with a single restriction endonuclease), it was necessary to confirm the orientation of the ligated insert. The purified samples (section 3.1.1.1.6) were PCR-amplified to identify the putative clonal plasmid DNA sample containing the insert DNA in the right direction. The PCR mixture (section 2.2.1.2.7) contained the forward primer (5'-d/AAACATATGGTGAGCAAGGGCGAGG-3') of the insert GFP gene and the reverse primer (5'-dCTTGGTCTTCAGGAACGGGGCAGCAC-3') of the STAT3βtc contained in the cloning vector. The mixtures were amplified as described in Table 2.4.

Each of the amplified samples (10.0 µL) was diluted to 12.0 µL with 6x loading dye and then analyzed by agarose gel electrophoresis (section 2.2.1.2.2). The amplified DNA fragments, representing the three different fusion genes were visualized under UV light. The amplification of the fusion genes, i.e, GFP-STAT3βtc (~ 2.5 kb), clearly suggested that the GFP insert gene was ligated into the cloning vector in the correct orientation.

3.1.1.1.8 Amplification of suspected clones (Maxiprep, Qiagen)

Following identification of the prospective clones, the expression construct were amplified through maxi-preparation (Qiagen, section 2.2.1.1.7). Aliquots (~ 50 µL) of XL1-Blue competent cells were transformed with plasmid DNA sample representing the prospective fluorescent STAT3βtc expression constructs. A control transformation was also prepared in parallel, see section 2.2.1.1.3. Following from transformation and subsequent mini- (5 mL) and then maxi-cultures (500 mL) (cultures supplemented with 100 µg/mL and 12.5 µg/mL ampicillin and tetracycline), the prospective clones were purified using the Qiagen Maxiprep Kit, section 2.2.1.1.7. The purified DNA samples were dissolved in 0.5 mL sterilized ddH₂O and their concentration estimated by absorbance at 260 nm, where A₂₆₀ = 1.0 for 50 µg/mL DNA. Restriction maps of the purified clones were also prepared, section 3.1.1.1.9.
3.1.1.1.9 Restriction mapping of prospective clones

Restriction maps for each of the fluorescent expression constructs were prepared with EcoRI and Ndel restriction endonuclease enzymes. Each construct was restricted with either EcoRI (to linearize the plasmid DNA) or Ndel (to release the insert GFP) or EcoRI and Ndel (to release both STAT3βtc and the insert GFP genes) in a separate 10 µL reaction. All reactions were prepared in a sterile 0.5 mL tube.

In preparing the EcoRI restriction reaction mixture, approximately 0.13 µg of the purified plasmid DNA was treated with 1x restriction endonuclease buffer H (Promega) (6.0 mM Tris-HCl (pH 7.5), 100 mM NaCl, 6.0 mM MgCl₂ and 1.0 mM DTT) and units EcoRI restriction enzyme. All 3 samples, i.e. eCFP-STAT3βtc, eGFP-STAT3βtc and eYFP-STAT3βtc, were treated as described above.

Similarly, the Ndel restriction reactions were prepared as follows; approximately 0.13 µg of each of the purified plasmid DNA samples was treated with 1x restriction endonuclease buffer D (6.0 mM Tris-HCl (pH 7.9), 150 mM NaCl, 6.0 mM MgCl₂ and 1.0 mM DTT) and 10 units Ndel restriction endonuclease.

In addition, approximately 0.13 µg of each of the purified plasmid DNA samples were restricted with both EcoRI and Ndel in a reaction containing the following ingredients; 1x restriction endonuclease Multi-Core™ buffer (Promega) (25.0 mM Tris-Acetate pH 7.5 (at 37°C), 100 mM CH₃CO₂K, 10.0 mM CH₃CO₂Mg, and 1.0 mM DTT), 6 units EcoRI restriction enzyme and 5 units Ndel restriction endonuclease.

Control restriction reactions were prepared in parallel. The first reaction contained approximately 0.13 µg cloning vector supplemented with 1x restriction endonuclease buffer H (Promega) and 12 units EcoRI restriction enzyme whilst the second was identical to the first, but also contained 10 units Ndel restriction endonuclease.

The samples were restricted (section 2.2.1.2.1) and then analyzed by agarose gel electrophoresis (section 2.2.1.2.2). The separated DNA fragments were visualized under UV light (λ = 312) to verify the integrity of the prospective clones. An image of the separated DNA fragments was captured with the Kodak DC120 digital camera (section 3.8). The image represents the restriction maps of the various expression constructs.
### 3.1.1.10 Nucleotide sequence analysis of selected clones

Following verification of the expression constructs through restriction mapping, approximately 2.5 to 3.0 μg (20.0 μL) of each of the purified DNA sample was transferred into sterile 1.5 mL tube. The samples were then sent for nucleotide sequencing (section 2.2.1.2.11).

### 3.1.2 Cloning pET-28c (+) -GFP expression constructs

In generating fluorescent vectors where the host protein can be fused to either the N- or the C-termini of the fluorescent protein, expression vectors pET-28c (+)-eCFP, pET-28c (+)-eGFP and pET-28c (+)-eYFP were generated by ligating the Ndel/Nhel restricted GFP PCR product into Ndel/Nhel restricted pET-28c (+) vector. Demonstrated below is a schematic representation of the cloning process.

![Cloning process diagram](image)

**Figure 3.2: Summary of the cloning procedure used to develop the expression constructs, pET-28c (+)-eCFP, pET-28c (+)-eGFP and pET-28c (+)-eYFP**

Primers were designed to introduce of Ndel and Nhel sites at the 5’ and 3’ ends of the GFP (—) insert DNA respectively via PCR-amplification. After amplification, the inserts DNA is purified, restricted with Ndel restriction endonuclease and further purified. The restricted GFP insert is then ligated into an acceptor pET-28a (+) cloning vector that has been restricted with both Ndel and Nhel. Following transformation, the cells are selected with the appropriate antibiotic (kanamycin) for a clone.
3.1.2.1 PCR-amplification of eCFP, eGFP and eYFP

A described in Figure 3.2, the cDNA sequences encoding the spectral variant of GFP, i.e., eCFP, eGFP and YFP, were PCR-amplified (section 2.2.1.2.7, Table 2.4) using the following primers; forward, 5'-dAAACATATGGTGAGCAAGGGCGAGG-3' (Tm: 56 °C, GC content (%): 59); reverse, 5'-dAAAAGCTAGCCTTGTACAGCTCGTCCATG-3' (Tm: 57 °C, GC content (%): 56). The primers were designed to introduce Ndel and Nhel restriction endonuclease sites at the 5' and 3' termini respectively. The PCR products were the purified (Qiagen QIAquick PCR purification kit, section 2.2.1.2.8) and then prepared prior to ligation.

3.1.2.2 Preparation of cloning vector and insert DNA by restriction digest

Approximately 2.0 μg of the cloning vector, pET-28c(+), was restricted in a 40 μL reaction supplemented with 1x restriction endonuclease Multi-Core™ buffer (Promega) (25 mM Tris-acetate, pH 7.5 (at 37 °C), 100 mM CH₃CO₂K, 10 mM Mg(CH₃COO)₂, 1 mM DTT), 15 units Ndel restriction endonuclease and 15 units Nhel restriction endonuclease enzyme. Further, the purified PCR inserts of the 3 GFP spectral variants were also prepared for ligation by restricting the products in a 40 μL reaction described as follows; 25 μL PCR products (i.e., either eCFP or eGFP or eYFP) were supplemented with 1x restriction endonuclease Multi-Core™ buffer (Promega) (25 mM Tris-acetate, pH 7.5 (at 37 °C), 100 mM CH₃CO₂K, 10 mM Mg(CH₃COO)₂, 1.0 mM DTT), 15 units Ndel restriction endonuclease and 15 units Nhel restriction endonuclease enzyme. All 3 restriction reactions were prepared in sterile 0.5 mL tubes.

The reactions were restricted (section 2.2.1.2.1) and then gel-purified (section 2.2.1.2.3). The concentrations of the samples, i.e., cloning vector, eCFP, eGFP and eYFP were calculated as 0.136, 0.064, 0.065 and 0.060 μg /μL respectively (GeneQuantII RNA/DNA calculator).

3.1.2.3 Ligation reaction

As previously described in section 2.2.1.2.10, "vector to insert" ratios 1: 3 (one part of vector to three parts of insert) and 1: 6 (one part of vector to six parts of insert) were prepared and then ligated to generate the required clones.

In generating pET-28c (+)-eCFP expression construct, we employed the ratio 1:3 and prepared a ligation reaction mixture (10 μL) containing the following; 200 ng cloning vector and 81 ng eCFP insert PCR product as well as 1x T4 DNA ligase buffer (Promega) (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM (DTT) and 1 mM ATP) and
3 units T4 DNA ligase enzyme. The ratio 1:6 was also prepared as follows: 200 ng cloning vector and 162 ng eCFP insert PCR product were supplemented with 1x T4 DNA ligase buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM DTT and 1 mM ATP) and 3.0 units T4 DNA ligase enzyme.

For the pET-28c (+)-eGFP expression constructs, the ratio 1:3 was adapted to generate the construct as follows; 200 ng cloning vector and 80.2 ng eGFP insert PCR product were supplemented with 1x T4DNA buffer and 3.0 units T4 DNA ligase enzyme. The ratio 1:6 was also prepared as follows; 200 ng cloning vector and 160.3 ng eGFP insert PCR product were supplemented with 1x T4 DNA buffer and 3.0 units T4 DNA ligase enzyme.

In building pET-28c (+)-eYFP expression construct, we again adapted the ligation ratios 1:3 and 1:6. For the ligation ratio 1:3; 200 ng cloning vector and 78.9 ng eYFP insert PCR product were supplemented with 1x T4 DNA buffer and 3.0 units T4 DNA ligase enzyme. The ratio 1:6 was also prepared as follows; 200 ng cloning vector and 157.8 ng eGFP insert PCR product were supplemented with 1x T4 DNA buffer and 3.0 units T4 DNA ligase enzyme.

Finally, two control reactions similar to the ligation reactions above were prepared in parallel. The first control reaction was prepared without the insert DNA, i.e., 200 ng cloning vector was supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme whilst the second control reaction was prepared minus both the insert gene and the T4 DNA ligase. The reaction contained 200 ng cloning vector and 1x T4 DNA ligase buffer. The control reactions were prepared to establish the efficiency of the restriction, and ligation processes.

Reactions were diluted to 10 μL, mixed gently before the initial room temperature incubation for 1 h followed by the 16 h incubation at 4 °C.

3.1.2.4 Transformation of XL1-Blue cells with ligated products

The ligated mixtures were used to transform 50 μL XL1-Blue E.coli cells, section 2.2.1.13. Transformed XL1-Blue E.coli colonies were cultured (37 °C) on LB-Agar plates containing 30 μg/mL kanamycin and 12.5 μg/mL tetracycline (3 plates per ligated reaction). Colonies were documented and subsequently used to generate mini-cultures (5 mL). The mini-cultures were purified (section 2.2.1.1.6) and then analysed by restriction digest.
3.1.2.5 Screening of potential clones by restriction analysis

The purified samples (representing all three expression constructs) were restricted with Ndel and Nhel (i.e. enzymes used to prepare the insert DNA) restriction endonuclease in a 10 µL reaction described as follows: 8.0 µL purified plasmid DNA was supplemented with 1x restriction endonuclease Multi-Core™ buffer (Promega) (25 mM Tris-acetate, pH 7.5 (37 °C), 100 mM CH₃CO₂K, 10 mM Mg(CH₃COO)₂, 1 mM DTT), 5 units Ndel restriction endonuclease and 5 units Nhel restriction endonuclease enzyme.

A control reaction containing approximately 0.6 µg cloning vector, pET-28c (+), 1x restriction endonuclease Multi-Core™ buffer (Promega) (25 mM Tris-acetate, pH 7.5 (at 37 °C), 100 mM CH₃CO₂K, 100 mM Mg(CH₃COO)₂, 1.0 mM DTT), 10 units Ndel restriction endonuclease and 10 units Nhel restriction endonuclease was prepared in parallel.

The reactions were restricted (section 2.2.1.2.1) and then analysed by agarose gel electrophoresis (section 2.2.1.2.2). Following separation, the gels were visualized using a UV transilluminator and then the suspected clones identified and documented.

The prospective clones were amplified by Maxiprep (Qiagen Maxiprep kit, section 2.2.1.1.7) in XL1-Blue competent cells. The concentrations of the purified DNA samples, i.e., pET-28c (+)-eCFP, pET-28c (+)-eGFP and pET-28c (+)-eYFP were determined by absorbance at 260 nm.

3.1.2.6 Restriction mapping of selected clones

All three purified expression constructs were analyzed by restriction mapping. Two test reactions were prepared, a reaction to linearize the plasmid DNA and another to release the ligated GFP insert. The final volume of the reaction was 10 µL.

The first reaction was prepared as follows, approximately 1.5 µg of the purified plasmid DNA, i.e., pET-28c (+)-eCFP or pET-28c (+)-eGFP or pET-28c (+)-eYFP, was supplemented 1x restriction endonuclease buffer 4 (NEB) (10 mM Tris-HCl, pH 7.9, 25 °C, 50 mM NaCl, 10 mM MgCl₂ and 1 mM DTT) and 5 units Ndel restriction endonuclease enzyme.

The second restriction reaction mixture was similar to the first, however, together with the Ndel restriction endonuclease, the reactions were further supplemented with 5 units Nhel restriction endonuclease enzyme. Both enzymes were used to clone the GFP insert into the pET-28c (+) cloning vector.
A control reaction, similar to those describe above, minus the restriction endonuclease enzymes were prepared in parallel. The reactions were then restricted (section 2.2.1.2.1) and then analysed using agarose gel electrophoresis (section 2.2.1.2.2). The gel was illuminated with a UV transilluminator and the image of the restricted samples captured using a Kodak 120 digital camera. The separation of the DNA fragments representing the restriction maps of the various constructs are shown in section 3.10.

The prospective clones were further analysed by nucleotide sequencing using the primers described in section 2.2.1.2.11. This was performed to ensure that no adventitious mutations had occurred during the cloning process.

3.1.3 Creating pET-28c (+) -STAT3βtc-GFP expression constructs

3.1.3.1 Ligating STAT3βtc into pET-28c (+) GFP cloning vector

The C-terminal-GFP expression constructs, pET-28c (+) -STAT3βtc-eCFP, pET-28c (+) -STAT3βtc-eGFP and pET-28c (+) -STAT3βtc-eYFP encoding the STAT3βtc-GFP chimaeric proteins were generated by ligating the Ncol/Ndel restricted STAT3βtc PCR product into the Ncol/Ndel restricted pET-28c (+)-GFP cloning vector. The restricted PCR product was cloned upstream to the GFP gene contained in the pET-28c (+)-GFP. Below is a schematic representation of the cloning procedure.
Figure 3.3: Summary of the cloning procedure used to develop the expression constructs; pET-28c (+)-STAT3ptc-eCFP, pET-28c (+)-STAT3ptc-eGFP and pET-28c (+)-STAT3ptc-eYFP.

Primers were designed to introduce Ncol and Nhel sites at the 5' and 3' ends of the STAT3ptc insert DNA respectively via PCR-amplification. After amplification, the inserts DNA is purified, restricted with Ncol and Ndel restriction endonuclease and further purified. The restricted STAT3ptc insert is then ligated into the cloning vector, pET-28c (+)-GFP (section 3.1.2) prepared with both Ncol and Ndel. Following transformation, the cells are selected with the appropriate antibiotic (kanamycin) for a clone.

3.1.3.2 PCR amplification of STAT3ptc insert gene

The cDNA encoding STAT3ptc was PCR-amplified (section 2.2.1.2.7, Table 2.4) in a reaction mixture containing the following primers; forward, 5'-dAAAA TATACCATGGGCCAGGCCAACC-3' (Tm: 53 °C, GC content (%): 59) and reverse 5'-dAAAACATATGTCCGCCTTCCAAACTGCATCAATG -3' (Tm: 57 °C, GC content (%) 40). The primers were designed to introduce Ncol and Ndel endonuclease restriction sites at the 5' and 3' termini respectively. Following amplification, the product was purified (Qiagen QIAquick PCR purification kit, section 2.2.1.2.8) and then prepared for ligation.

3.1.3.3 Preparation of cloning vector and insert DNA by restriction digest

The cloning vectors, i.e. pET-28c(+)–eCFP, pET-28c(+)–eGFP and pET-28c(+)–eYFP, and the STAT3ptc insert gene were restricted with Ncol and Ndel restriction endonucleases to generate the corresponding sticky ends.
Approximately 3.0 µg cloning vectors, i.e., either pET-28c(+)–eCFP or pET-28c(+)–eGFP or pET-28c(+)–eYFP, was restricted in a 30 µL reaction containing 1x restriction endonuclease buffer 4 (NEB) (10 mM Tris-HCl, pH 7.9 @ 25 °C, 50 mM NaCl, 10 mM MgCl₂ and 1 mM DTT), 15 units Ncol endonuclease enzyme and 15 units Ndel restriction endonuclease. Control reactions for each of the cloning vectors were also prepared in parallel, each containing similar ingredients as described above but no restriction endonuclease enzymes.

The insert DNA (STAT3βtc PCR product) was also prepared for cloning in a 30 µL reaction as follows; 24 µL STAT3βtc insert PCR product was supplemented with 1x restriction endonuclease buffer 4 (NEB) (10 mM Tris-HCl, pH 7.9 @ 25 °C, 50 mM NaCl, 10 mM MgCl₂ and 1 mM DTT), 15 units Ncol restriction endonuclease and 15 units Ndel restriction endonuclease enzyme. The reactions were restricted (section 2.2.1.2.1) and then gel-purified (section 2.2.1.2.3). Following purification, the concentrations of the samples including pET-28c (+) – eCFP, pET-28c (+) – eGFP, pET-28c (+) – eGFP and STAT3βtc were measured as 0.075, 0.040, 0.054 and 0.063 µg/mL respectively.

### 3.1.3.4 Ligation reaction

The prepared PCR product was ligated into the restricted cloning vector by adding the T4 DNA ligase and its corresponding buffer in a “1:3” or a “1:6” vector to insert ratio (section 2.2.1.2.10).

In generating the pET-28c(+)-STAT3βtc-eCFP expression construct, the ratio 1:3 reaction mixture was prepared as follows; approximately 150 ng cloning vector and 177.8 ng STAT3βtc insert PCR product were supplemented with 1x T4 DNA ligase buffer (Promega) (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM (DTT) and 1 mM ATP) and 3 units T4 DNA ligase enzyme. Further, ratio 1:6 was prepared as follows: approximately 150 ng cloning vector and 355.6 ng STAT3βtc insert PCR product was supplemented with 1x T4 DNA ligase buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM (DTT) and 1 mM ATP) and 3.0 units T4 DNA ligase enzyme.

In the case of pET-28c (+) –STAT3βtc- eGFP expression construct, the ratio 1:3 was adapted to generate the construct as follows; approximately 150 ng cloning vector and 177.8 ng STAT3βtc PCR insert was supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme. For ratio 1:6; approximately 150 ng cloning vector and 355.6 ng STAT3βtc insert DNA were supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme. The reactions were diluted to 15 µL with ddH₂O.
Two control reactions similar to those described above were prepared in parallel. The first reaction was prepared minus the insert DNA (i.e., ~ 150 ng cloning vector was supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme) whilst the second control reaction was prepared without the insert DNA and the T4 DNA ligase (i.e. 150 ng cloning vector was supplemented with 1x T4 DNA ligase buffer). The reactions were diluted to 10.0 µL with the appropriate volumes of sterile ddH₂O and then incubated at room temperature for 1 h and at 4 °C for 16 h.

The ligated mixture was used to transform DH5α E.coli cells (section 2.2.1.13) and then mini-cultured (5 mL) in a media supplemented with 30 μg/mL kanamycin. The plasmid DNA samples were purified as described in section 2.2.1.1.6.

3.1.3.5 Selection of potential clones by restriction analysis

The purified plasmid DNA samples were restricted with the cloning enzymes in a 10 µL reaction. For each expression construct, i.e., pET-28c (+) –STAT3βtc- eCFP, pET-28c (+) –STAT3βtc- eGFP and pET-28c (+) –STAT3βtc- eYFP, 8.0 µL purified plasmid DNA was supplemented with 1x restriction endonuclease buffer 4 (NEB) (10 mM Tris-HCl, pH 7.9 @ 25 °C, 50 mM NaCl, 10 mM MgCl₂ and 1 mM DTT), 5 units NcoI restriction endonuclease enzyme and 5 units NdeI restriction endonuclease. A control for each cloning vector was prepared in parallel. Each of the reactions contained 0.5 μg cloning vector, 1x restriction endonuclease buffer 4 (10 mM Tris-HCl, pH 7.9 @ 25 °C, 50 mM NaCl, 10 mM MgCl₂ and 1 mM DTT), 5 units NcoI restriction endonuclease and 5 units NdeI restriction endonuclease.

The reactions were restricted (section 2.2.1.2.1) and then evaluated on an agarose gel by electrophoresis (section 2.2.1.2.2). Prospective clones were identified and documented and subsequently amplified and purified by Maxi–preparation, section 2.2.1.1.7. The concentrations of the purified plasmid DNA samples were measuring the absorbance at 260 nm.

3.1.3.6 Restriction mapping of selected clones

To ensure that the right plasmid has been amplified, the clones were restricted in three different restriction endonuclease enzymes.

Reaction one was prepared to linearize the plasmid DNA construct at the 3' end of STAT3βtc and the 5' end of GFP; approximately 1.5 μg plasmid DNA was treated with
1x restriction endonuclease buffer 4 (NEB) (10 mM Tris-HCl, pH 7.9 @ 25 °C, 50 mM NaCl, 10 mM MgCl₂ and 1 mM DTT) and 7.5 units Ndel restriction endonuclease enzyme. In reaction two, a similar reaction to one was prepared and supplemented with 7.5 units Nhel restriction endonuclease to release the gene for GFP gene (i.e. either eCFP or eGFP or eYFP).

In reaction three, a similar reaction to reaction one was also prepared and supplemented with 7.5 units Ncol restriction endonuclease to release the gene for STAT3βtc. In the fourth reaction, the fusion gene (STAT3βtc-FP) was released from the cloning vector in a reaction containing the following ingredients: approximately 1.5 µg plasmid DNA, 1x restriction endonuclease buffer 4 (10 mM Tris-HCl, pH 7.9 @ 25 °C, 50 mM NaCl, 10 mM MgCl₂ and 1 mM DTT), 7.5 units Ncol restriction endonuclease and 7.5 units Nhel restriction endonuclease enzymes.

Finally, a control reaction containing the basic ingredients in reaction one (no restriction enzyme) was prepared with each expression construct. The reactions were restricted (section 2.2.1.2.1) and subsequently analysed by agarose gel electrophoresis (section 2.2.1.2.2). The gel was illuminated using a transilluminator UV light source λ= 312 nm and the image captured using a Kodak DC 120 206 M digital camera. The separated DNA fragments making up the restriction maps of the various expression constructs are shown in section 3.11. Furthermore, the purified samples were sequenced (section 2.2.1.2.11) to ensure that no mutation had occurred during the cloning process.

3.1.4 Creating pET-28c (+)-His-STAT3βtc expression vectors with C-tagged GFP

The His₅ tagged C-terminal GFP expression constructs; pET-28c(+)·His-STAT3βtc-eCFP, pET-28c(+)·His-STAT3βtc-eGFP and pET-28c(+)·His-STAT3βtc-eYFP encoding for the polyhistidine tagged His-STAT3βtc-GFP chimaeric His-STAT3βtc-GFP, were generated by inserting the Ndel restricted STAT3βtc PCR product into the Ndel restricted pET-28c (+)-GFP cloning vector. The restricted PCR product was cloned upstream to the GFP gene and downstream to the polyhistidine sequence in the pET-28c (+)-GFP cloning vector. Below is a schematic representation of the cloning procedure:
Primers were designed to introduce of Ndel sites at both the 5' and 3' ends of the STAT3(ttc (—)) insert DNA by PCR-amplification. After amplification, the insert DNA is purified, restricted with Ndel restriction endonuclease and further purified. The restricted STAT3(ttc insert is then ligated into an acceptor pET-28c (+)-GFP cloning vector that has been restricted with Ndel and dephosphorylated with a phosphatase. Following transformation, the cells are selected with the appropriate antibiotic (kanamycin) for a clone.

### 3.1.4.1 PCR amplification of STAT3(ttc insert gene

The cDNA encoding the STAT3(ttc gene was PCR-amplified with the forward (5'-dAAAA TATACTATGGCCAGGCAACC-3' (Tm: 52 °C, GC content: 52%)) and reverse (5'-dAAAA CATATGCCTCGCTTCCAAAAGCTGATCAATG-3' (Tm: 57 °C, GC content: 40%)) primers designed to introduce Ndel endonuclease restriction sites at the 5' and 3' termini respectively (section 2.2.1.2.7, Table 2.4). The PCR product was purified (section 2.2.1.2.8) and then prepared for restriction (section 3.1.4.2) prior to ligation.

### 3.1.4.2 Preparation of cloning vectors and insert DNA by restriction digest

The cloning vectors (i.e., pET-28c(+)-eCFP, pET-28c (+)-eGFP and pET-28c (+)-eYFP) as well as the STAT3(ttc insert DNA were restricted with Ndel restriction endonuclease in preparation for ligation.

![Diagram of the cloning procedure for developing expression construct pET-28c (+)-GFP cloning vector to generate pET-28c(+)-His-STAT3(ttc-eCFP, pET-28c(+)-His-STAT3(ttc-eGFP and pET-28c (+)-His-STAT3(ttc-eYFP expression constructs.](image-url)
The cloning vectors were prepared in a 30 μL reaction containing about 3.0 μg cloning vector, 1x restriction endonuclease buffer 4 (NEB) (10 mM Tris-HCl, pH 7.9 @ 25 °C, 50 mM NaCl, 10 mM MgCl₂ and 1 mM DTT) and approximately 15 units NdeI restriction endonuclease enzyme. Control reactions for each of the cloning vectors were also prepared in parallel with each containing similar ingredients as described above but no restriction endonuclease enzymes.

A 30 μL restriction reaction containing 24 μL STAT3βtc insert PCR product was supplemented 1x restriction endonuclease buffer 4 (10 mM Tris-HCl, pH 7.9 @ 25 °C, 50 mM NaCl, 10 mM MgCl₂ and 1 mM DTT) and 15 units NdeI restriction endonuclease enzyme was also prepared.

The reactions were restricted (section 2.2.1.2.1) and then gel-purified (section 2.2.1.2.3) using the Qiagen extraction kit. The purified DNA samples were quantified using a NanoDrop 2000 spectrophotometer from Thermo scientific (Wilmington, Delaware, USA) as 0.0180, 0.0432, 0.0051 and 0.0442 μg/μL for pET-28c(+)His–eCFP, pET-28c(+)His–eGFP, pET-28c(+)His–eYFP and STAT3βtc respectively.

### 3.1.4.3 Ligation reaction

The restricted insert DNA (STAT3βtc PCR product) was ligated into the prepared cloning vector by adding a T4 DNA ligase and its corresponding buffer using ratios 1:3 and 1:6 (vector to insert) described in section 2.2.1.2.10. On the basis of the concentration measured, the ligation reactions were prepared with reduced amounts of the cloning vector, i.e., 150, 75 and 150 ng, for pET-28c(+) His–STAT3βtc-eCFP, pET-28c(+) His–STAT3βtc-eGFP and PET-28c(+) His–STAT3βtc-eYFP respectively.

In developing pET-28c(+) His–STAT3βtc-eCFP expression construct, a 1:3 ratio was prepared in a 20 μL ligation reaction as follows; 150 ng cloning vector and 133.2 ng STAT3βtc insert DNA were supplemented with 1x T4 DNA ligase buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM (DTT) and 3.0 units T4 DNA ligase enzyme. The ratio 1:6 was also prepared as follows; 150 ng cloning vector and 266.4 ng STAT3βtc insert DNA were supplemented with 1x T4 DNA ligase buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM (DTT) and 3.0 units T4 DNA ligase enzyme.

Similarly, ratio 1:3 was adopted to develop the expression construct; pET-28c(+) His–STAT3βtc-eGFP, in a 25 μL ligation reaction as follows; 75 ng cloning vector and 133.2 ng STAT3βtc insert DNA was supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme. The ratio 1:6 was also prepared as follows; 75 ng cloning...
vector and 266.4 ng STAT3\(\beta\)tc insert DNA were supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme.

Further, in developing pET-28c (+) His–STAT3\(\beta\)tc-eGFP expression construct, we again adopted the ligation ratios 1:3 and 1:6. The ligation ratio 1:3 contained the following ingredients in a 20 \(\mu\)L reaction mixture; 150 ng cloning vector and 133.2 ng STAT3\(\beta\)tc insert DNA were supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme. The ratio 1:6 was also prepared as follows; 150 ng cloning vector and 266.4 ng STAT3\(\beta\)tc insert DNA were supplemented T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme.

Finally, two sets of control reactions similar to those described above were prepared in parallel. The first control reaction was prepared minus the STAT3\(\beta\)tc insert DNA and contained approximately 150 ng of the cloning vector (i.e., pET-28c (+)-His–eCFP, or pET-28c (+)-His–eYFP) supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme. The second control reaction was prepared minus the STAT3\(\beta\)tc insert DNA and the T4 DNA ligase. The reaction contained approximately 150 ng cloning vector supplemented with 1x T4 DNA ligase buffer. Due to insufficient cloning vector, no control reactions were prepared for the pET-28c (+) His–STAT3\(\beta\)tc-eGFP ligation. The reactions were incubated at room temperature for 1 h and then at 4 °C for 16 h. Following ligation, the reactions were used to transform DH5\(\alpha\) E.coli cells and then mini-cultured (5 mL) in preparation for Miniprep purification (section 2.2.1.1.6).

3.1.4.4 Selection of potential clones by restriction analysis

The purified plasmid DNA samples were restricted with the cloning enzymes in a 10 \(\mu\)L reaction. The restriction reaction contained the following ingredients; 8.5 \(\mu\)L purified plasmid DNA supplemented with 1x restriction endonuclease buffer 4 (10 mM Tris-HCl, pH 7.9 @ 25°C, 50 mM NaCl, 10 mM MgCl\(_2\) and 1 mM DTT) and 5 units Ndel restriction endonuclease enzyme.

Further, a control reaction containing the original cloning vectors were prepared in parallel. Each reaction contained approximately 0.5 \(\mu\)g cloning vector 1x restriction endonuclease buffer 4 (10 mM Tris-HCl, pH 7.9 @ 25°C, 50 mM NaCl, 10 mM MgCl\(_2\) and 1 mM DTT) and 5 units Ndel restriction endonuclease and were incubated at 37 °C for 1 h. The restricted samples were analysed by agarose gel electrophoresis to identify potential clones (section 2.2.1.2.2).
3.1.4.5 Screening potential clones by PCR

Since the insert STAT3βtc was ligated into the cloning vector non-directionally, it was necessary to confirm the orientation of the insert DNA by PCR-amplification. Thus, putative clonal plasmid DNA sample containing the insert DNA in the right orientation was PCR-amplified (section 2.2.1.2.7, Table 2.4) using the forward primer (5'-dTAAAATATACATATGGCCAGGCCAACC-3') for STAT3βtc cDNA and the reverse primer (5'-dAAAAATATACTTGTACAGCTCGTCCATG-3') for GFP cDNA. The amplified samples were verified by agarose gel electrophoresis (section 2.2.1.2.2). Prospective clones were amplified in DH5α cells and purified by Maxiprep, section 2.2.1.1.7. The concentrations of the samples were obtained by measuring absorbance at 260 nm. The plasmid DNA samples were sequenced by Eurofins using the T7 promoter primers (section 2.2.1.2.11).

3.1.4.6 Restriction mapping of selected clones

To generate a restrict map to further confirm the cloned expression constructs; we restricted the purified constructs with 3 different restriction endonuclease enzymes in a 15 µL reaction.

Reaction one was prepared to linearize the expression constructs at the 3' end of GFP using the Nhel restriction endonuclease enzyme; approximately 1.5 µg plasmid DNA was treated with 1x restriction endonuclease buffer 4 (10 mM Tris-HCl, pH 7.9 at 25 °C, 50 mM NaCl, 10 mM MgCl₂ and 1 mM DTT) and 7.5 units Nhel restriction endonuclease enzyme. In reaction two, a similar mixture to reaction one was prepared. The reaction contained all the above ingredients except the Nhel restriction endonuclease which was replaced with 7.5 units Ndel restriction endonuclease enzyme to release the cloned STAT3βtc gene. In reaction three, a similar reaction to one was also prepared and supplemented with 7.5 units Ndel restriction endonuclease to release STAT3βtc and GFP. By partial restriction, STAT3βtc-GFP fusion gene may be released.

Finally, a control reaction containing the ingredients in reaction one, minus the Nhel restriction enzyme was prepared for each expression construct. The reactions were restricted (section 2.2.1.2.1) and then analysed by agarose gel electrophoresis (section 2.2.1.2.2). Using a Kodak DC 120 206 M digital camera, the restriction maps of the various constructs were documented, section 3.13.
3.1.5 Creating pET-28c (+)-His–eCFP-TEV-eYFP construct

An expression construct consisting of the eCFP gene tethered to the eYFP gene with a TEV protease specific sequence was generated primarily to test the principle of FRET (concatemer expected to generate a steady and intense positive value for FRET) and to test the proteolytic activity of the TEV protease. Developing the expression construct involved ligating PCR-amplified eYFP gene into pET-28c (+) His–eCFP-TEV cloning vector. In generating pET-28c (+) - His-eCFP-TEV-eYFP, we used an existing cloning vector containing the eCFP and HIF1β genes tethered by the TEV recognition sequence, pET-28c (+) - His-eCFP-TEV-HIF1β.

Figure 3.5: Summary of the cloning procedure used to generate pET-28c (+) His–eCFP-TEV-eYFP expression construct. Primers were designed to introduce BamHI and NotI sites at the 5' and 3' ends of the eYFP (-) insert DNA by PCR-amplification. After amplification, the insert DNA is purified, restricted with both BamHI and NotI restriction endonucleases and further purified. The restricted eYFP insert is then ligated into an acceptor pET-28c (+)-eCFP-TEV cloning vector that has also been restricted with BamHI and NotI restriction endonucleases. Following transformation, the cells are selected with the appropriate antibiotic (kanamycin) for a clone.

3.1.5.1 PCR-amplification eYFP cDNA

The cDNA sequence encoding for eYFP fluorescent protein was amplified with a forward (5'-dAAAAGGATCCATGGTGAGCAAGGGCGAGG-3' (Tm: 61 °C, GC content (%) 64)) and reverse (5'-dAAAAGCGGCCGCCTTGTACAGCTCGTCCATG-3' (Tm: 65 °C, GC content (%) 66)) primers designed to introduce BamHI and NotI endonuclease
restriction sites at the 5' and 3' termini respectively (section 2.2.1.2.7, Table 2.4). The amplified product PCR product was purified (Qiagen QIAquick PCR kit, 2.2.1.2.8) and prepared for ligation.

### 3.1.5.2 Preparation of cloning vector and insert DNA by restriction digest

The cloning vector, pET-28c (+) His–eCFP-TEV-HIF1β and the insert PCR product, eYFP, were restricted with BamHI and NotI restriction endonuclease enzyme to prepare the reactants for ligation.

The cloning vector was restricted in a 30.0 μL reaction containing the following ingredients; 1.36 μg cloning vector, 1x restriction endonuclease buffer 3 (50 mM Tris-HCl, pH 7.9 at 25 °C, 100 mM NaCl, 10 mM MgCl₂ and 1 mM DTT), 15 units BamHI restriction endonuclease and 15 units NotI restriction endonuclease enzyme. In addition, the eYFP insert DNA was also restricted in a 30 μL reaction containing the following ingredients: 24.75 μL eYFP insert DNA, 1x restriction endonuclease buffer 3 (50 mM Tris-HCl, pH 7.9 @ 25 °C, 100 mM NaCl, 10 mM MgCl₂ and 1 mM DTT), 15 units BamHI restriction endonuclease enzyme and 15 units NotI restriction endonuclease.

A control reaction containing 0.25 μg of the cloning vector was also prepared in parallel. Both test and control reactions were restricted (section 2.2.1.2.1) and then gel-purified, section 2.2.1.2.3. The concentrations of the cloning vector and the eYFP insert DNA were determined using the GeneQuantII RNA/DNA calculator as 0.072 and 0.117 μg/μL respectively.

### 3.1.5.3 Ligation reaction

The restricted insert DNA was ligated into the restricted cloning vector by the addition of T4 DNA ligase and its corresponding buffer. We adapted ratios 1:3 and 1:6 to develop expression construct, pET-28c (+) His–eCFP-TEV-eYFP.

Ratio 1:3 was prepared in a 15 μL reaction as follows: approximately 200 ng cloning vector and 69.6 ng eYFP insert DNA were supplemented with 1x T4 DNA ligase buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM DTT and 1 mM ATP) and 3.0 units T4 DNA ligase enzyme. Ratio 1:6 was also prepared in a 15 μL reaction as follows; approximately 200 ng cloning vector and 139.02 ng were supplemented with 1x T4 DNA ligase buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM DTT and 1 mM ATP) and 3.0 units) T4 DNA ligase enzyme.
In addition to the test reactions, two other control reactions were prepared in parallel. The first reaction was prepared without the insert DNA and contained approximately 200 ng cloning vector supplemented with 1x T4 DNA ligase buffer and 3.0 units T4 DNA ligase enzyme. The second control reaction was prepared in the absence of the insert DNA and the T4 DNA ligase, thus the reaction contained; 200 ng cloning vector supplemented with 1x T4 DNA ligase buffer. Both test and control reactions were incubated at room temperature for 1 h and then at 4 °C for 16 h. The ligated products were then used to transform DH5α competent cells, section 2.2.1.1.3. Following transformation, colonies were cultured in 5 mL mini cultures and then purified by Miniprep, section 2.2.1.1.6.

### 3.1.5.4 Selection of potential clones by restriction analysis

The purified plasmid DNA samples (Miniprep DNA) were restricted with BamHI restriction endonuclease. The reactions contained the following ingredients; 8.5 μL (~0.3 μg) plasmid DNA, 1x restriction endonuclease buffer 3 (50 mM Tris-HCl, pH 7.9 at 25 °C, 100 mM NaCl, 10 mM MgCl₂ and 1 mM DTT) and 10 units BamHI restriction endonuclease. In parallel, a control reaction containing 0.34 pg cloning vector and 1x restriction endonuclease buffer 3 (50 mM Tris-HCl, pH 7.9 at 25 °C, 100 mM NaCl, 10 mM MgCl₂ and 1 mM DTT) was also prepared in parallel. Both test and control reactions were restricted (section 2.2.1.2.1) and then analysed by agarose gel electrophoresis (section 2.2.1.2.2).

Prospective clones were identified and documented. They were then amplified in DH5α cells and then purified by Maxiprep, section 2.2.1.1.7. The samples were also quantified by measuring absorbance at 260 nm. Also, the plasmid DNA samples were sequenced (section 2.2.1.2.11) to ensure that no adventitious mutation has occurred during cloning.

### 3.1.5.5 Restriction mapping of selected clones

To confirm that the amplified plasmid is correct, the pET-28c (+) His-eCFP-TEV-eYFP expression construct was restricted in 4 different reactions.

The first reaction describes the linearization of the expression construct; approximately 1.5 μg plasmid DNA was treated with 1x restriction endonuclease buffer 3 (50 mM Tris-HCl, pH 7.9 at 25 °C, 100 mM NaCl, 10 mM MgCl₂ and 1 mM DTT) and 5 units NotI restriction endonuclease enzyme. The second reaction was similar to the first, however, in addition to NotI, the reaction was further supplemented with 5 units Ndel restriction endonuclease enzyme to release the concatemer, eCFP-TEV-eYFP.
A third reaction similar to the second was also prepared and then supplemented with a third restriction endonuclease enzyme, Nhel, \textit{i.e.}, 5 units, to release the individual GFP variants (eCFP and eYFP) from the plasmid DNA. The fourth and final reaction contained all the ingredients in the first reaction, however, no restriction endonuclease was added. The reactions were were restricted (section 2.2.1.2.1) and then analysed by agarose gel electrophoresis (section 2.2.1.2.3). The image of the fractionated DNA fragments (restriction map) was captured, section 3.15.

3.2 RESULTS

3.2.1 Creating fluorescent STAT3\textsuperscript{ptc} expression constructs

Contrary to the commonly cited model of STAT3 signaling paradigm (Shuai, \textit{et al.}, 1994), numerous studies carried out \textit{in vitro} have suggested that cytoplasmic STAT3 may exist as part of a large poorly conformed molecular complex and/or an inactive dimer prior to activation via phosphorylation (Ndubuisi \textit{et al.}, 1999; Haan \textit{et al.}, 2000; Novak \textit{et al.}, 1998). Hence, in order to evaluate the conformational status of STAT3 \textit{in vitro} and to modulate its functional activities, fluorescent-based assays using fluorescent expression constructs containing a STAT3 mutant, STAT3\textsuperscript{ptc}, fused at either the 5' or the 3' end with spectral variants of GFP (\textit{i.e.} enhanced cyan, green and yellow fluorescent protein) were established. These set of expression constructs (both the N- and C- terminally fused STAT3\textsuperscript{ptc} constructs) were developed to evaluate which were the most useful for these studies.

3.2.1.1 pET-32a (+)-GFP-STAT3\textsuperscript{ptc} expression constructs

The fluorescent expression constructs containing a spectral variant of GFP and STAT3\textsuperscript{ptc} (Figure 3.7) was constructed by cloning the PCR-amplified GFP cDNA to the 5' end of the STAT3\textsuperscript{ptc} gene contained in a pET-32a (+) cloning vector (section 3.1.1). The fusion constructs are schematically represented in Figure 3.6.
Figure 3.6: Plasmid DNA map of GFP-STAT3βtc.
A representative plasmid map showing the relative positions of the GFP and STAT3βtc genes in the pET-32a (+) cloning vector.

Figure 3.7: Schematic representation of the GFP-STAT3βtc expression constructs.
Fluorescent STAT3βtc chimaeric proteins cloned and applied in the fluorescent-based assays in the study. The cloned GFP variants are colour coded as (— ), (— ) and (— ) representing enhanced cyan, green and yellow fluorescent proteins respectively. The parent protein, STAT3βtc, is colour coded (— ). All numbers refer to amino acid residues.

Following isolation of the expression constructs (i.e. pET-32a (+)-eCFP-STAT3βtc, pET-32a (+)-eGFP-STAT3βtc and pET-32a (+)-eYFP-STAT3βtc) from JM109 E.coli cells, the concentrations and quality of the DNA samples were determined. The concentrations were determined by measuring the absorbance of each DNA sample at 260 nm, thus yielding; 227, 135 and 250 µg/mL for pET-32a (+)-eCFP-STAT3βtc, pET-32a (+)-eGFP-STAT3βtc and pET-32a (+)-eYFP-STAT3βtc respectively. In addition, the quality of the expression constructs was evaluated by analyzing the DNA samples on an agarose gel via electrophoresis. The restriction data demonstrating restriction maps of the expression constructs is shown in Figure 3.8. The restriction study displayed shows the cloning was successful.

Lanes 2, 4 and 6, contains products generated after restricting the expression constructs with EcoRI restriction endonuclease. The restriction enzyme cleaves the construct at the 3' end of STAT3βtc to linearize the plasmid DNA and then reveal the approximate size of the expression construct upon electrophoresis. The experimental
data generated corresponded to the calculated molecular weights of the constructs, i.e.
approximately 7.92 kb.

Figure 3.8: Endonuclease restriction mapping of the pET-32a (+)-GFP-STAT3βtc expression constructs.
Expression constructs pET-32a (+)-eCFP-STAT3βtc, pET-32a (+)-eGFP-STAT3βtc and pET-32a (+)-eYFP-STAT3βtc were linearized with EcoRI restriction endonuclease to generate the approximate molecular weights of the linearized expression constructs as displayed in lanes 2, 4 and 6 respectively. Lanes 3, 5 and 7 contain DNA bands displaying the relative sizes of the DNA fragments generated after restriction with both EcoRI and Ndel restriction endonuclease enzymes. Arrows indicate the positions and the molecular weight of the DNA fragments relative to the 1kb DNA maker displayed in lane 1.

Also, lanes 3, 5 and 7 contain DNA fragments generated after restricting the expression constructs with both Ndel (cleaved the plasmid DNA at both the 5' and 3' ends of GFP) and EcoRI (cleaved the plasmid DNA at the 3' end of STAT3βtc) restriction endonuclease enzymes. The relative sizes of the DNA fragments following agarose gel electrophoresis corresponds to the calculated molecular weights of the pET-32a (+) cloning vector (~ 5.39 kb), the STAT3βtc (~ 1.80 kb) and the eGFP insert (~ 0.72 kb).

Plasmid DNA samples were also sequenced with the forward primer for the T7 promoter. Together with the restriction map, Figure 3.8, the fluorescent STAT3βtc expression constructs were confirmed.

3.2.2 pET-28c (+)-STAT3βtc-GFP expression constructs

3.2.2.1 pET-28c (+)-GFP expression constructs

In generating expression constructs with the GFP fused to the 3' end of the STAT3βtc gene (i.e. GFP fused to the C-terminus of STAT3βtc, Figure 3.9), the pET-28c (+)-GFP was initially constructed. Developing the constructs involved cloning the PCR-amplified
GFP cDNA into an empty pET-28c (+) cloning vector to generate the following expression constructs; pET-28c (+)-eCFP, pET-28c (+)-eGFP and pET-28c (+)-eYFP (section 3.1.3).

Upon isolation of the expression constructs from XL1-Blue *E.coli* cells, the concentrations and quality of the samples were determined and verified respectively. By absorbance, the concentrations of pET-28c (+)-eCFP, pET-28c (+)-eGFP and pET-28c (+)-eYFP were calculated as 262, 345 and 660 µg/mL respectively. Also, the quality of the plasmid DNA samples was verified via agarose gel electrophoresis. Demonstrated in Figure 3.10 are the restriction maps of the various expression constructs. The restriction data displayed confirms that the cloning was successful.
Figure 3.10: Endonuclease restriction mapping of the pET-28c (+)-GFP expression constructs.

Expression constructs pET-28c (+)-eCFP (A), pET-28c (+)-eGFP (B) and pET-28c (+)-eYFP (C) were electrophoresed as supercoiled plasmid DNA contained in lanes 2 in gels (A, B and C). Linearized with NdeI restriction endonuclease to generate the approximate molecular weights of the linearized expression constructs as displayed in lanes 3 (A, B and C). Lanes 4 (A, B and C) contains DNA bands displaying the relative sizes of the DNA fragments generated after restriction of the expression constructs with both NdeI and NheI restriction endonuclease enzymes. Arrows indicate the positions and the molecular weights of the DNA fragments relative to the 1kb DNA maker displayed in lane 1 (A, B and C).

Lane 2 (A, B and C) contain the supercoiled form (lower band) and the open circular form (upper bands) of the plasmid DNA constructs. All three construct displayed identical characteristics including molecular weight of approximately 4.5 kb.

In lane 3 (A, B and C) are the linearized form of the plasmid DNA constructs. Samples were restricted with NdeI restriction endonuclease (cleaves plasmid DNA construct at the 5' end of the GFP insert) to reveal their relative molecular weights i.e approximately 6.02 kb. The experimental molecular weight for all three constructs corresponded to the calculated molecular weights stated above. Also evident is the contamination by bacterial genomic DNA represented as a smear seen above and faintly below the linearized plasmid DNA band.

Lane 4 (A, B and C) contain products of the double restriction reaction. Plasmid DNA constructs were restricted with both NdeI (cleaves construct at the 5' end of the GFP insert) and NheI (cleaves construct at the 3' end of the GFP insert) restriction endonuclease enzymes to release the GFP insert. Both enzymes were used to clone the GFP PCR product. Following agarose gel electrophoresis, two DNA fragments with relative sizes corresponding to the cloning vector (~ 5.30 kb) and the GFP insert (~
0.72 KB) were observed. Also, the smeary patch observed above the band for the cloning vector may be contaminants from the bacterial genomic DNA.

In addition to the restriction maps, all three expression constructs were sequenced for any adventitious mutations. The resultant sequence data were found to be correct using the forward primer of the T7 promoter contained in the pET-28c (+) cloning vector.

### 3.2.2.2 pET-28c (+)-STAT3βtc-GFP expression constructs

Using the expression constructs described in section 3.2.2.1, PCR-amplified STAT3βtc cDNA was subcloned into the pET-28c (+)-GFP cloning vectors. The three expression constructs; pET-28c (+)-STAT3βtc-eCFP, pET-28c (+)-STAT3βtc-eGFP and pET-28c (+)-STAT3βtc-eYFP were created (section 3.1.3).

After the Maxiprep purification step, the concentrations and quality of the samples were determined. The concentrations of the plasmid DNA samples were determined by measuring the absorbance at 260 nm. With knowledge of the dilution factor, the concentrations were calculated as 360, 400 and 212 µg/mL. In addition, the quality of the purified plasmid DNA was evaluated by analysing the samples on an agarose gel via electrophoresis. Figure 3.11 (A, B and C) shows the restriction maps of the expression constructs.

![Figure 3.11: Endonuclease restriction mapping of the pET-28c (+)-STAT3βtc-GFP expression constructs](image-url)

Expression constructs pET-28c (+)-STAT3βtc-eCFP (A), pET-28c (+)-STAT3βtc-eGFP (B) and pET-28c (+)-STAT3βtc-eYFP (C) were electrophoresed as neat (unrestricted or intact) contained in lanes 2 of gels A, B and C. Lanes 3 contains linearized samples generated with Ndel restriction endonuclease, thus, revealing the actual molecular weights of the linearized expression constructs. Lanes 4 of gels A, B and C contains DNA bands displaying the relative sizes of the DNA fragments generated after restricting the expression constructs with Ndel and Nhel restriction endonuclease enzymes. The contents in lanes 5 of gels A, B and C are DNA bands showing the relative sizes of the DNA fragments generated after
restricting the expression constructs with Ncol and Ndel restriction endonuclease enzymes. Finally, lanes 6 of gels A, B and C contains samples demonstrating the relative sizes of the DNA fragments generated after restricting the expression constructs with Ncol and Nhel restriction endonuclease enzymes. Arrows indicate the positions and the molecular weights of the DNA fragments relative to the 1kb DNA maker displayed in lane 1 (A, B and C).

The restriction maps displayed suggest that the expression constructs (i.e. pET-28c (+)-STAT3βtc–eCFP, pET-28c (+)-STAT3βtc–eGFP and pET-28c (+)-STAT3βtc–eYFP) were successfully created.

Lane 2 (A, B and C) contains the unrestricted or supercoiled plasmid DNA expression constructs (lower band) and the open circular form (upper band). Except in lane 2 (B), where both band have not been properly resolved (may be due to contamination by genomic DNA), constructs displayed similar characteristics in their respective gels. They migrated to identical positions i.e. approximately 6.0 kb relative to the 1 kb DNA ladder.

Lane 3 (A, B and C) was used to demonstrate the relative molecular weight of the expression constructs. Constructs were linearized with Ndel restriction endonuclease enzyme (cleaved at the 3’ and the 5’ ends of STAT3βtc and GFP respectively) and then electrophoresed to positions corresponding to their calculated molecular weights i.e. approximately 7.82 kb.

The contents displayed in lane 4 (A, B and C) were products of the double restriction reaction. Constructs were treated with Ndel and Nhel (enzymes used to clone the GFP cDNA) restriction endonucleases to release the GFP insert. The Ndel and Nhel restriction endonuclease enzymes cleaved the plasmid DNA at the 5’ and 3’ ends of the GFP insert. The relative sizes of the two DNA fragments corresponds to their expect molecular weights i.e approximately 7.10 kb for pET-28c (+)-STAT3βtc and approximately 0.72 kb for GFP.

Also, the contents in lane 5 (A, B and C) are products generated after a double restriction reaction. The plasmid DNA samples were restricted with Ncol and Ndel (enzymes used to clone STAT3βtc cDNA) restriction endonuclease enzymes to release the STAT3βtc insert. The relative sizes of the DNA fragments corresponds to the calculated molecular weights of the pET-28c (+)-GFP cloning vector (~ 6.02 kb) and the STAT3βtc insert DNA (~ 1.80 kb).
Finally, contents analyzed in lanes 6 also displayed products of a double restriction reaction. The plasmid DNA constructs were treated with Ncol (cleaves constructs at the 5' end of STAT3βtc) and Nhel (cleaves construct at the 3' end of GFP) restriction endonucleases to release the fusion gene i.e. STAT3βtc-GFP, a further confirmatory study. Upon electrophoresis, the relative sizes of the DNA fragments were known to correspond to the expected molecular weights of the fusion gene (~ 2.52 kb) and the empty pET-28c (+) cloning vector (~ 5.30 kb). In addition to the restriction maps, the nucleotide sequence for the expression constructs were acquired to ensure that no adventitious mutations have occurred during cloning.

### 3.2.3 pET-28c (+)-His-STAT3βtc-GFP expression constructs

To facilitate purification, polyhistidine tagged isoform (i.e. pET-28c (+)-His-STAT3βtc-GFP) of the expression construct pET-28c (+)-STAT3βtc-GFP was developed (see section 3.4). The construct was developed by subcloning the STAT3βtc PCR product (STAT3βtc PCR-amplified to introduce Ndel restriction endonuclease sites at both 5' and 3' ends) into the Ndel site of the pET-28c (+)-GFP cloning vectors. This resulted in a product with a polyhistidine tag fused to the 5' end of the STAT3βtc. Two expression constructs containing His-STAT3βtc-eCFP and His-STAT3βtc-eYFP were successfully developed as shown in Figure 3.12.

![Figure 3.12: Schematic representation of the His-STAT3βtc-GFP expression constructs](image)

Polyhistidine tagged STAT3βtc fluorescent chimaeric proteins cloned in the study. The cloned GFP variants are colour coded as ( ) and ( ) representing enhanced cyan and yellow fluorescent proteins respectively. The polyhistidine tag at the 5' end of STAT3βtc is colour coded ( ) whilst the parent protein, STAT3βtc, is colour coded ( ).

After purification of the prospective clones, the concentrations of the plasmid DNA constructs were measured by absorbance and calculated as 245 μg/mL and 320 μg/mL for pET-28c(+)-His-STAT3βtc-eCFP and pET-28c(+)-His-STAT3βtc-eYFP respectively. The quality of the yield was evaluated by analyzing the constructs by agarose gel electrophoresis. Figure 3.12 is a restriction map confirming the expression constructs.
Figure 3.13: Endonuclease restriction mapping of the pET-28c (+)-His-STAT3βtc-GFP expression constructs.

Expression constructs pET-28c (+)-His-STAT3βtc-eCFP (A) and pET-28c (+)-His-STAT3βtc-eYFP were electrophoresed as supercoiled plasmid DNA contained in lane 2 (A and B). Lane 3 (A and B) contains linearized plasmid DNA constructs generated with Nhel restriction endonuclease to reveal the approximate molecular weights of the constructs. Lane 4 (A and B) contains DNA fragments generated with Ndel restriction endonuclease enzymes. Lane 5 (A and B) contains DNA fragments generated with both Ndel and Nhel restriction endonuclease enzymes. Arrows indicate the positions and the molecular weights of the DNA fragments relative to the 1kb DNA maker displayed in lane 1 (A and B).

Lane 2 (A and B) contains unrestricted or supercoiled DNA plasmid (lower band) DNA sample migrating with a molecular weight corresponding to approximately 4.5 kb. Also displayed is the open circular form or the nicked DNA sample (upper band) located above the supercoiled form. The observed smear may be due to contamination of the plasmid DNA constructs caused by the bacterial genomic DNA.

The contents in lane 3 (A and B) are the linearized form of the plasmid DNA constructs generated with Nhel restriction endonuclease. The restriction (Nhel cleaved the plasmid DNA construct at the 3' end of GFP) study revealed the relative molecular weight (~ 7.8 kb) of the constructs. The smear above the linearized sample may be due to contaminations from the bacterial genomic DNA.

Also, displayed in lane 4 (A and B) are the products of the restriction reaction generated with Ndel restriction endonuclease. Ndel was used to clone STAT3βtc; hence, by restricting the constructs with Ndel, STAT3βtc was released. Upon electrophoresis, the relative sizes of the two DNA fragments were revealed as 1.80 kb...
for the STAT3βtc insert and 6.02 kb for the pET-28c (+)-GFP cloning vector. Also displayed is the band located just above the cloning vector which may be due to partial restriction of the plasmid DNA construct. The band migrated with molecular weight corresponding to the linearized form.

In lane 5 (A and B), the products generated with both Ndel and Nhel restriction endonuclease is shown. The use of Ndel (cleaved the construct at both the 5' and 3' ends of the STAT3βtc insert DNA) and Nhel (cleaved the construct at 3' end of the GFP) restriction endonuclease generated DNA fragments corresponding to pET-28c (+) cloning vector (~ 5.3 kb), STAT3βtc-GFP fusion gene (~ 2.52 kb) and GFP gene (~ 0.72 kb). The presence of STAT3βtc-GFP fusion gene product may be due to incomplete restriction of the plasmid DNA by Ndel.

Finally, the expression constructs were sequenced and confirmed using the forward primer of the T7 promoter contained in the pET-28c (+) cloning vector.

3.2.4 pET-28c (+)-His-eCFP-TEV-eYFP expression constructs

The use of GFP-based fluorescent resonance energy transfer to investigate the STAT3βtc molecular interaction in vitro required the development of a FRET control expression construct (i.e. GFP fusion protein also known as a concatemer). In recognizing that the efficiency of energy transfer was dependant on the distance separating the two fluorophores (< 10 nm) (Cubitt et al., 1995), we developed a concatemer where the donor (eCFP) and the acceptor (eYFP) fluorophores were tethered by a spacer containing the TEV protease-specific recognition site (i.e E N L Y F Q G). This expression construct was developed to detect the proteolytic activity of the TEV protease based on the principle of FRET. The schematic representation of the construct is depicted in Figure 3.14.

![Figure 3.14: Schematic representation of His-eCFP-TEV-eYFP constructs](image)

Two GFP mutants linked by a TEV protease-specific sequence cloned and applied in the GFP-based FRET assays. The donor fluorophore (eCFP) is colour coded ( — ) while the acceptor fluorophore (eYFP) is colour coded ( — ). The linker sequence (TEV protease-specific recognition site) is also colour coded as ( — ).
The expression construction was developed by subcloning the PCR-amplified eYFP cDNA into the pET 28c (+)-eCFP-TEV cloning vector. The eYFP gene was fused to the 3' end of the TEV protease-specific recognition site contained in the cloning vector (see section 2.2.6). The cloning vector was originally developed to express the construct His-eCFP-TEV-HIF1β in a supporting study.

Upon isolation of the expression construct (pET-28c (+)-His-eCFP-TEV-eYFP) from DH5α E. coli cell, the concentration of the purified construct was determined (i.e. 325 µg/mL) by absorbance, and the quality, verified by analyzing the samples on an agarose gel via electrophoresis.

Following agarose gel electrophoresis, the restriction map displayed in Figure 3.14 was obtained. The data generated suggested that the concatemer was successfully cloned.

Lane 2 contains unrestricted or supercoiled DNA plasmid (lower band) DNA sample migrating with a molecular weight corresponding to approximately 4.5 kb. Also present is the open circular form (upper band) of the supercoiled plasmid DNA. This form is generated when the supercoiled DNA is damaged during the purification process.
Lane 3 contains the linearized form of the plasmid DNA expression construct, a product achieved with NotI restriction endonuclease. The result of this reaction revealed the relative molecular weight of the expression construct *i.e.* about 6.76 kb. The experimental molecular weight corresponds to the calculated data. As evident in lane 2, the open circular form is present as the upper band, above the linearized plasmid DNA.

In lane 4, the products of a double restriction study is displayed. The construct was restricted with both Ndel and NotI restriction endonuclease to release the GFP fusion gene (*i.e.* eCFP-TEV-eYFP), thus, Ndel restricted the plasmid DNA at the 5' end of the eCFP gene whilst the NotI restricted it at 3' end of the eYFP gene. The relative position of the two DNA fragments corresponds to their expected molecular weights *i.e.* approximately 1.46 kb for eCFP-TEV-eYFP and 5.30 kb for pET-28c (+) cloning vector.

The contents in lane 5 are products of a triple restriction reaction. The expression construct was restricted with Ndel, Nhel and NotI restriction endonucleases to release both eCFP (with Ndel and Nhel) and TEV-eYFP (with Nhel and NotI). The relative positions of the two DNA fragments correspond to their expected molecular weights *i.e.* approximately 0.72 kb for both eCFP and TEV-eYFP and 5.30 kb for pET-28c (+) cloning vector. The size of the TEV-protease-specific DNA sequence (~ 21 bp) had little or no influence on the relative positions of the two DNA fragments.

### 3.2.5 Expression of recombinant, fluorescent chimaeric proteins in *E.coli* cells

Among the various biological host systems for the production of heterologous proteins, the Gram−negative bacterium *Escherichia coli* has emerged as the most attractive and reliable because its genetic make-up is far better characterized than any other microorganism (Baneyx, 1999). Its ability to rapidly multiply to high density on a relatively inexpensive source of nourishment makes it a unique host for expressing large quantities of recombinant proteins, though, it is not certain that the recombinant protein will be expressed by the *E.coli* cell in high density, full length and in a biologically active form. Production of the protein is usually triggered with an inductant, isopropyl-β-D-1-thiogalactopyranoside (IPTG) which is non-hydrolysable lactose analogue used in the case of *lac* and *tac* promoters.

#### 3.2.5.1 Expression of recombinant, fluorescent STAT3βtc in *E.coli* cells

Following cloning and constructs validation, the various constructs were expressed on a small scale to evaluate the expression levels and the solubility of the recombinant protein. Since large amounts of the soluble fluorescent STAT3βtc was required for the
Edwin Nkansah

assay development and protein structural studies, it was imperative that the right expression condition was identified and applied.

As previously stated, solubility of the fluorescent STAT3βtc in *E. coli* was dependant on the order of fusion of the GFP reporter to STAT3βtc. Little or no soluble fraction of the C-terminally tagged fluorescent STAT3βtc (*i.e.* STAT3βtc-GFP) was generated upon lysis. The fluorescent STAT3βtc recombinant protein remained in the lysed pellet of the induced cells ([Figure 3.16](#) (i) and (ii)). Also, the different spectral variants of GFP fused to the STAT3βtc did not influence the level of soluble fractions present in the cell extract.

Further, expression of the STAT3βtc-GFP expression construct in various *E. coli* strains (*i.e.*, BL21 (DE3)’s Rosetta™, TKB1™, Origami™2 and CodonPlus™) and at varied temperatures (18, 21, 25, 30 and 37 °C) identified temperatures 21 °C and 25 °C as the optimum induction temperature. Although, the overall expression level was significantly improved in all *E. coli* strains, there was no obvious improvement in the solubility of the recombinant, STAT3βtc-GFP chimaeric protein as judge by the coomassie stained SDS-PAGE gel, [Figure 3.16](#).

[Figure 3.16]: Expression of the recombinant, STAT3βtc-GFP expression constructs in *E. coli* cell.

Following transformation of the different *E. coli* cells (*i.e.* Rosetta™ *E. coli* cell for the expression of unphosphorylated STAT3βtc-GFP and TKB1™ *E. coli* cell for the expression of phosphorylated STAT3βtc-GFP) with the individual plasmid of the various constructs, the cells were mini-cultured (5 mL) and then induced to express the recombinant, STAT3βtc-GFP protein. The cells were lysed (chemical lysis) and the lysate cleared by centrifugation. The resultant soluble cell extracts and the pellet of the lysed induce cells containing the insoluble STAT3βtc-GFP was separated on an SDS polyacrylamide (10%) gel. (i) Lanes 1, 2, and 3, represent soluble cell extracts from Rosetta™ *E. coli* cells expressing uSTAT3βtc-eCFP, uSTAT3βtc-eGFP and uSTAT3βtc-eYFP respectively; lane 4, molecular weight marker; lanes 5, 6, and 7, represent pellet of lysed induced Rosetta™ *E. coli* cells containing the insoluble uSTAT3βtc-eCFP, uSTAT3βtc-eGFP and uSTAT3βtc-eYFP migrating with an apparent molecular weight of 99 kDa respectively. (ii) 1, 2, and 3, represent soluble cell extracts from TKB1™ *E. coli* cells expressing pSTAT3βtc-eCFP, pSTAT3βtc-eGFP and pSTAT3βtc-eYFP respectively, lanes 4, 5, and 6, represent...
pellet of lysed induced TKB1™ E.coli cells containing the insoluble pSTAT3βtc-eCFP, pSTAT3βtc-eGFP and pSTAT3βtc-eYFP migrating with an apparent molecular weight of 99 kDa respectively; lane 7, molecular weight marker. Arrow denotes the position and the molecular weight of recombinant protein.

In a related study to generate soluble recombinant, STAT3βtc-GFP in the lysed cell extract, we fused a polyhistidine purification tag to the N-terminus of STAT3βtc to generate the recombinant, His-STAT3βtc-GFP expression construct. The idea was persuasive because a similar exercise in a previous study had significantly improved the expression levels and solubility of another recombinant fluorescent fusion protein. Disappointingly, however, very little or no improvement was observed following expression of the His-STAT3βtc-GFP expression construct. Though, there was a significant amount of expression of the recombinant protein (as judged by the fluorescent E.coli cells after IPTG induction), they remained insoluble and trapped in the pellet of the lysed induce Rosetta™ E.coli cells as shown in Figure 3.17.

Following transformation, the Rosetta™ cells were induced to express the recombinant His-STAT3βtc-GFP chimaeric protein. The cells were lysed (chemically) and the lysate cleared by centrifugation. The resultant soluble cell extract and the pellet of the lysed induced cells were separated on a 10 % SDS-polyacrylamide gel. (i) Lanes 1 and 2 represent soluble cell extracts from induced Rosetta™ E.coli cells expressing His-eCFP-uSTAT3βtc and His-eYFP-uSTAT3βtc respectively; lane 3 and 4, represent pellets of lysed induced Rosetta™ E.coli cells containing the insoluble His-eCFP-uSTAT3βtc and His-eYFP-uSTAT3βtc respectively; lane 5, contains pellet of the lysed uninduced Rosetta™ E.coli cells containing the His-eCFP-uSTAT3βtc expression construct. Arrow indicates the location and apparent molecular weight of the recombinant protein.
Conversely, we successfully expressed and obtained soluble fractions of the N-terminally tagged fluorescent STAT3βtc (i.e. GFP-STAT3βtc) in a variety of BL21 (DE3) *E.coli* strains. Recombinant GFP-STAT3βtc expression was induced at low temperature thus; induction was performed at 21 °C. A moderate level of expression was obtained relative to the expression levels of the original unlabeled STAT3βtc construct. **Figure 3.18**, demonstrates that, in addition to aiding protein solubility, the GFP tag on the N-terminus of STAT3βtc also served as a visual marker for STAT3βtc gene expression.

![Figure 3.18: Photographic representation of recombinant, uSTAT3βtc and GFP-uSTAT3βtc expressed in E.coli cells.](image)

The large scale expression of recombinant uSTAT3βtc, and eGFP-uSTAT3βtc in BL21 (DE3) Rosetta™ *E.coli* cells. Following expression of the recombinant protein, the cells were harvested by centrifugation in either 0.5 or 1 L centrifugation tube. (i); 4 x 0.5 L centrifuge tube containing harvested Rosetta *E.coli* cells that was induced to express non fluorescent uSTAT3βtc. (ii); 4 x 0.5 L centrifuge tube containing harvested Rosetta *E.coli* cells that was induced to express fluorescent eGFP-uSTAT3βtc. No irradiation was needed to observe fluorescence since the enhanced GFP mutant strongly auto fluoresce visible light.

The expressed GFP-STAT3βtc had the expected molecular weight of about 99 kDa. Following lysis, approximately 30 % of the GFP-STAT3βtc chimaeric protein was soluble. Thus, a significant percentage of the protein remained insoluble i.e. GFP-STAT3βtc remained incorporated in the pellet of the lysed induced cell (**Figure 3.19**).
Figure 3.19: Expression of the recombinant, GFP-STAT3\text{\textit{ptc}} expression constructs in \textit{E. coli} cell.

Following transformation of the different \textit{E. coli} cells (i.e. Rosetta\textsuperscript{TM} \textit{E. coli} cell for the expression of unphosphorylated GFP-STAT3\text{\textit{ptc}}) with the individual plasmid of the various constructs, the cells were mini-cultured (5 mL) and then induced to express the recombinant, GFP-uSTAT3\text{\textit{ptc}} chimaeric protein. The cells were lysed (chemical lysis) and the lysate cleared by centrifugation. The resultant soluble cell extracts (containing approximately 30 % of GFP-uSTAT3\text{\textit{ptc}}) and the pellet of the lysed induce cells containing the insoluble GFP-uSTAT3\text{\textit{ptc}} was separated on an SDS polyacrylamide (10%) gel. (i) Lanes 1, 2, and 3, represent soluble cell extracts from induced Rosetta\textsuperscript{TM} \textit{E. coli} cells expressing eCFP-uSTAT3\text{\textit{ptc}}, eGFP-uSTAT3\text{\textit{ptc}} and eYFP-uSTAT3\text{\textit{ptc}} respectively; lane 4, contains soluble cell extracts from un-induced Rosetta\textsuperscript{TM} \textit{E. coli} cells expressing eGFP-uSTAT3\text{\textit{ptc}}; lanes 5, 6, and 7, represent pellet of lysed induced Rosetta\textsuperscript{TM} \textit{E. coli} cells containing the insoluble eCFP-uSTAT3\text{\textit{ptc}}, eGFP-uSTAT3\text{\textit{ptc}} and eYFP-uSTAT3\text{\textit{ptc}} respectively; lane 8, contains pellet of the lysed un-induced Rosetta\textsuperscript{TM} \textit{E. coli} cells transformed with the eGFP-uSTAT3\text{\textit{ptc}} plasmid DNA. Arrow denotes the position and the molecular weight of recombinant protein. The recombinant GFP-uSTAT3\text{\textit{ptc}} migrated with an apparent molecular weight of 99 kDa.

Similar expression and solubility levels were observed for the phosphorylated GFP-STAT3\text{\textit{ptc}} (i.e GFP-pSTAT3\text{\textit{ptc}}) generated in the BL21 (DE3) TKB\textsuperscript{1} \textit{E. coli} cells. In addition, the biological activities of the soluble fractions of the recombinant, GFP-STAT3\text{\textit{ptc}} was retained following expression in all available \textit{E. coli} strains including the BL21 (DE3) TKB\textsuperscript{1} and BL21 (DE3) Rosetta\textsuperscript{TM} cells.
3.2.5.2 Expression of the spectral variants of GFP in *E. coli* cell

Following cloning and restriction mapping, the GFP expression constructs; pET 28c (+)-His-eCFP, pET 28c (+)-His-eGFP, and pET 28c (+)-His-eYFP were successfully expressed in BL21 (DE3) Rosetta™ *E. coli* cells. A higher level of expression was observed when induction was performed at low temperature. About 70 to 80 % soluble fraction of the fluorescent protein was obtained with an apparent molecular weight of 29 kDa following separation by SDS-PAGE. Very little insoluble fraction (incorporated in the pellet of the lysed induced cells) was observed.

3.2.6 Lysis of *E. coli* cells for the purification of the soluble recombinant, fluorescent chimaeric proteins

3.2.6.1 Extraction of the recombinant GFP-STAT3βtc chimaeric protein using sonication

The very first step in the purification of cytoplasmic, recombinant, GFP-STAT3βtc expressed in *E. coli* was to lyse the cell to release the protein. The cells expressing the recombinant protein were predominantly lysed by sonication (Becker *et al.*, 1998). Lysozyme was employed occasionally. However, the use of the enzyme was not encouraged because no significant improvement was observed in terms of the level of soluble GFP-STAT3βtc in the cell extract. Unless stated otherwise, cells (i.e. 0.5 or more grams of cells) were lysed by sonication to release recombinant, GFP-STAT3βtc. Following lysis, the lysate was cleared by centrifugation to generate the soluble cell extract and the pellet of the lysed cells. Aliquots were taken at each stage of the extraction process and then analysed by SDS-PAGE, Figure 3.20 (A) and (B). Details of the process are described in section 2.2.2.1.1.
Figure 3.20: Extraction of the recombinant GFP-uSTAT3\beta tc and GFP-pSTAT3\beta tc using sonication.

Aliquots taken during the extraction process were analysed by SDS-PAGE using 10 % polyacrylamide gel. A (i), (ii), (iii), represents eCFP-uSTAT3\beta tc, eGFP-uSTAT3\beta tc and eYFP-uSTAT3\beta tc extracted from Rosetta™ E.coli cells respectively; B (i), (ii) (iii), represents eCFP-pSTAT3\beta tc, eGFP-pSTAT3\beta tc and eYFP-pSTAT3\beta tc extracted fractions TKB1™ E.coli cells respectively. A,B, lane 1, un-induced GFP-STAT3\beta tc transformed cells; lane 2, induce GFP-STAT3\beta tc transformed cells; lane 3, pellet of sonicated induced cells; lane 4, soluble cell extract of sonicated induced cell. The arrow indicates the location and apparent molecular weight of the recombinant, GFP-STAT3\beta tc protein.

Although an amount of the recombinant, GFP-STAT3\beta tc was obtained in the soluble fraction after sonication, a significant percentage of the protein was still left insoluble and “trapped” in the pellet of the lysed induced cell (Figure 3.20 (A), (B) lane 3). Attempts to re-extract the extracted pellet by re-sonicating the re-suspended pellet of the previously lysed cells yielded very little extra soluble protein.

Also demonstrated in lane 2 of Figure 3.20 (A), (B), is the ability of IPTG to induce transcription of GFP-STAT3\beta tc gene. IPTG functions by binding to the lac\text{I} repressor and altering its conformation, which prevents the repression at the lac operator.

3.2.6.2 Extraction of the His-eCFP-TEV-eYFP chimaeric protein

Recombinant, His-eCFP-TEV-eYFP protein was expressed in BL21 (DE3) Rosetta™ E.coli cells at high levels following induction with IPTG at 21 °C (Figure 3.21). The brightly fluorescent E.coli cells were lysed by sonication and then cleared to generate the soluble cell extract and the pellet of the lysed cells.
Approximately 60 to 70% GFP fusion protein was soluble. However, a significant percentage was insoluble and remained trapped in the pellet of the lysed cells. Aliquots of the protein-buffer mixture were taken at each stage of the extraction process and then analyzed by SDS-PAGE to ensure that the correct protein had been expressed (Figure 3.22). The protein had the expected molecular weight of approximately 57kDa.
3.2.7 Western Bolt analysis of the extracted recombinant, GFP-STAT3βtc chimaeric protein

Prior to purification, the cell extracts containing the soluble GFP-STAT3βtc fraction were analysed by western blot to ensure that the appropriate full length recombinant protein had been expressed. Also, proteins expressed in BL21 (DE3) TKB1™ E.coli cells (i.e. GFP-pSTAT3βtc) were verified to ensure their phosphorylated state. An antibody generated against any phosphorylated tyrosine residue (PY20) or the phosphorylated tyrosine residue 705 in STAT3 (PY705) was applied. Unphosphorylated GFP-STAT3βtc (i.e. GFP-uSTAT3βt; usually expressed in Rosetta™ E.coli cells) was detected with antibodies either generated against STAT3 or against the GFP reporter. Figure 3.23 is a western blot demonstrating the detected GFP-STAT3βtc isoforms.

Figure 3.23: Western blot analysis of the GFP-STAT3βtc.

Soluble cell extracts from E.coli strains expressing GFP-uSTAT3βtc and GFP-pSTAT3βtc chimaeric proteins were separated on a 10% SDS-polyacrylamide gel and then transferred to nitrocellulose membrane for development with the various antibodies. (A): SDS-polyacrylamide gel containing; lane 1, unphosphorylated GFP-uSTAT3βtc expressed in Rosetta™ E.coli cell; lane 2, unphosphorylated GFP-uSTAT3βtc expressed in TKB1™ E.coli cell before tyrosine phosphorylation (TK); lane 3, phosphorylated GFP-pSTAT3βtc expressed in TKB1™ E.coli cell after tyrosine phosphorylation (TK). (B): (i), (ii), (iii) and (iv), represents nitrocellulose membranes probed with (1:5000-fold dilution) phosphotyrosine antibody (PY20), GFP antibody, STAT3 antibody (K-15) and STAT3 phosphotyrosine antibody (PY705) respectively. Membranes were developed with DAB and urea/H₂O₂ substrate. Arrows indicates the position and molecular weight the fluorescent STAT3βtc isoforms.
The blot confirmed the successful expression of the GFP-STAT3βtc. Immunoreactive bands in Figure 3.23 B (i) and (iv), demonstrated that phosphorylated fluorescent STAT3βtc was successfully expressed. When the Elk receptor kinase activity was not deliberately induced, we still observed phosphorylation of GFP-STAT3βtc as demonstrated by the immunoreactive band in lane 2 of Figure 3.23 B (i) and (ii). However, the immunoreactive band in lane 3 represent GFP-pSTAT3βtc proteins that were phosphorylated following induction of the Elk receptor kinase domain in the TKB1™E.coli cell. The multiple immunoreactive bands displayed in lane 2 may be due to the fact several proteins with different molecular weight may have been phosphorylated by the receptor kinase domain. Further, the location of the immunoreactive bands in lanes 1, 2 and 3 of Figure 3.23 B (ii) and (iii) clearly suggested that GFP protein is contained in the recombinant chimaeric protein (as demonstrated in (ii) with the anti-GFP antibody) and STAT3βtc is also part of the chimaeric protein (as demonstrated in (iii) with the anti-STAT3 K-15 antibody).

Activity assays performed with both GFP-STAT3βtc samples expressed in TKB1™ E.coli cell using samples before and after Elk kinase induction resulted in a similar outcome thus, also suggesting that the uninduced Elk kinase TKB1™ expressed fraction may in fact contain an active (phosphorylated) recombinant protein.

3.2.8 Purification of N-terminal GFP-STAT3βtc chimaeric protein (GFP-STAT3βtc)

Following E.coli lysis via sonication, the lysate was cleared by centrifugation. The generated cell extract was applied to a general two-step purification procedure that involved the use of ammonium sulphate to precipitate the fluorescent STAT3βtc (section 2.2.3.1.1) and then an anion exchange column to fractionate the re-suspended ammonium sulphate pellet using NaCl gradient (section 2.2.3.1.2). Unless stated otherwise, a third step involving the use of size exclusion chromatography was incorporated.

3.2.8.1 Ammonium sulphate precipitation

Following SDS-PAGE analysis of the various ammonium sulphate protein precipitates, the optimum percentage saturation was identified as 40 % (i.e. ammonium sulphate concentration of 1.76 M) for both the phosphorylated and unphosphorylated fluorescent STAT3βtc. As displayed in Figure 3.24 (A) and (B), a significant percentage (~ 50 to 60) of the fluorescent STAT3βtc was precipitated at 40 % ammonium sulphate saturation. In addition to the amount generated in the 40 % fraction, the fraction was much purer relative to those generated with 45, 50, 60 and 80 % ammonium sulphate.
saturation. Fractions generated with 10, 20 and 30 % ammonium sulphate predominantly contained contaminant with MW less than that of the target fluorescent STAT3\(\beta\)tc. Contaminants with relatively bigger MW were mostly contained 50 and 60 % fractions as evident in lanes 7 and 8 in both gels representing unphosphorylated and phosphorylated fluorescent STAT3\(\beta\)tc chimaeric proteins.

The streaking observed may be due to traces of ammonium sulphate and bacterial DNA fragments present in the loaded samples.

![Ammonium sulphate precipitate fractions](image)

**Figure 3.24:** Ammonium sulphate precipitate fractions analysed for recombinant, unphosphorylated (A) and phosphorylated (B) GFP-STAT3\(\beta\)tc chimaeric proteins. SDS-PAGE (10 %) gel is shown. Lane 1 (A and B), cell extract after lysis; lane 2 (A and b), ammonium sulphate protein precipitate with 10 % saturation; lane 3 (A and B), ammonium sulphate protein precipitate with 20 % saturation; lane 4 (A and B), ammonium sulphate protein precipitate with 30 % saturation; lane 5 (A and B), ammonium sulphate protein precipitate with 40 % saturation; lane 6 (A and B), ammonium sulphate protein precipitate with 45 % saturation; lane 7 (A and B), ammonium sulphate protein precipitate with 50 % saturation; lane 8 (A and B), ammonium sulphate protein precipitate with 60 % saturation; lane 9 (A and B), ammonium sulphate protein precipitate with 80 % saturation. All precipitations were generated at 0 °C to minimize non-reversible denaturation.

Fluorescence intensity measurements from each of the fractions were further used to complement the data displayed in **Figure 3.24** (A) and (B). By exploiting the absorbance of the GFP reporter fused to STAT3\(\beta\)tc, a convenient estimate of the fluorescent STAT3\(\beta\)tc content in each fraction was generated. As shown in **Figure 3.25**, the 40 % ammonium sulphate saturation precipitate recorded the highest emission intensity of approximately 10.5 \(\times\) 10^4 mAU at 480 nm for both unphosphorylated and phosphorylated GFP-STAT3\(\beta\)tc.
Figure 3.25: Demonstration of fluorescent STAT3βtc content present in the fractions generated upon ammonium sulphate precipitation.

The data represent the normalized fluorescence emission intensity generated with 5.0 µg of ammonium sulphate precipitate fractions. The excitation and emission filters used had wavelengths of 405 and 480 nm respectively. Both phosphorylated and unphosphorylated fluorescent STAT3βtc is represented in the data and colour coded as (— ) and ( —  )

Also, the apparent increasing fluorescence intensity between 50 to 60 % (i.e. from 3.6 to 4.3 x10⁴ mAU) may be due to precipitated GFP fluorophore alone generated as a result of proteolysis. Unlike other smaller proteins present in the cell extract, the beta barrel structure of the GFP fluorophore may only precipitate at a higher ammonium sulphate percent saturation.

3.2.S.2 Purification of fluorescent STAT3βtc present in 40 % fraction using anion exchange chromatography

The 40% ammonium sulphate saturated protein precipitate was redissolved and bound to the anion exchange column, the bound protein was eluted by a step-wise increment in salt (NaCl) concentration in the mobile phase (elution buffer E).

The contents of the eluted fractions were analysed by SDS-PAGE. The coomassie stained bands designated by the arrows and demonstrating the relative molecular weight (kDa) were identified as the fluorescent STAT3βtc chimaeric protein by immunoblot analysis (data not shown).
Figure 3.26: Purification of recombinant enhanced (cyan, green and yellow) unphosphorylated STAT3βtc chimaeric protein using anion exchange chromatography.

The SDS-PAGE (7%) used to analyse the eluates from the various spectral variants chimaeric proteins are shown. Upon generating the 40% ammonium sulphate protein precipitate, the protein precipitate was re-dissolved in buffer S and then filtered before binding to the anion exchange column. The bound protein was then eluted at various salt concentrations and then the eluates analysed and described as follows; lane 1(i, ii and iii), load (re-dissolved 40% ammonium sulphate protein precipitate); lane 2 (i, ii and iii), unbound protein described as “flow through”; lane 3(i, ii and iii), wash fraction containing non-specific interaction proteins; lane 4(i, ii and iii), 100 mM eluate; lane 5 (i, ii and iii), 150 mM eluate; lane 6 (i, ii and iii), 200 mM eluate; lane 7 (i, ii and iii), 300 mM eluate; lane 8 (i, ii and iii), 400 mM eluate; lane 9 (i, ii and iii), 1000 mM eluate.

As shown in Figures 3.26 (i, ii and iii) and 3.27 (i, ii and iii) for unphosphorylated and phosphorylated fluorescent STAT3βtc respectively, purer fractions were obtained with 100 and 150 mM NaCl. Thus, about 70% electrophoretic purity was achieved with buffer E supplemented with 100 and 150 mM NaCl for fluorescent STAT3βtc isoforms. With 200 mM salt, the remaining fluorescent STAT3βtc bound to the column was eluted.
Figure 3.27: Purification of recombinant enhanced (cyan, green and yellow) phosphorylated STAT3\textsuperscript{p} chimaeric protein using anion exchange chromatography.

The SDS-PAGE (7 %) used to analyse the eluates from the various spectral variants chimaeric proteins are shown. Upon generating the 40% ammonium sulphate protein precipitate, the protein precipitate was re-dissolved in buffer S and then filtered before binding to the anion exchange column. The bound protein was then eluted at various salt concentrations and then the eluates analysed and described as follows; lane 1(i, ii and iii), load (re-dissolved 40 % ammonium sulphate protein precipitate); lane 2 (i, ii and iii), unbound protein described as “flow through”; lane 3(i, ii and iii), wash fraction containing non-specific interaction proteins; lane 4(i, ii and iii), 100 mM eluate; lane 5 (i, ii and iii), 150 mM eluate; lane 6 (i, ii and iii), 200 mM eluate; lane 7 (i, ii and iii), 300 mM eluate; lane 8 (i, ii and iii), 400 mM eluate; lane 9 (i, ii and iii), 1000 mM eluate.

In addition, as evident in lanes 7, 8, 9 and 10 (i.e. lanes contains samples generated with 300, 400 and 1000 mM NaCl) for both Figures 3.26 and 3.27, subsequent increase in the salt content of the elution buffer generated eluates containing contaminant house-keeping proteins with relatively higher MW. This purification data was generated with an elution buffer with pH 7.0, thus a unit below the binding buffer (i.e. pH 8.0). The drop in pH between the binding buffer and the elution buffer significantly improved the purification process.

Both the unphosphorylated and the phosphorylated GFP-STAT3\textsuperscript{p} demonstrated similar elution profile indicating that the single phosphate group on the phosphorylated STAT3\textsuperscript{p} isoform did not confer any selective advantage during anion exchange purification. In addition, the elution profiles generated for STAT3\textsuperscript{p} fused to the different GFP spectral variant (i.e. data i, ii and iii for both Figures 3.26 and 3.27) were similar, suggesting that like the phosphate group, no selective advantage was conferred by the mutations in the various fluorophores.

The purified 100 and 150 mM fraction were label and stored at 4 °C for short term storage (i.e. protein used within 4 weeks of purification, no significant degradation
observed) or diluted with glycerol to 10 % w/v and then stored at -80 °C (i.e. long-term storage).

The purified unphosphorylated and phosphorylated fluorescent STAT3βtc fractions were applied in both structural (i.e. gel filtration and Native-PAGE/Fergusson plot to characterize the active and inactive fluorescent STAT3βtc) and functional (ELISA, PEMSA and FRET) assays described in subsequent sections.

### 3.2.8.3 Ni²⁺ Sepharose™ affinity purification of recombinant His-eCFP-TEV-eYFP fusion protein

Following lysis of the bacterial cells by sonication, the cell extract was bound to Ni²⁺ Sepharose resin and the His-eCFP-TEV-eYFP fusion protein eluted by step-wise increment in imidazole concentration in the mobile phase. As shown in Figure 3.28, about 80 to 90 % electrophoretic purity was achieved with 200 and 300 mM imidazole supplemented equilibration buffer. A near homogeneous His-eCFP-TEV-eYFP fraction was obtained with 400 mM imidazole. As evident in lane 2, the column’s dynamic binding capacity was not exceeded since no unbound His-eCFP-TEV-eYFP was observed at 57 kDa.

---

Figure 3.28: Purification of recombinant His-eCFP-TEV-eYFP

An SDS-PAGE (10%) used to analyse the contents in each eluate is shown. Upon lysing the bacterial cells, the lysate was cleared and then the cell extract was applied to the column. The bound materials were eluted with equilibration buffer supplemented with increasing concentrations of imidazole. The eluates were then analyzed by SDS-PAGE and described as follows: lane 1, cell extract applied to HisTrap column; lane 2, unbound material described as "flow through"; lane 3, wash fraction containing materials engaged in non specific interactions; lane 4, 100 mM eluate; lane 5, 200 mM eluate; lane 6, 300 mM eluate; lane 7, 400 mM eluate; lane 8, 500 mM eluate.
Also in lane 3, the 30 mM imidazole present in the wash buffer was not sufficient to elute the target protein, however, a significant percentage of the contaminating proteins were eluted. The 200, 300 and 400 mM eluates containing the His-eCFP-TEV-eYFP fusion protein were pooled and dialyzed against dialysis buffer described in section 2.2.3.1.4. The fusion protein was either stored at 4 °C for a short term (i.e. protein used within 4 weeks after purification without significant degradation) or diluted with glycerol to 10 % w/v and then stored at -80 °C (i.e. long-term storage).

The prepared fusion protein was principally used in a FRET-based assay to demonstrate FRET and the subsequent disruption of FRET due to the proteolytic activities of the TEV protease. In addition to the FRET-based assay, the susceptibility of the TEV linker in the purified GFP fusion protein to the TEV protease was tested and demonstrated by SDS-PAGE, section 7.4.

3.2.9 Spectroscopic properties of the fluorescent chimaeric proteins

3.2.9.1 Spectroscopic properties of GFP-STAT3βtc chimaeric proteins

Although it was apparent that the GFP reporter contained in the GFP-STAT3βtc construct was functional as a result of the whole-cell fluorescence, their excitation and emission spectra (excited at 425 nm) were obtained to confirm the retention of the fluorescent properties of the GFP construct. In addition, the emission spectrum of the His-eCFP-TEV-eYFP was also obtained with an excitation wavelength of 425 nm.

3.2.9.1.1 Excitation and emission spectra obtained from enhanced (Cyan, Green and Yellow) GFP-STAT3βtc chimaeric proteins

Measurements were taken in a 3.5 mL disposable cuvette as describe in section 2.9.1. The excitation and emission spectra for eCFP-uSTAT3βtc, eGFP-uSTAT3βtc and eYFP-uSTAT3βtc were obtained with approximately 2.0 nM GFP-STAT3βtc, i.e., Figure 3.29 (A) for excitation and Figure 3.29 (B) for emission. STAT3βtc exhibited no fluorescence in this range (data not shown).
Figure 3.29: Normalized absorption (A; dashed lines) and emission spectra (B; solid lines) for GFP-STAT3βtc chimaeric protein (A), excitation spectra of eCFP-uSTAT3βtc, eGFP-uSTAT3βtc, and eYFP-uSTAT3βtc represented by colour coded broken lines as (———), (-----) and (------) were obtained with emission wavelengths of 475, 509 and 527 nm respectively. (B), the emission spectra of eCFP-uSTAT3βtc, eGFP-uSTAT3βtc, and eYFP-uSTAT3βtc represented by colour coded solid lines as (———), (-----) and (------) were obtained with excitation wavelengths of 430, 493 and 515 nm respectively.

The excitation spectra for the various GFP-uSTAT3βtc chimaeric proteins were unique, with their peak excitation defined as an average intensity of eCFP between 430.5 to 431.5 nm, eGFP between 493.5 to 494.5 nm and eYFP between 513.5 to 514.5 nm. The eCFP-uSTAT3βtc spectrum exhibited two excitation peaks, a major peak at 434 nm and a minor peak at about 453 nm. In addition, as evident in Figure 3.29 (B), two emission peaks were apparent, a major peak at 475 nm and a minor peak at 496 nm.

The observations concurred with the literature account of the spectral characteristics of the eCFP fluorophores (Cubitt et al., 1995). The peak emissions were defined as the average intensity of the eCFP between 474.5 to 475.5 nm, eGFP between 509.5 to 510.5 nm and eYFP between 525.5 to 526.5 nm. Single emission peak profiles were generated for eGFP and eYFP reporters fused to STAT3βtc, an observation similar to that reported in literature (Patterson, et al., 2001).

3.2.9.1.2 Spectral characteristic of His-eCFP-TEV-eYFP

Following purification (section 2.2.3.1.4), the emission spectra of the GFP fusion protein (excited at 425 nm) was obtained to confirm the retention of the spectral properties of the expressed fusion protein. A shown in Figure 3.30, the emission spectra of 1 μM solution of His-eCFP-TEV-eYFP was obtained. Two pronounced emission peaks were apparent; a minor peak at 475 nm, representing the eCFP...
fluorophore and a major peak at 527 nm, representing the eYFP fluorophore. The TEV protease specific linker was not fluorescent in this range.

![Normalized fluorescence emission spectra of His-eCFP-TEV-eYFP fusion protein](image)

Figure 3.30: Normalized fluorescence emission spectra of His-eCFP-TEV-eYFP fusion protein

The emission spectra (excitation 425 nm) non-radiative energy transfer from an excited eCFP fluorophore to the eYFP fluorophore by means of intermolecular long range dipole-dipole coupling.

The spectrum shows a marked decrease in the eCFP emission with a concomitant increase in the eYFP emission at around 527 nm. The generated increase was significantly greater than the emission intensity produced by direct excitation of 1 μM eYFP, suggesting that the two fluorophore are in close proximity and in the range required for energy transfer to occur (typically < 10 nm). Using the ChembioOffice software, the extended sequence length of the TEV protease-specific linker was estimated to be 3.8 nm.
3.3 DISCUSSION

The principal objectives of the work described in this chapter were the construction, expression, extraction and purification of recombinant, fluorescent STAT3βtc chimaeric protein. These objectives were largely achieved since the relevant expression construct of fluorescent STAT3βtc were cloned and successful expressed in the relevant E.coli strains. Other peripheral constructs such as the positive control expression construct for FRET assays, pET-28c (+)-eCFP-TEV-eYFP were developed and successfully expressed in E.coli cells. Following expression of the constructs, the proteins were extracted by sonication and then purified to near homogeneity either by anion exchange chromatography via ammonium sulphate precipitation in the case of the fluorescent STAT3βtc constructs or by Ni²⁺ Sepharose™ affinity purification in the case of the eCFP-TEV-eYFP fusion protein referred to as concatemer.

3.3.1 Construction, expression and extraction of fluorescent STAT3βtc expression constructs.

The critical role of STAT3 in oncogenesis demonstrated in many types of human tumour cell lines and clinical tumour specimen as well as the potential benefits of inhibiting its signalling have turned STAT3 into a valid target for drug development (Bowman et al., 2000; Turkson et al., 2000), and thus have intensified the search for inhibitors that would modulate it activities. Since STAT3 dimerization is a crucial occurrence in its activation, impeding dimerization is predicted to compromise STAT3 signalling and the influence it has on a large number of cellular functions.

As such, to investigate drug-like small molecules capable of interfering with dimer formation, we designed a fusion protein containing the N-terminally truncated spliced variant of STAT3, STAT3βtc (Becker et al., 1998) fused at either the C- or the N-terminus with spectral variants of the GFP (Figures 3.8 and 3.9). The fusion of the GFP reporter to either the N- or the C-termini of the STAT3βtc was performed because it was impossible to predict whether the order of fusion would disturb the biological functions of the parent protein. As illustrated in the previous sections in this chapter, the construction, the expression and the purification of the expression constructs are clearly detailed.

Following cloning of the N-terminus GFP tagged STAT3βtc expression constructs (i.e. pET 32a (+)-GFP-STAT3βtc) via restriction/ligation technique, the constructs were verified by restriction mapping using the restriction endonuclease enzymes used for the
cloning. The clones were further confirmed by oligonucleotide DNA sequencing to highlight any mutations that might have occurred during the cloning process. Likewise, the C-terminus GFP tagged STAT3\(\beta\)tc expression constructs (*i.e.* pET 28c (+)-STAT3\(\beta\)tc-GFP) constructed via restriction/ligation cloning were also verified through restriction mapping and oligonucleotide DNA sequencing. Following verification, the expression constructs (*i.e.* both the N- and the C-terminus GFP tagged STAT3\(\beta\)tc) were appropriate for expression in *E. coli* responsive to the T7 polymerase based expression system (pET vector).

Bacterial expression of fluorescent STAT3\(\beta\)tc of *E. coli* was under the control of the lac Operon. The addition of the molecular mimic of alolactose, IPTG (non-metabolized inducer) allowed the controlled expression of the GFP-STAT3\(\beta\)tc gene downstream of the T7 RNA polymerase binding site (Figure 3.20 (A), (B)).

However as demonstrated in the previous sections, expression of a soluble recombinant, fluorescent STAT3\(\beta\)tc chimaeric protein was dependent on the order of fusion of the GFP reporter. Though a significant amount of the recombinant, GFP-STAT3\(\beta\)tc chimaeric protein was insoluble and therefore incorporated in inclusion bodies, approximately 30 to 40 % was obtained in the soluble cell extract (Figure 3.19). In contrast, although STAT3\(\beta\)tc-GFP was expressed in large amounts, at various expression conditions and in various *E. coli* cell strains (*i.e.* variant of BL21 (DE3) including Rosetta™, TKB1™, Origami™2 and CodonPlus™), almost no soluble fraction was obtained, thus near 100 % of the recombinant, fluorescent STAT3\(\beta\)tc remained insoluble and therefore retained in inclusion bodies (Figure 3.16 (i)(ii)). The incorporation of the recombinant proteins into inclusion bodies was thought due to aggregation of dysfunctional (misfolded) fluorescent STAT3\(\beta\)tc chimaeric protein. It is usually speculated that they form when the non-native monomeric recombinant protein molecules self-assembly into satellites of stable, but growing aggregates which are thermodynamically favourable (Fink, 1998).

Furthermore, since other members of the group had managed to express non-fluorescent polyhistidine tagged STAT3\(\beta\)tc in milligram amounts, a polyhistidine tagged fluorescent STAT3\(\beta\)tc construct (His-STAT3\(\beta\)tc-GFP) was also developed. Though successfully expressed in large quantities as judged by the coomassie stained band, little or no soluble fraction was generated following lysis and SDS-PAGE analysis (Figures 3.12 and 3.17). Although similar fusion constructs have been designed and developed by other research groups, expression of the constructs were performed in mammalian cell lines. Moreover, the expression construct contained the full length
Edwin Nkansah

STAT3 with the GFP reporter fused to the C-terminus (Schroder et al., 2004; Kretzschmar et al., 2004).

Relative to STAT3; the alternative splice variant, STAT3β is shortened by 48 amino acids at the carboxyl terminal transactivation domain (Caldenhoven et al., 1996). The truncation is reported to have conferred a significant functional advantage on the β variant such as improved dimer stability (Park et al., 2000). Furthermore, the reported reduction in the dephosphorylation rate and the concomitant increase in the DNA-binding activity (Moriggi et al., 1996) suggested that the STAT3β was indeed the ideal variant for our binding studies. Furthermore, equilibrium sedimentation studies by Braunstein and colleagues also revealed that the N-terminally truncated STAT3β isoform, STAT3βtc, was predominantly monomeric (Braunstein et al., 2003) and could be expressed in large quantities (Becker et al., 1998), thus making it beneficial to structural studies and large scale drug screen.

Though STAT1 and STAT3 are mostly activated in response to cytokines and growth factors (Darnell et al., 1994; Vinkemeier et al., 1996) in vivo, Shuai et al (Shuai et al., 1996) successfully activated the N-terminally truncated STAT1β fragment in vitro using the tyrosine kinase activity of the immunoprecipitated EGF receptor from A431 cells and resulting in yields of up to 75 %. Following from Shuai approach, Becker et al (Becker et al., 1998) unsuccessfully attempted to activate STAT3βtc in vitro with the immunoprecipitated tyrosine kinase receptor, but managed to activate the protein in vivo with conditions similar to Shuai’s. In this study, the method of Becker et al. was used to activate STAT3βtc by phosphorylating the Y705 using the tyrosine kinase domain of the Elk receptor (a member of the Eph receptor family) co-expressed with STAT3 in the E. coli BL21 (DE3) cell line TKB1™.

The TKB1™ cell line is a tyrosine kinase (TK) derivative of the BL21 (DE3) E.coli strain that carries a plasmid encoding an inducible tyrosine kinase gene (pTK). The gene is under the control of the trp promoter whose activity is inhibited by tryptophan (an amino acid found in yeast extract and tryptone in Luria bertani (LB) medium). Therefore, protein activation is not expected in the IPTG induced medium, but in the tryptophan-free TK induction medium (Studier et al., 1986; Weiner et al., 1994).

Following from Becker's approach for pSTAT3βtc, recombinant, GFP-pSTAT3βtc chimaeric protein was successfully expressed in the active form. Activation of the recombinant, GFP-STAT3βtc was confirmed by western blot using both the pY20 and pY705 antibodies (Figure 3.23 B (i), (iv)). It was perhaps surprising when it was revealed following western blot analysis that the control aliquots collected from the
Edwin Nkansah

IPTG induced and TK induced media were reactive to both anti-phosphotyrosine antibodies, i.e., pY20 and pY705. The observation suggested that the recombinant protein samples were both phosphorylated at the Y705 residue (Figure 3.23 B (i) (iv); lanes, 1 and 2). Functional studies (DNA binding studies) of the protein samples further confirmed that both pre- and post TK treated aliquots were activated.

In contrast, the inactive (unphosphorylated) GFP-uSTAT3βtc constructs were generally expressed in the BL21 (DE3) Rosetta™ E.coli strain lacking the kinase. Full length constructs, represented by the immunoreactive bands corresponding to the predicted size of the constructs are demonstrated in Figure 3.23 B (ii); lanes 3.

3.3.2 Purification of fluorescent STAT3βtc

The process of purification was started by lysing the E.coli host cells to release the cytoplasmic fluorescent STAT3βtc. Cells were lysed predominantly by sonication. Though during sonication, heat was generated which may be detrimental to the chimaeric protein, it was the obvious choice because it was convenient, reliable and reproducible. Lysis with lysozymes was not encouraged since the treatment usually added extra contaminating proteins to the cell extract which may include unwanted proteolytic enzymes.

After lysis, the lysate was usually cleared by centrifugation and the resultant soluble fraction subjected to a stepwise salting-out process with ammonium sulphate. Prior to fractionation by ammonium sulphate, fragmented nucleic acid generated through sonication was originally removed with polyethylenimine (PEI, cationic polymer) as describe by Becker, et al. However, significant amount of the recombinant protein was lost to the PEI-DNA complex precipitate. This was evident by the green fluorescent nature of the supposedly whitish PEI-DNA complex precipitate. In addition, downstream applications including functional assays (e.g. ELISA) performed with PEI-treated protein sample generated severely compromised results compared to that obtained with PEI-free protein samples. Similarly, PEI-treated STAT3βtc (not labelled with GFP) applied in ELISA assays generated data with significantly high signal to noise ratio. Also, deoxyribonuclease 1 (DNASE1) (non-specifically hydrolyzes DNA to generate 5(a)-phosphodinucleotide and oligonucleotide end products) was also employed, however, spectrophotometric measurements (absorbance at 260 nm) revealed that almost 100 % DNA of the cellular DNA was either removed or degraded at the end of the purification phase, i.e., ammonium sulphate precipitation (~ 30%) and anion exchange chromatography (~ 60-65 %).
As evident in Figure 3.24, at 40% ammonium sulphate saturation, a significant percentage of the recombinant protein formed aggregates and subsequently precipitated along with other contaminating proteins of relatively lower molecular weight. Both active and inactive isoforms of the protein generated similar data. Also obvious was the fact that at a relatively lower percentage saturation of ammonium sulphate – perhaps up to 45 %, the generated protein precipitate contained proteins samples with molecular weight less than or equal to the target protein while proteins with a molecular weight greater than the target protein were generally precipitated at a considerably higher ammonium sulphate saturation between 50 to 80%. This observation was further confirmed with the histogram represented as Figure 3.25, where at 40 % saturation, both GFP-STAT3Btc isoforms generated the highest and fairly similar emission intensities (~10,000 mAU), with relatively lower intensities recorded for other fractions. Hence, on the basis of their relative hydrophilicity, GFP-STAT3Btc was separated from the crude cell extract by a gradual saturation of the extract with ammonium sulphate.

The resultant ammonium sulphate precipitate pellet (i.e. 40 % saturation) was subjected to further purification through anion exchange chromatography. Several ion exchange columns including the Vivapure C, D, Q and S columns were evaluated to identify the appropriate column required to obtain a homogeneous mixture of the GFP-STAT3Btc construct. In line with the acidic nature (theoretical pi, approximately 6.26) of the GFP-STAT3Btc protein, the Q column, pre-packed with positively charged quaternary ammonium ion was used to generate a approximately 70 to 80 % homogeneity. The recombinant protein was bound to the column with a higher pH (8.0) and then washed and eluted with a corresponding lower pH (7.0) buffer supplemented with NaCl, up to concentrations of 1 M. The drop in pH by one unit permitted the separation of a major contaminating protein that migrated (SDS-PAGE) with an apparent molecular weight of approximately 60 kDa, thus purity of the target protein was significantly enhanced when it was eluted with a pH 7.0 buffer. Reasonably pure fractions of fluorescent STAT3Btc isoforms were obtained in the 100, 150 and 200 mM eluates. Moreover, relatively larger contaminants were eluted with 300 400 and 1000 mM NaCl respectively (Figures 3.26 and 3.27).

The final step in the purification process employed the use of a size exclusion column (either Superose™ column; S6 or S12) to further purify the pooled 100, 150 and 200 mM eluates. Though this purification step was not extensively employed, it was specifically used to characterize the phosphorylated and unphosphorylated GFP-STAT3Btc recombinant protein isoforms, chapter 4.
3.3.3 GFP-STAT3βtc retains spectral properties

The spectral properties of the various spectral variants of the GFP reporter fused to the N-terminus of the STAT3βtc were retained following extraction and purification of the GFP-STAT3βtc isoforms. The rigid and robust structure of the GFP fluorophore, (Ormø et al., 1996), allowed it to withstand up to 0.1 % SDS and temperatures up to 65 °C.

Further, the excitation and emission spectrum of the GFP-STAT3βtc constructs (i.e. eCFP-STAT3βtc, eGFP-STAT3βtc and eYCPF-STAT3βtc), Figure 3.29 (A) (B), were in agreement with those reported by Miyawaki and colleagues as well as Patterson and colleagues. The observation did not only confirm that their spectral properties GFP variants were intact, but also confirmed that the fusion did not affect the folding of the GFP protein.

In summary, C- and N-terminal GFP tagged STAT3βtc expression constructs were successfully developed and expressed in E.coli cells, though little or no soluble fraction of the C-terminal GFP tagged construct (STAT3βtc-GFP) was obtained. In contrast, soluble fractions of the N-terminal GFP tagged STAT3βtc (GFP-STAT3βtc) construct was obtained. Further, an in vivo activation procedure using the TKB1 E.coli strain allowed the easy preparation of active or phosphorylated N-terminal GFP tagged STAT3βtc (GFP-pSTAT3βtc). The phosphorylated and unphosphorylated protein samples were responsive to their respective antibodies. A simple three-step purification protocol resulted in production of about 80% homogeneous GFP-STAT3βtc protein in mg amounts. Spectroscopic analysis of the purified proteins suggested that various GFP-STAT3βtc constructs, i.e., eCFP-STAT3βtc, eGFP-STAT3βtc and eYFP-STAT3βtc, retained their spectral properties.
4 STRUCTURAL CHARACTERIZATION

4.1 Characterization of phosphorylated and unphosphorylated GFP-STAT3βtc isoform

In the classical STAT3 signaling pathway, activation of STAT3 by interleukin-6 (IL-6) is followed by the formation of phosphorylated STAT3 dimers from the cytoplasmic pool of unphosphorylated STAT3 monomers (Shuai et al., 1994). These phosphorylated STAT3 dimers translocate to the nucleus and bind DNA response elements to initiate transcription of target genes (Wegenka et al., 1993). The ability to translocate into the nucleus and to bind DNA is strongly attributed to dimerization which in-turn is strongly dependent on tyrosine phosphorylation (Chen et al., 1998; Becker et al., 1998; Schindler et al., 1992). In recent years, however, a number of studies have questioned the classical model of STAT3 signaling. Based on co-immunoprecipitation experiments, some researchers have suggested that prior to their activation, STAT3 exist as dimers (Novak et al., 1998; Haan et al., 2000). Furthermore, Ndubuisi's and other research groups suggested that only a small fraction of cytoplasmic STAT3 exist as monomers, with majority, not only associated with each other but also with additional scaffolding proteins. These multimeric ‘STATosomes’ were described as large and poorly characterized molecular mass complex(es) ranging in size from 200 kDa to 2 mDa (Stancato et al., 1996; Haan et al., 2000; Ndubuisi et al., 1999). In addition, equilibrium sedimentation studies have shown that unphosphorylated STAT3 homodimers are reasonably stable, with Kd value less than 10 nm. Moreover, a similar study using N- and C-terminally truncated mutant of STAT3, STAT3βtc (Becker et al., 1998) revealed that a significant fraction existed as monomers (0.2 to 1.0 mg/mL), suggesting a significant decrease in dimer stability (Braunstein et al., 2003). This revelation gives credence to speculations that the amino-terminal domain (and/or the carboxyl-terminal transcriptional activation domain) generally stabilizes homodimerization of unphosphorylated STAT3 proteins (Kisseleva et al., 2002; Darnell et al., 1997).

Consistent with this, in our effort to understand and characterized the conformational changes that would be associated with generating phosphorylated and unphosphorylated GFP-STAT3βtc in E.coli, we have employed distinct, but complementary methods, such as native gel electrophoresis (i.e. Ferguson Plots (Shuai et al., 1994), size exclusion chromatography, circular dichroism and mass spectrometry. Studies were performed with purified preparations of the GFP-STAT3βtc recombinant protein isoforms.
4.1.1 Native Triton X-100 Polyacrylamide Gel Electrophoresis (Native PAGE)

In preparation for a successful fractionation to evaluate the structural conformation of the purified GFP-STAT3ptc isoforms by native Triton X-100 PAGE, certain experimental conditions deemed critical were investigated. It was imperative that the protein sample to be analyzed on the native gel remained soluble before and during electrophoresis. Thus, it was important that the samples including the GFP-STAT3ptc did not form aggregates or multimer via inter-chain disulphide linkage. The optimum working concentrations of Triton X-100 and DTT was investigated. Further, we investigate the effect of freeze/thaw cycles on the conformation of the fluorescent STAT3ptc.

4.1.1.1 Solubilizing effect of Triton X-100 on GFP-STAT3ptc isoforms

The effect of Triton X-100 was evaluated at different concentrations. Thus, the reaction mixture described in section 4.1.1.2.1 was supplemented with 0.05, 0.10, 0.15 and 2 % Triton X-100. Two sets of the reaction mixtures were prepared. The first set was incubated at room temperature (~ 24 °C) and the second set was incubated at 37 °C. Both sets of reactions were incubated for 1h. Following incubation, the reactions were treated as described in section 4.1.1.2.2.

4.1.1.2 Reducing GFP-STAT3ptc isoforms at varied temperatures

In addition to Triton X-100, the optimum working conditions of dithiothreitol (DTT) was evaluated. As described by Shaw and colleagues (Shaw et al., 2003), STAT3 dimers stabilized by inter-chain disulphide linkages was observed following SDS-PAGE analysis of cell lysate. Thus, the conditions necessary to generate the optimum reduction activity of DTT in our reaction mixtures were evaluated. As described in section 4.1.1.2.1, two sets of reactions containing either unphosphorylated or phosphorylated GFP-STAT3ptc and supplemented with DTT to either 40 or 200 mM DTT were prepared. The first set was made up of ten reactions, 5 containing GFP-uSTAT3ptc and the remaining 5 containing GFP-pSTAT3ptc. Reactions were supplemented with 40 mM DTT. The preparations were then paired, each containing GFP-uSTAT3ptc and GFP-pSTAT3ptc. The second set of reactions was similar to the first set, except that, the concentration of DTT was increased to 200 mM. In all, 2x 5 pairs of reactions were prepared. Each pair was incubated for 1 h at either 0 or 4 °C or 24 °C or 37 °C or 50 °C. Upon incubation the reactions were treated as described in section 4.1.1.2.2.
4.1.1.3 Disruptive effect of freeze/thaw treatment on GFP-STAT3ptc

In addition to chemical disruption of GFP-STAT3βtc multimers or aggregates; formation of which may be directed by inter-chain disulphide linkages, disruption by freeze/thaw treatment (physical means) in a reduced environment was evaluated. Eight reactions containing either GFP-uSTAT3βtc or GFP-pSTAT3βtc were prepared and divided into 4 pairs, each containing GFP-uSTAT3βtc and GFP-pSTAT3βtc reaction mixtures, section 4.1.1.2.1. The first pair of reactions had no DTT added whilst the second pair was supplemented with 100 mM DTT. The third and fourth pairs were subjected to 15 min of freeze/thaw treatment in absence and presence of 100 mM DTT respectively. Following incubation for 30 min, the reactions were treated as described in section 4.1.1.2.2. After electrophoresis, the gels were illuminated with blue light and then the images of the separated GFP-STAT3βtc proteins captured with a Kodak 120 206 M digital camera.

4.1.1.2 Native Triton X-100 Polyacrylamide Gel Electrophoresis (Native PAGE)/Ferguson Plots

Electrophoretic mobility studies have reported that a direct relationship exist between the molecular mass of a protein and its relative migration through a number of gels that differ in their acrylamide concentration (Shuai et al., 1994). By using 5, 6, 7 and 8% Triton X-100 native polyacrylamide gel, we attempted to characterize purified phosphorylated and unphosphorylated GFP-STAT3βtc chimaeric proteins isoforms in terms of dimers and monomers respectively. Their relative migrations compared with that of a series of protein standards; i.e., ovalbumin (43 kDa); bovine serum albumin monomer (66 kDa); conalbumin (75 kDa); adolase (158 kDa); ferritin (440 kDa) (Gel Filtration Calibration Kit HMW, GE Healthcare, Buckinghamshire, UK), electrophoresed simultaneously in a 5, 6, 7 and 8% Triton X-100 native gels were computed and used to generate curves from which the molecular weight of the GFP-STAT3βtc isoforms were estimated.

As described in section 2.2.5.3.2, 5, 6, 7 and 8 percent Triton X-100 polyacrylamide gel solutions. Upon polymerization, the gels were assembled in the buffer reservoir, section 2.2.5.3.1., and then equilibrated, section 2.2.5.3.2. The wells were flushed with fresh ice-cold reservoir electrode buffer and then the samples loaded.

4.1.1.2.1 Sample preparation

The protein standards; ovalbumin (43 kDa); Bovine Serum Albumin monomer BSA (66 kDa); conalbumin (75 kDa); adolase (158 kDa); ferritin (440 kDa) were dissolved to
mg/mL, 4 mg/mL, 3 mg/mL, 4 mg/mL and 0.3 mg/mL respectively with the calibration buffer (100 mM Tris-HCl, pH 8.5, 100 mM NaCl).

Into each protein sample, i.e., the phosphorylated and the unphosphorylated GFP-STAT3βtc (~ 8 µg), 5 µL 6x Triton X-100 reaction buffer (360 mM Tris-HCl, pH 8.8, 60 % (v/v) glycerol and 12 % (w/v) Triton X-100) was added and then supplemented with the appropriate concentration of DTT (final DTT concentration used was specific to a particular study, however, 40 mM was usually employed). The preparation was subsequently made up to 30 µL with ddH₂O. The sample was mixed gently and then centrifuged at 13,000 xg for 15 s to drain the mix to the bottom of the tube as well as sediment particulates and insoluble materials. Prior to electrophoresis, the sample was incubated at room temperature for 30 to 60 min. The preparation was either centrifuged at 13,000 xg for 15 s before loading or loaded directly after incubation depending on the study. However, a sample was usually subjected to further centrifugation before electrophoresis. Wells were flushed with fresh ice-cold reservoir buffer prior to loading.

4.1.1.2.2 Electrophoresis of non-denatured protein samples

Into each well, i.e., 5 mm x 1mm, 7.5 µL (~2 µg) of the reaction mix was loaded using a long loading tip. One lane per gel contained bromophenol blue loading dye representing the dye front. The loaded samples were electrophoresed at 100 V simultaneously until the bromophenol blue dye had migrated approximately 2/3 to 3/4 down the length of the gel. The run was stopped and the separated GFP-STAT3βtc isoforms visualized under blue light. The image of the separation was captured using the Kodak DC120 206 M digital camera. The gel was subsequently stained with coomassie brilliant blue for 30 to 60 min and then de-stained with ddH₂O (~ 30 mL, exchanged every 30-40 min) to generate a distinctively stained protein bands. The gels were then subjected to analysis section 4.1.1.2.3.

4.1.1.2.3 Evaluation of molecular weight using Ferguson plot

Before drying the gels, the relative migration of the purified GFP-STAT3βtc isoforms were documented and compared with those of the several protein standards electrophoresed simultaneously on the 5, 6, 7, and 8% gels.

The relative migrations, i.e., distance in centimeter (cm), were documented using a pair of calipers and a standard ruler. The mean distance between the base of each well to the leading (a) and the trailing (b) end of each stained protein band was documented. Figure 4.1 describes how the distance travelled by each of the protein samples was generated from the gels.
Using the collected data, the percent acrylamide concentration was plotted on the x-axis and the relative migration ($R_m$), represented as $100[\log (R_m \times 100)]$, was plotted on the y-axis (Braunstein et al., 2003). According to Ferguson, negative slopes for each of the protein samples, i.e., both protein standards and the GFP-STAT3$\beta$t isoforms; represent the retardation coefficient for each sample.

Further, a plot was constructed with the retardation coefficients ($K_r$) on the y-axis and the molecular weights of protein standards on the x-axis. From the plot, the molecular weights of both GFP-STAT3$\beta$t isoforms were estimated.

4.1.1.3 Size exclusion chromatography (SEC)

To further explore the oligomeric state of the GFP-STAT3$\beta$t isoforms, SEC was employed. Unlike Native-PAGE, section 4.1.1.2, where the GFP-STAT3$\beta$t isoforms were separated based on their intrinsic charges and to some extent their hydrodynamic size, fractionation by SEC through Superose™ columns was primarily based on the sizes of the GFP-STAT3$\beta$t isoforms. In this study, the Superose™ 6 10/300 GL Tricon™ SEC columns connected to the ÄKTA™ purifier system was used. The molecular weight of the GFP-STAT3$\beta$t isoforms were determined by comparing their elution volume parameter, gel-phase distribution coefficient ($K_{av}$), with those generated for the high molecular weight calibration standards including; ovalbumin (43 kDa); bovine serum albumin monomer (66 kDa); conalbumin (75 kDa); adolase (158 kDa) and thyroglobulin (669 kDa).
4.1.1.3.1 Preparation of Superose™ 6 column for gel filtration chromatography

Prior to connecting the column to the ÄKTA™ Purifier system, we ensured that there was no air bubbles trapped in the tubing by flushing them with filtered and degassed ddH_2O. The column was then connected to the purifier system and then together with the pump, flushed with 2.5 column volumes (60 mL) of ddH_2O at 0.2 mL/min at 4 °C. The column was subsequently equilibrated with filtered and degassed equilibration buffer (60 mL); 20 mM HEPES, pH 7.0, 100 mM NaCl, 1 mM DTT and 1 mM EDTA at 0.5 mL/min at 4 °C. After equilibrating the column, the purifier system was prepared to fractionate at 0.5 mL/min with 0.5 mL fraction collection.

4.1.1.3.2 Sample preparation

Pre-purified (sections 2.2.3.1.1 and 2.2.3.1.2) phosphorylated or unphosphorylated GFP-STAT3ptc was made up to 150 µL (~ 240 µg) using the equilibration buffer supplemented with 1 mM DTT. In addition, the protein standards (see section 4.1.1.2.1) were also made up to 150 µL (~ 100-150 µg) from their stock concentrations of 4 mg/mL, 4 mg/mL, 3 mg/mL, 4 mg/mL and 5 mg/mL for ovalbumin (43 kDa), bovine serum albumin monomer (66 kDa), conalbumin (75 kDa), aldolase (158 kDa) and thyroglobulin (669 kDa) respectively with the equilibration buffer. The buffer was also supplemented with 1 mM DTT. The samples were prepared in a 0.5 mL sterile tube. Following preparation, the tubes were centrifuged for 15 s at 13,000 xg at room temperature. The samples were incubated at room temperature for 30 min. Before loading, the samples were centrifuged extensively, i.e., 90 s at 13,000 xg at room temperature to completely sediment particulate and insoluble debris.

4.1.1.3.3 Determination of elution volume (V_e)

100 µL (< 2 % of the geometric column volume (V_c)) of the individual protein standards were loaded into the injection loop (200 µL) and then fractionated on the column. No eluates were collected. The injection loop was manually flushed with 5 mL equilibration buffer after each run. The order of loading is as follows; ovalbumin (43 kDa), bovine serum albumin monomer (66 kDa), conalbumin (75 kDa), aldolase (158 kDa) and thyroglobulin (669 kDa).

Upon documenting the elution volumes of the protein standards, those of the phosphorylated and unphosphorylated GFP-STAT3ptc isoforms were also obtained. A total of 36 0.5 mL fractions were collected for each isoform. The eluates were collected into 1.5 mL tubes, labeled according to the elution volumes and then stored at -20 °C for future SDS-PAGE and western blot analysis.
The elution volume was defined as the volume of the eluent from the point of injection to the center of the elution peak, see Table 4.3.

4.1.1.3.4 Determination of void volume (V₀)

As stated by the manufactures of the Gel Filtration Calibration Kit, the elution volume for blue Dextran 2000 is equivalent to the void volume (V₀) of the column. Hence, a 1 mg/mL (0.5 mL) blue Dextran 2000 solution was prepared with equilibration buffer. The solution was agitated until the Blue Dextran was completely solubilized. The solution (100 μL) was loaded onto the column to determine the void volume. The void volume was calculated from the first eluted peak (A280 nm) from Blue Dextran. No eluate was collected.

4.1.1.3.5 Preparation of calibration curve

The gel-phase distribution coefficient (Kav) values for each of the protein standards and as well as the GFP-STAT3βtc isoforms were obtained from their elution volumes (Ve), section 4.1.1.3.3 and the void volume (V₀) of the column, section 4.1.1.3.4. Hence to calculate the Kav values for the protein standards, we applied the equation below:

\[
K_{av} = \frac{V_e - V_o}{V_c - V_o}
\]

Equation 4.1: Gel-phase distribution coefficient where Vc = column void volume, Vₑ = elution volume, Vc = geometric column volume.

With the Kav values, Table 4.3, and the molecular weights (MW) of the protein standards, a calibration curve of Kav versus logarithm MW of each of the protein standards was generated. From the Kav values for the phosphorylated and unphosphorylated GFP-STAT3βtc isoforms, their corresponding molecular weights were determined from the calibration curve, Figure 4.7.

4.2 RESULTS

4.2.1 Analysis of purified GFP-STAT3βtc by native Triton X-100 PAGE

Prior to the application of the GFP-uSTAT3βtc and GFP-pSTAT3βtc to functional assays, both isoforms were structurally characterized to evaluate protein stability and changes in the conformation of active (phosphorylated) and inactive (unphosphorylated) forms of purified fluorescent STAT3βtc. Both isoforms were
analyzed by Native Triton X-100 PAGE/Ferguson Plots and gel filtration chromatography.

To evaluate protein solubility, complex formation and structural or conformational stability of GFP-STAT3\(\beta\)tc, native Triton X-100 acrylamide gel analysis was employed. We treated the GFP-STAT3\(\beta\)tc isoforms under various conditions before separation by the Native Triton X-100 PAGE. By supplementing the reaction buffer and the resolving gels with Triton X-100, the solubilisation effect of the non-ionic detergent on the GFP-STAT3\(\beta\)tc isoforms was evaluated. This was demonstrated in Figure 4.2 (A and B), where, two identical sets of reactions treated with increasing percentage concentration of Triton X-100 were incubated for 1 h at either 24°C or 37°C.

Though surprising, both GFP-uSTAT3\(\beta\)tc (monomeric) and GFP-pSTAT3\(\beta\)tc (putative dimer) migrated to adjacent locations on the native gel. Further, no obvious structural and physical alterations to either GFP-uSTAT3\(\beta\)tc or GFP-pSTAT3\(\beta\)tc was observed following native-Triton X-100 PAGE analyses of the samples pre-incubated at 24°C. In contrast, samples pre-incubated with Triton X-100 at 37°C displayed significant aggregation or precipitation that was apparent in the sample wells containing the unphosphorylated GFP-STAT3\(\beta\)tc (GFP-uSTAT3\(\beta\)tc) isoform. The protein aggregates were obvious in the wells of lanes 2 to 5 but not in lane 1 of Figure 4.2 (B) (sample preparation analyzed in lane 1 contains no Triton X-100).

To confirm at temperatures greater than 24 °C; GFP-uSTAT3\(\beta\)tc precipitates and forms aggregates in the presence of Triton X-100, preparations of GFP-uSTAT3\(\beta\)tc and GFP-pSTAT3\(\beta\)tc supplemented with 0.1 % Triton X-100 and then pre-incubated at 0, 4, 24, 37 and 50°C were prepared and the separated as captured in Figure 4.2 (C).
Figure 4.2: Disruption and solubilisation of fluorescent STAT3ptc.

Analyses were performed on a Native Triton X-100 gel (7%). Preparations containing either GFP-uSTAT3ptc (1 to 5) or GFP-pSTAT3ptc (6 to 10) were incubated with increasing percentage concentration of Triton X-100 (i.e. from 0 to 2%) at room temperature (A) and at 37 °C (B) for 1hr. The reactions were loaded into sample wells washed with ice-cold tank buffer and separated by Native Triton X-100 PAGE. Surprisingly, both GFP-uSTAT3ptc (supposedly monomeric) and GFP-pSTAT3ptc (putative dimer) migrated to adjacent locations in the gel. In gel C, preparations containing GFP-uSTAT3ptc (u) and GFP-pSTAT3ptc and supplemented with 0.1% Triton X-100 were incubated for 1h at 0 °C, 4 °C, 24 °C (room temperature), 37 °C and 50 °C and then separated by native Triton PAGE. In gel D, reactions containing either GFP-uSTAT3ptc or GFP-pSTAT3ptc and supplemented with 0.1% Triton X-100 were prepared. Also, reactions were either supplemented with or without 100 mM DTT prior to several cycles of freeze/thaw treatment before separation on a native Triton X-100 gel. Preparations analysed in lanes 1 to 4 contained GFP-uSTAT3ptc and those analyzed in lanes 5 to 8 contained GFP-pSTAT3ptc. Samples in lanes 1 and 5 were incubated in reaction buffer with no DTT, whilst, samples in lanes 2 and 6 were supplemented with 100 mM DTT. Samples in lanes 3 and 7 were subjected to 2 cycles of freezing (15 min at -80 °C) and thawing (15 min at 24 °C) treatment during incubation in a reaction with no DTT, whilst, samples in lanes 4 and 8 were subjected to 2 cycles of freezing (15 min) and thawing (15 min at 24 °C) treatment during incubation in a reaction buffer containing 100 mM DTT. Samples were separated by native Triton X-100 PAGE. The resultant gels following electrophoresis were illuminated with blue light and then the images capture with a Kodak C 120 digital camera.
As expected, significant levels of precipitation were observed for the GFP-uSTAT3βtc preparations pre-incubated at 37 °C and 50 °C (Figure 4.2 (C), lanes 7 and 9). Near complete precipitation was observed for the sample pre-incubated at 50 °C. In contrast, as evident in lanes 2, 4, 6, 8 and 10 of Figure 4.2 (C), no precipitation was observed for the GFP-pSTAT3βtc preparations pre-incubated at the various temperatures.

To improve GFP-uSTAT3βtc solubility and minimize precipitation in order to facilitate resolution on the Triton X-100 gel, preparations containing GFP-uSTAT3βtc supplemented with or without 100 mM DTT were pre-incubated and simultaneously subjected to 2 cycles of freezing/thawing treatment. As shown in lane 1 of the Figure 4.2 (D), the absence of DTT resulted in significant precipitation and poor resolution of the GFP-uSTAT3βtc. The precipitate survived 2 cycles of freeze (-80 °C)/thaw (24 °C) treatment (lane 3) but not after supplementing the reaction mixture with 100 mM DTT (lane 4). A similar observation was apparent in lane 2, when the preparation was supplemented with 100 mM DTT but not subjected to the freeze/thaw treatment.

In the case of GFP-pSTAT3βtc, it was obvious that the 100 mM DTT employed had little or no effect on the conformation and resolution of the protein. However, tighter band were obtained for the 100 mM DTT treated samples (lanes 6 and 8) as opposed to the non-DTT treated samples in lanes 5 and 7. The freeze/thaw treatment had no obvious effect on the GFP-pSTAT3βtc isoform.

In a related study, 40 mM DTT was sufficient to decrease precipitate generated from a preparation containing approximately 2 μg of the GFP-uSTAT3βtc isoform. Thus, the most probable explanation for the extensive precipitation of the GFP-uSTAT3βtc isoform at temperatures greater than 24°C may be due to the formation of inter-chain disulphide bonds which is also consistent with the ability of 40 mM DTT to reduce the complex. However, when the DTT content in identical preparations was increased five-fold to 200 mM, GFP-uSTAT3βtc was resolved into 3 distinct bands, suggesting that the excess DTT might have encouraged multimerization of GFP-uSTAT3βtc proteins or the complexation of GFP-uSTAT3βtc with another contaminating protein of different or identical charge and size. These observations are displayed in Figure 4.2.

4.2.2 Characterization of GFP-uSTAT3βtc and GFP-pSTAT3βtc by Native-PAGE (TX)

Non-denaturing Triton X-100 polyacrylamide gel electrophoresis (PAGE) was one of the techniques used to explore the oligomeric state of the purified, GFP-uSTAT3βtc and GFP-pSTAT3βtc chimaeric proteins. Native Triton X-100 does not denature proteins or dissociate protein complexes. Mobility of the recombinant protein isoforms
is dependent on their mass and charge ratios, i.e., their hydrodynamic size and their electric charge.

Several electrophoretic mobility studies have shown that a direct relationship exists between the molecular weight of a protein and its relative migration through a series of gels differing in acrylamide content (Shuai et al., 1994 and Braunstein et al., 2003). Thus, the plot of log (migration) versus acrylamide concentration is known as the “Ferguson Plot” (Ferguson, 1964).

Ferguson noticed that the logarithm of a protein's mobility varied linearly as a function of the acrylamide content of the gel used to determine the mobility. The resultant slope of the relationship generated a parameter described as the retardation coefficient ($K_r$). Thus, the relative mobility of the purified phosphorylated and unphosphorylated GFP-STAT3βtc was compared with that of a series of standard protein markers in 5, 6, 7 and 8% native Triton X-100 gels. This is shown in Figure 4.3 (A, B, C and D), where A, B, C and D represent 5 %, 6 %, 7 % and 8 % native Triton X-100 gels.

Following staining and then subsequent gel-background destaining, the relative migrations ($R_m (R_m)$) of the protein samples (i.e. purified fluorescent STAT3βtc isoforms included) were measured as described in section 4.1.1.2.3 and then documented in Table 4.1. The table contains the data describing the mean distances travelled by the series of standard protein markers and the GFP-STAT3βtc isoforms in the native gels.
Figure 4.3: Migration of the purified fluorescent STAT3ptc isoforms with standard protein markers on native gels with different acrylamide content

Native polyacrylamide gels containing 5% (A), 6% (B), 7% (C) and 8% (D) acrylamide containing Triton X-100 (0.8 v/v). Fluorescent STAT3ptc together with the standard protein markers were treated with 40 mM DTT and 0.05% Triton X-100 and then incubated at room temperature for 1h. Upon incubation the samples (~ 2.0 µg/well) were loaded on the gels and then electrophoresed until the dye front of had migrated approximately 3/4 to 6/7 down the length of the gels. The gels were illuminated with blue light and an image capture and then stained with coomassie instant blue dye as to produce the image shown. The contents of the lanes in each of the four gels are described as follows; lane 1(A, B, C and D), ovalbumin (43 kDa); lane 2 (A, B, C and D), bovine serum albumin monomer (66 kDa); lane 3 (A, B, C and D), conalbumin (75 kDa); lane 4(A, B, C and D), aldolase (158 kDa); lane 5(A, B, C and D), ferritin (440 kDa); lane 6 (A, B, C and D), GFP-uSTAT3ptc (u kDa); lane (A, B, C and D), GFP-pSTAT3ptc (p kDa). Relative migrations of the various proteins were estimated in centimetres (cm) and then used to generate retardation coefficients (K_r).
### Table 4.1: Mean distance travelled by the standard protein markers and the purified, GFP-STAT3ptc Protein sample | $R_m$ in 5% gel (cm) | $R_m$ in 6% gel (cm) | $R_m$ in 7% gel (cm) | $R_m$ in 8% gel (cm) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ovalbumin (43)</td>
<td>6.60</td>
<td>5.50</td>
<td>4.80</td>
<td>4.30</td>
</tr>
<tr>
<td>BSA (67)</td>
<td>3.50</td>
<td>2.30</td>
<td>2.30</td>
<td>1.80</td>
</tr>
<tr>
<td>Conalbumin (75)</td>
<td>2.20</td>
<td>1.70</td>
<td>1.50</td>
<td>1.30</td>
</tr>
<tr>
<td>Aldolase (158)</td>
<td>0.50</td>
<td>0.20</td>
<td>0.10</td>
<td>0.05</td>
</tr>
<tr>
<td>GFP-uSTAT3βtc</td>
<td>2.10</td>
<td>1.20</td>
<td>0.20</td>
<td>0.10</td>
</tr>
<tr>
<td>GFP-pSTAT3βtc</td>
<td>2.70</td>
<td>1.50</td>
<td>0.30</td>
<td>0.15</td>
</tr>
</tbody>
</table>

The table provides information of the molecular weights of the standard protein markers and their relative migrations (cm) in the various native Triton X-100 acrylamide gels.

From Table 4.1, the data representing the relative migrations of the individual protein samples in the various gels (i.e. with varied acrylamide concentrations) were incorporated specifically as $100[\log (R_m * 100)]$. The resultant data is shown in Table 4.2.2.

### Table 4.2: Relative migrations of the standard protein markers and the GFP-STAT3βtc isoforms expressed as $100[\log (R_m * 100)]$

<table>
<thead>
<tr>
<th>% acrylamide</th>
<th>Ovalbumin</th>
<th>BSA</th>
<th>Conalbumin</th>
<th>Aldolase</th>
<th>GFPuSTAT3βtc</th>
<th>GFPpSTAT3βtc</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>281.90</td>
<td>254.40</td>
<td>234.20</td>
<td>169.80</td>
<td>232.20</td>
<td>243.10</td>
</tr>
<tr>
<td>6</td>
<td>274.00</td>
<td>236.10</td>
<td>223.00</td>
<td>130.10</td>
<td>207.90</td>
<td>216.00</td>
</tr>
<tr>
<td>7</td>
<td>267.60</td>
<td>235.20</td>
<td>217.60</td>
<td>84.50</td>
<td>130.10</td>
<td>139.70</td>
</tr>
<tr>
<td>8</td>
<td>262.80</td>
<td>225.50</td>
<td>209.60</td>
<td>39.70</td>
<td>100.00</td>
<td>117.60</td>
</tr>
</tbody>
</table>

The table provides information of the relative migrations (cm) of the various protein samples in the several native gels with varied acrylamide concentration.

A graph of the relative migrations versus the percent acrylamide concentrations was plotted. From the graph, a linear series of retardation coefficients ($K_r$) were generated for the purified GFP-STAT3βtc isoforms and the standard protein markers. Figure 4.4 describes the graph from which the series of retardation coefficients were generated.
Figure 4.4: The relative migrations of the purified GFP-uSTAT3βtc, GFP-pSTAT3βtc and the standard protein markers in the series of native Triton X-100 polyacrylamide gels

The relative migration ($R_m$) of the protein standards and the purified fluorescent STAT3βtc proteins in a series of native Triton X-100 gels. Purified samples of GFP-uSTAT3βtc and GFP-pSTAT3βtc and four other protein standards of known molecular weights were simultaneously separated on a 5, 6, 7 and 8% native Triton X100 polyacrylamide gel. Following electrophoresis, the fluorescent STAT3βtc isoforms in the various gels were illuminated with blue light to confirm their location and then stained with coomassie blue dye as demonstrated (x, GFP-uSTAT3βtc; +, GFP-pSTAT3βtc; ♦, ovalbumin (43 kDa); ■, BSA monomer (66 kDa); â, conalbumin (75 kDa); x, aldolase (158 kDa). The percentage of acrylamide in the gels was plotted on the x axis and the $R_m$, represented as $100[\log (R_m x 100)]$ was plotted on the y axis. The generated negative slopes represented the retardation coefficient's ($K_r$).

On the basis of Ferguson plot analysis, the retardation coefficient ($K_r$) for the individual protein samples was directly proportional to their molecular weights. Hence, a graph of $K_r$ versus molecular weights (kDa) of the protein standards was constructed and used to evaluate the molecular weights of the GFP-STAT3βtc isoforms, Figure 4.5.

Figure 4.5: Evaluation of the molecular weights of GFP-uSTAT3βtc and GFP-pSTAT3βtc using the Ferguson plot

Since retardation coefficient ($K_r$) relates linearly with molecular weight, $K_r$ data generated from fig 4.1.2 was plotted on the y axis versus known molecular weights on the x axis. The retardation coefficient of
unphosphorylated and phosphorylated GFP-STAT3βtc were evaluated and correlated with molecular weights of 173.73 and 167.54 kDa respectively.

The Ferguson plot therefore suggests that the purified unphosphorylated and phosphorylated GFP-STAT3βtc, i.e., GFP-uSTAT3βtc and GFP-pSTAT3βtc respectively, migrated through the various native Triton X-100 gels with apparent molecular weights of 173.73 and 167.54 kDa respectively.

The data suggests that the relative mobility of the GFP-STAT3βtc isoforms in the 4 different gels were somewhat similar, Figure 4.4, and that the supposedly monomeric GFP-uSTAT3βtc was migrating through the Triton X-100 gels with the apparent molecular weight of a dimeric GFP-pSTAT3βtc, even though, it was expressed in an E.coli strain devoid of the tyrosine kinase domain of the Elk receptor.

A similar observation was made with the unlabelled and purified STAT3βtc isoforms following separation of the samples on the four different gels, suggesting that either the unphosphorylated isoforms were indeed migrating as inactive dimers or a charge to mass ratio effect.

4.2.3 Characterization of purified recombinant phosphorylated and unphosphorylated fluorescent STAT3βtc by size exclusion chromatography

To further explore the oligomeric state of the GFP-STAT3βtc isoforms, the Superose™ 6 column connected to an Äkta Purifier system was adapted and used to facilitate the characterization of the purified, recombinant, proteins. By using column chromatography, the intrinsic charge effect on fractionation was eliminated.

In addition to purification, gel filtration chromatography has evolved as a quick and reliable analytical technique extensively used to determine molecular weights and sizes of recombinant proteins molecules. The idea is based on the well-researched ability of the gel filtration matrix to sieve protein molecules according to size.

Recognizing the possibility, the GFP-STAT3βtc isoforms and four other standard protein markers were investigated on the Superose™ 6 column driven by the Äkta Purifier system, section 4.1.1.3. The molecular weight of the GFP-STAT3βtc isoforms were determined by comparing their elution volume (Ve) parameter, the gel-phase distribution coefficient (Kav) with those of the standard protein markers; ovalbumin (43 kDa); conalbumin (75 kDa); aldolase (158 kDa) and thyroglobulin (669 kDa). Table 4.3
contains data describing the elution volumes and the gel-phase distribution coefficients of the GFP-STAT3βtc isoforms, *i.e.*, GFP-pSTAT3βtc and GFP-uSTAT3βtc, and the various standard protein markers fractionated on the S6 column.

<table>
<thead>
<tr>
<th>Standard Protein Marker</th>
<th>Elution Volume (mL)</th>
<th>Gel-phase distribution coefficient (Kav)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ovalbumin (43)</td>
<td>17.10</td>
<td>0.59</td>
</tr>
<tr>
<td>Conalbumin (75)</td>
<td>16.71</td>
<td>0.56</td>
</tr>
<tr>
<td>Aldolase (158)</td>
<td>15.98</td>
<td>0.51</td>
</tr>
<tr>
<td>Thyroglobulin (669)</td>
<td>12.71</td>
<td>0.31</td>
</tr>
<tr>
<td>GFP-pSTAT3βtc</td>
<td>14.87</td>
<td>0.42</td>
</tr>
<tr>
<td>GFP-uSTAT3βtc</td>
<td>16.31</td>
<td>0.52</td>
</tr>
</tbody>
</table>

In a similar pattern to the protein standards, the GFP-STAT3βtc isoforms eluted as single distinct peaks with elution volumes 14.87 mL and 16.31 mL for GFP-pSTAT3βtc and GFP-uSTAT3βtc respectively, *Figure 4.6 (A),(B).*

With the elution volumes parameter (Ve) of the individual protein samples as well as the void volume (Vo) of the S6 column, *i.e.*, ~ 7.4 mL, the gel-phase distribution coefficients (Kav) for the individual protein samples were calculated using the equation described in section 4.1.1.3.5.
Figure 4.6: Size exclusion chromatograms of purified, recombinant, GFP-pSTAT3ptc and GFP-uSTAT3ptc using Superose™ 6 column

Purified preparation of GFP-pSTAT3ptc (---; dimeric) and GFP-pSTAT3ptc (—; monomeric) was applied to an AKTA purifier system and then fractionated by size exclusion chromatography on a Superose™ 6 (56) column calibrated with 4 standard protein markers of known molecular weights (i.e. ovalbumin, 43 kDa; conalbumin, 75 kDa; aldolase, 158 kDa; thyroglobulin, 669 KDa). The elution volume (mL) is plotted on the x axis the absorbance (mAU) is plotted 280 nm is plotted on the y axis. (A) A chromatogram showing elution volumes for dimeric GFP-pSTAT3ptc (i.e. 14.87 mL) and monomeric GFP-uSTAT3ptc (i.e. 16.31 mL). (B) Chromatograms of GFP-pSTAT3ptc and GFP-uSTAT3ptc superimposed to demonstrate their identical width. From the elution volumes of the fractionated protein samples and the void volume of the column, the gel-phase distribution coefficient for the entire standard protein markers and the test samples (i.e. phosphorylated and unphosphorylated GFP-STAT3ptc) were obtained.

By plotting a calibration curve of the gel-phase distribution coefficient (Kav) versus the logarithm of the molecular weight of the various standard protein markers (log MW), the oligomeric state of the phosphorylated and the unphosphorylated GFP-STAT3ptc was determined. As shown in the plot in Figure 4.7 the molecular weights of the GFP-STAT3ptc isoforms were derived. Based on the curve, the gel-distribution coefficients for GFP-pSTAT3ptc and GFP-uSTAT3ptc are indicated and correlate with the molecular weights of 251 and 100 kDa respectively.

As anticipated, the observations suggested that phosphorylated GFP-STAT3ptc (GFP-pSTAT3ptc) eluted as a putative dimer. The elution volume from which the gel-phase distribution coefficient was generated was fairly consistent with the notion that after tyrosine phosphorylation, STAT3 homo- or heterodimerize by reciprocal interaction between the SH2 domain of one STAT3 protein molecule and the tyrosine phosphorylated tail segment of another STAT3 protein molecule (Darnell et al., 1994; Schindler et al., 1995).
Figure 4.7: Chromatographic calibration curve for the various standard protein markers and GFP-pSTAT3βtc and GFP-uSTAT3βtc generated with Superose™ 6 column

The gel-phase distribution coefficient ($K_a v$) derived from the elution volumes ($V_e$) of the standard protein markers with known molecular weights was plotted on the y-axis versus logarithm of known molecular weight on the x-axis. The $K_a v$ values for GFP-pSTAT3βtc and GFP-uSTAT3βtc are highlighted and correlate with antilog MW of 2.4 (~ 250 kDa) and 2.0 (~ 100 kDa) respectively.

Although the experimental weight was about 50 kDa more than the theoretical molecular weight (~ 200 kDa), SDS- PAGE and western blot (primary antibody used was anti-STAT3 phosphotyrosine PY705) analysis of the peak fractions (i.e. 13.5 to 17.5 mL) did not only demonstrate that the major protein content of the peak fractions were GFP-STAT3βtc, but also, confirmed that the protein was successfully phosphorylated by the Elk receptor kinase domain (see Figure 4.3.3 (A)). Even though the filtration was performed under reducing conditions, the most probable explanation for the differences in molecular weight may be due to less accurate size estimation at >100 kDa, or that the GFP-pSTAT3βtc protein molecule may be in complex with another protein of lower molecular weight through inter-chain disulphide linkage.

In the case of the unphosphorylated GFP-STAT3βtc (GFP-uSTAT3βtc), the observation suggested that the purified recombinant protein eluted as a monomers with an apparent molecular weight of 100 kDa. This observation was relatively consistent with the theoretical molecular weight of the fusion protein calculated as 99 kDa. The data was further confirmed by SDS-PAGE (i.e. fractions 13.5 to 17.5 mL) and western blot (primary antibody used was the anti-STAT3 K-15) and then suggested that the fractions collected for major peak contained about 85 to 90 % pure recombinant GFP-uSTAT3βtc chimaeric proteins (as judged by Coomassie Stain).
The retention of the phosphorylated fluorescent STAT3βtc on the size exclusion column was compared with that of the unphosphorylated fluorescent STAT3βtc, which was expressed in an E.coli strain without the tyrosine kinase domain of the Elk receptor (Figure 4.8 (A) and (B)). Thus, based on the gel-phase distribution coefficients, the retention profiles for GFP-pSTAT3βtc and GFP-uSTAT3βtc were equivalent to 250 and 100 kDa respectively (Figure 4.7). This suggested that GFP-uSTAT3βtc can form stable homodimers in addition to remaining as a monomer.

Thus, as illustrated by the retention profiles for both isoforms, it was evident in lane 3, Figure 4.8 A and B, that the dimeric GFP-pSTAT3βtc eluted quicker with the least residence time on the column, i.e., with less elution buffer (~ 1 mL) than the monomeric GFP-uSTAT3βtc fusion protein.

In a related study, a similar elution profile was obtained for the unlabelled and purified STAT3βtc isoforms. The active isoform, pSTAT3βtc, eluted as a dimer with an
Edwin Nkansah

apparent molecular weight of 199 kDa whilst the inactive isoform, uSTAT3βtc, also eluted with an apparent molecular weight of 79 kDa (data not presented).

4.3 DISCUSSION

The principal objective of the work described in this chapter was to structurally characterize the recombinant GFP-STAT3βtc isoforms, i.e., phosphorylated (active) and unphosphorylated (inactive). This was largely achieved through Size Exclusion Chromatography (SEC). GFP-pSTAT3βtc and GFP-uSTAT3βtc were characterized as a dimer and a monomer respectively. Although attempts were made to resolve the GFP-STAT3βtc isoforms into dimers and monomers by Native PAGE /Ferguson plot, both isoforms were resolved as a dimer. The technique was not really satisfactory due to its dependence on the intrinsic charge as well as the hydrodynamic size of the isoform. Further, aggregation induced precipitation of the inactive GFP-uSTAT3βtc was successfully evaluated using clear Native-PAGE.

4.3.1 GFP-uSTAT3βtc multimerizes into aggregates via inter-chain disulphide bridging in response to increasing temperature

Numerous studies have suggested that STAT3 dimerizes prior to phosphorylation (Li et al., 2004; Braunstein et al., 2003; Schröder et al., 2004; Kretzschmar et al., 2004). It’s been suggested that prior to activation, STAT3 exist as a poorly characterized large molecular mass complex(es) that range from approximately 200 kDa to 2 mDa (Stancato et al., 1996; Haan et al., 2000; Novak et al., 1998; Ndubuisi et al., 1999). Consistent with this, our effort to evaluate the conformational changes associated with our fluorescent expression construct containing the N- and C-terminally truncated STAT3 mutant, STAT3βtc led us to conclude that inactive GFP-STAT3βtc multimerizes into aggregates through inter-chain disulphide bridging. This finding corroborates that of Li et al, which demonstrated the presence of inactive STAT3 complexes including dimers using SDS-PAGE and suggested formation directed by inter-chain disulphide linkage.

The very first evidence that suggested that purified, recombinant, inactive GFP-uSTAT3βtc forms aggregates through inter-chain disulphide linkage was observed following analysis of the protein sample with Native-Poly-Acrylamide Gel supplemented with a non-ionic surfactant, Triton X-100 (TX), up to 0.079 %. The use of TX in the Tris-Glycine gel did not only retain the conformation of the recombinant proteins, but also reduced their electrophoretic mobility and distribution coefficients in acrylamide gel (Hamana et al., 1975). However, increasing the percentage concentration of the TX in the protein preparations up to 2 % did not significantly influence aggregation of the
GFP-STAT3βtc protein molecules, particularly GFP-uSTAT3βtc and therefore may not have played any role in the smooth migration of the fluorescent chimaeric protein molecules into the native gels.

Consequently, we investigated whether DTT (40 mM) would either prevent GFP-STAT3βtc from clustering into aggregate or reduce GFP-STAT3βtc aggregates into their original soluble entities. Significant reduction in aggregation was observed in the fraction incubated at room temperature (24 °C) following separation by native-gel electrophoresis. In contrast, profoundly high levels of GFP-uSTAT3βtc aggregation was observed following separation of the 37 °C treated samples (Figure 4.2 (A), (B)). Thus, it was plausible that aggregation of the GFP-uSTAT3βtc was temperature dependent and and stabilized by the intra- and/or inter chain disulphide (SS) linkage. Recognising the possibility that temperature was playing a critical role the in GFP-uSTAT3βtc aggregate formation, we applied Native-PAGE to investigate identical preparations of both the active and inactive GFP-STAT3βtc isoforms incubated at 0 °C, 4 °C, 24 °C, 37 °Cand 50 °C. Based on Figure 4.2 (C), aggregation of GFP-uSTAT3βtc increased significantly with temperature especially for the 37 °C and 50 °C treated samples. We observed complete aggregation for the 50 °C treated sample. In contrast, no significant aggregation was observed for the active isoform, GFP-pSTAT3βtc, even at 50 °C. These observations suggested that the conformation of the two isoforms may play a role in the temperature induced aggregation. Based on the reported crystal structure of activated STAT3β homodimer (Becker et al., 1998) together with the compact mobility conformation displayed in the TX-gel in Figure 4.2, it was acceptable to suggest that the cysteine residues responsible for SS bond formation remains buried (solvent inaccessible), even at 50 °C and therefore unable to participate in intra- and inter-chain disulphide linkage within and without the GFP-STAT3βtc molecules, hence, little or no aggregation was observed for the phosphorylated isoform. The observation also suggested that activation via phosphorylation conferred thermal stability.

Conversely, we suspect the loosely arranged unphosphorylated isoform, GFP-uSTAT3βtc, gradually unravels (unfold) with increasing temperature (heat stress) and consequently exposes buried hydrophobic side chains including the thiol groups of the cysteine residues, thereby encouraging non-specific protein-protein interactions including intra- and inter-chain SS linkage leading to oligomers of GFP-uSTAT3β being formed as evident in the wells displayed in Figure 4.2 (C).

Moreover, the possibility that the SS linkage implicated in GFP-uSTAT3βtc aggregation may be stabilizing the unravelled protein complex was investigated and found to be
consistent with the ability of DTT to reduce. Thus, a set of GFP-uSTAT3βtc preparations supplemented with different concentrations of DTT, up to 1 M was incubated at 37 °C prior to separation on a TX-native gel. The level of aggregation decreased with increasing DTT up to concentration of 200 mM where we observed multimers of the GFP-uSTAT3βtc protein bands in the gel, thus suggesting the formation of a DTT adduct, i.e., the two sulphur atoms of DTT forms disulphide bonds with the different sulphur atoms of the cysteine residues contained in the GFP-uSTAT3βtc protein; in such cases, DTT cannot cyclise since it has no remaining free thiol groups (Zhang et al., 1991).

Furthermore, the presence of the GFP-uSTAT3βtc aggregates suggested that the protein did not regain its native structure upon cooling since the analysis by native gel was carried out at 4 °C.

### 4.3.2 Characterizing phosphorylated and unphosphorylated GFP-STAT3βtc using Native-PAGE (TX)/Ferguson plot

Recognising the possibility that the purified preparations of the phosphorylated and unphosphorylated GFP-STAT3βtc isoforms could be fractionated into dimers and monomers respectively through using TX native gel electrophoresis, a study was carried out to compare the oligomeric states of both isoforms. Ferguson had suggested that the logarithm of a protein's mobility varied linearly as a function of the gel acrylamide concentration used to determine the mobility. Hence, the relative migration of purified active and inactive GFP-STAT3βtc was compared with that of a series of standard protein marker in 5 %, 6 %, 7 % and 8 % native TX gels (Figure 4.4) from which a series of linear retardation coefficients ($K_r$) were derived (Figure 4.5). It is suggested that the generated $K_r$ values for each protein was directly proportional to its molecular mass.

On the basis of the Ferguson plots, it was suggested that the phosphorylated GFP-pSTAT3βtc migrated with an apparent molecular weight of 167.54 kDa whilst the unphosphorylated isoform (GFP-uSTAT3βtc) was also fractionated with an apparent molecular weight of 173.73 kDa. Although surprising (in relation to the supposedly monomeric GFP-uSTAT3βtc), it was not unprecedented in the field of signal transduction that unactivated proteins (monomers) may dimerize prior to activation (Carr et al., 2001; Chow et al., 2001). In relation to the active isoform, the experimental molecular weight was no surprise since it was expected to migrate through the various gels as an active and a stable homodimer. Though the inactive isoforms (GFP-uSTAT3βtc) migrated apparently as a dimer, this was not consistent with data...
generated by other techniques that suggested that STAT3\(\beta\)tc was predominantly monomeric (Becker et al., 1998; Braunstein, et al., 2003). Although some reports speculate that the eCFP and eYFP proteins may exist as weak dimers, it was generally accepted that the spectral variants of the GFP predominantly existed as monomers (Cubitt et al., 1995; Heim et al., 1995; Miyawaki et al., 1999).

Hence, it was reasonable to speculate that because electrophoretic mobility of the purified preparations in the native gel was dependent on both the electric charge and hydrodynamic size, and therefore it was not possible to fractionate the GFP-STAT3\(\beta\)tc isoforms into monomers and dimers using this method. Both isoforms migrate with almost identical charge to mass ratio, hence our inability to use Native-PAGE to differentiate the isoforms.

### 4.3.3 Characterizing phosphorylated and unphosphorylated GFP-STAT3\(\beta\)tc using Size Exclusion Chromatography (SEC)

Becker et al fractionated active and inactive STAT3\(\beta\)tc into monomer and dimer respectively using a gel filtration column (Becker et al., 1998). Through this approach, Size Exclusion Chromatography (SEC) was employed to fractionate the active and inactive GFP-STAT3\(\beta\)tc into dimers and monomers respectively. Unlike Native-PAGE, SEC eliminates the effect the intrinsic charge on the protein will have on fractionation. Beside purification, gel filtration chromatography has been extensively used to determine molecular weight and size of proteins (Braunstein et al., 2003). The technique is based on the well-researched ability of the gel matrix to sieve protein molecules according to size.

Following calibration of the Superose™ 6 (S6) column with four different protein standards (known molecular weights), both isoforms of the GFP-STAT3\(\beta\)tc were fractionated on column. Both active and inactive isoforms eluted as single peaks with distinct elution volumes, i.e. 14.87 and 16.31 ml respectively (Figure 4.6 (A)). From the elution volumes, the gel-phase distribution coefficients (Kav) for both isoforms were generated. The molecular weights for the GFP-pSTAT3\(\beta\)tc and GFP-uSTAT3\(\beta\)tc isoforms were estimated at 251 and 100 kDa respectively using the plot described in Figure 4.7. The estimated sizes suggested that the phosphorylated isoforms eluted as a dimer (~ 251 kDa) whilst the unphosphorylated isoform eluted as a monomer (~ 100 kDa). Unlike the GFP-pSTAT3\(\beta\)tc which eluted with an experimental molecular weight of approximately 251 kDa (i.e. about 51 kDa more that the theoretical molecular weight 200 kDa), the GFP-uSTAT3\(\beta\)tc was eluted into a volume corresponding to 100 kDa, i.e. about 1 kDa more than the theoretical molecular weight of the unphosphorylated
isoform. This confirmed that structurally, the phosphorylated isoform was either made up of two GFP-uSTAT3βtc protein molecules or two GFP-uSTAT3βtc molecules in association with another protein molecule of about 50 kDa or a GFP-uSTAT3βtc molecule covalently linked to another protein molecule that is about 150 kDa. Though plausible, the last two arguments were not entirely likely since prior to SEC, the protein had been purified by anion exchange chromatography. Therefore, the difference in the experimental data for the phosphorylated isoform, GFP-pSTAT3βtc, may be attributed to decreased accuracy above 100 kDa, and possibly conformational constraints encountered by the protein during fractionation, thereby marginally increasing the residence time of the protein in the column and subsequently increasing the volume of the buffer required to elute the recombinant protein. Furthermore, it was also possible to speculate that the GFP-pSTAT3βtc co-expressed with the Elk receptor catalytic domain may not all dimerize through the reciprocal phosphotyrosyl-SH2 domain interaction.

Further confirmatory studies including SDS-PAGE and western blot were performed on the fractions of the major peaks for the phosphorylated and unphosphorylated GFP-STAT3βtc isoforms, Figure 4.8. The protein contents of the major peaks for both isoforms were immunoreactive to their respective antibodies such as; the anti-STAT3 PY705 antibody in the case of GFP-pSTAT3βtc and the anti-STAT3 K-15 antibody in the case of the GFP-uSTAT3βtc, Figure 4.8. The elution profile as displayed in Figures 4.3.1, also suggested that the phosphorylated isoforms eluted with about 1 mL less volume than the unphosphorylated isoform.

Furthermore, on the basis of the principle of SEC, smaller molecules (smaller than the pore size of the gel matrix) enter the porous particles and then fractionate with a longer path and a longer transit time than larger molecules (larger than the pore size) which undergo total exclusion. Thus, as demonstrated by the retention profiles for both isoforms, it was evident in lane 3, Figure 4.8 A and B, that the dimeric GFP-pSTAT3βtc eluted quicker with the least residence time on the column, i.e., with less elution buffer than the monomeric GFP-uSTAT3βtc fusion protein.

In summary, the purified recombinant GFP-uSTAT3βtc was characterized as monomers and GFP-pSTAT3βtc as dimers, via Size Exclusion Chromatography (SEC). The Native-PAGE/Ferguson plot approach was less clearly able to structurally characterize the chimaeric proteins isoforms. Nonetheless, the approach was successfully applied to gather compelling evidence that suggested that unlike the phosphorylated isoform, the unphosphorylated isoforms was highly susceptible to
temperature dependent aggregation mediated by inter-chain disulphide linkage. Due to reports that have implicated reactive oxygen species as activators of STAT3 \textit{in vivo} (Carballo \textit{et al.}, 1999; Simon \textit{et al.}, 1998), the redox sensitivity of the unphosphorylated isoform described in the chapter would be a subject for further study together with its functional relevance.
5 BINDING SPECIFICITY OF GFP-STAT3βtc

5.1 Enzyme-Linked Immuno-Sorbent Assays (ELISA)

Enzyme Linked Immuno-Sorbent Assay (ELISA) is an assay format similar to the antibody sandwich principle. Assay development begins by immobilizing the capture molecule (e.g. peptide, DNA and antibodies) specific to the target protein on the microtiter plate surface. After the immobilization step, the unbound probe is removed by washing and the plate blocked with a blocking agent (e.g. salmon sperm DNA or BSA). The target protein is then incubated with the immobilized probe, which then captures or interacts with the target protein. After washing off the unbound materials, the captured protein can be detected conventionally by probing the plate with a primary antibody specific to the captured protein and then a secondary antibody (usually conjugated to HRP) specific to the primary antibody. After the washing to remove unbound detection antibodies, the substrate solution (e.g. SIGMAFAST™ OPD; o-phenylenediamine dihydrochloride) is added and a colour develops in proportion to the amount of bound secondary antibody and for that matter captured target protein. The intensity of the colour is measured after a period of incubation.

Whilst developing the laborious conventional method, a fluorescent-based approach which is quicker and less laborious was also investigated.

After binding the GFP tagged STAT3βtc protein to the immobilized probe, unbound materials were aspirated and the wells were washed to remove traces of the unbound materials including the GFP-STAT3βtc chimaeric protein. The captured GFP-STAT3βtc was then measured by fluorescence. Based on the excitation and emission wavelengths of the fluorescent tag fused to the captured GFP-STAT3βtc, the appropriate filters were applied to generate the needed response, i.e., the fluorescence intensity was proportional to the captured GFP tagged STAT3βtc. By applying both formats, the GFP-STAT3βtc isoforms were assayed.

On the basis of the assay described above, the biological functionality of the purified, recombinant GFP-STAT3βtc protein was evaluated by a microtiter plate format. Using both the conventional ELISA and the fluorescence-based microtiter plate assay formats, both phosphorylated (active) and unphosphorylated (inactive) GFP-STAT3βtc were assayed with an immobilized biotinylated dsM67 (an immobilized modified c-fos sis inducible enhancer) and an biotinylated surrogate phosphotyrosyl peptide (a phosphotyrosyl peptide derived from the interleukin-6 receptor subunit gp130)
respectively. The reactions were carried out in black 96 well flat-Bottom high binding plate purchased from Greiner Bio-One, Stonehouse, UK. The biotinylated probes (i.e. dsM67 or Phospho-peptides) were immobilized using streptavidin.

5.1.1 Preparation of Streptavidin coated plates

The frozen streptavidin powder (Sigma-Aldrich, UK) was reconstituted with ddH$_2$O to a final concentration of 10 mg/mL to generate the mother stock solution. A working stock of 1 mg/mL was generated by further diluting the stock with 20 mM potassium phosphate buffer (pH 6.5). Both streptavidin stock solutions were stored at -80 °C until use.

In preparing the streptavidin coating solution, 50 µL of 1 mg/mL streptavidin solution was diluted to 10 mL with 20 mM K$_2$HPO$_4$ buffer (pH 6.5) to a final concentration of 5.0 µg/mL. Into each well of the flat-bottom high binding plate, 100 µL of the 5 µg/mL streptavidin solution was dispensed. The plate was sealed and incubated at 37°C with high humidity for 16 h.

Following incubation, the coating solution was discarded and the plate washed 3x with 200 µL/well of PBS. The unoccupied binding sites were either blocked with just blocking buffer (200 µL 7% (w/v) BSA in PBS (181 mM NaCl, 2.68 mM KCl, 4.28 mM Na$_2$HPO$_4$ (pH7.3), 1.46 mM KH$_2$PO$_4$)) or blocked with the blocking buffer supplemented with the appropriate biotin-labeled probe for 1h at room temperature. The spent solution was aspirated and the plate subsequently washed 3x with PBS solution, 200 µL/well.

For storage, the blocking solution was aspirated, but the plates were not washed with PBS. Plates were sealed and stored at 4°C until future use. Plates must be used in 8 weeks.

5.1.2 Evaluation of streptavidin coating

After washing the plates, 100 µL of 0, 100, 200, 400 and 800 pg/well of Biotin-HRP conjugate was incubated with either 250 or 500 or 1000 ng/well streptavidin immobilized to generate the solid phase. The plate was then incubated at room temperature for either 60 or 120 min with rocking. Following immobilization, the unbound biotin-HRP solution was aspirated and the plate washed 3x with 200 µL/well PBS. Fresh SIGMAFAST™ o-Phenylenediamine dihydrochloride/Hydrogen peroxide solution was prepared, i.e., OPD/H$_2$O$_2$; 1 OPD tablet and 1 H$_2$O$_2$ dissolved in 20 mL ddH$_2$O, and 100 µL was dispensed into each well. The plate was sealed with a plate.
seal and then enforced with an aluminium foil before incubated at room temperature (24 °C) for 25 min with rocking to promote colour change. The change in colour was measured using the PerkinElmer EnVision plate read 450 nm. A typical 300 ng/well streptavidin gave approximately 0.1 absorbance reading when measured at OD\textsubscript{450 nm} whilst a 1000 ng/well gave absorbance readings ranging from 0.8 to 1.0 mAU at an OD\textsubscript{450 nm}.

5.1.3 Immobilization of biotin-dsM67 DNA on streptavidin coated plate

Blocking solution was aspirated and residual solution removed from each well by inverting the plate and blotting it on a paper towel. A 120 μM stock solution of biotinylated dsM67 DNA was diluted to 2.88 μM with PBS. The 2.88 μM biotin-dsM67 PBS solution was further diluted to generate 300 pmol working solution (1.5 mL) using PBS. The working solution, i.e., 300 pmol, was serially diluted with PBS to generate 1.5 mL of the following amounts: 0.003, 0.01, 0.3, 0.1, 1, 3, 10, 30 and 100 pmol. The solutions were prepared by diluting 0.5 mL 300 pmol dsM67 DNA stock to 1.5 mL with PBS to generate 100 pmol DNA stock. The solution was mixed by inversion and then 0.5 mL was subsequently diluted to 1.5 mL with PBS to generate 30 pmol of the dsM67 DNA solution. The resultant 30 pmol solution was mixed by inversion. The 30 pmol dsM67 DNA solution and subsequent solutions were diluted in a similar manner.

After developing the assay format, the dsM67-PBS solution was dispensed, 100 μL/well (assay was performed in duplicates). The plate was then sealed and incubated at room temperature for 1 h with rocking.

After incubation, the dsM67-PBS solution was aspirated and the plate washed 3x with PBS, 200 μL/ well. It was re-blocked with blocking solution for 45 min at room temperature and then re-washed 3x with PBS after aspirating the blocking solution. The plate was sealed in readiness for the capture experiment.

5.1.3.1 Microtiter plate assay to demonstrate the capture of phosphorylated GFP-STAT3βtc (GFP-pSTAT3βtc) by immobilized dsM67 DNA

The binding functionality of the purified, recombinant GFP-pSTAT3βtc was assayed with different amounts of the immobilized dsM67 DNA, section 5.1.3. The analyte, GFP-pSTAT3βtc (3.0 μM), was prepared for capture in a binding buffer (20 mM HEPES, pH7.9, 40 mM KCl, 5 mM MgCl\textsubscript{2}, 0.03 mM EDTA, 0.03 mM EGTA, 1.0 mM DTT and 8 % (w/v) glycerol) for 20 min at room temperature while rocking.

173
GFP-pSTAT3βtc (~ 150000 pmol) was dispensed 50 µL/well (assay was performed in duplicate and strictly according to the assay layout). The plate was sealed with a plate seal and the binding reaction incubated at room temperature for 1 h with rocking (Figure 5.1).

After incubation, the unbound protein solution was aspirated and the plate gently blotted with paper towels. It was then washed 3x with PBS (200 µL/well) and again blotted with paper towels to remove excess washing buffer. Finally, freshly prepared PBS, 100 µL/well, was dispensed into the bound wells and the plate read to measure captured GFP-pSTAT3βtc, sections 5.1.5.1 and 5.1.5.2.

5.1.4 Immobilization of biotin-phosphotyrosyl peptide on streptavidin coated plate

Following plate evaluation, section 5.1.2, the biotin-labeled surrogate phosphotyrosyl peptide (biotin-pYLPQT-NH₂) was immobilized via biotin-streptavidin coupling. A 5 mM stock solution of the peptide was reconstituted with PBS to a 1.5 mL volume of 300 pmol phosphotyrosyl peptide. The 300 pmol peptide solution was serially diluted with PBS to generate 1.5 mL solutions containing 0.003, 0.01, 0.3, 0.1, 1, 3, 10, 30 and 100 pmol of the biotin labeled phosphotyrosyl peptide. The solutions were prepared as described in section 5.1.3.

The biotin labeled peptide-PBS solution was dispensed, 100 µL/well (assay was performed in duplicates and according to the assay layout). The plate was sealed and the binding reaction incubated at room temperature for 1 h with rocking.
After incubation, the unbound peptide-PBS solution was aspirated and the plate washed 3x with PBS (200 μL/ well). It was re-blocked with blocking solution for 45 min at room temperature. The spent blocking solution was subsequently aspirated and the plate washed 3x with PBS. The plate was then sealed in readiness for the analysis.

5.1.4.1 Microtiter plate assay to demonstrate the capture of unphosphorylated GFP-STAT3βtc (GFP-uSTAT3βtc) by immobilized phosphotyrosyl peptide

The binding functionality of the purified, recombinant, GFP-uSTAT3βtc was evaluated with different amount of immobilized phosphotyrosyl peptide, H-pYLPQTV-NH2, section 2.9.4. The GFP-uSTAT3βtc (3.0 μM), was prepared for the binding reaction in a binding buffer (20 mM HEPES, pH7.9, 40 mM KCl, 5 mM MgCl2, 0.03 mM EDTA, 0.03 mM EGTA, 1.0 mM DTT and 8 % (w/v) glycerol) for 20 min at room temperature whilst rocking.

Following preparation, 50 μL (~ 150000 pmol) GFP-uSTAT3βtc was dispensed into each well (assay was performed in duplicate and strictly according to the assay layout). The plate was sealed with a plate seal and the binding reaction incubated at room temperature for 1 h with rocking (Figure 5.2).

After binding, the unbound protein solution was aspirated and the plate gently blotted with paper towels. The plate was then washed 3x with PBS (200 μL/ well) and further blotted with paper towels to remove the excess washing buffer. Finally, freshly prepared PBS was dispensed, 100 μL/ well, and the plate read to measure the capture of the unphosphorylated GFP-uSTAT3βtc by the immobilized phosphotyrosyl peptide, sections 5.1.5.1 and 5.1.5.2.
5.1.5 Detecting captured GFP-STAT3\textit{ptc}

5.1.5.1 Detection by intrinsic GFP fluorescence

By measuring the fluorescence from the GFP tag, the amount of GFP-STAT3\textit{ptc} captured by the immobilized probes, \textit{i.e.}, either dsM67 or H-pYLPQTV-NH\textsubscript{2} was quantified. Upon washing and stabilizing the wells with freshly prepared PBS (100 \mu L/ well), the plate was sealed and then equilibrated for 5 min at room temperature. The seal was gently removed and the plate read at excitation and emission wave lengths; 450 and 485 nm for eCFP tagged chimaeric proteins, 485 and 509 nm for eGFP tagged chimaeric proteins and 509 and 535 nm for eYFP tagged chimaeric proteins respectively. The data output from the plate reader was manually exported into an excel spread sheet, and the extent of capture was related to the fluorescence intensity. Following from the fluorescence-based read-out, the plate was prepared via conventional ELISA and then re-read, section 5.1.5.2.

5.1.5.2 Detection by conventional ELISA colorimetry

After aspirating the stabilizing PBS solution, the captured protein was incubated with the primary antibody (100 \mu L/ well; STAT3 (K15): sc-483 rabbit polyclonal anti-STAT3 antibody, 1: 2000-fold dilution) prepared in a blocking solution at room temperature for 45 min with rocking. The plate was aspirated and then washed 3x with PBS (200 \mu L/ well). It was then blotted with paper towel. The protein/antibody complex was subsequently incubated with the secondary antibody (100 \mu L/ well; ECL Rabbit IgG, HRP-Linked Whole Ab (from donkey) 1: 2000-fold dilution) also prepared in a blocking buffer for a further 45 min at room temperature. The unbound secondary antibody solution was aspirated and the plate was washed 3x with PBS (200 \mu L/ well). The plate was then blotted to remove excess wash solution.

After probing the captured protein with the primary and secondary antibodies, the OPD/H\textsubscript{2}O\textsubscript{2} substrate (tablet set dissolved in 20 mL \textit{ddH}_2O to generate 0.4 mg/mL OPD and 0.4mg/mL urea hydrogen peroxide and 50 mM phosphate-citrate, pH5.0) mix was prepared and dispensed into the wells (100 \mu L/ well). The plate was sealed with a plate sealer and then enforced with aluminium foil (substrate is hygroscopic and light sensitive). The reaction was incubated at room temperature for 20 min with rocking.

Since the fluorescence read-out, immunoreactions and OPD/H\textsubscript{2}O\textsubscript{2} substrate treatment stages were performed in a black plate (black plates provides reduced background signal that enhances assay sensitivity with minimized cross talk and prevent light piping between wells), the substrate reaction mixture (brownish yellow) was transferred into
corresponding well of a clear low binding plate to be read after colour development. The plate was read at 450 nm. The data output from the plate reader was manually exported into an excel spread sheet, where the extent of capture was related to absorbance.

5.1.6 Inhibition of GFP-STAT3βtc/probe interaction

5.1.6.1 Inhibition of GFP-pSTAT3βtc/dsM67 DNA complex with either dsM67 DNA (unlabeled) or phosphotyrosyl peptide (unlabeled)

The assay was prepared in clear non-binding plates before being transferred into the black plate containing the immobilized capture probe (dsM67 DNA).

The different concentrations of the unlabeled dsM67 DNA (inhibitor) was prepared from a 300 μM stock solution. Using the stock DNA solution, 200 μM dsM67 DNA working solution was prepared by diluting with 100 μM NaCl solution. Subsequently; 0.02, 0.06, 0.2, 0.6, 2.0, 6.0, 20 and 60 μM dsM67 DNA solutions were prepared by serial dilutions.

The unlabeled phosphotyrosyl peptide inhibitor was prepared from a 50 mM stock containing 100% DMSO. A 30 mM working solution was initially prepared by diluting with 100% DMSO. Further dilutions of the 30 mM working solution were performed to generate 1, 3 and 10 mM phosphotyrosyl peptide solutions using 100% DMSO. By using 1:50-fold dilutions format, 0.02, 0.06, 0.2, 0.6, 2.0, 6.0, 20 and 60 μM peptide solutions were prepared with each containing approximately 2% DMSO. PBS was applied as the diluent.

The analyte, GFP-pSTAT3βtc (100 nM), was prepared with binding buffer and then pre-incubated with each of the serially diluted unlabeled dsM67 DNA preparations in the following reaction; 30 μL GFP-pSTAT3βtc (30 nM), 20 μL binding buffer and 50 μL unlabeled dsM67 DNA (0.01, 0.03, 0.1, 0.3, 1.0, 3.0, 10.0 and 30.0 μM) solution. The protein-DNA mixture was mixed gently to create homogeneity. The plate was then sealed with a plate seal and then incubated for 30 min at room temperature with shaking. After incubation, the protein-DNA complex mix was transferred into the assay plate (black plate coated with 300 ng/well biotin labeled dsM67 DNA) whilst strictly maintaining the assay layout and then incubated with the immobilized dsM67 DNA. The plate was re-sealed and the mixture incubated for 30 min at room temperature while shaking.
In a similar manner, GFP-pSTAT3βtc was pre-incubated either various unlabeled phosphotyrosyl peptide solutions prior to incubation with the immobilized dsM67 DNA. Upon incubation, the unbound protein mix was aspirated and the plate washed 3x with PBS (200 μL/ well) to remove residues. The bound complex was stabilized with freshly prepared PBS (100 μL/ well) for 5 min at room temperature.

The plate was read for fluorescence using the PerkinElmer EnVision plate reader using excitation and emission wave lengths of 485 and 509 nm respectively. The data output from the plate reader was manually exported into an excel spread sheet and the extent of inhibition was related to the fluorescence intensity.

5.1.6.2 Inhibition of GFP-uSTAT3βtc/phosphotyrosyl peptide complex with either dsM67 DNA (unlabeled) or phosphotyrosyl peptide (unlabeled)

The assay mix was prepared in the clear non-binding plate then transferred into a black 96 well plate containing immobilized biotinylated phosphotyrosyl peptide (300 ng/ well). Both of the unlabeled probes, *i.e.*, dsM67 DNA and phosphotyrosyl peptide, were prepared as described in section 5.1.6.1.

The analyte, GFP-uSTAT3βtc (100 nM) was prepared for complex formation with the binding buffer and then pre-incubated with each of the serially diluted and unlabeled probes in the reaction as follows; 30 μL GFP-uSTAT3βtc (30 nM), 20 μL binding buffer and 50 μL unlabeled probe (0.01, 0.03, 0.1, 0.3, 1.0, 3.0, 10.0 and 30.0 μM). The protein-probe mix was gently homogenized. The plate was sealed with a plate seal and the reaction mix incubated for 30 min at room temperature with shaking. After incubation, the protein-probe complex mixture was transferred into the assay plate (black plate coated with 300 ng/well biotin labeled -pYLPQTV-NH₂ peptide) whilst strictly maintaining the assay layout and then incubated with the immobilized phosphotyrosyl peptide for a further 30 min at room temperature while shaking.

After incubation, the unbound protein mixture was aspirated and the plate washed 3x with PBS (200 μL/ well) to remove trace of the unbound materials. The bound complex was stabilized with freshly prepared PBS (100 μL/ well) for 5 min at room temperature. The plate was subsequently read for fluorescence using the PerkinElmer EnVision plate reader with excitation and emission wave lengths of 485 and 509 nm respectively. The data output from the plate reader was manually exported into an excel spread sheet and the extent of inhibition was related to the fluorescence intensity.
5.2 RESULTS

5.2.1 Enzyme-Linked Immuno-Sorbent Assay (ELISA) formats

5.2.1.1 Evaluation of streptavidin coating

Before using a streptavidin coated plate for functional assays, the batch was validated by assaying an identical plate prepared concurrently with biotin via HRP-labelled biotin. The in-house coating and control assay was performed routinely to ensure that control values fall within the pre-established ranges. This routine control test provided valuable information about the quality of the plate and the assay materials. In Figure 5.3, a typical plate layout for a 96 well streptavidin coated plate is displayed.

![Figure 5.3: Maxisorp plate format for plate validation.](image)

A 96 well clear maxisorp plate is used. First, the capture antibody (streptavidin) was bound to the microtiter plate to create the solid phase described as follows; AB1-12, 250 ng/ well; CD1-12, 500 ng/ well; EF1-12, 1000 ng/well. Unbound materials' including excess streptavidin molecules was aspirated and then the plate was blocked. Following blocking, the biotin-HRP (i.e. serially diluted as displayed; 0, 100, 200, 400, 600 and 800 pg/well) was bound to the immobilized streptavidin. The biotin-HRP was incubated with the solid phase streptavidin for 60 min (AF, 1-6) and for 120 min (AF, 7-12). The plate was washed and 100 μL of the substrate solution (i.e. 1 OPD tablet and 1 urea/H₂O₂ in 20 mL ddH₂O) dispensed into the wells. Incubation of the plate on a shaker for 30 min, the substrate was converted into a soluble end product by the HRP that is orange-brown in colour and that was read on a PerkinElmer EnVision at 450 nm. It must be noted that the colour development was directly proportional to the amount of bound biotin- HRP. A typical data for plate validation is displayed in Figures 5.4.

As anticipated and subsequently demonstrated in Figure 5.4, binding of biotin HRP to the immobilized streptavidin increased with increasing amount of immobilized streptavidin. The trend was similar in all three cases (i.e. 250, 500 and 1000 ng/well) even though it was much more pronounced when wells coated with 500 and 1000 ng...
streptavidin were incubated with different concentrations of biotin-HRP; 0, 100, 200, 400, 600 and 800 pg/well. No significant signal was observed for the 250 ng/well even with increasing biotin-HRP up to 800 pg/well.

![Graph showing absorbance at 450 nm versus mass of biotinylated HRP per well.](image)

**Figure 5.4: Validation of streptavidin coating using biotin-HRP**

The assay was performed using a Maxisorp 96 well plate. The plate was coated with streptavidin as described in Figure 5.3 and then incubated with different concentrations of biotin-HRP (i.e. 0, 100, 200, 400, 600 and 800 pg/well) either over 60 min or 120 min. The data obtained from the PerkinElmer EnVision plate reader was plotted as the absorbance at 450 nm versus the mass of biotinylated HRP per well. The graph shows the mean specific absorbance obtained from triplicate samples and is a representation of n = 5 independent coatings.

From the biotin-HRP calibration curve performed concurrently with the validation plate, the binding efficiency of the biotin-HRP to the immobilized streptavidin was determined. In most case, however, the value was close to 0.5, suggesting that the integrity of the coating was similar to that of the commercially available pre-coated Greiner BioOne 96 well plates.

In a related study, the incubation period of the streptavidin-biotin-HRP complexation reaction was increased to 120 min, after which the captured biotin-HRP was measured. No significant variations relative to the former study (incubation period of 60 min) was observed, suggesting that 60 min for streptavidin-biotin-HRP incubation was sufficient for biotin-HRP capture.

**5.2.1.2 Analysis of the biological functionality of GFP-STAT3βtc isoforms**

Small molecules that binds to the STAT3 SH2 domain and down regulate downstream signalling by either preventing STAT3 activation or disrupting active dimers could be an
invaluable tool for the inhibition of tumour formation dependent on the constitutive activation of STAT3. To this end, microtiter plate assays were developed and used to evaluate the biological functionalities of the purified, recombinant, GFP-STAT3βtc isoforms.

In determining the biological functionality of the unphosphorylated isoform, the GFP-uSTAT3βtc was assayed with the pY motif of an immobilized biotinylated phosphotyrosyl peptide a surrogate that mimics STAT3βtc docking at the gp130 receptor docking site and dimerization. Further, the DNA-binding functionality of the phosphorylated isoform was investigated by capturing the purified, recombinant GFP-pSTAT3βtc protein with an immobilized biotin labelled hSIE oligonucleotide (dsM67 DNA) probe. Following from the binding study, the assay was further adapted to demonstrate the ability of the phosphotyrosyl peptide to inhibit DNA binding by disrupting the GFP-pSTAT3βtc dimer.

The GFP-uSTAT3βtc/pY peptide and the GFP-pSTAT3βtc/dsM67 DNA cell free binding assays were considered likely to be the best compromise for targeting and disrupting the cellular functions of the STAT3.

5.2.1.2.1 Detection of GFP-uSTAT3βtc bound to the immobilized biotin-phosphotyrosyl peptide.

Following immobilization of the varying concentration of biotinylated phosphotyrosyl hexapeptide (i.e. biotin-P-pY-LPQTV-NH₂) derived from the interleukin-6 receptor subunit gp130 with high affinity for STAT3 SH2 domain (Stahl et al., 1995; Ren et al., 2003), the purified, recombinant, GFP-uSTAT3βtc was assayed to mimick STAT3-receptor docking via the SH2 domain of the GFP-uSTAT3βtc and the phosphorylated tyrosine motif of the immobilized surrogate peptide. Following incubation and then subsequent washes, the plate was measured using the PerkinElmer EnVision plate reader to determine the amount of captured or "docked" GFP-uSTAT3βtc protein. The amount captured was measured initially by fluorescence, via the GFP reporter fused to the STAT3βtc and confirmed by conventional ELISA via HRP-coupled secondary antibody and the OPD/urea/H₂O₂ substrate. As demonstrated in Figures 5.5 and 5.6 the cellular functionality of the purified, recombinant GFP-uSTAT3βtc was retained and successfully assayed.
Figure 5.5: Detection of bound GFP-uSTAT3βtc to immobilized biotin-phosphotyrosyl peptide surrogate by fluorescence

The assay was performed in black high binding 96 well plates for the fluorescence read-out format. Increasing amounts (i.e. 0.03, 0.1, 0.3, 1, 3, 10, 30, 100 and 300 pmol / well) of the biotin-PpY-LPQTV-NH₂ was immobilized to the microtiter plate to create the solid capture phase. Following removal of unbound material, the plate was blocked and then assayed with 150 pmol GFP-uSTAT3βtc per well. The reaction was incubated at room temperature and then aspirated. The plate was then washed to remove traces of unbound GFP-uSTAT3βtc. The immobilized PpY-LPQTV-NH2-GFP-uSTAT3βtc complex was stabilized with fresh PBS and then read via fluorescence. The excitation and emission filters used were the FITC’s 485 and 535 nm filters respectively. A graph of the mean fluorescence intensity (y-axis) against the amount of immobilized PpY-LPQTV-NH2 probe (x-axis) was plotted to demonstrate proportional of GFP-uSTAT3βtc captured by the immobilized PpY-LPQTV-NH2 probe.

The first 3 set of data were controls demonstrating the extent at which their mean fluorescence intensity at 509 nm could influence the binding study. The first control (i.e. PP (-)/F-uST3 (-)) demonstrated the background noise in a well devoid of a capture probe and the analyte. The second control (i.e. PP (-)/F-uST3 (+)) was performed to demonstrate non-specific binding of the GFP-uSTAT3βtc to the streptavidin coated well while the third control (i.e. PP (+)/F-uST3 (-)) demonstrated the mean fluorescence intensity from the immobilized phosphotyrosyl peptide (300 pmol/well). Further, as displayed and anticipated, the immobilized phosphotyrosyl surrogate peptide captured and retained the GFP-uSTAT3βtc in a dose-dependent manner, up to amounts of approximately 3 pmol /well. The increasing mean fluorescence intensity suggested a gradual increase in the amount of GFP-uSTAT3βtc being captured, which is also consistent with the increasing amount of the immobilized biotin-phosphotyrosyl peptide available. The strong increase in the mean fluorescence signal relative to the controls was indicative of the biological functionality of the purified, recombinant protein. However, no significant increase in binding was observed upon
incubating GFP-uSTAT3\beta tc in the well containing 30, 100 and 300 pmol of the immobilized biotin-phosphotyrosyl peptide. This observation suggested saturation.

After acquiring binding data via fluorescence, further confirmatory reading of the capture of the GFP-uSTAT3\beta tc protein by the immobilized phosphotyrosyl peptide was obtained by conventional ELISA, acquired at 450 nm. As demonstrated in Figure 5.5, Figure 5 also showed that GFP-uSTAT3\beta tc was captured and retained by the biotinylated peptide. Even though a drop in signal was observed relative to the fluorescence read-out, the binding profile was similar to that demonstrated in Figure 5.5. Again, the \( k_d \) was not calculated since the exact amount of the immobilized phosphotyrosyl peptides was not known.

The capture of the unphosphorylated GFP-uSTAT3\beta tc did not only demonstrate the binding kinetics of the protein, but it also showed that uSTAT3\beta tc was still functional as a chimaeric protein.
5.2.1.2.2 Detection of GFP-pSTAT3βtc bound to the immobilized biotin-DNA oligonucleotide probe.

The high affinity biotinylated modified c-fos sis inducible enhancer dsDNA oligonucleotide sequence (SIE or M67) containing the recognition site of STAT3 was immobilized following serial dilution. The purified, recombinant, GFP-pSTAT3βtc was applied and assayed \textit{in vitro} to mimic active STAT3 dimer binding to the DNA consensus sequence (TTCCCGTAA) via the DNA binding domain of the pSTAT3βtc \textit{in vivo}. Following incubation and the subsequent washes, the plate was measured using the PerkinElmer EnVision plate reader to determine the amount of captured GFP-pSTAT3βtc chimaeric protein. The captured GFP-pSTAT3βtc was initially measured by fluorescence, \textit{via} the GFP tag fused to the STAT3βtc and then by conventional ELISA with HRP-conjugated second antibody and OPD/urea/H$_2$O$_2$ substrate. As shown in \textbf{Figures} 5.7 and 5.8, the DNA-binding activity of active pSTAT3βtc was demonstrated through the purified, recombinant, GFP-pSTAT3βtc chimaeric protein.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure5_7.png}
\caption{Detection of bound GFP-pSTAT3βtc to immobilized biotin-dsM67 oligonucleotide probe by fluorescence}
\end{figure}

\textbf{Figures} 5.7 and 5.8, the DNA-binding activity of active pSTAT3βtc was demonstrated through the purified, recombinant, GFP-pSTAT3βtc chimaeric protein.
As described in the previous section, the first 3 set of data were controls to demonstrate the extent at which the fluorescence read-out could influence the binding study. The first control experiment (i.e. dsM67 (-)/F-uST3 (-)) was to demonstrated the noise level of the wells with no capture probe and the analyte. The second control (i.e. dsM67 (-)/F-uST3 (+)) was performed to demonstrate non-specific binding of the GFP-pSTAT3βtc to the streptavidin coated well whilst the third control (i.e. dsM67 (+)/F-uST3 (-)) demonstrated the mean fluorescence intensity from the immobilized biotinylated double stranded DNA (dsM67) probe (300 pmol/well). In all 3 controls, no significant signal was generated that could influence fluorescence signals obtained from the test read-out. As anticipated, the active GFP-pSTAT3βtc demonstrated binding to the immobilized dsM67 oligonucleotide probe. Increasing mean fluorescence intensity representing increasing dsM67 DNA/GFP-pSTAT3βtc complex formation was observed upon incubating the recombinant protein with increasing amount of the immobilized DNA probe, thus suggesting high affinity of GFP-pSTAT3βtc to the oligonucleotide consensus sequence contained within the immobilized probe. Moreover, the strong increase in the mean fluorescence intensity (up to 10 pmol/well) was indicative of the retention of the biological functionality of the purified, recombinant protein GFP-pSTAT3βtc. The decreasing mean fluorescent intensity observed for the 100 and 300 pmol/well data points may be due either steric hindrance or conformational constraints as a result of molecular crowding.

Consequently, a confirmatory binding data was acquired via conventional ELISA by measuring absorbance of the substrate, OPD/H₂O₂, at 450 nm. The data obtained (Figure 5.8) further demonstrated the affinity of the protein for the native DNA sequence as well as the stability of the GFP-pSTAT3βtc/dsM67 DNA complex following the various washing steps. Furthermore, the binding profile relative to that obtained for the fluorescence read-out, Figure 5.7, was similar. Again, the $k_d$ was not calculated since the exact amount of immobilized oligonucleotide probe was not known.
5.2.1.3 Disruption of GFP-uSTAT3βtc/phosphotyrosyl peptide or GFP-uSTAT3βtc/dsM67 DNA complexes using immobilized biotin-phosphotyrosyl peptide in vitro

Following the GFP-uSTAT3βtc/phosphotyrosyl binding studies demonstrated in section 5.2.1.2.1, it was confirmed that immobilized biotin-phosphotyrosyl peptide conveniently "pulled down" or captured the GFP-uSTAT3βtc recombinant protein via the pY-SH2 domain interactions. To inhibit capture of the GFP-uSTAT3βtc by the immobilized peptide probe as well as investigate the selectivity of the SH2-phosphotyrosyl interaction, we prepared a premixed reaction containing increasing concentrations the free and unlabeled version of the immobilized phosphotyrosyl peptide in a competition assay. In another preparation, GFP-uSTAT3βtc was also incubated with increasing concentrations of a free and unlabelled dsM67 oligonucleotide DNA probe. The study was to investigate the relative affinity of the protein for both probes. Both sets of pre-incubated reactions were assayed against a constant amount of the immobilized phosphotyrosyl peptide probe. Following incubations and subsequent washes, the plate was measured and the fluorescence read-out presented as percentage inhibition as shown in Figure 5.9.
Figure 5.9: Percentage inhibition of biotin-phosphotyrosyl/GFP-uSTAT3\(\beta\)tc interaction.

The assay was performed in a flat bottom high binding black plate containing the biotinylated phosphotyrosyl peptide surrogate immobilized on a streptavidin coated plate (300 ng /well). The GFP-uSTAT3\(\beta\)tc (~ 30 nM /well) premixed with different concentrations of either the non biotin phosphotyrosyl peptide or the non biotin dsM67 oligonucleotide was pre-incubated and then incubated on the plate with the biotin immobilized phosphotyrosyl peptide. The plate was handled as described in the section 5.1.4. The plate was measured using the fluorescence read-out format. The generated mean fluorescence intensity for each data point was normalized to generate the proportion of bound GFP-uSTAT3\(\beta\)tc. From this, the percentage inhibition was calculated relative to the positive control data point (i.e. GFP-uSTAT3\(\beta\)tc samples devoid of inhibitors and representing 100 % binding to probe). A plot of % inhibition (y-axis) versus inhibitors (\(\mu\)M) (x-axis) (i.e. inhibitors colour coded as displayed in data chart) was generated to demonstrate selectivity, specificity and affinity of the inhibitors to the GFP-uSTAT3\(\beta\)tc protein.

Significant percentage inhibition was observed in the preparations containing the increasing concentrations of the free and unlabelled phosphotyrosyl peptide inhibitor. The highest percent inhibition \(i.e\) approximately 50 to 55 % was obtained with 30 and 100 \(\mu\)M of the unlabelled phosphotyrosyl peptide. This showed that the free phosphotyrosyl peptide specifically binds to the GFP-uSTAT3\(\beta\)tc SH2 domain and inhibit capture by the biotin labelled immobilized phosphotyrosyl peptide version. Also, inhibition levelled off between the 30 \(\mu\)M to 100 \(\mu\)M range suggesting saturation. The \(IC_{50}\) value for the unlabelled phosphotyrosyl peptide inhibitor was estimated at 20 ± 5 \(\mu\)M.

In a related study, the relative affinity of the GFP-uSTAT3\(\beta\)tc protein to the immobilized phosphotyrosyl peptide and a free unlabelled dsM67 oligonucleotide DNA was compared in a premixed reaction. Following binding and subsequent plate measurement, a significant inhibition of GFP-uSTAT3\(\beta\)tc binding to the immobilized peptide probe was recorded with increasing concentration of free and unlabelled
dsM67 probe up to 30 μM. Approximately 80 to 85 % inhibition was observed with 30 μM of the unlabelled dsM67 DNA probe as opposed to 50 to 55 % generated with a similar concentration of the unlabelled phosphotyrosyl peptide inhibitor probe. This suggested that the GFP-uSTAT3βtc preferentially binds to the free dsM67 oligonucleotide DNA thus, promoting a conformational rearrangement of the GFP-uSTAT3βtc protein and therefore eliciting the best inhibitory data (IC_{50} = 0.1 ± 0.05 μM) relative to the peptide probe. The protein/DNA complex may have resulted in a conformation that restricts access of the phosphotyrosyl motif of the immobilized probe to the SH2 domain of GFP-uSTAT3βtc. To verify this assertion, GFP-uSTAT3βtc premixed with different concentration of salmon sperm DNA was pre-incubated and then incubated with the immobilized phosphotyrosyl peptide to obtain an inhibition curve. No significant inhibition was observed, corroborating the earlier assertion that the dsM67 DNA sequence specifically encourages conformational changes that inhibit phosphotyrosyl/SH2 interaction.

5.2.1.4 Disruption of GFP-pSTAT3βtc/phosphotyrosyl peptide or GFP-pSTAT3βtc/dsM67 DNA complexes using immobilized biotin-dsM67 DNA in vitro

As described in section 5.2.1.2.2, the DNA-binding functionality of the phosphorylated pSTAT3βtc as a chimaeric protein, GFP-pSTAT3βtc, was conveniently demonstrated by the "capture" of the purified, recombinant, GFP-pSTAT3βtc with the immobilized biotinylated dsM67 oligonucleotide DNA probe. To disrupt active GFP-pSTAT3βtc dimers and therefore prevent binding of the protein to the immobilized DNA sequence, two sets of premixed reactions were prepared. The first set contained GFP-pSTAT3βtc premixed with different concentrations of competing and unlabelled cfsM67 oligonucleotide DNA sequence. The second set contained GFP-pSTAT3βtc premixed with increasing concentrations of the unlabelled phosphotyrosyl peptide inhibitor. Both sets of premixed reactions were pre-incubated prior to incubation with the immobilized biotinylated c/sM67 oligonucleotide DNA probe. Following incubations and the subsequent washes, the plate was measured and the mean fluorescence read-out interpreted as percentage inhibition shown in Figure 5.10. From the data, it was possible to determine the concentration of the inhibitor at which DNA binding is inhibited by 50 %, referred to DB_{50}.

The data generated from the competitive assay, recorded significant inhibition of the DNA-binding functionality of the GFP-pSTAT3βtc protein with increasing concentrations of the free and unlabelled corresponding DNA. The highest percentage inhibition, i.e., 100 % was obtained with 0.3 μM of the unlabelled DNA, thus generating a competition curve that showed saturation at concentrations above 0.3 μM. The DB_{50}
for the competitive inhibitor DNA was estimated at 0.06 ± 0.015 μM, a value suggesting that the phosphorylated pSTAT3ptc contained in the purified, recombinant GFP-pSTAT3ptc chimaeric protein did not only have a strong affinity for the native promoter sequence contained in the DNA probe, but also retained the integrity of the DBD regardless of the harsh *in vitro* conditions.

![Graph](image)

Figure 5.10: Percentage inhibition of biotin-dsM67 DNA - GFP-pSTAT3ptc interaction as a function of inhibitor concentration

The assay was performed in a flat bottom high binding black plate containing the biotinylated dsM67 oligonucleotide DNA immobilized on a streptavidin coated plate (300 ng /well). The GFP-pSTAT3ptc (~ 30 nM /well) premixed with different concentrations of either the non biotin dsM67 oligonucleotide or the non biotin phosphotyrosyl peptide was pre-incubated and then incubated with the immobilized biotin phosphotyrosyl peptide. The plate was handled as described in section 5.1.3. The plate was measured using the fluorescence read-out format. The generated mean fluorescence intensity for each data point was normalized to generate the proportion of bound GFP-pSTAT3ptc. From this, the percentage inhibition was calculated relative to the positive control data point (i.e. GFP-pSTAT3ptc samples devoid of inhibitors and representing 100 % binding to probe). A plot of % inhibition (y-axis) versus inhibitors (μM) (x-axis) (i.e. inhibitors colour coded as displayed in data chart) was generated to demonstrate selectivity, specificity and affinity of the inhibitors to the GFP-pSTAT3ptc protein.

In the second inhibitory study, the relative affinity of the activated protein, GFP-pSTAT3ptc, to both the free and unlabelled phosphotyrosyl peptide probe as well as the labelled and immobilized dsM67 DNA probe was measured. The essence of the peptide inhibitor was to inhibit the GFP-pSTAT3ptc DNA-binding functionality and therefore indirectly demonstrate their ability to disrupt the active GFP-pSTAT3ptc dimer by associating with the GFP-pSTAT3ptc and destabilizing the pSTAT3ptc/pSTAT3ptc dimer molecule thereby forming the inactive GFP-pSTAT3ptc/phosphotyrosyl peptide heterocomplexes.
Hence, following the fluorescence read-out, the percentage inhibition was calculated and shown in Figure 5.10. The inhibition gradually increased with increasing concentration of the unlabelled phosphotyrosyl peptide inhibitor. At 100 μM peptide inhibitor concentration, the highest percentage inhibition of approximately 70 % was achieved, generating an estimated DB \(_{50}\) of 37 ± 3 μM. This suggested weaker affinity of the protein to the unlabelled phosphotyrosyl peptide inhibitor relative to the immobilized dsM67 DNA probe. However, the observed percentage inhibition suggested that at higher concentrations, the phosphotyrosyl peptide inhibitor may be disrupted the GFP-pSTAT3βtc dimer via their SH2 domain, thus shifting the reaction equilibrium from active GFP-pSTAT3βtc dimers to inactive GFP-pSTAT3βtc/peptide heterocomplexes. The GFP-pSTAT3βtc/peptide interactions may therefore distort the protein’s conformation and reduce affinity for the immobilized dsM67 DNA. Alternatively, the unlabelled peptide inhibitor was simply sterically interfering with the DNA-binding activity of GFP-pSTAT3βtc.

To investigate the specificity of the peptide inhibitor, the GFP-pSTAT3βtc was premixed and then pre-incubated with different concentrations of the free and unlabelled phosphotyrosyl peptide inhibitor, H-pY-LKTKFI-NH\(_2\), corresponding to the core native C-terminal STAT3 SH2 domain binding sequence GSAA-PpYLKTKFIIC (Mao et al., 2005). The peptide was taken as a standard since the sequence surrounding the Y\(^{705}\) had been shown to bind to the STAT3 SH2 domain in the phosphorylated STAT3 crystal structure (Becker et al., 1998). Following incubation with the immobilized DNA probe and subsequent fluorescence measurement, the data showing percentage of inhibition was generated. A very strong affinity for the immobilized dsM67 DNA was demonstrated by the GFP-pSTAT3βtc protein, suggesting a much weaker inhibition potency of this peptide even at 100 μM relative to the former peptide inhibitor derived from the gp130 receptor. The data was consistent with earlier reports that suggested that the gp130 based phosphotyrosyl peptide has significantly higher affinity for the STAT3 SH2 domain than the latter phosphotyrosyl peptide (Schust et al., 2003; Ren et al., 2003).

5.3 DISCUSSION

The main objectives of the work described in this chapter were to evaluate the biological functionality of the purified, recombinant, GFP-STAT3βtc isoforms and to develop an easy, quick, robust and sensitive fluorescence-based microtiter plate assay to measure the functionality of the GFP tagged STAT3βtc constructs. The objectives were largely achieved since purified fractions of the phosphorylated and unphosphorylated GFP-STAT3βtc isoforms were successfully captured with either
Edwin Nkansah

Biotinylated dsM67 oligonucleotide DNA or biotinylated gp130 derived phosphotyrosyl peptide immobilized on streptavidin coated surface respectively. The captured or bound fluorescent chimaeric protein was initially detected by fluorescence and then by conventional ELISA. Further, through competitive inhibitions assays, the IC$_{50}$ values for both unlabeled dsM67 DNA and phosphotyrosyl peptide inhibitors were estimated.

### 5.3.1 Streptavidin coated plates

Black, flat-bottom high binding plates were successfully coated with streptavidin reconstituted in 20 mM potassium phosphate buffer, pH 6.5. Even though 300 ng/well (300 ng/well per 100 μL) streptavidin coating was sufficient to generate a reasonable binding signal, wells coated with 1000 ng (per 100 μL) of streptavidin was considered to have attained its optimum surface capacity (Figure 5.4), thereby generating significantly higher signals corresponding to about 10-fold signal improvement.

Coating the well with streptavidin at 37 °C did not necessary increase the binding capacity of streptavidin per well, although the binding kinetics were slightly enhanced when streptavidin coating was performed at 37 °C. However, pH of the streptavidin vehicle had a significant effect on the adsorption efficiency of streptavidin onto the polystyrene plate. The most efficient adsorption was obtained with acidic buffers (specifically pH 5.5 - 6.5). When coating buffers of higher pH-values (> 7.0) were applied, the adsorption efficiency of the streptavidin was significantly reduced. The most efficient adsorption among the buffers with a constant pH (6.5) was achieved with potassium phosphate buffer; yielding 40 to 45 % improvement in the adsorption capacity, relative to other buffers (HEPES- or Tris-based buffers).

Using potassium phosphate buffer (pH 6.5) as the optimum buffer for streptavidin, and subsequent incubation of plates at 37 °C (high humidity) for 14 h resulted in improved adsorption properties of the streptavidin and allowed the coating process to be carried out with a shorter incubation time (~ 6 h) rather than the commonly reported time of 14 h.

Unoccupied binding sites were usually blocked with BSA contained in PBS buffer. Blocking of the wells improved the stability and the biotin-HRP binding property of the coated streptavidin.

The efficiency of the coating was usually evaluated with the biotin-HRP conjugate. Typically, 400 pg/wells biotin-HRP generated a binding signal of approximately 0.5 mAU at 450 nm following incubation (60 min) on a streptavidin surface coated with 1000 ng/well (i.e. 100 μL per well).
5.3.2 Immobilized capture probes

Different concentrations of the capture probes (i.e. either the biotinylated dsM67 oligonucleotide DNA or the biotinylated phosphotyrosyl peptide) were successfully immobilized on the streptavidin coated surface (1000 ng/well) after 60 min of incubation at room temperature and with rocking. Significant improvement in the stability of the immobilized capture probe was achieved with phosphate buffered saline as the vehicle for the capture probes, generating 30 to 40 % increment in the immobilization capacity of the probe, relative to other buffers (HEPES or Tris-based buffers). The immobilization temperature for the capture probes was optimum at room temperature (~ 20 to 22 °C). However, the binding kinetics was enhanced by increasing incubation temperature to 37 °C.

5.3.3 Specific interaction of the recombinant GFP-uSTAT3βtc with the immobilized phosphotyrosyl peptide

By applying the immobilized phosphotyrosyl peptide probe, the capture of the purified recombinant GFP-uSTAT3βtc mimicked the transient docking of STAT3 via its SH2 domain at the phosphotyrosyl motif (pY905) of the activated gp130 receptor in the cytosolic region of the cell (Haan, et al., 1999; Greenlund et al., 1994). To evaluate the functionality of the STAT3-SH2 domain contained in the unphosphorylated fluorescent STAT3βtc (GFP-uSTAT3βtc), the purified, recombinant, monomeric protein was captured by the immobilized pY motif contained in the hexapeptide sequence based on the known STAT3 docking site of the gp130 (i.e. EGMPKSpYLPQTVRQ; the underlined sequence represent the immobilized biotinylated sequence while the phosphotyrosyl motif is bolded) (Stahl et al., 1995; Haan et al., 1999; Lehmann et al., 2005). The phosphotyrosyl peptide probe, H-pYLPQTV-NH₂, has been shown to have high affinity for the STAT3-SH2 domain (Ren et al., 2003) and therefore has been applied successfully in other assay formats including Fluorescence Polarization (FP) assays to screen small molecules that binds to the STAT3-SH2 domain (Schust et al., 2004). In addition, the phosphotyrosyl peptide has been successfully used as an inhibitor in the PEMSA and FRET assays described in chapters 6 and 7 respectively.

Using both the fluorescence-based and the conventional ELISA read-out formats, we detected a dose-dependent increase in the level of the GFP-uSTAT3βtc constructs captured by the immobilized phosphotyrosyl peptide up to peptide levels of approximately 3 pmol/well, after which the increase in the amount of the immobilized phosphotyrosyl peptide had no effect on the level of protein captured or bound (see Figures 5.5 and 5.6). The observed signal suggested that the immobilized peptide specifically targets and “recruits” the recombinant, GFP-uSTAT3βtc protein;
Edwin Nkansah

corroborating well known knowledge that gp130 docking sequence had a high affinity for the STAT3 protein (Stahl et al., 1995). In addition, the capture of the purified protein demonstrated that the integrity of the binding pocket of the GFP-uSTAT3βtc construct was retained in the fluorescent construct and not compromised during the extraction and purification processes.

Further, the observed saturation for the immobilized peptide as recorded by both the fluorescence and the antibody based microtiter plate assays indicates that the streptavidin coated wells could sufficiently accommodate up to 3 pmol of the biotinylated phosphotyrosyl peptide per well, and thus suggesting that the immobilized peptide could only capture up to 3 pmol (~ 0.002%) of the GFP-uSTAT3βtc protein when incubated together. Though the amount captured was negligible relative to the applied protein (per well), it must be noted that the applied protein was not a completely homogeneous mixture and so contained contaminants in the Bradford assay used to estimate the concentration of the GFP-uSTAT3βtc protein. Hence, the amount of the applied protein may not necessarily be the exact concentration of the target GFP-uSTAT3βtc protein applied.

As demonstrated by the identical titration profile for both the fluorescence-based and the conventional (ELISA) microtiter plate assays, it was obvious that the less expensive and less laborious fluorescence-based assay format was a more convenient approach to measure the biological functionality of the protein being investigated. Hence, the format was further explored to characterize the phosphotyrosyl peptide in relation to its ability to specifically target and capture the GFP-uSTAT3βtc via the pY-SH2 domain interaction in a competition assay. The inhibitory potency of the free (unlabelled) corresponding phosphotyrosyl peptide isoform confirmed that the interaction between the phosphotyrosyl peptide and the GFP-uSTAT3βtc protein molecule was indeed mediated by the pY-SH2 domain interactions (Figure 5.9). More interestingly, the study was in good agreement with the knowledge that, the pY motif contained in the immobilized hexapeptide sequence present in the interleukin-6 signal-transduction receptor subunit, gp130, is indeed crucial for STAT3 "recruitment" and subsequent activation through phosphorylation (Gerhardt et al., 1996).

5.3.4 Specific interaction of the recombinant GFP-pSTAT3βtc with the immobilized dsM67 oligonucleotide DNA

Activation of the STAT3 molecules via phosphorylation results in the formation of active homo- or heterodimers that renders them competent to undergo nuclear translocation, where they bind to specific promoter sequences and then regulate gene expression
As discussed in chapter 1, phosphorylation of STAT3 and the subsequent dimerization of STAT3 protein molecules through the phosphotyrosyl/SH2 domain interaction is essential for the DNA binding activities of the protein (Shuai et al., 1994; Wen et al., 1996). To this end, the binding functionality of the highly conserved DNA-binding domain contained in the purified, recombinant, GFP-pSTAT3βtc protein constructs was evaluated using the biotin labelled high affinity c-fos sis-inducible element (hSIE; M67) oligonucleotide DNA sequence (Wagner et al., 1990). The functionality was measured by the fluorescence-based and then the conventional ELISA read-out formats. The phosphorylated, GFP-pSTAT3βtc was incubated with different concentrations of the dsM67 DNA and the extent of capture the GFP-pSTAT3βtc measured.

The presence of the immobilized dsM67 oligonucleotide DNA displayed a dose-dependent increase in the amount of GFP-pSTAT3βtc protein captured by the dsM67 DNA probe, up to concentrations of 3 pmol/ well (Figures 5.7 and 5.8). The observation was consistent with some of the reported biological functions of the active STAT3βtc including the phosphorylation being a pre-requisite for STAT3 DNA-binding. In addition, the formation of the biotin-dsM67/GFP-pSTAT3βtc complex suggested that GFP-STAT3βtc construct co-expressed with the Elk receptor tyrosine kinase domain (expression in the TKB1™ E.coli strain) did not only successfully dimerize following phosphorylation but also retained its DNA binding capability (Becker et al., 1998). To test the selectivity of this interaction, the unphosphorylated, monomeric isoform, GFP-uSTAT3βtc, was incubated with the immobilized dsM67 DNA. No significant binding was detected in both read-out formats. This was consistent with the STAT3 signalling paradigm, i.e. phosphorylation and subsequent dimerization being a prerequisite for DNA-binding. However, the observation was proven otherwise in the PEMSA assays described in chapter 6.

Further, GFP-uSTAT3βtc was pre-incubated with different concentrations of unlabeled dsM67 DNA prior to incubation with the immobilized phosphotyrosyl peptide applied in section 5.3.3. The presence of the dsM67 DNA generated a strong dose-dependent inhibition of the GFP-uSTAT3βtc binding to the immobilized phosphotyrosyl peptide (Figure 5.9). This suggested that the GFP-uSTAT3βtc protein had a higher affinity for dsM67 DNA inhibitor and therefore was promoting the formation of GFP-uSTAT3βtc/dsM67 DNA complexes that may be conformationally constrained, hence preventing GFP-uSTAT3βtc from binding to the pY motif of the immobilized phosphotyrosyl peptide probe. Also, the dsM67 DNA may be interacting with the immobilized peptide probe, hence hindering GFP-uSTAT3βtc from binding to the
immobilized peptide. Though the latter is less likely, binding of the high affinity dsM67 DNA had been studied in the PEMSA and the FRET assays (see chapter 7) to induce (GFP-uSTAT3βtc)/dsM67 DNA interactions. Nevertheless, we suggested that a transient association may have occurred between the immobilized dsM67 DNA and the GFP-uSTAT3βtc protein.

5.3.5 H-pYLPQTV-NH2 and dsM67 DNA significantly inhibits GFP-uSTAT3βtc binding to the pY motif of the immobilized biotin-H-pYLPQTV-NH2 peptide sequence

Through competition assays, we were able to confirm reports in literature that suggested that the complex formed between the immobilized phosphorylated peptide and the GFP-uSTAT3βtc protein was indeed mediated by the pY- SH2 domain interaction (Schust et al., 2003; Turkson et al., 2001). The binding activity of the unphosphorylated GFP-uSTAT3βtc protein was significantly reduced following pre-incubation of the protein with the peptide inhibitor (unlabelled phosphotyrosyl peptide isoform) prior to binding to the immobilized biotinylated phosphotyrosyl peptide isoform. The percentage inhibition was recorded as 25 %, generated with equimolar concentrations of the GFP-uSTAT3βtc and the phosphotyrosyl peptide inhibitor. In a related study, approximately 95 % inhibition was observed when the unlabelled phosphotyrosyl peptide inhibitor was substituted for the unlabelled dsM67 oligonucleotide DNA. The observation suggested that, though unphosphorylated, the uSTAT3βtc containing construct, GFP-uSTAT3βtc, demonstrated higher affinity (~ 4-fold) to its recognition sequence contained in the unlabelled dsM67 DNA sequence relative to the pY motif of the phosphotyrosyl peptide probe.

5.3.6 H-pYLPQTV-NH2 and dsM67 DNA disrupts GFP-pSTAT3βtc binding to the STAT3 recognition sequence contained within the immobilized biotinylated dsM67 oligonucleotide DNA

To confirm that the GFP-pSTAT3βtc actually binds to and forms a stable complex with the high affinity dsM67 DNA as reported in literature (Park et al., 1996 and Shuai et al., 1994), a competition assay was prepared. Although a drop in the binding activity (i.e. percentage inhibition of about 18 % in an equimolar preparation containing the GFP-pSTAT3βtc and the phosphotyrosyl peptide inhibitor) of the active GFP-pSTAT3βtc was observed when pre-incubated with the unlabelled phosphotyrosyl peptide, H-pYLPQTV-NH2, prior to incubation with the immobilized biotinylated dsM67 DNA, a complete inhibition of the biotin-dsM67/GFP-pSTAT3βtc interaction was recorded when the unlabelled phosphotyrosyl inhibitor was substituted for the unlabelled dsM67 DNA (Figure 5.10). The observed 100 % inhibition further confirmed the affinity of STAT3 for its native recognition sequence contained in dsM67 oligonucleotide DNA. Furthermore,
the data was consistent with earlier reports that suggested that phosphorylation and subsequent dimerization through phosphotyrosine-SH2 domain interactions was essential for DNA binding (Shuai et al., 1994). The small percentage of inhibitory activity (i.e. ~ 18%) recorded for the unlabelled phosphotyrosyl peptide inhibitor further demonstrated the stability of the active GFP-pSTAT3βtc dimer. The phosphotyrosyl peptide was less able to compete with the SH2 domain contained in the GFP-pSTAT3βtc dimer than that contained in the inactive GFP-uSTAT3βtc monomer.

In summary, conventional ELISA was successfully adapted to a simple, efficient and less laborious fluorescence-based microtiter plate assay to evaluate the biological functionality of the purified, recombinant fluorescent labelled STAT3βtc isoforms, GFP-STAT3βtc. Both active and inactive purified GFP-STAT3βtc, i.e., GFP-pSTAT3βtc and GFP-uSTAT3βtc, were successfully captured or bound to immobilized dsM67 DNA and immobilized phosphotyrosyl peptide receptor surrogate respectively. The bound fluorescent protein molecules were detected by fluorescence at the wavelength emitted by the GFP reporter linked to the STAT3 protein. Subsequent confirmation by HRP-coupled antibody was not routinely used, since the fluorescent binding profiles were identical to that generated by the conventional ELISA read-out.

Further, the microtiter plate assay was used to generate inhibition data that demonstrated the relative affinities of the GFP-STAT3βtc isoforms for the peptide and DNA probe used in the study. The assay was robust even in the presence of up to 10% DMSO. Thus, these and other features of the assay makes it an appropriate alternative to conventional ELISA for measuring the biological functionalities of fluorescently tagged recombinant proteins.
6 GFP-STAT3βtc/dsDNA COMPLEX IN GELS

6.1 Electrophoretic Mobility Shift Assay (EMSA)

EMSA is a simple, quick and sensitive method for analysing sequence-specific DNA binding proteins in complex with double stranded DNA (Fried and Crothers, et al., 1981). It is also used to analyze protein-protein, protein-DNA and protein-peptide interactions. EMSA is based on the idea that protein bound DNA migrates slower than non-bound DNA in a native polyacrylamide or agarose gel, resulting in a "shift" in migration of the labelled DNA. EMSA was applied to qualitatively and quantitatively characterize GFP-STAT3βtc and GFP-STAT3βtc-DNA complexes in the presence or absence of either a phosphotyrosyl peptide inhibitor or a small molecule inhibitor.

Activation of STAT3 through tyrosine phosphorylation is sufficient for STAT3 dimers to acquire DNA-binding capabilities (Lutticken et al., 1994; Yuan et al., 1994; Shuai et al., 1993; Sasse et al., 1994). Although, numerous studies have been carried out to support this model, very little is known about the biological relevance of the unphosphorylated STAT3 in relation to drug discovery. Recently, however, a number of studies have provided evidence to suggest that uSTAT3 exist as a dimer prior to activation (Braunstein et al., 2003; Novak et al., 1998; Haan et al., 2000). Moreover, Chatterjee-Kishore and colleagues have also reported that unphosphorylated STAT1 binds to the DNA recognition sequence pSTA1, i.e., either as a monomer or a dimer (Chatterjee-Kishore et al., 2000). These studies suggested that uSTAT3 may have other biological functions including DNA binding capabilities, thus we applied our GFP-STAT3βtc construct to investigate the binding activity of the STAT3βtc contained in the construct.

Recognizing this possibility, we have applied EMSA to demonstrate mobility shift between complexes of the purified GFP-pSTAT3βtc and GFP-uSTAT3βtc and the STAT- specific double stranded M67 DNA (a high affinity mutant of the sis-inducible element (hSIE) of the human c-fos gene) (Wagner et al., 1990). Further, EMSA was used to investigate the relative stabilities of the of the protein-DNA complexes formed between either the GFP-pSTAT3βtc or the GFP-uSTAT3βtc and the dsM67 DNA using the STAT3 SH2 domain-specific phosphotyrosyl peptides (a surrogate receptor derived from the interleukin-6 receptor subunit gp130) and some small molecules.

With GFP reporter serving as a "tracking device", the relative mobility shift between the GFP-pSTAT3βtc and the GFP-uSTAT3βtc as well as DNA bound GFP-pSTAT3βtc and DNA bound GFP-uSTAT3βtc was captured after illuminating the gel with blue light (~
(470 nm) from a Safer Imager™ (Invitrogen™, UK). In addition, separated protein samples may be transferred onto appropriate membrane materials and then detected either by western blot using primary antibodies against STAT3 or by southern blot using Streptavidin-Horseradish peroxidase conjugate.

6.1.1 EMSA in polyacrylamide gel

EMSA was performed under non-reducing conditions using Bio-Rad mini gel system. Unless otherwise stated, 5 % polyacrylamide gels were used for the analysis. Prior to gel preparation, the glass plates were prepared as described in section 2.2.5.1.1.

6.1.1.1 Preparation of Tris-Boric-EDTA (TBE) polyacrylamide gel

A 5 % 0.25x TBE, polyacrylamide gel was used in the EMSA study. The gel was prepared in a 50 mL falcon tube as follows; 375 µL filtered and degassed 10x TBE buffer (1000 mM Tris, pH 8.2, 900 mM boric acid and 100 mM EDTA), 2.5 mL ProtoGel (37: 5: 1 acrylamide to bis-acrylamide stabilized solution), 2.5 % (w/v) glycerol and 0.012 mL N,N,N',N'-tetramethylethylene diamine (TEMED) was made up to 15.0 mL with ddH₂O. Upon adding 0.12 mL 0.1 mg/mL APS to the gel mix, polymerization was initiated. The mixture was poured gently into the space between the two glass plates until it was completely filled. The appropriate comb (10 wells; 5mm x 1mm or 15 wells; 3 mm x 1mm) were positioned. No stacking gel was required. The gel was allowed to polymerize for 1 h at room temperature and a further 2 h min at 4 °C. After the gel was sufficiently polymerized, the plates were assembled as described in section 2.2.5.3.1. The pre-chilled 0.25x TBE (~ 1L) was poured into the buffer reservoir. The combs were gently removed and the wells were immediately flushed with the chilled 0.25x TBE tank buffer. The electrophoresis tank was packed in ice and the gel was pre-run at 200 V for 2 h while preparing samples in section 6.1.4.

6.1.2 EMSA in agarose gels

EMSA using agarose was performed under non-reducing conditions using Bio-Rad Wide Mini-Sub cell GT, 15 x 10 cm gel system. Unless stated otherwise, 2 % (w/v) agarose gel was used for the analysis. Prior to gel preparation, the gel tray (15 x 7 cm) and comb (30 wells; 1.50 x 2.69 mm, 20 wells, 1.50 x 4.84 mm) were rinsed with 70 % ethanol and then ddH₂O. The accessories were wiped dry and gel tray subsequently sealed at both ends with an autoclave tape in preparation for the molten agarose. The tray was place on a white tissue paper.
6.1.2.1 Preparation of Tris-Boric-EDTA (TBE) agarose gel

The 2 % (w/v) TBE agarose gel was prepared in a 100 mL sterile glass bottle as follows; 1.5 mL 10x filtered and degassed TBE buffer (1000 mM Tris, pH8.2, 900 mM boric acid and 100 mM EDTA) and 1.2 g agarose powder was made up to 60 mL with ddH₂O. The agarose suspension was swirled to mix and then boiled for 1 min to melt the agarose powder. The molten agarose solution was then left to cool (~ 50 °C) on the bench. The molten solution was slowly transferred into the gel tray-tank and to about 0.5 cm depth. The appropriate comb was subsequently inserted whilst ensuring that bubbles were pushed away to the sides using disposable tips. The gel was allowed to set for 1 h. Upon setting, the comb was gently and the gel was transferred into Bio-Rad Wide Mini-Sub cell GT system. The wells were flushed with chilled 0.25x TBE reservoir buffer. The electrophoresis tank was packed in ice and the gel subsequently run at 200 V for 2 h while preparing samples in section 6.1.4.

6.1.3 Preparation of binding reactions

6.1.3.1 Preparation of unbound GFP-STAT3\(\beta\)tc control

Only purified, recombinant GFP-STAT3\(\beta\)tc isoforms were used in the EMSA study. A typical "protein only" reaction mix contained the following in the order of addition; appropriate volume of sterile ddH₂O to dilute reaction mix to 15 µL, 3.0 µL 5x binding buffer (100 mM HEPES, pH 7.9, 200 mM KCl, 5 mM MgCl₂, 0.15 mM EDTA, 0.15 mM EGTA, 5.0 mM DTT and 40 % (w/v) glycerol), 0.5 µg salmon sperm DNA, 2.5 µg BSA (5 mg/mL) and 3 pM GFP-STAT3\(\beta\)tc (~5 µg). The reaction mixture was stirred gently with the tip of the 20 µL pipette (tip) and then centrifuged for 15 s at 13,000 xg. The mix was subsequently incubated at room temperature while rocking for a minimum of 20 min.

6.1.3.2 Preparation of GFP-STAT3\(\beta\)tc/dsM67 DNA complex reaction

All reactions were prepared in 0.5 mL sterile tubes. A typical EMSA protein-DNA complex reaction (15 µL) was prepared using a 1:1 protein-DNA molar ratio. The order of addition of reaction components is as follows; sterile ddH₂O (appropriate volume to dilute reaction mixture to 15 µL), 3.0 µL 5x binding buffer (100 mM HEPES, pH 7.9, 200 mM KCl, 5 mM MgCl₂, 0.15 mM EDTA, 0.15 mM EGTA, 5.0 mM DTT and 40 % (w/v) glycerol), 0.5 µg salmon sperm DNA, 2.5 µg BSA (5 mg/mL), 3 µM GFP-STAT3\(\beta\)tc (~5 µg) and either 3 µM or 9 µM of Biotin-labeled or unlabeled dsM67 DNA. The reaction mix was stirred gently with the 1 µL pipette tip and then centrifuged for 15 s at 13,000 xg. The mixture was subsequently incubated at room temperature for a minimum of 10 min while shaking.
6.1.3.2.1 Determination of the molar ratio for GFP-STAT3ptc/dsM67 complex formation

Reactions containing the following ingredients were prepared in a sterile 0.5 mL tube to a volume of 15 μL; either 0 or 0.3 or 0.9 or 3.0 or 9.0 μM dsM67 was incubated with 3 μM of either GFP-pSTAT3ptc or GFP-uSTAT3ptc. Other ingredients used includes; sterile ddH₂O (appropriate volume to dilute reaction mixture to 15 μL), 3.0 μL 5x binding buffer (100 mM HEPES, pH7.9, 200 mM KCl, 5 mM MgCl₂, 0.15 mM EDTA, 0.15 mM EGTA, 5.0 mM DTT and 40 % (w/v) glycerol), 0.5 μg salmon sperm DNA and 2.5 μg BSA (5 mg/mL). The reaction mix was stirred gently with the 1 μL pipette tip and then centrifuged for 15 s at 13,000 xg. The mixture was subsequently incubated at room temperature for 1 h while rocking.

A total of 10 reactions were prepared, 5 for each isoform of GFP-STAT3ptc chimaeric protein. The first of each 5 of the reactions contained no dsM67 DNA. The remaining 4 reaction contained increasing concentration of dsM67 DNA.

Upon incubation, the reaction was further centrifuged for 15 s at 13,000 xg and then loaded into either the 0.25x TBE polyacrylamide (5 %) or agarose (2 %) gel and subsequently electrophoresed as described in section 6.1.4.

6.1.3.2.2 Analysis of the time needed for GFP-STAT3ptc/dsM67 DNA complex formation

Reactions containing the following ingredients were prepared in sterile 0.5 mL tube to a total volume of 15 μL; sterile ddH₂O (appropriate volume to dilute reaction mixture to 15 μL), 3.0 μL 5x binding buffer (100 mM HEPES, pH 7.9, 200 mM KCl, 5 mM MgCl₂, 0.15 mM EDTA, 0.15 mM EGTA, 5.0 mM DTT and 40 % (w/v) glycerol), 0.5 μg salmon sperm DNA, 2.5 μg BSA and 1.5 μM GFP-pSTAT3ptc or GFP-uSTAT3ptc (~ 2.5 μg). The reaction mix was stirred gently with the 1 μL pipette tip and then centrifuged for 15 s at 13,000 xg. The mixture was then incubated at room temperature for 10 min.

A total of 10 reactions were prepared, 5 for each isoform GFP-STAT3ptc. Reaction number 5 of each sets of reactions were incubated for 60 min with 1.5 μM dsM67 DNA. Reactions 4, 3, 2 and 1 were also incubated with 1.5 μM dsM67 DNA for 30, 15, 10 and 5 min respectively prior to electrophoresis. The reactions were incubated at room temperature while rocking. They were further centrifuged for 15 s at 13,000 xg, loaded onto the 0.25x TBE polyacrylamide (5 %) gel and then electrophoresed as described in section 6.1.4.
6.1.3.3 Preparation of the complex between GFP-STAT3ptc and the phosphotyrosyl peptide

Here, the dsM67 DNA was substituted with one of the following phosphotyrosyl peptides: either H-pYLPQTV-NH$_2$ (Ren et al., 2003) or Acetyl-pYLPQ (Turkson et al., 2001) or H-pYLTKFI-NH$_2$ (Turkson et al., 2001) (where “p” represents phosphorylation). The peptides were prepared and stored (-80 °C) in 100 % dimethyl sulfoxide (DMSO). As a control, a final concentration of 500 μM was prepared in a 15 μM reaction mix containing the following ingredients; ddH$_2$O (appropriate volume to dilute reaction mixture to 15 μL), 3.0 μL 5x binding buffer (100 mM HEPES, pH7.9, 200 mM KCl, 5 mM MgCl$_2$, 0.15 mM EDTA, 0.15 mM EGTA, 5.0 mM DTT and 40 % (w/v) glycerol), 0.5 μg salmon sperm DNA, 2.5 μg BSA and 3.0 μM GFP-STAT3ptc (~ 5.0 μg) isoforms, i.e., either the phosphorylated or the unphosphorylated isoforms. The reaction mixtures were stirred gently with the 1 μL pipette tip and then centrifuged for 15 s at 13,000 xg. The mix was then incubated at room temperature for a minimum of 10 min.

Following incubation, the reactions were further centrifuged for 15 s at 13,000 xg, loaded onto TBE agarose or polyacrylamide gel and then electrophoresed as described in section 6.1.4.

6.1.3.4 Inhibition or disruption of GFP-STAT3ptc/dsM67 DNA complex with a phosphotyrosyl peptide inhibitors

To determine the degree of inhibition by the phosphotyrosyl peptide inhibitor, the order of incubation of the hot probe (dsM67) DNA and the cold probe (phosphotyrosyl peptide) were investigated. Inhibition reactions containing either 100 or 200 μM H-pYLPQTV-NH$_2$ phosphotyrosyl peptide inhibitor were prepared.

The first set (duplicate) of reactions contained the following; ddH$_2$O (appropriate volume to dilute reaction mixture to 15 μL), 3.0 μL 5x binding buffer (100 mM HEPES, pH 7.9, 200 mM KCl, 5 mM MgCl$_2$, 0.15 mM EDTA, 0.15 mM EGTA, 5.0 mM DTT and 40 % (w/v) glycerol), 0.5 μg salmon sperm DNA, 2.5 μg BSA and 3.0 μM GFP-STAT3βtc. The mixtures were pre-incubated with either 100 or 200 μM peptide inhibitor for 1 h at room temperature prior to incubation with dsM67 DNA (3 μM). The reactions were further incubated for 1 h at room temperature while rocking.

The second preparations were similar to the first except that both the hot probe (dsM67) DNA and the cold probe (peptide inhibitor) were incubated together with the reaction mixture for 1 h at room temperature.
The third and final set of reactions were also similar to the first reactions except that the preparations were pre-incubated with the hot probe, 3 μM (dsM67) DNA for 1 h at room temperature prior to incubation with either 100 or 200 μM phosphotyrosyl peptide inhibitor. The mixture was further incubated for 1 h.

The preparations were gently stirred with a 1 μL pipette tip prior to incubation and then centrifuged for 15 s at 13,000 xg. The reactions was loaded onto TBE agarose or polyacrylamide gel and subsequently electrophoresed as described in section 6.1.4.

6.1.3.4.1 Disruption of pre-formed GFP-STAT3βtc/dsM67 DNA complex using the different forms of phosphotyrosyl peptide inhibitors

All the three peptides listed in section 2.8.4.3 were applied in this study. Thus, 8 times 15 μL inhibition reactions supplemented with either H-pYLPQTV-NH₂ or Acetyl-pYLPQ or H-pYLKTKFI-NH₂ to 1, 3, 10, 30, 100, 150, 300 and 500 μM were prepared. Each reaction mix contained the following ingredients; ddH₂O (appropriate volume to dilute reaction mixture to 15 μL), 3.0 μL 5x binding buffer (100 mM HEPES, pH 7.9, 200 mM KCl, 5 mM MgCl₂, 0.15 mM EDTA, 0.15 mM EGTA, 5.0 mM DTT and 8 % (w/v) glycerol, 0.5 μg salmon sperm DNA, 2.5 μg BSA, 3 μM dsM67 DNA and 3 μM GFP-STAT3βtc isoforms. The mixtures were gently stirred and then centrifuged for 15 sec at 13,000 xg. They were then incubated at room temperature for 1 h with rocking.

Upon incubation, the preparations were supplemented with the various peptide inhibitors to the appropriate concentrations, i.e., 1, 3, 10, 30, 100, 150, 300 and 500 μM. They were mixed, centrifuged and then further incubated at room temperature for 1 h. Following incubation, the preparations were separated as described in section 6.1.4.

6.1.4 Electrophoresis of protein, protein/DNA and protein/peptide complexes

Prior to loading, the reservoir buffer in the buffer dam was replaced with a fresh ice-chilled buffer. The wells were also flushed with the same buffer. Into each well, 7.5 μL (~2.5 μg GFP-STAT3βtc) of the reaction mixture was loaded per well. No loading dye was used. The samples were electrophoresed for 2 h at 150 volts, i.e., until the protein only, protein-DNA complex and free-biotin-labeled dsM67 have migrated approximately 1/4, 2/3 and 3/4 down the length of the 5 % gel respectively. At completion, electrophoresis was stopped, and the gel system disassembled. The protein or protein-DNA complex was visualized under blue light and the image capture using a Sony Cyber Shot digital camera. The image was documented and the gel either stained with coomassie brilliant blue dye or the binding reactions transferred via either western blot.
(protein) or southern blot (DNA) onto nitrocellulose (Hybond-C Extra) or nylon membrane (Hybond-N Extra).

6.1.4.1 Electrophoretic transfer of binding reactions

Upon successful fractionation of the binding reactions, the samples were eluted onto either a nitrocellulose or a nylon membrane from the polyacrylamide gel (Towbin et al., 1979). The transfer of protein from polyacrylamide gel was performed in a BioRad Mini Protean® 3 Western Trans-blot system. The transfer system was prepared as described in section 2.2.5.2.

6.1.4.2 Preparation for western blotting (Nitrocellulose membrane)

The following accessories were prepared prior to blotting; the Bio-Ice cooling unit was filled with ddH₂O and frozen at -20 °C until ready to use, 800 mL transfer buffer (31 mM Tris, pH8.5, 192 mM glycine, 20 % methanol diluted to 800 mL with ddH₂O), 4 x 3 MM paper and 1 x nylon Hybond-C nitrocellulose membrane was cut to the dimension of the gel. The gel, nitrocellulose membrane, 3MM paper and fiber pads were soaked in the transfer buffer for 20 min. The sandwiched was prepared in a tray partially filled with the transfer buffer as follow as described in section 2.2.5.2.1.

Upon preparation of the sandwich-cassette, it was placed in the module and then the whole unit transferred into the tank. The Bio-Ice cooling unit was also added. The tank was then completely filled with transfer buffer. The lid was positioned and tank was packed in ice to avoid overheating. The transfer was performed for 2 h at 150 mA and 100 V.

Upon completion, the sandwich was disassembled and the membrane removed for development, section 2.2.5.2.1. The system was cleaned with a detergent and rinsed with ddH₂O.

6.1.4.3 Preparation for southern blotting (nylon membrane)

The following accessories were prepared prior to blotting; the Bio-Ice cooling unit was filled with ddH₂O and frozen at -20 °C until ready to use, 1 L pre-chilled 0.25x TBE buffer (~ 2 °C). Also, 4x 3 MM paper and 1 x nylon Hybond-N® membrane were cut to the dimension of the gel. The gel, nitrocellulose membrane, 3MM paper and cleaned fiber pads were soaked in the 0.25x TBE buffer for 10 min. The sandwiched was prepared in a tray partially filled with the tank buffer as described in section 2.2.5.2.1.

Upon preparation of the sandwich-cassette, it was placed in the module and then the whole unit transferred into the tank. The Bio-ice cooling unit was also added. The tank
was then completely filled with the transfer buffer and then packed in ice. The lid was positioned and then the transfer performed for 40 min at 300 mA.

Following transfer, the unit was disassembled and the membrane gently removed with a clean forceps at the edges. It was then transferred onto a fresh 3MM paper with the gel-membrane interface facing up. The membrane was blotted with the 3MM paper (~1 min) and then developed as described in section 6.1.4.4.

### 6.1.4.4 Development of positively charged nylon membrane

Prior to staining the transferred biotin labeled dsM67 DNA probe, i.e. free and protein bound, with streptavidin-Horseradish peroxidase conjugate, the membrane was spotted with 0.5 µL 3 µM biotin labeled dsM67 DNA at the top left corner. It was then baked at 37 °C for 15 to 20 min and then gently wrap in a Saran™ cling film.

The transferred DNA was cross-linked by exposure to a UV light source, 312 nm, for 5 min. The membrane was subsequently unwrapped and then either stored dry and in a dark place between two clean 3MM papers at room temperature or was developed to stain the cross-linked DNA. Development of the membrane is described in section 6.1.4.5.

### 6.1.4.5 Detecting biotin-labeled dsM67 oligonucleotide DNA

The cross-linked membrane was transferred into a clean petri dish. It was blocked with 20 mL of blocking buffer (2 mg/mL BSA and 0.025 % Tween 20 in PBS; 181 mM NaCl, 2.68 mM KCl, 4.28 mM Na₂HPO₄ (pH 7.3), 1.46 mM KH₂PO₄) for 25 min with gentle rocking. Whilst blocking, 20 mL conjugate/blocking buffer solution containing Streptavidin-Horseradish peroxidase conjugate in a 1:10000-fold dilution was prepared. The blocking buffer was replaced with the freshly prepared conjugate/blocking solution. The membrane was incubated in the conjugate/blocking buffer solution for 30 min at room temperature with gentle rocking. Following incubation, the membrane was transferred into a fresh petri dish and washed 2x with 20 mL of wash buffer (0.025 % (v/v) Tween in PBS) with gentle rocking for 5 min. The membrane was transferred into a fresh petri dish and then washed 1x with PBS solution. It was then carefully blotted with a paper towel to remove excess PBS. The transferred biotinylated dsM67 oligonucleotide DNA was subsequently stained either by ECL or by using the SIGMA FAST™ 3, 3'-Diaminobenzidine (DAB) tablets as described in section 2.2.5.2.2.

After staining and subsequent documentation, the membrane was preserved between 3 MM papers.
6.2 RESULTS

6.2.1 Electrophoretic Mobility Shift Assay (EMSA)

As previously discussed in chapters 1, activation of STAT3 via phosphorylation at the Y705 residue is required for DNA binding, thus implying that dimer formation through pY-SH2 domain interaction is essential for DNA binding. Extensive research into STAT3 activities resulted in a well accepted signalling model in which activation (i.e. phosphorylation) was considered a prerequisite for active dimer formation and DNA binding. Enormous amount of genetic and biochemical data have been generated in the past to support this signalling paradigm (Shuai et al., 1993; Shuai et al., 1994; Schaefer et al., 1997; Kisseleva et al., 2002). However, no group have reported on how activation via phosphorylation capacitates STAT3 DNA binding. To this end, conventional EMSA was adopted into a Protein Electrophoretic Mobility Shift Assay (PEMSA) to investigate the DNA binding ability of the phosphorylated and unphosphorylated STAT3ptc protein represented by the GFP-STAT3ptc construct. Furthermore, the assay was used to learn more about the conformational changes in STAT3 induced by the protein-probe interaction as well as the stability of the complex formed.

6.2.1.1 Demonstration of the differential DNA binding activities between the fluorescent STAT3ptc isoforms and the dsM67 DNA

By using either the phosphorylated or the unphosphorylated isoforms of fluorescent STAT3ptc and the biotinylated, double stranded, modified high affinity c-fos sis inducible element (biotin-dsM67 DNA) oligonucleotide DNA, we have successfully demonstrated protein-DNA binding interactions in native polyacrylamide (5%) as well as agarose gel (2%).

Following purification, fluorescent STAT3ptc isoforms (i.e. GFP-pSTAT3ptc and GFP-uSTAT3ptc) were incubated with or without increasing concentrations of the biotin-dsM67 oligonucleotide DNA at room temperature. Both sets of reactions were electrophoresed on 5 % native polyacrylamide gel at 4 °C. The gels were illuminated with blue light (~ 470 nm) and then the image captured with a Sony Cyber Shot Digital camera (Figure 6.1). The GFP tag fused to the N-terminus of STAT3 served as the “tracking device” with which the mobility shift of the GFP-STAT3ptc-dsM67 DNA complex relative to either the protein alone or dsM67 DNA alone samples. The image was quantified using the ImageJ software and the binding constant for both GFP-pSTAT3ptc/dsM67 and GFP-uSTAT3ptc/dsM67 complexes were derived from plots of GFP-STAT3ptc band intensity versus dsM67 oligonucleotide DNA concentration.
As demonstrated by Figure 6.1, the presence of the modified high affinity *c-fos sis* inducible oligonucleotide DNA sequence (dsM67) resulted in a dose-dependent shift in the mobility of the GFP-STAT3τtc/dsM67 DNA complex relative to the GFP-STAT3τtc only samples. The purified, recombinant GFP-pSTAT3τtc was able to bind to its recognition sequence to generate a significant band shift (Figure 6.1, lanes 2, 3, 4 and 5) relative to the control sample contained in lane 1. As expected, GFP-pSTAT3τtc demonstrated strong affinity and specificity to the dsM67 DNA probe with a binding constant estimated as (300± 50 nM), following quantification of the protein bands (not shown).

Although unprecedented among the STAT family of transcription factors, it was nonetheless surprising to observe binding between the unphosphorylated GFP-uSTAT3τtc isoform and the dsM67 DNA probe. The GFP-uSTAT3τtc demonstrated high affinity and specificity for the dsM67 DNA which was represented as the shift in the mobility of the GFP-uSTAT3τtc/dsM67 DNA complex (Figure 6.1, lane 8, 9 and 10) relative to the unbound GFP-uSTAT3τtc control sample in lane 6. Although the binding constant of the GFP-uSTAT3τtc was about 3-fold higher in value (i.e 900± 50 nM) relative to the active GFP-pSTAT3τtc isoform, its propensity to completely bind to the dsM67 DNA and form a stable complex was clearly evident. These observations are...
novel relative to the DNA binding ability of the inactive uSTAT3βtc. The observation is consistent with earlier held assertions that activation via tyrosine phosphorylation increases its affinity for the DNA by increasing STAT3 dimer stability (Zhang et al., 1995). Alternatively, activation via phosphorylation and subsequent dimer formation may be a prerequisite for nuclear importation but not necessary for DNA binding.

To further investigate whether the mobility shift observed for both phosphorylated and unphosphorylated GFP-STAT3βtc was indeed attributed to the complex formed with the dsM67 DNA, the GFP-STAT3βtc isoforms were incubated with or without biotin labelled dsM67 DNA and then electrophoresed on a 5% native polyacrylamide gel. Following electrophoresis, the fractionated samples including; dsM67 only, GFP-STAT3βtc only and GFP-STAT3βtc/dsM67 DNA complex, were electrophoretically transferred onto either a Hybond™ C membrane (probed with Anti-STAT3 K-15 antibody) or a Hybond™ N membrane (probed with Streptavidin-conjugated HRP) (Figure 6.2 (i), (ii) and (iii)).

![Figure 6.2: Purified, recombinant, GFP-STAT3βtc isoforms bind strongly to biotinylated dsM67 DNA](image)

Following preparations of the binding reactions containing the GFP-STAT3βtc isoforms incubated with or without biotin-dsM67 DNA, they were separated on a 5% native-polyacrylamide gel (i; protein bands illuminated at 470 nm and image captured with a Sony digital Cybershot camera) and then transferred onto either Hybond™-C Extra (ii; membrane probed and developed with the Anti-STAT3 (K-15) antibody and the DAB substrate respectively) or Hybond™-N+ membrane (iii; membrane probed and developed with the Streptavidin-HRP and the DAB substrate respectively). Lane 1, free biotinylated-dsM67 DNA; lane 2, GFP-uSTAT3βtc (unphosphorylated) only; lane 3, GFP-pSTAT3βtc (phosphorylated) only; lane 4, GFP-uSTAT3βtc in complex with biotin-dsM67 DNA; lane 5, GFP-pSTAT3βtc in complex with biotin-dsM67 DNA.
As demonstrated in Figure 6.2 (ii), the lanes containing just the GFP-STAT3βtc isoforms (lanes 2, 3, 4, 5) and those containing the GFP-STAT3βtc isoforms in complex with the biotin-dsM67 DNA were immunoreactive to the STAT3 (K-15) antibody. Also, streptavidin-HRP was used to detect the shift in mobility between the free biotin-dsM67 DNA (see Figure 6.2 (iii), lane 1) and the bound biotin-dsM67 DNA (DNA in complex with the GFP-STAT3βtc isoforms) (Figure 6.2 (iii), lane 4 and 5). As illustrated by the immunoreactive bands, the binding of the GFP-STAT3βtc isoforms to the biotin labelled dsM67 DNA probe was specific, evident by the upper streptavidin-HRP reactive band in Figure 6.2 (iii), lanes 4 and 5. The lower immunoreactive bands near the bottom of the gel represent the free biotin-dsM67 DNA (see (iii), lane 1) and the excess biotin-dsM67 DNA (see (iii) lane 4 and 5) used in the binding reaction.

6.2.1.2 Disrupting GFP-STAT3βtc/dsM67 DNA complex with a phosphotyrosyl peptide

To understand the conditions that modulate the protein/DNA interaction and to evaluate the stability of the complex in preparation for small scale screening of small molecules capable of interfering with STAT3 DNA-binding activity, an inhibition assay was developed to either inhibit or disrupt GFP-STAT3βtc/dsM67 DNA complex using a number of STAT3-binding phosphotyrosyl peptides. Hence, the phosphotyrosyl peptides including; H-pY-LKTKFI-NH$_2$, derived from the native C-terminal region of the STAT3 SH2 domain binding sequence (Becker et al., 2000; Park et al., 2000; Turkson et al., 2001) and H-pY-LPQTV-NH$_2$, derived from the interleukin-6 receptor subunit, gp130 (Ren et al., 2003 Stahl et al., 1995) were initially examined in the assay. The prediction was that, the pY (phosphorylated tyrosine residue) group contained in both phosphotyrosyl peptide H-pY-LKTKFI-NH$_2$ or H-pY-LPQTV-NH$_2$ would bind to the SH2 domain of the GFP-STAT3βtc isoforms through the conventional pY-SH2 interaction and therefore either inhibit the formation or disrupt the GFP-STAT3βtc/dsM67 DNA complex thereby reversing the mobility shift of the protein/DNA complex back to the unbound state (GFP-STAT3βtc only). To investigate the hypothesis, purified, recombinant, GFP-pSTAT3βtc isoforms was premixed and pre-incubated with or without different concentration of the phosphotyrosyl peptides inhibitor, H-pY-LPQTV-NH$_2$, prior to incubation with biotin-dsM67 DNA probe for EMSA. Similarly, purified, recombinant, GFP-uSTAT3βtc was also pre-mixed and pre-incubated with different concentrations of either of the following phosphotyrosyl peptides; H-pY-LPQTV-NH$_2$ or H-pY-LKTKFI-NH$_2$, prior to incubation with biotin-dsM67 DNA probe for EMSA. As demonstrated in Figure 6.3 (A), (B), (C), the reaction products were clearly fractionated on a 5 % native-polyacrylamide gel.
Figure 6.3: Disruption of GFP-STAT3βtc/dsM67 DNA-binding by a peptide

The purified, recombinant, GFP-pSTAT3βtc and GFP-uSTAT3βtc were preincubated with different concentrations of the phosphotyrosyl peptide for 1 h prior to incubation with the dsM67 DNA probe and separated by native-acrylamide gel electrophoresis. (A), GFP-pSTAT3βtc was preincubated with 1 to 40 μM H-pY-LPQTV-NH₂ (lanes 2 and 4 to 8) or without peptide (lanes 1 and 3). (B), GFP-uSTAT3βtc was preincubated with 1 to 40 μM H-pY-LKTKFI-NH₂ (lanes 2 and 4 to 8) or without peptide (lanes 1 and 3). (C), GFP-uSTAT3βtc was preincubated with 1 to 40 μM H-pY-LKTKFI-NH₂ (lanes 2 and 4 to 8) or without peptide (lanes 1 and 3). Lanes 3 to 8 contained samples incubated with the dsM67 DNA probe. The arrows indicate the positions of the unbound and bound GFP-STAT3βtc protein samples in the gel.

It was obvious that the purified, recombinant, GFP-STAT3βtc isoforms analysed were functional, demonstrated by the complete shift in the mobility of protein/DNA (GFP-STAT3βtc/dsM67 DNA) complex, lane 3, relative to the unbound control sample contained in lane 1.

The presence of the H-pY-LPQTV-NH₂ peptide inhibitor at increasing concentrations did not result in a decrease in the level of GFP-pSTAT3βtc bound to the dsM67 DNA, suggesting that the peptide inhibitor may not be capable of disrupting either GFP-pSTAT3βtc dimer or the GFP-pSTAT3βtc/dsM67 DNA complex or could not inhibit the active GFP-pSTAT3βtc dimer from binding to the DNA, at concentrations up to 40 μM.
Edwin Nkansah

(i.e. an ~ 13-fold increased concentration relative to either the GFP-pSTAT3βtc or dsM67). In a subsequent experiment, DNA binding could still be observed when the peptide inhibitor was used in a 333-fold molar excess, thus demonstrating the strong affinity of the phosphorylated GFP-pSTAT3βtc protein for the DNA probe. No obvious shift in the mobility of the GFP-pSTAT3βtc incubated with peptide inhibitor alone was observed (lane 2).

In contrast, pre-incubating GFP-uSTAT3βtc with the H-pY-LPQTV-NH₂ peptide (B) resulted in a slight dose-dependent decrease in the amount of the GFP-uSTAT3βtc protein binding to the dsM67 DNA probe. This was evident in lane 8 of Figure 6.3 (B), suggesting that, the excess phosphotyrosyl peptide targeted to the SH2 domain of uSTAT3βtc specifically blocked the GFP-uSTAT3βtc dimer formation and therefore DNA binding. The affinity of the phosphotyrosyl peptide to SH2 domain of the GFP-uSTAT3βtc protein was estimated by the intensity the GFP fluorescence in lane 8 relative to that in lane 3.

In another experiment, the GFP-uSTAT3βtc was pre-incubated with the pY-LKTKFI-NH₂ peptide prior to incubation with the dsM67 DNA probe, Figure 6.3 (C). No obvious inhibition of the GFP-uSTAT3βtc binding to DNA was apparent, up to concentration of 40 μM (lane 8). Even at 200 μM, the STAT3 derived PY-LKTKFI-NH₂ did not produce any significant inhibition of the GFP-uSTAT3βtc DNA binding activity, thus suggesting that the inhibitory potency of the peptide was weak relative to that of the gp130 derivative, H-pY-LPQTV-NH₂.

Even though the H-pY-LPQTV-NH₂ inhibitor demonstrated greater inhibitory or disruptive power in dissociating the GFP-uSTAT3βtc/dsM67 DNA complex, it showed no significant activity towards disrupting the active GFP-pSTAT3βtc dimer either in the presence or the absence of the dsM67 oligonucleotide DNA probe. No activity was observed even after prolonged incubation, up to 22 h, an indication of the greater stability of the GFP-pSTAT3βtc dimer and GFP-pSTAT3βtc/dM67 DNA complex. However, in relation to GFP-uSTAT3βtc protein, the data suggested that the H-pY-LPQTV-NH₂ peptide was essential for the reduction in the GFP-uSTAT3βtc DNA binding activity, though, no similar effect was observed for the H-pY-LKTKFI-NH₂ peptide. The observation also indicated the significance of the peptide sequence surrounding the phosphorylated tyrosine (pY; modified amino acid).
6.2.1.3 Investigating the order of addition of the components for inhibition of the GFP-STAT3βtc/dsM67 DNA complex formation

Building on the observations in section 6.2.1.2, it was investigated whether the order of addition of the phosphotyrosyl peptide probe relative to the dsM67 DNA probe would influence the inhibitory potency of the phosphotyrosyl peptide inhibitor. The gp130 derived phosphotyrosyl peptide was used, since it generated the best inhibitory effect. Reactions were incubated for either 2 or 22 h. The first set of reactions was prepared by premixing the purified, recombinant, GFP-uSTAT3βtc with or without the c/sM67 DNA probe and then incubated for either 2 or 22 h prior to incubation with the phosphotyrosyl peptide at two different concentrations. The second set of reactions contained the GFP-uSTAT3βtc protein premixed in reverse i.e. with or without different concentrations of the phosphotyrosyl peptide for 2 or 22 h, prior to incubation with the dsM67 oligonucleotide DNA probe. Both sets of reaction were analyzed on native acrylamide gels. The gel was illuminated and the image captured after electrophoresis (Figure 6.4).

The shift in mobility of the GFP-uSTAT3βtc/dsM67 DNA complex (see lane 3; upper band) relative to the GFP-uSTAT3βtc protein (see lane 1; upper band) confirmed that the purified, recombinant, GFP-uSTAT3βtc protein was functional. In addition, the shift further demonstrates the stability of the GFP-uSTAT3βtc/c/sM67 DNA complex over time since the shift could be observed with samples incubated for over 24 h at 4°C. In addition to corroborating the strong specific interaction between the GFP-uSTAT3βtc and the dsM67 DNA probe, the stability observed will prove useful for an assay suitable for screening small molecules.

The diffused upper band in lanes 5 and 7 were from samples premixed with 100 μM and 200 μM i.e., 33-fold and 66-fold molar excess, H-pY-LPQTV-NH₂ peptide respectively, prior to incubation with the dsM67 DNA. The diffuse continuum of the bands, relative to the compact band in lane suggested inhibition of the DNA-binding activity of the GFP-uSTAT3βtc by the phosphotyrosyl peptide. Though the band shift was not so distinctly reversed (i.e. a shift to the GFP-uSTAT3βtc unbound state displayed by the upper band in lane 1), the inhibitory effect of the peptide was evident in the partially shifted bands.
Figure 6.4: Peptide probe, H-pY-LPQTV-NH₂, disrupts DNA binding activity of GFP-uSTAT3βtc.

Recombinant, GFP-uSTAT3βtc was premixed either with the phosphotyrosyl peptide or the dsM67 DNA probe, for either 2 h or 22 h prior to incubation for a further 22 h or 2 h with either the dsM67 DNA probe or the phosphotyrosyl peptide inhibitor respectively. The samples were then electrophoresed on a 5% native polyacrylamide gel. Lane 1, contains GFP-uSTAT3βtc only; lane 2, contains GFP-uSTAT3βtc premixed with the 200 μM phosphotyrosyl peptide; lane 3, contains GFP-uSTAT3βtc premixed with dsM67 DNA probe; lane 4, contains GFP-uSTAT3βtc premixed with the dsM67 DNA probe prior to incubation with 100 μM phosphotyrosyl peptide; lane 5, contains GFP-uSTAT3βtc premixed with 100 μM phosphotyrosyl peptide prior to incubation with dsM67 DNA probe; lane 6, contains GFP-uSTAT3βtc premixed with the dsM67 DNA probe prior to incubation with 200 μM of the phosphotyrosyl peptide; lane 7, contains GFP-uSTAT3βtc premixed with 200 μM of the phosphotyrosyl peptide prior to incubation with the dsM67 DNA. The arrows indicate the identity and location of the separated samples.

In contrast to lanes 5 and 7, lanes 4 and 6 contained the purified GFP-uSTAT3βtc premixed with the dsM67 DNA probe, prior to incubation with either 100 μM or 200 μM phosphotyrosyl peptide inhibitor. This order of incubation generated a significant dose-dependent shift in the mobility of the GFP-uSTAT3βtc, thus suggesting that the conformation assumed by the GFP-uSTAT3βtc in complex with the dsM67 DNA may have provided alternative access for the peptide to the GFP-uSTAT3βtc SH2 domain. The peptide/SH2 domain interaction may have destabilized the preformed STAT3/DNA complex and forced their dissociation giving an inactive GFP-uSTAT3βtc/peptide heterocomplex. However, since there was no obvious band shift demonstrated by the GFP-uSTAT3βtc/peptide complex (i.e. upper band in lane 2) compared to the unbound GFP-uSTAT3βtc, lane 1, it is not clear whether the peptide remains bound to STAT after dissociation from the DNA. This could be tested with biotinylated peptide detected with streptavidin-HRP.
6.2.1.4 Investigating the inhibitory activity of the phosphotyrosyl peptide inhibitors in agarose gels

The assays were fractionated on a 2% agarose gel by electrophoresis. Unlike the 5% polyacrylamide gels, the resolution of the protein samples on 2% agarose gels was much better in terms of a greater mobility shift, but the bands were more diffuse. Both isoforms of GFP-STAT3βtc were used in the assay. Phosphotyrosyl peptide inhibitors including: H-pY-LPQTV-NH₂, Acetyl-pY-LPQ and H-pY-LKTKFI-NH₂, were assayed for their ability to disrupt DNA-binding of GFP-STAT3βtc isoforms.

Even though the potency of the gp130 phosphotyrosyl peptide had been demonstrated above in relation to the unphosphorylated isoforms, it was also important to investigate other peptides as inhibitors of the DNA binding activity of active GFP-STAT3βtc. Hence, on the basis of studies conducted by Turkson and colleagues, the Y-1 position of the gp130 derived phosphotyrosyl peptide probe was investigated. The acetylated peptide, Acetyl-pY-LPQ-NH₂, was synthesised and assayed with the GFP-STAT3βtc isoforms to evaluate its inhibitory properties.

The purified, recombinant GFP-STAT3βtc protein was premixed with or without the dsM67 DNA probe, prior to incubation with different concentrations of peptides, H-pY-LPQTV-NH₂, Acetyl-pY-LPQ-NH₂ or H-pY-LKTKFI-NH₂.

Similarly, purified recombinant, GFP-pSTAT3βtc protein was premixed with or without the dsM67 DNA probe, prior to incubation with the different concentrations of peptides, H-pY-LPQTV-NH₂, Acetyl-pY-LPQ or H-pY-LKTKFI-NH₂. Both sets of reactions were analyzed by EMSA (Figures 6.5, 6.6 and 6.7) to demonstrate the shift in mobility of the GFP-STAT3βtc isoforms as a result of the targeted disruption of the protein dimer and therefore impede DNA binding.
Figure 6.5: Demonstration of the disruptive effect of the H-pY-LPQTV-NH₂ peptide on the pre-formed GFP-STAT3βtc/dsM67 DNA complex

The purified, recombinant GFP-STAT3βtc isoforms were preincubated with or without dsM67 DNA probe, prior to incubation with different concentrations of the H-pY-LPQTV-NH₂ peptide. The samples were separated on a 2% agarose gel by electrophoresis. Lanes 1 to 10 contain GFP-uSTAT3βtc preincubated with or without dsM67 DNA probe prior to incubation with or without 3 to 500 µM H-pY-LPQTV-NH₂ phosphotyrosyl peptide; lanes 11 to 19 contain GFP-pSTAT3βtc preincubated with or without the dsM67 DNA prior to incubation with or without 3 to 1000 µM H-pY-LPQTV-NH₂ phosphotyrosyl peptide. The positions and identity of the components and complexes are indicated with an arrow.

As expected, both isoforms of the GFP-STAT3βtc were functional, demonstrated by the mobility shift in lanes 3 and 13 for GFP-uSTAT3βtc and GFP-pSTAT3βtc respectively. The recombinant proteins were functional in all 3 assays demonstrated in lane 3 of Figures 6.6 and 6.7. Further, the presence of the phosphotyrosyl peptide inhibitor in lanes 4 to 10 of Figures 6.5 and 6.6 resulted in a dose-dependent decrease in the proportion of GFP-uSTAT3βtc bound to the dsM67 DNA probe and therefore a simultaneous shift of the of the GFP-uSTAT3βtc to the unbound state. The observation was consistent with reports that the sequence, H-pY-LPQ, specifically interact with the SH2 domain of the GFP-uSTAT3βtc chimaeric protein. The disruptive effect, however, yielded a near identical DB₅₀ value (i.e. 30±15 µM) for both phosphotyrosyl peptides (i.e. H-pY-LPQTV-NH₂ and Acetyl-pY-LPQ-NH₂) assayed with GFP-uSTAT3βtc (Figures 6.5 and 6.6, lanes 1 to 10) thus, suggesting that the type of group occupying the Y-1 position may not be essential in disrupting unphosphorylated GFP-uSTAT3βtc/dsM67 DNA complex.
Figure 6.6: Demonstration of the disruptive effect of the Ac-pY-LPQ-NH$_2$ peptide on the preformed GFP-STAT3ptc/dsM67 DNA complex

The purified, recombinant GFP-STAT3ptc isoforms were preincubated with or without dsM67 DNA probe, prior to incubation with different concentrations of the Ac-pY-LPQ-NH$_2$ peptide. The samples were separated on a 2 % agarose gel via electrophoresis. Lanes 1 to 10 contains GFP-uSTAT3ptc preincubated with or without dsM67 DNA probe prior to incubation with or without 3 to 500 pM Ac-pY-LPQ-NH$_2$ phosphotyrosyl peptide; lanes 11 to 19 contains GFP-pSTAT3ptc preincubated with or without the dsM67 DNA prior to incubation with or without 3 to 1000 pM Ac-pY-LPQ-NH$_2$ phosphotyrosyl peptide. The positions and identity of the samples are indicated with an arrow.

Unlike the H-pY-LPQTV-NH$_2$, the acetylated derivative showed some dose-dependent reduction in the level of GFP-pSTAT3ptc bound to the dsM67 DNA at concentrations, up greater than 150 pM (Figure 6.6).

Though the STAT3-derived phosphotyrosyl peptide containing the H-pY-LKTKFI-NH$_2$ motif has been reported to bind to the STAT3 SH2 domain and inhibit dimerization and therefore DNA-binding (Turkson et al., 2001), no significant inhibition effect was observed upon incubation of the pre-formed complex with the different concentrations of the phosphotyrosyl peptide. The observation suggested that the peptide may be binding to the GFP-uSTAT3ptc SH2 domain with a very low affinity and therefore would require high concentrations of the peptide (i.e. at concentrations > 500 pM) to effect the disruption of the protein/DNA complex and give the signature band shift (lanes 7 to 10 of Figure 6.7). Further, these observations agreed with findings from other laboratories where a fluorescein labelled peptide was used in a Fluorescence Polarization (FP) assays to demonstrate STAT3 binding and inhibition. Very weak binding was observed even at significantly higher concentrations (Schust et al., 2004).
Figure 6.7: Demonstration of the disruptive effect of the H-pY-LKTKFI-NH\textsubscript{2} peptide on the preformed GFP-STAT3\textbeta tc/dsM67 DNA complex

The purified, recombinant GFP-STAT3\textbeta tc isoforms were preincubated with or without dsM67 DNA probe, prior to incubation with the different concentrations of the H-pY-LKTKFI-NH\textsubscript{2} peptide. The samples were separated on a 2 % agarose gel via electrophoresis. Lanes 1 to 10 contains GFP-uSTAT3\textbeta tc preincubated with or without dsM67 DNA probe prior to incubation with or without 3 to 500 pM H-pY-LKTKFI-NH\textsubscript{2} phosphotyrosyl peptide; lanes 11 to 19 contains GFP-pSTAT3\textbeta tc preincubated with or without the dsM67 DNA prior to incubation with or without 3 to 1000 pM H-pY-LKTKFI-NH\textsubscript{2} phosphotyrosyl peptide. The positions and identity of the samples are indicated with an arrow.

Unlike in the GFP-uSTAT3\textbeta tc, no significant inhibitory activity was observed when the phosphotyrosyl peptides (lanes 11 to 19 of Figures 6.5, 6.6 and 6.7) were assayed for their ability to disrupt the active GFP-pSTAT3\textbeta tc dimer and GFP-pSTAT3\textbeta tc/dsM67 DNA complex. However, a decreasing intensity of the shifted band representing the GFP-pSTAT3\textbeta tc/dsM67 DNA complex incubated with 150, 300, 500 and 1000 pM Acetyl-pY-LPQ-NH\textsubscript{2} phosphotyrosyl peptide was observed (lanes 16, 17, 18 and 19 of Figure 6.6). This suggested that the acetylation of STAT3 docking motif may have significantly enhanced the affinity of the peptide for SH2 domain of the GFP-pSTAT3\textbeta tc dimer and therefore partially disrupt the dimer and subsequently reduce its affinity for the dsM67 oligonucleotide DNA. Also, the observation that the corresponding hexapeptide bearing a free N-terminus (i.e. H-pY-LPQTV-NH\textsubscript{2}) assayed in Figure 6.5 lanes 14 to 19, produced no significant disruption of the GFP-pSTAT3\textbeta tc dimer corroborates earlier findings that the free Y-1 position (i.e. the N-terminus of the phosphorylated tyrosine residue) results in decreased affinity of the phosphotyrosyl peptide for the GFP-pSTAT3\textbeta tc SH2 domain.
6.2.1.5 Comparative inhibition of the DNA-binding activity of GFP-uSTAT3βtc using either peptide or non-peptide small molecule inhibitors

In Figures 6.5, 6.6 and 6.7, electrophoretic mobility shift assays were used to demonstrate how peptide-based STAT3 inhibitors designed to target the STAT3-SH2 domain effectively blocked dimerization and therefore the DNA-binding activity of the purified, recombinant, GFP-STAT3βtc isofoms, especially, GFP-uSTAT3βtc in vitro. Non-peptide, drug-like small molecules are known to bind to the STAT3-SH2 domain and directly block STAT3 dimer formation and DNA-binding activity. These have also been tested to demonstrate their potency in vitro. In addition, to the phosphotyrosyl peptide probe, H-pY-LPQTV-NH₂, two other non-peptide small molecule inhibitors of STAT3; STA-21 (known to block constitutive STAT3 signalling by impeding STAT3 DNA-binding and dimerization as well as STAT3 dependent luciferase activity, Song et al., 2005) and RH06 (synthesized in-house, known to inhibit STAT3 signalling) were assayed for their ability to impede GFP-uSTAT3βtc DNA-binding using EMSA.

Following purification, recombinant, GFP-uSTAT3βtc was preincubated (1 h) with or without dsM67 DNA (3 μM) prior to incubation (1 or 16 h) with different concentrations of either H-pY-LPQTV-NH₂, or RH06 or STA-21. After incubation, the reaction mixture where electrophoresed on a 2 % agarose gel, Figure 6.8.

![Figure 6.8: Comparative inhibitory of GFP-STAT3βtc/dsM67 DNA complex using a peptide- or a non-peptide-based inhibitors probes](image)

After purification, recombinant GFP-uSTAT3βtc was mixed with or without dsM67 oligonucleotide DNA prior to incubation with increasing concentrations (up to 1000 μM) of either H-pY-LPQTV-NH₂, or RH06 or STA-21. Upon incubation, the reactions were electrophoresed on a 2% agarose gel by electrophoresis. Lane 1, contains recombinant, GFP-uSTAT3βtc chimaeric protein; lane 2, contains recombinant, GFP-
uSTAT3βtc incubated with the hSIE (dsM67) DNA; lanes 3-8, contains recombinant, GFP-uSTAT3βtc incubated with the hSIE (dsM67) DNA prior to incubation with different concentrations of the H-pY-LPQTV-NH$_2$ phosphotyrosyl peptide; lanes 9-14, contains recombinant, GFP-uSTAT3βtc incubated with the hSIE (dsM67) DNA prior to incubation with different concentrations of the RH06 non-peptide inhibitor; lanes 15-20, contains recombinant, GFP-uSTAT3βtc incubated with the hSIE (dsM67) DNA prior to incubation with different concentrations of the STA-21 non-peptide inhibitor. The positions and identities of the protein/protein-DNA/protein-inhibitor bands are indicated with arrows.

This suggested that the concentration of the phosphotyrosyl peptide inhibitor corresponding to the 50 % reduction in the GFP-uSTAT3βtc DNA-binding activity (DB$_{50}$) was approximately 30 μM. A complete mobility shift to the unbound state was observed following exposure to 100 μM of the phosphotyrosyl peptide. In contrast, incubation of the pre-formed GFP-uSTAT3βtc/dsM67 DNA complex with increasing concentrations of the non-peptide RH06 small molecule inhibitor resulted in no inhibitory effect, up to concentrations of 1000 μM (Figure 6.8, lanes 9-14). Complete shift representing the protein/DNA complex was evident even with 1000 μM small molecule inhibitor. On the other hand, the STA-21 inhibitor, 500 μM, generated a reduction in the level of GFP-uSTAT3βtc in complex with the dsM67 DNA probe (Figure 6.8, lanes 15-20). This observation implied that the STA-21 may be binding to the proposed binding pocket and disrupting GFP-uSTAT3βtc/dsM67 DNA complex at significantly higher concentrations, up to 1000 μM (lanes 19 and 20) and therefore results in a band shift. Moreover, the concentration of the small molecule inhibitor corresponding to 50% decrease in GFP-uSTAT3βtc DNA-binding activity (DB$_{50}$) was recorded as approximately 500 μM. On the contrary, however, no reduction in the level of DNA bound GFP-uSTAT3βtc was observed following incubation at room temperature for 2 h.

6.3 Discussion

The prime objective of the work described in this chapter was the application of the purified, recombinant GFP-STAT3βtc isoforms in a novel protein mobility shift assay which will now be referred to as PEMSA (Protein Electrophoretic Mobility Shift Assay). Significantly, PEMSA was used to demonstrate the DNA binding functionality of the GFP-STAT3βtc isoforms represented as a significant mobility shift relative to the unbound samples. In addition, PEMSA was used to demonstrate that the surrogate receptor phosphotyrosyl peptide, H-pYLQTV-NH$_2$, and its truncated acetylated derivative, Ac-pYLQ-NH$_2$ significantly inhibits the DNA-binding activity of GFP-uSTAT3βtc chimaeric protein. Relevant binding constants and DB$_{50}$ values were derived from the PEMSA assays.
6.3.1 Protein Electrophoretic Mobility Shift Assay (PEMSA)

PEMSA is a simple and a sensitive method for demonstrating sequence-specific DNA binding activity of fluorescent tagged DNA binding protein, e.g. STAT3\textsubscript{βtc}. The assay was applied to demonstrate the DNA-binding capabilities of both the active and inactive GFP-STAT3\textsubscript{βtc} isoforms as demonstrated by a significant mobility shift. Further, it was also employed to demonstrate dissociation of the GFP-STAT3\textsubscript{βtc}/DNA complex, as demonstrated by the mobility shift of the unbound protein as judged by the fluorescent reporter. Unlike conventional EMSA, PEMSA is based on the protein mobility shift between the unbound and the DNA bound GFP-STAT3\textsubscript{βtc} preparations. The mobility shift between the unbound (slow migrating) and the DNA bound GFP-STAT3\textsubscript{βtc} (fast migrating) was principally due to the conformational change induced by the bound high affinity SIE oligonucleotide DNA (M67) and the increased negative charge due to the phosphate backbone of the DNA thus leading to a differential migration between the unbound and the bound GFP-STAT3\textsubscript{βtc} through the gel matrix.

The fractionation system used was either a non-denaturing polyacrylamide gel or an agarose gel. It was assumed that since the hydrodynamic sizes of the unbound protein and the DNA bound complex varied considerably, it would allow the differential fractionation of both samples (i.e. DNA bound or unbound), though the excess negative charges provided by the bound DNA may influence the rate of migration of both samples.

6.3.2 Differential DNA binding activity of the phosphorylated and the unphosphorylated recombinant, GFP-STAT3\textsubscript{βtc} isoforms

PEMSA was used to demonstrate differential migration and DNA binding activity the of the purified, phosphorylated and unphosphorylated GFP-STAT3\textsubscript{βtc} protein. Comparing the binding properties of the two isoforms revealed a marked quantitative difference in the DNA binding activities of the phosphorylated and unphosphorylated fluorescent STAT3\textsubscript{βtc} isoforms. Though activation via tyrosine phosphorylation is commonly reported to be essential for the DNA-binding activity of STAT3 (Shuai et al., 1994), it was demonstrated in this chapter that the unphosphorylated isoform of STAT3 represented by the GFP-\textsuperscript{u}STAT3\textsubscript{βtc} bound to and formed a stable complex with the high affinity SIE DNA oligonucleotide probe (dsM67 DNA). In addition, it was demonstrated that the affinity of the protein to the dsM67 DNA was approximately 3x less (i.e. \(\sim 0.9\ \mu M\)) compared to that observed for the corresponding phosphorylated isoform, \(\sim 0.3\ \mu M\) (Figure 6.1).
The differences in the DNA binding activity between the active and inactive GFP-STAT3ptc isoforms molecules may be due to the dimer stability conferred on the active isoform by the reciprocal phosphotyrosyl-SH2 domain interaction. Even though the high resolution structures of the DNA bound STAT1 and STAT3 (Chen et al., 1998; Becker et al., 1998) suggest that the mutual binding of the individual STAT molecules to the DNA sequence was possible even without the phosphotyrosyl/SH2 domain interactions (the structure suggests that each STAT molecule of the DNA bound STAT dimer makes contact to only half of the palindromic DNA sequence), the activated form of the proteins were reported to show considerably greater DNA specificity and dimer stability in vivo (Zhang et al., 1995). This is consistent with the observation in Figure 6.1, where purified, recombinant, GFP-uSTAT3ptc chimaeric protein formed a stable complex with dsM67 DNA probe, albeit with a lower binding constant relative to the phosphorylated GFP-pSTAT3ptc isoform. Consistent with this observation was the report that unphosphorylated STAT1 (monomeric) binds to its recognition sequence by contacting one half of the palindromic GAS DNA sequence and that the affinity of the uSTAT1 monomer or dimer to its recognition sequence was much less than that for the pSTAT1 homodimer (Chatterjee et al., 2000).

Recognising the protein mobility shift generated as a result of the GFP-STAT3ptc/dsM67 DNA complex formation, the specificity of both GFP-STAT3ptc isoforms to the high affinity dsM67 DNA recognition sequence was demonstrated by southern blot by using biotin labelled dsM67 DNA probe (Figure 6.2). Immunoreactive bands generated with the streptavidin-HRP confirmed the specificity of both GFP-STAT3ptc isoforms to the cis-recognition sequence. Likewise, an immunoreactive band generated with the STAT3 K-15 antibody also confirmed the presence of the STAT3 transcription factor in the fluorescent protein constructs. Furthermore, other confirmatory studies were performed to demonstrate the sequence specificity of the GFP-uSTAT3ptc isoforms. Thus, by incubating the recombinant protein with 100-fold excess of a randomly chosen oligonucleotide DNA sequence and then analysed by PEMSA, no obvious mobility shift was observed for both treated isoforms. Thus, suggesting that no significant interaction occurs between the recombinant protein and either a specific unrelated DNA sequence or a non-specific (sonicated salmon sperm) DNA sequence.

6.3.3 Disruption of the DNA binding activity of the purified, recombinant GFP-STAT3ptc by phosphotyrosyl peptide

As an essential step in STAT3 activation, dimer formation between two STAT molecules presents an attractive target to impede STAT3 DNA-binding and subsequent
transcriptional activity (Shuai et al., 1994; Turkson et al., 2001; Turkson et al., 2004). A phosphotyrosyl-based peptidomimetics approach to either block or disrupt GFP-STAT3βtc dimer formation and DNA binding in vitro was explored, both for proof-of-concept testing as well as to develop PEMSA as general method and specifically for the identification of leads for STAT3 drug discovery.

The binding activities of purified, recombinant, GFP-uSTAT3βtc and GFP-pSTAT3βtc were assayed with phosphotyrosyl peptides such as; H-pY-LPQTV-NH$_2$ (Stahl et al., 1995) (derived from the interleukin-6 receptor subunit gp130) or Ac-pY-LPQ-NH$_2$ (acetylated tripeptide derivative of H-pY-LPQTV-NH$_2$) or H-pY-LKTKFI-NH$_2$ (derived from the core C-terminal STAT3 SH2 domain sequence surrounding the Y$^{705}$ residue), up to concentrations of 500 and 1000 μM for GFP-uSTAT3βtc and GFP-pSTAT3βtc respectively. The relative disruptive effect of the peptide, H-pY-LPQTV-NH$_2$, was also compared with the non-peptide, drug-like small molecules including; RH06 and STA-21, up to concentrations of 1000 μM.

Though some degree of inhibition was reported using the classical inhibition paradigm for EMSA, that is pre-incubating GFP-uSTAT3βtc with the phosphotyrosyl peptide inhibitor prior to incubation with the dsM67 DNA oligonucleotide (Turkson et al., 2001; Park, et al., 2000; Song et al., 2004), a much more significant improvement in the inhibitory potency of the phosphotyrosyl peptide was achieved for PEMSA when the GFP-uSTAT3βtc was preincubated with the dsM67 DNA prior to incubation with the phosphotyrosyl peptide. This order of incubation, though unconventional for normal EMSA was applied and used to generate most of the data discussed in this chapter (Figure 6.3).

Even though Turkson et al reported the inhibition of active STAT3α activity using the STAT3 derived phosphotyrosyl peptide, H-pY-LKTKFI-NH$_2$, we were unable to demonstrate his finding with our recombinant protein construct which contains the beta (β) splice variant rather than the STAT3α used in his study (Figure 6.7). Our inability to reproduce the inhibition data reported by Turkson and colleagues using PEMSA was most likely due to the STAT3 variant used in this study. Since the STAT3 (i.e. STAT3α and STAT3β) isoforms possess the same DNA-binding domain, they would be expected to have similar DNA-binding activities. However, several reports have demonstrated a significant difference in the DNA binding activity of the STAT3 isoforms in vitro, i.e., STAT3α and STAT3β. The reports suggest that the active STAT3β (which has the C-terminal 55 amino acid residues of the STAT3α replaced with a unique 7 amino acid residue sequence) shows a significantly greater DNA-binding affinity than
the full length STAT3α (Schaefer et al., 1995; Park et al., 1996; Schaefer et al., 1997). The enhanced affinity is attributed to the dimer stability of the STAT3β isoform which is purported to be due to the absence of the C-terminal region unique to the STAT3α. Park and colleagues demonstrated that the C-terminal region confers a negative effect on dimer stability and concomitantly, the DNA binding activity of the STAT3α isoform. These, they argued were as a result of the net negative charge conferred by the C-terminal 55 amino acid residues (i.e. 8 Asp and Glu versus 1 Arg; net charge −7), which was likely to either exert electrostatic repulsion or induce conformational changes at the C-terminal region of the STAT3α isoform, thereby, compromising the stability of the active STAT3α dimer and its DNA binding activity. Thus, it was concluded that the striking difference in dimer stability between the STAT3α and the STAT3β isoforms contributed significantly to the susceptibility of the active -3α isoform to disruption by the phosphotyrosyl peptide inhibitor used by Turkson and colleagues (Turkson et al., 2001). In addition to the extraordinary dimer stability of the STAT3β isoform, the poor inhibitory activity demonstrated by the H-pY-LKTKFI-NH₂ phosphotyrosyl peptide was in agreement with reports by other groups that suggested that the peptide binds to the STAT3 SH2 domain with low affinity and that may only show activity only at significantly higher concentrations (Schust et al., 2004). Moreover, other supposedly higher affinity phosphotyrosyl peptides, assayed with the activated GFP-pSTAT3βtc/dsM67 DNA complex showed no significant decrease in the amount of protein bound to the DNA.

Consequently, other peptide probes derived from the STAT3 docking site of the gp130 receptor, H-pY-LPQTV-NH₂ (Y904), and its acetylated N-terminus derivative, Acetyl-pY-LPQ-NH₂ were also investigated. Reports suggested that relative to the former phosphotyrosyl peptide described above, the derivatives of the gp130 subunit showed significantly higher affinity (133-fold) for the STAT3 SH2 domain (Ren et al., 2003). However, no inhibition of the DNA binding activity of the activated GFP-pSTAT3βtc construct was observed. On the contrary, a dose-dependent decrease in the level of GFP-uSTAT3βtc bound to the dsM67 DNA probe was observed when the gp130 derived phosphotyrosyl peptide derivatives were assayed for their ability to inhibit the binding activity of the unphosphorylated GFP-STAT3βtc isoform (Figure 6.5 and 6.6). Near complete inhibition was demonstrated via the mobility shift of the GFP tagged STAT3βtc from the bound to the unbound state when assayed with approximately 150 μM of the phosphotyrosyl peptide. The DB₅₀ value for both the H-pY-LPQTV-NH₂ and its acetylated tripeptide derivative was determined as 30±15 μM.
The inhibitory activity of the H-pY-LPQTV-NH$_2$ was compared with that of the non-peptide, cell permeable, drug-like small molecule inhibitors such as RH06 and STA-21. Unlike RH06, which was still in the developmental stages at the time of writing, STA-21 had been identified and applied in a number of assays as a control compound. It was identified through computational screening of potential small molecule inhibitors of STAT3. Though Song and colleagues reported that STA-21 significantly inhibited the STAT3 biological functions including STAT3 dimerization, DNA binding and nuclear translocation as well as STAT3 regulated genes such as Bcl-xL and cyclin D1, it showed slight inhibition of GFP-uSTAT3βtc DNA-binding activity. The small molecule (STA-21) did demonstrate a dose-dependent disruption of the GFP-STAT3βtc/dsM67 DNA complex, up to concentrations of 1000 μM after 2 h incubation. Moreover, an increase in percentage inhibition by 5-fold was observed following incubation at 4 °C for 16 h (Figure 6.8). This observation suggested that the inhibitory effect might be induced by a slow conformational change in the protein molecule following interaction with the inhibitor. Upon binding to GFP-uSTAT3βtc, most probably close to the SH2 domain, it disrupts the GFP-uSTAT3βtc dimer and therefore inhibits DNA binding. It must be noted that the binding pocket of the STA-21 is currently unproven.

Following the demonstration in this chapter that the gp130 derived phosphotyrosyl peptide significantly inhibits unphosphorylated GFP-uSTAT3βtc's DNA-binding activities in vitro, we attempted to suggest a model for the interactions between the gp130 derived phosphotyrosyl peptide and the GFP-uSTAT3βtc protein molecules or the GFP-uSTAT3βtc/dsM67 DNA complex. Though by SEC, we demonstrated that GFP-uSTAT3βtc predominantly exists as monomers, we speculated that the presence of the high affinity SIE oligonucleotide DNA may encourage the formation of inactive GFP-uSTAT3βtc dimers and therefore the formation of the GFP-uSTAT3βtc/dsM67 DNA complex. By introducing the higher affinity gp130 phosphotyrosyl peptide targeted to the GFP-uSTAT3βtc SH2 domain, the pY-SH2 interaction may induce a conformational change in the protein leading to the disruption of the inactive dimer and the subsequent release of the dsM67 DNA target. The disruption of the protein/DNA complex may result in either the formation of GFP-uSTAT3βtc/phosphotyrosyl peptide heterocomplex or GFP-uSTAT3βtc/GFP-uSTAT3βtc (dimer) or GFP-uSTAT3βtc (monomers) with a concomitant mobility shift. An essential feature of this model is the role of the phosphorylated tyrosine - Src Homology 2 domain (pY-SH2) interaction, since the disruptive effect of all the gp130 derived peptide inhibitors required the tyrosine to be phosphorylated. Though this was not verified with an unphosphorylated peptide probe, there was enough evidence (data shown in chapter 5; ELISA) to suggest that the disruptive effect of the phosphotyrosyl peptide was indeed mediated
by the pY motif. The observations were consistent with findings reported by groups that characterized the interaction between the modified tyrosine residue (pY) and the SH2 domain of STAT through phospho-peptidomimetic probes (Shuai et al., 1994; Zvelebil et al., 1995).

These findings demonstrate that unlike the STAT3 derived phosphotyrosyl peptides, the gp130 derived peptide including tetrapeptides bind to the SH2 domain contained within the unphosphorylated GFP-uSTAT3beta protein construct with higher affinity and a favourable conformation, and disrupts the GFP-uSTAT3beta/dsM67 DNA complex. The essential peptide sequence required to produce the specific interactions with the STAT3 SH2 domain can be reduced to a tripeptide, represented as XpYL (where X is a group that occupies the pY-1 position, pY is the phosphorylated tyrosine residue and L is leucine and positioned at the at the pY+1 position). Though H-pY-LPQTV-NH$_2$ and its acetylated tripeptide, Ac-pY-LPQ-NH$_2$ demonstrated near identical disruptive effect on the unphosphorylated GFP-uSTAT3beta/dsM67 DNA complex, the acetylated derivative further demonstrated a dose-dependent decrease in the level of phosphorylated GFP-pSTAT3beta bound to the dsM67 DNA probe with about 50% reduction in the level of DNA bound at peptide inhibitor concentration of 500 μM. These observations indicate that the acetylated derivative was capable of adopting a conformation similar to that of the H-pY-LPQTV-NH$_2$ during the pY-SH2 interaction.

Further, it clearly demonstrated the relevance of the Y-1 position, corroborating findings from other groups that the interaction between the pY of the peptide and the STAT3 SH2 domain was dependent on the moiety at the Y-1 position (Batzer et al., 1995). As a result of the strong ionic interaction between the positively charged amino acid residue R$_{609}$ in the STAT3 SH2 binding pocket and the negatively charged phosphate group of the pY, it was speculated that the absence of a neutral group (e.g. G, Ac, P) at the Y-1 position may encourage the free N-terminus (i.e. NH$_3^+$), to engage in unfavourable interactions with either the R$_{609}$ or itself through interactions between the free N-terminus and the pY thereby, distorting the bound conformation of the phosphotyrosyl peptide.

The relevance of the Y+1 position was not readily evident since both the gp130 and the STAT3 derived phosphotyrosyl peptides contained leucine residues (L) at the Y+1 position but exhibit different affinities. However, several studies have stressed the relevance of the “L” residue in pY-SH2 interaction (Campbell et al., 1994; Batzer et al., 1995). Even though, Batzer and colleagues reported that three to five amino acids residues C-terminal to pY was required for the pY-SH2 domain interaction, the inhibition studies conducted with the acetylated gp130-derived peptide suggested that
the residue at position Y+4 may not be relevant. The amino acid residue at the Y+3, glutamine (Q), seems to be relevant for STAT3 recruitment and subsequent activation.

In conclusion, we have applied PEMSA to demonstrate the DNA binding functionality of both purified, isoforms of GFP-STAT3βtc chimaeric proteins through a protein/DNA mobility shift. In both cases, the binding to the DNA was specific and resulted in stable GFP-pSTAT3βtc/dsM67 DNA or GFP-uSTAT3βtc/dsM67 DNA complexes. In addition, the assay was used to demonstrate the specificity of the STAT3-specific phosphotyrosyl peptides (especially the gp130 based peptide derivatives) to the STAT3 SH2 domain through inhibiting or disrupting dimer formation and concomitant DNA binding demonstrated as a protein mobility shift.
7 FRET-BASED ANALYSIS OF GFP-STAT3βtc/GFP-STAT3βtc INTERACTIONS

7.1 Fluorescent Resonance Energy Transfer (FRET)

Green fluorescent protein-based fluorescent resonance energy transfer has emerged as a crucial technique in the study of protein-protein interaction in the post-genomic era. It is rapidly assuming credence as an important technique to study molecular interactions (van Rheenen et al., 2004). FRET, is a non-radiative transfer of energy from an excited chromophore (donor; i.e. eCFP) to another chromophore (acceptor; i.e. eYFP) by intermolecular long range dipole-dipole coupling (Cardullo, 2007), Figure 7.1.

![Jablonski diagram demonstrating electronic transition in FRET](image)

Figure 7.1: Jablonski diagram demonstrating electronic transition in FRET

The bold horizontal lines ($S_0$, $S_1$ and $S_2$) represent the electronic energy levels whilst the thin horizontal lines represent the different vibrational energy levels within the $S$ state. Irradiation of a fluorophore generates transition from the ground state ($S_0$) to a higher energy state ($S_1$ or $S_2$) and to any vibrational level ($10^{15}$ s). Upon rapid internal conversions to the lowest vibrational level in the excited state ($10^{10}$ s), energy is lost relatively slowly ($10^8$ s) by the emission of a photon (fluorescence, F). FRET is only possible if the donor fluorophore is in close proximity with the acceptor fluorophore with sufficient spectral overlap.

Also crucial for an effective energy transfer is the distance between the donor and acceptor fluorophore which must be less than 10 nm or 100 Å (R), an indication that FRET can serve as a molecular ruler for estimating distances within 10 nm (Stryer and...
Haugland, 1967). Surprisingly, however, these distances are within the confines of conventional protein dimensions and similar to that demonstrated by multimeric protein complex or complexes present in biological systems (Stryer, 1978; Sheng and Hoogenraad, 2007). Furthermore, the orientation of the two chromophores relative to each other must be favourable to generate a good transition dipole, Figure 7.1. FRET allows the GFP mutant fluorescent tags to be fused to the parent protein at either the N- or C-termini.

FRET has allowed activities such as real-time protein-protein interactions, protein conformational changes or post-translational modifications in vivo in a non-invasive manner to be monitored (Truong et al., 2001). In addition, in vitro FRET has been employed to study equilibrium binding constant $K_d$ of protein-protein interactions as well as biochemical cascades based on steady-state and time resolve (Martin et al., 2008).

![Figure 7.2: Schematic representation of the basic principles of FRET](image)

The two putative protein molecules (L) bearing a donor (eCFP) or an acceptor (eYFP) fluorophore is depicted and may be encircled by a sphere representing the maximum distance requirement for FRET. (A), no FRET when the fluorophore FRET pair is more than 100 Å apart. (B), FRET, when the fluorophore FRET pair is in molecular contact (less than 100 Å) as a result of a favourable transition dipole orientation generated by the interacting interface between the parent molecules. (C), weak FRET, when the fluorophore FRET pair is less than 100 Å apart, but is not in a favourable transition dipole orientation.

Thus, on the basis of the studies conducted by Becker and colleagues and corroborated by Braunstein and colleagues, it was revealed that unlike uSTAT3, uSTAT3βtc predominantly exist as monomers prior to phosphorylation, steady-state FRET assays were used to investigate the association of the purified, recombinant, GFP-uSTAT3βtc molecules in vitro. Further, we investigated the effect of STAT3 SH2-specific phosphotyrosyl peptide and STAT3-specific oligonucleotide DNA (dsM67 DNA) as well as Mg$^{2+}$ ions on eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly in solution (Wagner et al., 1990; Novak et al., 1998). The choice of FRET pair used, i.e., donor;
eCFP and acceptor; eYFP, was based on the study conduct by Tsien (Tsien, 1998; Miyawaki et al., 2003) where the pair demonstrated a large spectral overlap with negligible excitation cross-talk (Figure 7.1) and thus a high $R_0$. Unless stated otherwise, FRET was measured predominantly by the efficiency of energy transfer referred to as FRET efficiency (E), section 7.1.3.

7.1.1 Absorption and fluorescent measurement

Samples were measured in a 3.5 mL Acryl/Acrylic 10 x 10 x 45 mm disposable cuvette (Sarstedt, 10 mm path length). Protein concentrations (FRET pair) were determined via intrinsic GFP absorbance, section 2.2.4.2. Excitation and emission spectra of the FRET pair, Figure 3.29, were obtained using a single sample unit PerkinElmer LS 55 Luminescence spectrometer (excitation 425 nm, 5 nm slit width, 1 nm interval, 1 s integration, spectra corrected for instrument response) with Xenon arc lamp. Unless stated otherwise, the FRET mix was always excited at 425 nm to minimize indirect excitation of the acceptor fluorophore. Fluorescence and FRET measurements were performed at room temperature.

7.1.2 Preparation for FRET

7.1.2.1 Preparing eCFP-uSTAT3βtc (donor)/eYFP-uSTAT3βtc (acceptor) FRET pair assay mixture

Unless stated otherwise, FRET was measured predominantly via the efficiency of energy transfer. To obtain the FRET spectra for eCFP-uSTAT3βtc (donor)/eYFP-uSTAT3βtc acceptor FRET pair, the purified fluorescent STAT3βtc chimaeric protein was dialyzed against a higher pH buffer and high salt buffer (buffer D), i.e., 20 mM HEPES pH 8.5, 100 mM NaCl, 1 mM EDTA, and 1 mM DTT. Upon dialysis, the concentrations of the FRET pair were determined (typically 3.0 μM). Assay mix containing the FRET pair in the following donor to acceptor molar ratios; 0.1:1, 0.3:1, 1:1, 1:3 and 1:10, were prepared in 1.5 ml sterile tubes. The mixture was then diluted to 0.5 mL with the high pH and salt buffer. The reactions were duplicated. They were thoroughly stirred, centrifuged briefly at 13,000 xg for 15 s and then incubated for 30 min at room temperature with rocking. After incubation, the FRET reaction mixture was transferred into the 3.5 mL cuvette and then diluted to 3 mL with FRET buffer (lower pH and lower salt); 50 mM HEPES pH 7.2, 50 mM NaCl, 1 mM EDTA, 5 mM MgCl$_2$ and 1 mM DTT and 5% (w/v) glycerol. The final concentrations of the FRET pair constructs were 0.034 μM (donor; ~ 10 μg) and 0.034 μM (acceptor; ~ 10 μg) for the 1:1 donor to acceptor molar ratio mix. The cuvette was sealed with parafilm and inverted 3x to obtain a homogeneous mixture. The mixture was checked for bubbles before incubating for either 2 h at room temperature or 16 h at 4 °C. Incubation was carried
out with rocking. After incubation, FRET fluorescence was obtained after direct excitation of the FRET assay mixture, section 7.1.3. The data output from the spectrophotometer was manually exported into an excel spread sheet that relates the donor/acceptor molar ratios to the emission intensities of the either the donor (475 nm) or the acceptor (527 nm).

7.1.2.1.1 Stabilizing effect of dsM67 DNA in eCFP-uSTAT3βtc (donor)/eYFP-uSTAT3βtc (acceptor) interactions measured by FRET

Using the ratio 1:3, i.e., eCFP-uSTAT3βtc (0.034 μM): eYFP-uSTAT3βtc (0.102 μM), (optimum FRET signal), the extent of protein-protein interaction was investigated with molar excess of the dsM67oligonucleotide DNA stabilizer (10x). Assay mixture containing donor to acceptor molar ratio 1:3 was prepared in a 1.5 mL sterile tube. The mixture was thoroughly stirred and then centrifuged at 13,000 xg for 15 s. The reaction was then supplemented with dsM67 DNA solution to a final concentration of 2.04 μM after diluting the assay mix to 0.5 mL with the high pH and high salt buffer (dialysis buffer D, section 7.1.2.1).

In parallel, a similar reaction mixture containing no dsM67 DNA was also prepared. Both test and control reactions were centrifuged at 13,000 xg for 15 s and then incubated for 30 min at room temperature with rocking. The binding reactions were transferred into the 3.5 mL cuvette and then diluted to 3.0 mL with the FRET buffer. The cuvettes were sealed with parafilm and then gently inverted 3x to ensure homogeneity. The reactions were then incubated for either 2 h at room temperature or 16 h at 4 °C with rocking. Upon incubation, FRET fluorescence was obtained after direct excitation of the assay mixture, section 7.1.3. The data output from the spectrophotometer were manually exported into an excel spread sheet and relates the stabilizing effect of the dsM67 DNA in complex with the FRET pair assembly.

7.1.2.1.2 Stabilizing effect of divalent cation (Mg²⁺) in eCFP-uSTAT3βtc (donor)/eYFP-uSTAT3βtc (acceptor) interactions measured by FRET

Following verification of the appropriate donor to acceptor molar ratio required for optimum FRET signal, i.e., 1:3 eCFP-uSTAT3βtc (0.034 μM): eYFP-uSTAT3βtc (0.102 μM), the assembly of the FRET pair was stabilized by Mg²⁺ from magnesium chloride (MgCl₂). An assay mix containing donor to acceptor molar ratio 1:3 was prepared in a 1.5 mL sterile tube. The mixture was thoroughly stirred and then centrifuged at 13,000 xg for 15 s. The reaction was then supplemented with MgCl₂ solution to a final concentration of 5 mM after diluting the assay mix to 0.5 mL with the high pH and high salt buffer (dialysis buffer, section 7.1.2.1).
Also, similar reaction mixtures containing no Mg\(^{2+}\) ions were prepared in parallel. Both test and control reactions were centrifuged at 13,000 xg for 15 s and incubated for 30 min at room temperature with rocking. The reactions were transferred into the 3.5 mL cuvette and then diluted to 3.0 mL with the FRET buffer. The cuvettes were sealed with parafilm and then gently inverted 3x to ensure homogeneity. The reactions were then incubated on a rocker for either 2 h at room temperature or 16 h at 4 °C. Following incubation, FRET fluorescence was obtained after direct excitation of the assay mixture, section 7.1.3. The data output from the spectrophotometer was manually exported into an excel spread sheet to relate the stabilizing effect of the divalent cation, Mg\(^{2+}\), in the eCFP-uSTAT3\(\beta_{tc}/eYFP-uSTAT3\(\beta_{tc}\) assembly and FRET.

7.1.2.1.3 Stabilizing effect of glycerol in eCFP-uSTAT3\(\beta_{tc}\) (donor)/eYFP-uSTAT3\(\beta_{tc}\) (acceptor) interactions measured by FRET

With knowledge of the appropriate donor to acceptor molar ratio, i.e., 1:3 eCFP-uSTAT3\(\beta_{tc}\) (0.034 μM): eYFP-uSTAT3\(\beta_{tc}\) (0.102 μM), the stabilizing effect of glycerol (0 to 50 %) on the FRET pair assembly was investigated. Six identical assay mixtures containing the FRET pair were prepared in 1.5 mL sterile tubes. The mixtures were thoroughly stirred and centrifuged at 13,000 xg for 15 s. They were incubated for 10 min at room temperature and then diluted to 0.5 mL with the high pH and high salt buffer (dialysis buffer), section 7.1.2.1.

Control reaction (6x) containing only the eCFP-uSTAT3\(\beta_{tc}\) protein molecules (~ 10 μg) were also prepared in parallel. The reactions were diluted to 0.5 mL with the high pH and high salt dialysis buffer after 10 min incubation. Both test and control reactions were centrifuged at 13,000 xg for 15 s and incubated for 30 min at room temperature with rocking. They were transferred into the 3.5 mL cuvette and then diluted to 3.0 mL with the FRET buffers prepared to generate; 0, 5, 10, 15, and 30 % glycerol (v/v). The cuvettes were sealed with parafilm and then inverted 3x to ensure homogeneity. The reaction was checked again to ensure that the mixture was indeed homogenous. They were then incubated on a rocker for either 2 h at room temperature or 16 h at 4 °C. After incubation, FRET fluorescence was obtained after direct excitation of the assay mix, section 7.1.3. The data output from the spectrophotometer was manually exported into an excel spread sheet to relate the stabilizing effect of glycerol on the eCFP-uSTAT3\(\beta_{tc}/eYFP-uSTAT3\(\beta_{tc}\) assembly and FRET.

7.1.2.1.4 Ionic strength and eCFP-uSTAT3\(\beta_{tc}\) (donor)/eYFP-uSTAT3\(\beta_{tc}\) (acceptor) interactions measured by FRET

The relationship between the solubility of the FRET pair and the efficiency of energy transfer was investigated by varying the concentration of NaCl in the FRET buffer from
0 to 500 mM. Assay mix (6x) containing the donor and acceptor FRET pair in a 1:3 molar ratio was prepared in a 1.5 mL sterile tube. The mixtures were thoroughly stirred and briefly centrifuged at 13,000 xg for 15 s. They were then stabilized for 10 min at room temperature and then diluted to 0.5 mL with the high pH and high salt buffer (dialysis buffer), section 7.1.2.1.

Further, control reactions containing only the eCFP-uSTAT3βtc donor molecules (~10 μg) were also prepared in parallel. The reactions were diluted to 0.5 mL with the high pH and high salt dialysis buffer after 10 nm incubation. Both test and control reactions were centrifuged at 13,000 xg for 15 s and then incubated for 30 min at room temperature with rocking. They were transferred into the 3.5 mL cuvette and then diluted to 3.0 mL with the FRET buffer prepared to generate final NaCl concentrations of 16.6, 50, 100, 150, 300, and 500 mM. The cuvettes were sealed with parafilm and then inverted 3x to ensure homogeneity. The mixtures were then incubated on a rocker for either 2 h at room temperature or 16 h at 4 °C. Following incubation, FRET fluorescence was obtained after direct excitation of the assay mixture, section 7.1.3. The data output from the spectrophotometer was manually exported into an excel spreadsheet that demonstrated the effect of increasing salt (NaCl) concentration on the eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly and stability.

7.1.2.2 Preparing control assay for measuring FRET by donor fluorophore quenching (FRET efficiency)

To measure FRET through donor quenching, donor only reactions (without energy transfer acceptor) were prepared in parallel. The reactions were prepared to measure the fluorescence intensity of the donor molecule relative to the corresponding sample contained in the FRET reaction mixture. The reactions were prepared in sterile 1.5 ml tubes. Into each tube, reactions containing approximately; 1, 3, 10, 30 and 10 μg eCFP-uSTAT3βtc (donor) were transferred. The samples were diluted to 0.5 mL with the high pH and high salt buffer. Each of the control reaction corresponded to the donor content in FRET pair reactions described in section 7.1.2.1. The reactions were briefly centrifuged at 13,000 xg for 15 s and then incubated at room temperature for 30 min. They were then transferred into 3.5 mL cuvettes and subsequently diluted to 3 mL with FRET buffer to final concentrations of 0.0034, 0.0102, 0.0340, 0.1020 and 0.3400 μM. The cuvettes were sealed with parafilm and the reaction inverted 3x to ensure homogeneity. The mixture was checked for bubbles before incubating on a rocker for either 2 h at room temperature or 16 h 4 °C. The fluorescence intensities of the donor only reactions were obtained after direct excitation of the assay mixture at 425 nm, section 7.1.3. The data output from the spectrophotometer was manually exported into
Edwin Nkansah

an excel spread sheet that demonstrated a linear increase of fluorescence at 475 nm with increasing concentration of eCFP-uSTAT3βtc chimaeric protein.

7.1.2.3 Preparing control assay for measuring FRET by sensitized acceptor emission upon donor excitation (FRET ratio)

To measure FRET through sensitized acceptor emission, acceptor only reactions (without energy transfer donor) were prepared in parallel to section 7.1.2.1. The reactions were prepared to measure the direct excitation of the acceptor relative to their corresponding FRET sample containing the FRET pair. The reactions were prepared in sterile 1.5 ml tubes. Into each tube 1, 3, 10, 30 and 100 μg eYFP-uSTAT3βtc (acceptor) was transferred. The samples were diluted to 0.5 mL with the high pH and high salt buffer. The concentrations of the acceptor FRET molecule contained in the control reactions corresponds to that contained in the FRET reaction samples, section 7.1.2.1. The reactions were centrifuged at 13,000 xg for 15 s and then incubated at room temperature for 30 min. They were transferred into 3.5 mL cuvettes and then diluted to 3 mL with FRET buffer to final concentrations of 0.0034, 0.0102, 0.0340, 0.1020 and 0.3400 μM. The cuvettes were sealed with parafilm and the reaction inverted 3x to ensure homogeneity. The mixtures were checked for bubbles before incubating for either 2 h at room temperature or 16 h at 4 °C. Incubation on rocker was carried out with rocking. The fluorescence intensities at 527 nm were obtained after direct excitation of the assay mixture, section 7.1.3. The data output from the spectrophotometer was manually exported into an excel spread sheet that demonstrated a linear increase in fluorescence at 527 nm with increasing concentration of eYFP-uSTAT3βtc chimaeric protein.

7.1.2.4 Preparing control assay for non-specific interactions between donor (eCFP) and acceptor (eYFP) fluorophores

To control for non-specific interactions, eCFP (not fused to uSTAT3βtc) was titrated into constant eYFP-uSTAT3βtc (~10 μg). Control reactions containing the eCFP (donor) and eYFP-uSTAT3βtc (acceptor) in mass ratios 0.1:1.0, 0.3:1.0, 1.0:1.0, 3.0:1.0 and 10.0:1.0 were prepared in sterile 1.5 mL tube. The mixtures were diluted to 0.5 mL with the high pH and high salt buffer. Each of the control reaction corresponded to the FRET reaction samples in section 7.1.2.1. The reactions were centrifugationed at 13,000 xg for 15 s and then incubated at room temperature for 30 min. They were transferred into 3.5 mL cuvettes and then diluted to 3 mL with FRET buffer to final concentrations of 0.0034, 0.0102, 0.0340, 0.1020 and 0.3400 μM for the eCFP protein. The cuvette was sealed with parafilm and the reaction inverted 3x to ensure homogeneity. The mixture was checked for bubbles before incubating for either 2 h at room temperature or 16 h at 4 °C. The FRET fluorescence was obtained after direct
excitation of reaction mixture, section 7.1.3. The data output from the spectrophotometer was manually exported into an excel spreadsheet that demonstrated the effect of non-specific interaction on the actual FRET generated through protein/protein assembly between the parent protein molecules, i.e., uSTAT3\beta\text{tc}/uSTAT3\beta\text{tc}.

7.1.2.5 Preparing control assay to demonstrate FRET using eCFP-TEV-eYFP chimaeric protein

The purified polyhistidine tagged eCFP-TEV-eYFP concatemer, i.e. eCFP tethered to eYFP by a short a peptide linker sequence sensitive to the TEV protease, was used to demonstrate the principle of FRET via the proteolytic activity of the TEV protease. The cleavage assay was performed by preparing a series 50 µL reactions containing increasing concentration of the TEV protease, i.e., 2.5, 5.0, 7.5, 10.0, 12.5, 25.0, 50.0 and 100 µg/mL, in a sterile 1.5 mL tube. Each reaction was supplemented with the purified GFP fusion protein (~ 12.5 µg) and then diluted with the cleavage buffer (50 mM Tris-HCl (pH 8.0), 0.5 mM EDTA and 1 mM DTT) to 50 µL. The mixture was stirred and centrifuged at 13,000 x g for 15 s. A control reaction containing the concatemer (~ 12.5 µg) but no TEV protease was prepared in parallel.

Both test and control reactions were incubated at 4 °C for 16 h with rocking. After incubation, 10 µL aliquots of the reaction mixture were removed and transferred into sterile 0.5 mL tubes and immediately boiled at 95 °C for 10 min. The samples were kept frozen at -20 °C until the addition of 6 x SDS loading dye and DTT in readiness for SDS-PAGE and western blot analysis, sections 2.2.5.1 and 2.2.5.2.

The remaining 40 µL proteolytic reaction mixtures were transferred into the 3.5 mL cuvettes and then diluted to 3.0 mL with FRET buffer. The cuvettes were sealed with parafilm and inverted 3x to ensure a homogeneous reaction mix. The reactions were then incubated at room temperature for 15 min whilst ensuring that no bubbles were present. Upon incubation, FRET fluorescence was obtained after direct excitation of proteolytic reaction mixture at 425 nm, section 7.1.3. The data output from the spectrophotometer was manually exported into an excel spreadsheet that demonstrated FRET and the loss of FRET as a result of the proteolytic activity of the TEV protease.

7.1.3 FRET measurement

In preparing for FRET measurements, the lamp in PerkinElmer LS55 luminescence spectrometer was pre-warmed for 15 min. Before measurement, the assay samples
were checked to ensure that no bubbles were present. Using the FL WinLab software, the FRET fluorescence for the test and control experiments preparations were measured simultaneously by direct excitation of the reaction mixtures at 425 nm (donor). The emission spectra for both test and control samples were scans from 410 nm to 600 nm. Measurements were taken in duplicates. Following scanning, samples were re-checked to ensure that bubbles have not been inadvertently introduced which may affect the outcome of the FRET assay. From the generated spectra, i.e., test and control, the relative emission intensities were identified at 475 nm for eCFP only or eCFP-uSTAT3βtc without FRET and at 527 nm for FRET-based sensitized emission of eYFP only or eYFP-uSTAT3βtc.

To determine FRET efficiency (E), the donor emission intensity at 475 nm for the test (FRET pair; donor (D) and acceptor (A), F_D(A) and control (corresponding donor (D) only, F_D) preparations were measured. With the values of F_D(A) and F_D normalized to their respective concentrations of donor, the efficiency of energy transfer between a single D - A FRET pair was calculated according to the below:

\[
E = 1 - \frac{F_{DA}}{F_D} = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6}
\]

Equation 7.1: FRET efficiency where \(F_{DA}\) = the intensity of the donor molecule in the presence of a functional acceptor (test), \(F_D\) = the intensity of the donor molecule in the absent or deactivation of the acceptor molecule (control), \(R\) = the distance between the D and A fluorophore, \(R_0\) = the distance parameter calculated from the spectroscopic and mutual dipole orientational parameters of D and A or is the radius at which FRET efficiency is 50% ( Förster radius).

For FRET ratio, the ratio of the acceptor emission intensity at 527 nm due to FRET and direct excitation (\(F_{AD}\)) and the acceptor emission intensity at 527 nm due to direct excitation (\(F_A\)) was determined. A FRET ratio of 1 or less meant no FRET whilst a FR greater than 1 meant FRET had occurred.

\[
FR = \frac{F_{AD}}{F_A}
\]

Equation 7.2: FRET ratio where \(F_{AD}\) = the acceptor emission intensity at 527 nm due FRET and direct excitation and \(F_A\) = the acceptor emission at 527 nm due to direct excitation respectively.
Further, ratiometric evaluation of the acceptor and donor emission intensity was randomly used to determine FRET, i.e., by finding the ratio of the peak intensities at 527 nm (acceptor) and 475 nm (donor):

\[
FRET = \frac{A_{Em.}(527 \text{ nm})}{D_{Em.}(475 \text{ nm})}
\]

Equation 7.3: FRET where \(A_{Em.}\) = the emission peak intensity of the acceptor at 527 nm and \(D_{Em.}\) = the emission peak intensity of the donor at 475 nm.

The data output from the spectrometer was manually exported into an excel spreadsheet and then presented in the form of column charts.

7.2 RESULTS

7.2.1 Fluorescent Resonance Energy Transfer (FRET).

Though recent findings have suggested that full length STAT3 predominantly exist in the cytoplasm as dimers or multimers with a \(K_d\) value less than 10 nM, prior to phosphorylation (Ndubuisi et al., 1999; Braunstein et al., 2003), gel filtration chromatography performed on purified, recombinant, uSTAT3\(\beta\)tc (expressed in E.coli BL21 Rosetta™ E.coli cells) has indicated that the protein exist predominantly as monomeric, correlating with previous studies conducted by Becker and colleges (Becker et al., 1998).

To investigate the protein/protein interaction between STAT3 protein molecules in the inactive state, STAT3\(\beta\)tc was explored in vitro in order to understand the conditions that modulate STAT3/STAT3 protein assembly as well as attempt to identify small molecules capable of interfering with these interactions.

Consequently, the uSTAT3\(\beta\)tc protein represented by the GFP-uSTAT3\(\beta\)tc chimaeric protein was used in the FRET assay. On the basis of the three dimensional structure of the DNA bound STAT3\(\beta\), Figure 1.3, the C-terminus GFP- tagged STAT3\(\beta\)tc construct (STAT3\(\beta\)tc-GFP) may have allowed efficient energy transfer between the donor and acceptor domains of the interacting chimaeric protein molecules as a result of the close proximity of the STAT3\(\beta\) C-termini. However, no significant soluble fraction of the uSTAT3\(\beta\)tc-GFP expression construct was obtained after construction and expression (sections 3.1.3 and 3.2.2), thus suggesting that the chimaeric protein was non-functional (Figure 3.16). In contrast, soluble fractions of the GFP-STAT3\(\beta\)tc constructs were generated after lysis, Figure 3.1.9. Furthermore, since the three dimensional
structure of DNA bound unphosphorylated STAT3 dimer was unknown and that of the non-DNA bound unphosphorylated STAT3 dimer (Ren et al., 2008) was not identical to that described by Becker et al in Figure 1.3, it was therefore appropriate to use GFP-uSTAT3βtc construct to investigate more about the oligomerization state of unphosphorylated STAT3 in vitro.

The dimerisation status of the GFP-uSTAT3βtc chimaeric protein was characterized by gel filtration chromatography (section 4.1.1.3) and shown to be monomeric (section 4.3). The interactions between the monomeric GFP-uSTAT3βtc molecules were assayed using the eCFP-uSTAT3βtc and the eYFP-uSTAT3βtc species in the FRET-based assay. FRET between the GFP variants was demonstrated and quantified as the efficiency of energy transfer using Equation 7.1.

7.2.2 Demonstration of FRET

Generally FRET occurs when the donor and the acceptor fluorophores are within 10 nm of each other; proteins within that range usually interact directly with each other (van Rheenen et al., 2004). To demonstrate FRET between the donor (eCFP) and acceptor (eYFP) fluorophores as a control, the His-eCFP-TEV-eYFP fusion protein with both fluorophores tethered to the recognition sequence of tobacco etch virus (TEV) protease (section 3.1.5 and 3.2.4) was designed and developed. Assuming an extended linker conformation, the distance between the two fluorophores would be approximately 3.8 nm. We initially wanted to verify that steady-state FRET could be demonstrated in vitro with the concatemer and that the signal would be abolished when the concatemer was treated with the functional TEV protease.

Thus following purification, the recombinant, GFP fusion protein was mixed with or without different concentrations of the TEV protease in the cleavage buffer, and then incubated at 4 °C for 16 h. This would result in an expected change of FRET signal which should be clearly detectable.

Aliquots (i.e. 40 μL) of the reaction mixtures were diluted to 3.5 mL with FRET buffer and the emission scan acquired by exciting the reaction mixture at 425 nm (eCFP fluorophore). As shown in Figure 7.3 (A) (B) and (C), FRET was demonstrated between the donor and acceptor fluorophore and disrupted with increasing concentration of the TEV protease.

The emission spectrum of the intact recombinant, His-eCFP-TEV-eYFP fusion protein (Figure 7.3 (A)) clearly demonstrate that excitation of the donor at 425 nm results in
the emission of the donor (emission wavelength; 475 nm) and of the acceptor (emission wavelength; 527 nm) thus suggesting FRET. Upon adding different concentrations of the recombinant, TEV protease, the emission intensities of the donor fluorophore gradually increased with a corresponding decrease in the emission intensities of the acceptor fluorophore (Figure 7.3 (B)).

![Figure 7.3: Disruption of FRET between covalently linked eCFP and eYFP GFP mutants by the recombinant TEV protease](Image)

Following incubation of the recombinant, His-eCFP-TEV-eYFP GFP fusion protein with different concentrations of the TEV protease, the mixture was diluted with FRET buffer and then scanned in a 1 mm path length cuvette with the excitation wavelength of 425 nm. (A): Emission spectrum of the GFP fusion protein demonstrating FRET before addition of the recombinant TEV protease. (B): Emission spectra of the GFP fusion protein showing decreasing FRET following pre-incubation with increasing concentrations of the recombinant TEV protease (0 to 100 μg/mL). (C): Emission spectrum of the GFP fusion protein showing no significant FRET. The data from the spectrophotometer was exported to an excel spread sheet and the emission intensities plotted against wavelength. The emission spectra displayed are the mean specific emissions from triplicate readings and is a representative of n = 2 independent assays.
This observation suggested that FRET did occur when the donor and acceptor GFP mutants were in close proximity (i.e., physically linked in the concatemer). Therefore, by comparing the emission spectra (A) and (C), we confirmed that the disappearance of FRET was due to the increasing presence of the TEV protease. The remaining emission signals observed in (C) for both the donor (475 nm) and acceptor (527 nm) fluorophores represent the direct excitations of the fluorophores. Further, a good correlation between the relative emission ratios and the concentration of the TEV protease was also observed. A 4.23-fold increase in the eYFP/eCFP emission ratio was achieved upon near complete cleavage of the GFP-fusion proteins. In addition, the efficiency of energy transfer between the two GFP mutants contained in the His-eCFP-TEV-eYFP GFP fusion protein was evaluated from the emission intensities (475 nm) of the donor fluorophore recorded in Figure 7.3 (A) and (C) and then calculated as approximately 50 ± 5%. This suggested that the GFP fusion protein could be employed as reliable and efficient positive control construct during subsequent FRET measurements.

Further confirmation of the proteolytic activities of the TEV protease and therefore FRET loss was demonstrated by SDS-PAGE. Here, the remainder of the cleavage reactions (~10 µL) was diluted with 6x SDS loading buffer, boiled and then separated on a 10% SDS polyacrylamide gel (Figure 7.1.2.2).

Figure 7.4: Demonstrating proteolysis of the His-eCFP-TEV-eYFP GFP fusion protein with the TEV protease

The recombinant GFP concatemer was incubated with increasing concentrations of the TEV protease, prior to incubation at 4°C for 16 h. Aliquots of the reaction mixtures were analyzed on a 10% polyacrylamide gel (i) and then transferred unto a nitrocellulose membrane (ii). The images show the gradual disappearance of the low mobility species (uncleaved substrate; His-eCFP-TEV-eYFP) and the concomitant appearance of the high mobility species (cleaved product; GFP mutants). (i) (ii) lane M, contains pre-stained protein marker; lanes 1 to 9 contains 0 to 100 µg/mL TEV protease. (ii) Development with anti-polyhistidine polyclonal antibody (1:5000-fold dilution) using the DAB and urea/H₂O₂ substrate. Arrows indicate the position of the concatemer and its degradation products respectively.
As shown in Figure 7.4 (i), the decreasing density of the His-eCFP-TEV-eYFP GFP fusion protein (57 kDa) bands with increasing concentrations of the TEV protease resulted in the concomitant increase in the density of the band representing the GFP mutants (29 kDa). The densities of the coomassie stained bands representing the GFP fusion proteins in lane 9 relative to lane 1 suggested that the proteolytic activities of the TEV protease was responsible for the simultaneous decrease and increase of the His-eCFP-TEV-eYFP fusion protein and the GFP variants (i.e. mixture if His-eCFP and eYFP proteins) respectively. This observation did not only suggest proteolytic cleavage of the tether linking the GFP fluorophores by the TEV protease but, also suggested that the TEV protease may have experienced minimum or no steric hindrance to complete proteolysis. Also demonstrated is the representative immunoblot (ii) showing immunoreactive bands corresponding to the predicted molecular weights of the His-eCFP-TEV-eYFP fusion protein and its degradation product following proteolysis. Though no coomassie stained band for the degradation products was evident in lane 1 of (i), a significant corresponding immunoreactive band was evident in lane 1 (ii), this suggested very limited proteolysis occurred following lysis. Further, since the molecular weight of the polyhistidine tagged, recombinant, TEV protease is approximately 27 kDa, it co-migrated with the immunoreactive bands representing the degradation product (29 kDa) hence, the apparently enhanced densities of those bands.

7.2.3 Steady-state FRET assays between eCFP-uSTAT3βtc (donor) and eYFP-uSTAT3βtc (acceptor) protein species.

Using the method described in section 7.1.2, STATSptc - STAT3ptc protein-protein interaction was demonstrated using the FRET-base homogeneous assays. By using the fluorescently tagged STATSptc protein molecules, the donor tagged species (eCFP-uSTAT3βtc) was titrated into the acceptor tagged species (eYFP-uSTAT3βtc) whilst monitoring the relative change in the emission intensity peak of the donor. Other control parameters including that for non-specific interactions and direct excitation of the donor were monitored in parallel. The decreasing donor emission intensity (at 475 nm) of the test sample (containing donor and acceptor protein species) relative to that of the control sample (containing donor only) was considered proportional to the number of fluorescent STAT3βtc FRET pairs interacting. The differences in the donor emission intensities were evaluated to generate the efficiency of energy transfer between the fluorophores fused to the interacting STAT3βtc molecules and therefore quantified using the equation describe in section 7.1.3. The FRET efficiency was validated with that generated from the GFP fusion cleaved with the recombinant TEV protease, to model the decrease in FRET efficiency.
Following purification, the recombinant, acceptor species eYFP-uSTAT3βtc was mixed with different concentrations of the donor species eCFP-uSTAT3βtc. In parallel, a set of control reactions containing similar concentrations of the donor species were also prepared. Both test and control reactions were thoroughly mixed and then diluted to 0.5 mL with the buffer D (pH 8.5), prior to incubation at room temperature for 30 min. The mixtures were subsequently incubated for 2 h at room temperature.

The emission spectra of the titrations of the donor species (1, 3, 10, 30, 100 and 340 μg) into the acceptor species (~10 μg) were obtained following excitation of the reaction mixture at 425 nm. Similarly, the emission spectra for the control reactions containing only donor species (demonstrating the measure of the direct excitation of the donor species) were also obtained, revealing a linear increase in the emission peak at 475 nm. The donor emission peak intensities (475 nm) for test and control mixtures were recorded and calculated as the efficiency of energy transfer between the donor and acceptor fluorophore fused to the interacting STAT3βtc protein molecules. The difference between the emission peak intensity for the control samples (donor only sample) and that of the test samples (donor and acceptor) was employed as a measure for the occurring energy transfer. A representative FRET efficiency assay is demonstrated in Figure 7.5.

![Figure 7.5: Typical steady-state FRET efficiency assay.](image)

Following titration of the donor species into the acceptor species, the reaction mixture was diluted with FRET buffer to 3.0 mL and then excited at 425 nm following incubation at room temperature for 2 h. From the generated emission spectra, the intensities at 475 nm for the test and the corresponding control samples were recorded and used to generate the efficiency data. A plot of the FRET efficiencies against the concentrations of the donor species eCFP-uSTAT3βtc was generated. The control reactions represent the direct excitation of the donor species. A plot of the FRET efficiencies of the titrations versus the
concentration (µM) of the eCFP-uSTAT3tc was produced. The data displayed is one representative experiment out of more than 5. The calculated FRET efficiencies can be represented in percentages.

Increasing FRET efficiency was recorded with increasing donor until the acceptor species was interacting with 340 µg of the donor species. The drop in efficiency after the peak was due to molecular overcrowding caused by the donor species, thus encouraging nonspecific interaction among the donor species as a result of the bias in favour of the 10-fold molar excess of the donor species present in the reaction mixture. The highest FRET efficiency data point was obtained from the reaction mixture containing 102 nM (~ 100 µg) and 34 nM (~ 30 µg) donor and acceptor species respectively.

To control for non-specific interactions between the eCFP and eYFP tags, different concentrations of the eCFP fluorophore (not fused to the STAT3tc) was titrated into the eYFP-uSTAT3tc acceptor protein molecule (identical quantities as the test sample describe above). A linear increase in the emission peak at 475 nm was observed. The slope of the curve was almost identical to that obtained for the control reaction that represented direct excitation of the donor species. The observed data therefore suggested that the measured FRET signal was principally due to interactions between the STAT3tc protein molecule and not from the interactions between the fluorophore tags (Figure 7.6). Thus, any nonspecific interaction was negligible.
Figure 7.6: Control curves for the steady-state FRET assay.

The emission peak intensities of the eCFP-uSTAT3βtc (donor species) control samples for either direct excitation or nonspecific interactions was obtained from the spectra following excitation of the samples at 425 nm. Controls for direct excitation, different concentrations of the donor species (i.e. 3.4, 10.2, 34, 102 and 340 nM) corresponding to that present in the test samples was diluted to 0.5 mL with buffer D, prior to dilution to 3.0 mL with the FRET buffer. Controls for nonspecific interaction, different concentration of the eCFP fluorophore (not fused to STAT3βtc) was titrated into the acceptor species (eYFP-uSTAT3βtc) and the resulting mixture diluted to 0.5 mL, prior to subsequent dilution to 3.0 mL before reading the emission peak intensities. A plot of emission peak intensities at 475 nm versus the concentrations of the donor species was generated to produce the slopes of both curves. One representative experiment out of 5 is shown.

Though the generated FRET efficiency chart was reproducible and a characteristic of protein-protein interaction, the detected signals were somewhat transient since FRET was lost after extensive incubation or agitation of the reaction mixture. These observations may suggest that one or both of the recombinant, GFP-uSTAT3βtc FRET pair were either precipitating out of solution or were forming multimeric aggregates through inter-chain disulphide linkages between the uSTAT3βtc protein molecules contained in the FRET construct (section 4.2.1). To reduce this effect and obtain sustained FRET signals, we investigated some of the components of the FRET buffer (section 7.2.4).

7.2.4 Investigating selected ingredients of the FRET buffer that may improve the efficiency of energy transfer between FRET pair

Together with the ionic strength, the percent glycerol content in the FRET buffer was investigated to identify the optimum concentrations necessary to encourage sustained interactions between the fluorophore labelled STAT3βtc chimaeric proteins. Firstly, we investigated the effect of the ionic strength by manipulating the concentrations of the NaCl present in the FRET buffer. Thus, following purification and functional analysis, 6x reactions containing 34 μM eYFP-uSTAT3βtc (acceptor species) mixed with 102 μM (molar ratio of donor: acceptor is 3:1) eCFP-uSTAT3βtc (donor species), prior to dilution to 0.5 mL with buffer D. The reactions were subsequently diluted to 3.0 mL with FRET buffer supplemented with or without NaCl to concentrations of; 0, 50, 100, 150, 300 and 500 mM. In addition, control reactions containing just the donor species and diluted to 3.0 mL with FRET buffer supplement to the above NaCl concentrations were also prepared in parallel. The choice of donor to acceptor ratio (i.e. 3:1) was identified as the ratio necessary to produce optimum energy transfer. Following incubation at room temperature for 2 h, the emission spectra for both the test and control reactions were obtained with an excitation wavelength of 425 nm. From the emission peak...
intensities recorded, the efficiency of energy transfer between the donor and the acceptor species was calculated for each data point (Figure 7.7).

As demonstrated in Figure 7.7, the efficiency of energy transfer between the interacting FRET pairs increased gradually with increasing NaCl concentrations up to 300 mM (~ 55% FRET efficiency), where the efficiency dropped significantly to about 5%. Similarly, a linear increase in the emission peak intensities (475 nm) for the control reactions was observed up to the reaction supplemented with 500 mM NaCl. The drop in the FRET efficiency may suggest quenching of the GFP fluorophore by the excess chloride ions present in the buffer. Some variants of the GFP, especially the yellow fluorescent protein (i.e. H148Q) have been reported to display a progressive decrease in intensity following exposure to chloride ions at higher concentration (Jayaraman et al., 2000). Furthermore, a decreased interaction between the GFP-uSTAT3βtc protein molecules may have accounted for the loss of FRET signal at high salt concentrations.

![Figure 7.7: Demonstration of the effect of NaCl on the efficiency of energy transfer](image)

FRET pair mixtures containing 102 μM and 34 μM for the eCFP-uSTAT3βtc (donor species) and the eYFP-uSTAT3βtc (acceptor species) respectively were incubated in the FRET buffers supplement with or without NaCl up to a final concentration of 500 mM. Similarly, a set of control reactions containing 34 μM of eCFP-uSTAT3βtc (donor species) only were also incubated in FRET buffers supplemented with or without NaCl up to 500 mM. Following incubation, both sets of mixtures were excited at 425 nm to obtain the emission peak intensities from which the efficiency of energy transfer between the interactions STAT3βtc was calculated. A plot of the FRET efficiency versus the concentration gradient of NaCl was generated. Displayed is one representative experiment out of four. The FRET efficiency signals can be represented in percentages.
Although it is reported that protein-protein interactions show little or no dependence on the concentration of NaCl salt (Dumetz et al., 2007), the highest FRET efficiency value was acquired with samples supplemented with 300 mM NaCl. The observation reflected attractive interaction between the uSTAT3βtc molecules that may have been encouraged by the NaCl content in the buffer. However, subsequent studies were performed with buffer supplemented with 50 mM NaCl. The data displayed in Figure 7.7 suggested that, at 50 mM the efficiency of energy transfer was approximately 20%, a percentage found to be very stable and reproducible.

In furtherance to improving and sustaining FRET signals, we also investigated the effect of glycerol on the efficiency of energy transfer. We hypothesized that increasing the viscosity of the FRET buffer will minimize tumbling of the FRET pair and therefore stabilize eCFP-uSTAT3βtc/eYFP-uSTAT3βtc interactions which will therefore result in an improved energy transfer between the FRET fluorophores. Since the concentration of glycerol needed to generate optimum energy transfer was not known, a set of reactions supplemented with or without glycerol (5%) were prepared.

Two sets of reactions, similar to that described in section 7.2.3 were prepared; the first set of reactions contained 0% glycerol whilst the second set was supplemented with 5% glycerol. Following incubation for 2 h, the reactions were read with as previously described in section 7.2.3 and the resultant emission peak intensities calculated as the efficiency of energy transfer (E). As shown in Figure 7.8, the effect of 5% glycerol was evident.

**Figure 7.8:** Demonstrate the attractive interactive effect of 5% glycerol in eCFP-uSTAT3βtc/eYFP-uSTAT3βtc protein interaction. Following titration of increasing concentrations of the donor species into the acceptor species, the reaction mixtures were diluted to 3.0 mL with FRET buffers supplemented with or without glycerol. Similarly, control reactions containing different concentration of the donor species diluted with or without FRET buffer supplemented with or without glycerol. From the emission spectra of the titrations, the emission peak
intensities for both samples prepared with or without 5% (w/v) glycerol were calculated. The calculated FRET efficiency was plotted against the concentration of the donor species. One representative experiment out of two is shown.

As established in previous titrations, the efficiency of energy transfer was highest with the 3:1 donor to acceptor molar ratio (i.e. approximately 22%) preparations. Similarly, the generated FRET efficiencies for the titrations supplemented with the 5% (w/v) glycerol were generally higher relative to those obtained with 0% glycerol. The negative FRET efficiency (i.e. ~ -0.07%) recorded for the glycerol treated sample containing 340 mM eCFP-uSTAT3tc may suggest bias due to donor overcrowding, thus a case of nonspecific interactions between donor species. The next step was to identify the optimum percentage glycerol needed to generate the highest FRET efficiency as well as to investigate the possibility of nonspecific interactions between donor species as a result of overcrowding. Hence, FRET reactions containing 102 and 34 nM eCFP-uSTAT3tc and eYFP-uSTAT3tc (i.e. 3:1 donor to acceptor molar ratio) were prepared respectively. The reactions were pre-incubated in buffer D and then in FRET buffers supplemented with up to 30% glycerol. Following incubation at room temperature for 2 h, the reactions were excited at 425 nm. Using the emission spectra, the emission peak intensities were used to generate the efficiency of energy transfer displayed in Figure 7.1.4.3.
Though a gradual increase in FRET efficiency was recorded with increasing percentage glycerol up to 15 %, it was also observed that it was approaching impossible to generate a homogeneous mixture as a result of the elevated viscosity of the reaction mix due to the increasing glycerol content, i.e., up to 15 %. At concentrations above 5 %, the elevated viscosity of the FRET buffer was apparent and this may have resulted in FRET efficiencies that may be artefacts. However, the FRET efficiency data obtained with preparations supplemented with glycerol to 5 % was considered appropriate since they were robust and reproducible.

7.2.5 Investigating the co-factors that influence FRET between eCFP-uSTAT3βtc and eYFP-uSTAT3βtc FRET pair species

Through immunoprecipitation, Novak and colleagues also demonstrated the existence of pre-activated dimer of STAT3 molecules, formation of which was independent of the conventional phosphotyrosyl-SH2 interaction which follows cytokine activation in cells. The dimer formed required the presence of a divalent ion such as Mg$^{2+}$ ion for stability, which according to Novak, resulted in the dimer withstanding denaturing agents. Hence, on the basis of that, it was hypothesized that by modulating the divalent ions interacting with the GFP-uSTAT3βtc protein molecules prior to assay preparation, this might modulate dimer formation, and may encourage the formation of stable dimers between FRET pairs that will lead to sustained FRET signals.

Here, preparations similar to those described in Figure 7.8; section 7.2.4, were prepared. Upon mixing the FRET pairs, the preparations were pre-incubated in buffer D, prior to incubation with FRET buffers supplement with or without 5 mM MgCl$_2$. After incubation, the mixtures (both test and controls set-up) were excited at 425 nm, to generate the emission spectra and therefore the emission peak intensities. As displayed in Figure 7.10, supplementing the FRET buffer with the divalent cation generally enhanced the efficiency of energy transfer between the interacting FRET pairs.
Figure 7.10: Demonstrate the stabilizing effect of Mg$^{2+}$ ion on GFP-uSTAT3btc protein-protein interaction and therefore FRET efficiency

Two sets of test reactions containing the eYFP-uSTAT3btc (acceptor species; ~ 34 nM) incubated with different concentrations of the eCFP-uSTAT3btc (donor species) where pre-incubated with buffer D, prior to incubation with FRET buffer supplemented with or without 5 mM MgCl$_2$. Similarly, equivalent concentrations of the eCFP-uSTAT3btc only control reactions were prepared in parallel. Following incubation, the FRET reactions were excited at 425 nm to generate the emission peak intensities from the emission spectra. The efficiency of energy transfer was then calculated using the formula described in section 7.1.3. A plot of the FRET efficiency versus the concentrations of the donor species was generated. The data displayed is one representative experiment out of three.

As expected, the highest FRET efficiencies for both the Mg$^{2+}$ ion treated (~ 25 %) and untreated (~ 17 %) preparations were obtained with samples containing 102 and 34 nM for eCFP-uSTAT3btc and eYFP-uSTAT3btc FRET species respectively. Though the stabilizing influence of the divalent cation could not be clearly demonstrated with other functional assays, the effect was reversed with 10x concentrated chelating agent, EDTA (50 mM), in a preparation containing the FRET species. The observation suggested that the Mg$^{2+}$ cations was required to for either encouraging GFP-uSTAT3btc protein-protein interactions or stabilizing the dimers formed by the FRET species.

On the basis of the DNA-binding studies demonstrated in the previous chapters, where the GFP-uSTAT3btc isoform was shown to bind to and form a stable complex with the high affinity SIE DNA (dsM67 DNA), it was hypothesized that the presence of the high affinity SIE DNA sequence (dsM67 DNA) in the FRET preparations may lead to a conformational re-arrangement of the FRET species around the dsM67 oligonucleotide DNA. This interaction between the FRET species and the dsM67 DNA sequence would provide stability for the FRET pairs and therefore result in a sustained FRET signals.
Following from the structural (chapter 4) and some functional (chapters 5 and 6) characterization of the FRET donor and acceptor species, FRET reactions were prepared by mixing 102 and 34 nM of the eCFP-uSTAT3\beta tc and the eYFP-uSTAT3\beta tc FRET species respectively, prior to incubation in the buffer D. The preparations were further diluted to 3.0 mL with the FRET buffer supplemented with or without 340 nM dsM67 oligonucleotide DNA. Similarly, donor only control reactions containing 102 nM were prepared in parallel. From the resultant emission spectra, the emission peak intensities were obtained for both the test and the control samples. The FRET efficiency for each sample was then calculated.

As expected, the presence of the high affinity SIE DNA sequence in the FRET buffer resulted in an enhanced FRET signal calculated as the efficiency of energy transfer. Following treatment with dsM67 DNA, the efficiency of energy transfer between the interaction FRET pair was calculated as 26 % relative to that obtained from the non dsM67 treated samples (approximately 12 %).

![Figure 7.11](image-url)  
**Figure 7.11:** Demonstrating the effect of the high affinity dsM67 oligonucleotide DNA on GFP-uSTAT3\beta tc protein interactions and therefore the efficiency of energy transfer.  
FRET reactions (2x) containing 102 nM of the donor species (eCFP-uSTAT3\beta tc) and 34 nM of the acceptor species (eCFP-uSTAT3\beta tc) were initially incubated in the buffer D prior to incubation in FRET buffer supplemented with or without 340 nM dsM67 oligonucleotide DNA. Similarly, 2x “donor only” control reactions treated with or without 340 nM dsM67 DNA were prepared in parallel. Following incubation, the reactions were excited at 425 nm to obtain the emission spectra. From the emission spectra, the corresponding emission peak intensities were obtained and therefore the FRET efficiencies calculated. A plot of the FRET Efficiency versus the reactions profiles was constructed as depicted above. One representative experiment out of three is shown.

In addition to the FRET efficiencies displayed in **Figure 7.11,** subsequent excitation of the samples after 24 h of incubation at 4 °C generated FRET efficiency data which was
fairly identical to that demonstrated in Figure 7.11. Thus, not only did the high affinity dsM67 DNA encouraged eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly between the FRET pairs, but also dsM67 DNA provided stability for the interacting FRET pair and therefore reproducible FRET efficiency.

7.2.6 Disruption of FRET signals using STAT3-SH2 specific phosphotyrosyl peptide

In support of previous chapters (chapters 5 and 6), it was hypothesized that by applying the STAT3-SH2 domain specific phosphotyrosyl peptide in the FRET assay, the peptide would block interactions between the FRET pair and therefore leads to the loss of FRET which would then highlight the relevance of the SH2 domain in unphosphorylated STAT3 dimer formation.

To demonstrate the specificity of the phosphotyrosyl peptide inhibitor for the STAT3 SH2 domain, FRET was demonstrated between a fluorescein-labeled STAT3-SH2 specific phosphotyrosyl peptide; 5-carboxyfluorescein-G-pY-LPQTV-NH₂ or FAM-G-pY-LPQTV-NH₂ (excitation; 492 nm and emission; 518 nm) and the eCFP-uSTAT3βtc protein molecules.

The FRET efficiencies obtained from titrating different concentrations of the donor species, i.e., eCFP-uSTAT3βtc; 3.4, 10.2, 34.0, 102.0 and 340 nM, into any of the acceptor species, i.e., either the 5-carboxyfluorescein-G-pY-LPQTV-NH₂ (34 nM) or the eYFP-uSTAT3βtc (34 nM) was shown in Figure 7.12. As illustrated in the data, FRET was observed in both preparations and the efficiency of energy transfer was highest in reactions containing 3-fold molar excess of the donor species. The FRET efficiency was 34 % and 36 % for preparations containing 34 nM eYFP-uSTAT3βtc or 5-FAM-G-pY-LPQTV-NH₂ respectively. Indeed, a positive FRET efficiency was calculated for the mixture containing 1:1 donor to acceptor molar ratio. In demonstrating FRET, especially between the eCFP-uSTAT3βtc protein and 5-FAM labelled phosphotyrosyl peptide assembly, the interaction further confirmed the specificity of the peptide probe as well as the functionality of the donor species. Moreover, the FRET signal observed did not only suggest that the eCFP fluorophore and the 5-FAM dye were in close proximity (< 10 nm) but also they were on average oriented favourably relative to each other and therefore allowed the efficient transfer of energy between the fluorophore and the dye molecule.

Further, confirmatory studies, Figure 7.12 (B), were carried out to support the data shown in Figure 7.12 (A). Consequently, the ability of a corresponding unlabelled
phosphotyrosyl peptide, H-pY-LPQTV-NH₂, to inhibit binding between the eCFP-uSTAT3βtc and the labelled acceptor species, 5-FAM-G-pY-LPQTV-NH₂, was investigated to demonstrate the specificity of the phosphotyrosyl peptide sequence. Thus, FRET preparations containing 102 nM eCFP-uSTAT3βtc and 34 nM 5-FAM-G-pY-LPQTV-NH₂ were incubated with or without different concentrations of unlabelled phosphotyrosyl peptide probe prior to incubation in the FRET buffer.

Similarly, FRET reactions containing 102 nM eCFP-uSTAT3βtc and 34 nM eYFP-uSTAT3βtc were also prepared and then incubated with or without H-pY-LPQTV-NH₂ phosphotyrosyl peptide (~ 340 nM) prior to incubation in FRET buffer. Control reactions containing the appropriate concentrations of the donor species were prepared in parallel. Following incubation and acquisition of the emission spectra, the FRET efficiencies were calculated and displayed in Figure 7.12 (B) (C).

As expected, the presence of the unlabeled peptide resulted in a dose-dependent decrease in FRET efficiency (Figure 7.12 (B)), with approximately 40 % reduction at the inhibitor peptide concentrations between 33 to 66 nM. The reduction in the FRET signals did not only confirm that the calculated FRET efficiency displayed in Figure 7.12 (A) was real but also confirmed that the interaction between the 5-FAM labelled peptide and the eCFP-uSTAT3βtc donor species was indeed mediated by the phosphotyrosyl-SH2 domain interaction.
Figure 7.12: Demonstrating FRET and the disruption of FRET signals.

Comparative FRET assay were prepared by mixing eCFP-uSTAT3βtc with different concentrations of either eYFP-uSTAT3βtc or 5-FAM-G-pY-LPQT-V-NH₂ prior to incubation with the FRET buffer (A). Competitive binding assay prepared by incubating reaction mixtures containing FRET pair, eCFP-uSTAT3βtc and 5-FAM-G-pY-LPQT-V-NH₂, as well as a corresponding unlabeled competitor phosphotyrosyl peptide probe (H-pY-LPQT-V-NH₂) (B). Inhibition assay prepared by incubating the FRET reactions containing FRET pair, eCFP-uSTAT3βtc and eYFP-uSTAT3βtc, supplemented with or without dsM67 DNA and their corresponding inhibitory reactions using the unlabeled phosphotyrosyl peptide probe applied in (B) (C). Corresponding control reactions containing the donor species were prepared for all three studies. Reactions were excited at 425 nm and the emission peak intensities read off at 475 nm. (A), the calculated FRET efficiencies was plotted against the concentrations of the titrant (eCFP-uSTAT3βtc; donor species). (B), a plot of the calculated FRET efficiencies versus the concentration of the peptide inhibitor was constructed. (C), a plot of the calculated FRET efficiencies against the reaction profiles involved in the study. Each data is one representative experiment out of three.

Further, as demonstrated in Figure 7.12 (C), 340 nM of the phosphotyrosyl peptide probe significantly disrupted the interaction between the GFP-uSTAT3βtc FRET pair and therefore the efficiency of energy transfer. The observations did not only suggest that phosphotyrosyl peptide targeted to the STAT3-SH2 domain effectively blocked eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly but also confirmed that the STAT3 SH2 domain was very much involved in mediating STAT3/STAT3 interactions (as also reported by Haan et al., 1999).

In summary, Figure 7.13, the eCFP-uSTAT3βtc/eCFP-uSTAT3βtc assembly has been demonstrated via FRET and through the efficiency of energy transfer between the GFP tags fused to the N-terminus of uSTAT3βtc protein molecules. Further, that stabilizing effect of cofactors such as Mg²⁺ ions and the dsM67 DNA sequence on the interacting FRET pair molecules was evident in the efficiency of energy transfer between the FRET pair. Furthermore, FRET has been shown to occur between the eCFP-uSTAT3βtc species and the 5-carboxyfluorescein labelled phosphotyrosyl peptide
Edwin Nkansah

(target to the STAT3 SH2 domain) as demonstrated the inhibition of that interaction with a corresponding unlabeled phosphotyrosyl peptide probe.

Figure 7.13: Quantitative Steady-state FRET efficiencies between interacting eCFP-uSTAT3βtc and eYFP-uSTAT3βtc FRET pairs

FRET assays demonstrate the efficiency of energy transfer between interacting GFP-uSTAT3βtc FRET pairs (i.e. reactions generally contains 102 nM eCFP-uSTAT3βtc and 34 nM eYFP-uSTAT3βtc) contained in a reaction either supplemented with or without additives such as Mg$^{2+}$ or dsM67 DNA or phosphotyrosyl peptide inhibitor. The emission peak intensities (475 nm) were obtained following excitation at 425 nm. A plot of the FRET efficiencies versus the reaction profile was constructed. Displayed is a chart containing representative experiments analysed using FRET. The FRET efficiency is displayed at percentages.

From the data displayed in Figure 7.13, significant FRET signals represented as FRET efficiency was detected in five sets of data; including the positive control construct, eCFP-TEV-eYFP (~ 50 %); FRET pair, eCFP-uSTAT3βtc and eYFP-uSTAT3βtc (~ 12 % ); FRET pair supplemented with Mg$^{2+}$ ions (~ 18 %); FRET pair supplemented high affinity SIE DNA sequence (~ 26 %); FRET pairs supplemented with both Mg$^{2+}$ ions and the high affinity SIE DNA sequence (33%). A negative FRET efficiency (~ -58 %) was calculated for the reaction supplement with the phosphotyrosyl peptide inhibitor targeted to the STAT3-SH2 domain.

The efficiency of the energy transfer calculated for the positive control was approximately 50 %, about 4-fold stronger than that generated from the FRET reaction containing just the FRET pairs (~ 12. %). This observation suggested that distance between the fluorophores of the interacting FRET pair may be significantly greater than the positive control (tether ~ 3.8 nm). In addition, other factor such as unfavourable
orientation of the fluorophores relative to the interacting FRET pairs may have played a role in the reduced FRET efficiency obtained.

By supplementing FRET reactions with Mg²⁺ ion and/or dsM67 DNA, the calculated FRET efficiencies appreciated to 18 %, 26 % and 33 % respectively. These observations indicated that the additives may have induced changes in the interacting FRET pairs thereby either stabilizing the assembly formed or minimizing the distance between the fluorophores involved in the radiationless energy transfer.

Finally, the observed loss of FRET signals (E ≈ −58 %) in the FRET pair preparation supplemented with molar excess of the peptide inhibitor relative to the positive control (preparation containing FRET pair minus phosphotyrosyl inhibitor) (E ∼ 12 %), further confirmed the relevance of the STAT3 SH2 domain in STAT3/STAT3 interactions as well as the sequence specificity of the phosphotyrosyl peptide probe.

### 7.3 DISCUSSION

The principal objective of the work described in this chapter was to investigate interactions between the unphosphorylated uSTAT3βtc molecules contained in the GFP-uSTAT3βtc fusion protein through FRET. We developed fluorescent STAT3βtc constructs with the GFP fluorophore fused to either the N- or the C-terminus to test the four possible dimer combinations (i.e. N-N or N-C or C-N or CC) to promote efficient energy transfer. Though we were successful in purifying the N-terminus GFP tagged STAT3βtc isoforms, purification of the C-terminus isoforms proved more difficult due to protein solubility issues. However, by using the N-terminal GFP tagged STAT3βtc constructs; FRET was successfully demonstrated between the assembled eCFP-uSTAT3βtc/eYFP-uSTAT3βtc FRET pair molecules. Further, the interacting FRET pair molecules were stabilized to promote the efficient transfer of energy between the GFP fluorophores of the interacting molecules using stabilizing agents including the dsM67 DNA sequence as well as Mg²⁺ ions. Moreover, FRET between the interacting molecules was disrupted with the gp130 derived phosphotyrosyl peptide targeted to the SH2 domain of the GFP-uSTAT3βtc molecule.

Part of the work required a positive control to be constructed to quantify the FRET signal between the FRET pair. A concatemer was successfully cloned and purified for the quantitative FRET study.
7.3.1 Protein-Protein Interaction between two GFP-uSTAT3βtc molecules is demonstrated by FRET

Techniques including co-immunoprecipitation have been used to demonstrate the existence of heterodimeric complexes of STAT1 with STAT2 or STAT3 and homodimeric complex of STAT3 before activation via cytokine stimulation (Stancato et al., 1996; Novak et al., 1998; Haan et al., 2000). Novak and colleagues demonstrated that the stability of unphosphorylated full length STAT3 complex was assured only in the presence of divalent cations and that the conformation of the inactive complex may be different to that of its activated isoform. By using gel filtration chromatography, Ndubuisi reported the identification of a large mass of multi-protein complex protein to which inactive STAT3 was associated ('STATosomes'). Therefore, the study suggested that the observed inactive STAT3 dimers generated by co-precipitation may not necessarily be as a result of direct STAT3-STAT3 interactions, but may be due the incorporation of many STAT3 molecules into the observed STATosome. Due to the uncertainty of the conformation of inactive STAT3 before activation as well as the invasive nature of the techniques that have been previously employed, Kretzschmar and colleagues as well as Schröder and colleagues used FRET and BRET respectively to demonstrate the existence of non-activated full length STAT3 dimers in living cells. Several research groups have used RET- based in vitro assays to investigate protein-protein interactions. Particularly, SUMOylation (SUMO pathway) and the SUMO protease activities have been studied extensively by FRET and BRET (Martin et al., 2007; Kim et al., 2009). Further, Green and Alberola applied FRET in vitro to develop a sensor for ERK activity, following from Xu and colleagues findings working on Caspase activity sensor (Xu et al., 1998; Taw et al., 2001).

Though the crystal structures of the unphosphorylated STAT1 and 5, suggested that the core fragment dimer interface was formed from the amino acid residues of the CCD and DBD, the observation was not necessary applicable to STAT3 (Mao et al., 2005; Neculai et al., 2005) since, the same interface was largely absent in STAT3 (Ren et al., 2008). Together with the data from Multi-Angle Light Scattering (MALS) analysis, Ren and colleagues concluded that the N- and C-terminally truncated STAT3α construct (residues 127-688) was primary monomeric in solution.

To this end, a homogeneous cell-free assay approach through FRET was investigated to attempt to address the question. Becker et al reported that the STAT3 construct, N-terminally truncated STAT3βtc (residues 127-722), used in our study was predominantly monomeric (Becker et al., 1998). By employing analytical ultracentrifugation, Braunstein and colleagues also confirmed Becker's observations by
reporting that the STAT3\textsuperscript{\beta\textsubscript{tc}} construct had a significant increase of the presence of monomeric STAT3 compared to the full-length STAT3\textsuperscript{\alpha} (Braunstein \textit{et al}., 2003). On the other hand, Haan \textit{et al} demonstrated that the purified, recombinant STAT3-SH2 domain could be induced to form stable dimers at physiological pH (i.e. 7.5), but existed as monomers under acidic conditions. Also, Ren \textit{et al} demonstrated that full-length STAT3 was dimeric in low salt buffer as well as in PBS (pH 7.4), however, he observed that it was partly dimeric in high-salt buffer, an indication of a weakened dimer (Ren, \textit{et al}., 2008). Hence, on the basis of these reports, protein-protein interaction between the uSTAT3\textsuperscript{\beta\textsubscript{tc}} contained in the GFP-uSTAT3\textsuperscript{\beta\textsubscript{tc}} construct was investigated in a homogeneous assay. Furthermore, pH 7.2 rather than pH 8.5 was used since the more physiological pH resulted in a stronger eCFP-uSTAT3\textsuperscript{\beta\textsubscript{tc}}/eYFP-uSTAT3\textsuperscript{\beta\textsubscript{tc}} interaction. Hence, by using the N-terminally truncated STAT3\textsuperscript{\beta} isoform of the full length STAT3, FRET was demonstrated between the purified, recombinant eCFP-uSTAT3\textsuperscript{\beta\textsubscript{tc}} and eYFP-uSTAT3\textsuperscript{\beta\textsubscript{tc}} FRET pair.

7.3.2 The GFP fusion construct demonstrates FRET

FRET was demonstrated quantitatively with the GFP fusion protein designed to assay proteolytic activity of the TEV protease. The fusion protein contained the TEV recognition sequence (~38 Å) tethering the C- and N-termini of the eCFP and eYFP fluorophores respectively. The construct showed a 4.23-fold increase of emission ratio (475/527) when treated with the purified, recombinant TEV protease. The observation was in agreement with studies conducted with thrombin and caspase-3 probes, yielding 4.6- and 5-fold increase in emission ratio’s respectively (Zhang \textit{et al}., 2004; Luo \textit{et al}., 2001). Other GFP fusion constructs containing the weakly fluorescent BFP and GFP fluorophores have been previously reported to generate only 1.9- and 3-fold increase in emission ratio’s (Mitra \textit{et al}., 1996; Heim \textit{et al}., 1996), thus suggesting that our choice of the FRET pair, \textit{i.e}., eCFP and eYFP, represented a better candidate for the GFP FRET-based assays (Miyawaki \textit{et al}., 1997; Tsien, 1998).

Furthermore, a significant change in the emission ratio was recorded for the eCFP-TEV-eYFP construct, enforcing its appropriateness for the direct determination of enzyme kinetics. By titration, a smooth dose response curve for the TEV protease was observed, Figures 7.3 and 7.4. A significant increase in the eCFP emission peak (475 nm) was observed with 10 μg/mL TEV protease. Though intact residual GFP fusion constructs were evident after treatment with 100 μg/mL TEV protease, Figure 7.4, a FRET efficiency value of 50% was generated (Figure 7.3 (A), (B), (C)). Since FRET is distance-dependent (typically 10-100 Å) and the recommended Förster distances for studying biological macromolecules ranges from 20 Å to 90 Å, the recorded 50 %
FRET efficiency was sufficient energy transfer for the study. The value obtained was fairly consistent with that reported by Han and colleagues which was approximately 60%. Also, by using shorter thrombin recognition sequence (LVPR) between the GFP fluorophores, Zhang and colleagues could extend the linker with a GS amino acid sequence (GS sequence is considered the most favourable linker repeat due to the flexibility it confers) at the N- and the C- terminals, i.e., GSLVPRGS. This resulted in a 100% FRET efficiency. Even though Zhang et al. applied a linker that was a residue longer than the TEV recognition sequence (ENLYFQQ) used in this study, they still reported an enhanced efficiency of about 50% relative to that reported with the eCFP-TEV-eYFP construct, thus suggesting that the thrombin linker sequence may enable the GFP fluorophores to assume a more favourable orientation to allow efficient transfer of energy. Further, the G-S repeat also adds flexibility to the peptide linker, hence, commonly used in linker design (Argos, 1990).

7.3.3 FRET between eCFP-uSTAT3βtc and eYFP-uSTAT3βtc

Following direct excitation of the homogenous FRET mixture at 425 nm, the decrease in the donor emission at 475 nm (donor quenching due to resonance energy transfer to the acceptor molecule) and also the concomitant increase in acceptor emission at 527 nm were both used as direct measures of the efficiency of energy transfer between the interacting purified, recombinant uSTAT3βtc molecules fused to the eCFP and eYFP fluorophores at the N-terminus. Preparations including donor only reactions (assay without energy transfer acceptor that controls for direct excitation of the donor molecule) were applied as controls to contrast the FRET signals measured (Figure 7.5). Similarly, control reactions for non-specific interactions between the fluorescent tags yielded a slope value (2.92) close to that obtained for direct excitation of the donor species (2.67), suggesting that non-specific interaction is negligible and that the measured FRET was indeed due to the interaction between uSTAT3βtc molecules contained in the FRET construct (Figure 7.6).

Using a competition-based approach to probe the specificity of the FRET signals, we observed 12% FRET efficiency with equimolar mixture of the donor and acceptor FRET constructs. The value was enhanced by 54% with 3-fold molar excess of the donor construct. On the other hand, a significant drop (approximately 90%) in the efficiency of energy transfer was observed with 10-fold molar excess of the donor interaction partner. On one hand, the observations suggested that the FRET efficiency did not only scale with the distance and the orientation between the FRET pair, but also with the probability that the donor species was associated with an acceptor species. On the other hand, the observed drop in FRET efficiency in response to the molar excess
of the donor species could be attributed to non-specific aggregation of the GFP-STAT3ptc molecules induced by molecular overcrowding. A similar data was obtained with molar excess of the unlabelled uSTAT3ptc molecules that occupy the protein/protein interaction interface (data not shown).

Since the distance requirement for efficient energy transfer between the two fluorophores should be less than 100 Å (or < 10 nm), FRET signals would only be observed when the GFP-fused parent proteins were in close proximity. Hence, on that basis, the observations suggested that the equilibrium between the assembled fluorescent STAT3ptc molecules, i.e. between two donor molecules (eCFP-uSTAT3ptc), two acceptor molecules (eYFP-uSTAT3ptc) and between a donor and an acceptor molecule (eCFP-uSTAT3ptc and eYFP-uSTAT3ptc) was created with equimolar concentrations of the FRET pair. However, it was suggested that the generated equilibrium was disturbed in the presence of excess donor species (e.g. 3 to 10-fold molar excess of donor species relative to acceptor species), thereby resulting in a bias in favour of the eCFP-uSTAT3ptc/eCFP-uSTAT3ptc (no FRET) and eCFP-uSTAT3ptc/eYFP-uSTAT3ptc (enhance FRET up to 3-fold molar excess) assemblies. This effect may have accounted for the transient nature of the interaction. Even though our results agree with studies conducted by other research groups that carried out the investigation in living cells, we could not confidently state that the observed FRET signals were as a result of dimer formation between a simple FRET pair since we could not ignore the existence of higher-order aggregates or complexes, as had been suggested by Ndubuisi and colleagues (Ndubuisi et al., 1999) and also demonstrated in chapter 4. Moreover, the FRET assay could not be used for absolute structural characterization, hence, could not be applied to distinguish between FRET signals generated from either dimers or higher-order complexes.

However, it was clear that the change in pH from 8.5 to 7.2 encouraged protein/protein assembly in a favourable conformational orientation and therefore resulted in the recorded FRET signals.

7.3.4 FRET is stabilized in the presence of certain ingredients.

Contents of the FRET buffer was investigated to elicit the optimum conditions necessary to improve and sustain FRET signal. Since it was unclear which factors would contribute to generating a sustained FRET signal, we investigated the ionic strength as well as the viscosity of the FRET buffer via the sodium chloride (NaCl) and the glycerol content of the homogeneous FRET mixture. Since efficient energy transfer
requires close proximity of the interacting FRET pair, the experiments were limited by the separation achieved with the N-terminal fusions.

Several *in vitro* FRET studies have used a fixed concentration of NaCl in their investigations, thus ignoring the effect of the ionic strength on protein-protein interaction and the subsequent FRET signal measured. Hence, based on the solubility issues associated with the fluorescent STAT3βtc construct, we investigated protein-protein interaction under different salt concentrations (*i.e.* 0 to 500 mM NaCl). A gradual increase in the efficiency of energy transfer, up to concentrations of 300 mM NaCl was recorded (Figure 7.7). Though no significant difference in the FRET efficiency was recorded for the 100 (32 %) and 150 mM (33 %) salt buffer, near doubling (*i.e.* 57 %) of the FRET efficiency value was obtained with the reaction supplemented with 300 mM NaCl. The increase in the FRET efficiency could be attributed to improved solubility and concomitant increase in protein-protein interaction between FRET constructs. However, in 500 mM NaCl, the FRET pair recorded a steep drop in FRET efficiency to approximately 5 %. The observation indicated that high ionic strength of the buffer affected interactions between FRET pair that may have resulted in the reduction in the efficiency of energy transfer between the FRET pair molecules.

In a related study, we investigated the stability of the assembled eCFP-uSTAT3βtc/eYFP-uSTAT3βtc FRET complex in the homogenous reaction mixture by applying the crowding effect. Macromolecular crowding has been reported to enhance protein structure and functions including protein-protein interactions in cells (Minton, 2005; Ellis et al., 2001). In addition, McPhie and colleagues as well as McNulty and group suggested that the crowding effect improves protein stability and prevent aggregate formation respectively (McPhie, et al., 2006; McNulty, et al., 2006). Hence, based on these findings, glycerol was used to create crowding effect to mimic the cellular environment. Significant improvement in the FRET efficiency was recorded for the reactions treated with 5 % glycerol (Figure 7.8), with the highest FRET efficiency (*i.e.* near 80 % increment) recorded for the reaction containing 3-fold molar excess (*i.e.* 102 µM) of the donor species. Increasing the percentage glycerol content to 10 or even 15 % increased the FRET efficiency values to 68 and 77 % respectively. However, at 30 % glycerol, a drop in FRET efficiency was observed, thus suggesting that at concentrations above 15 %, glycerol were interfering with the FRET pair assembly. Hence, it was suggested that the FRET constructs might have assumed a much more compact and stable conformation in the presence of up to 15 % glycerol, corroborating that the crowding effect encourages protein-protein interactions and therefore efficient energy transfer between the FRET pair. It is suggested that the end to end distance
between proteins inside cells are relatively shorter than the purified counterpart in solution (Ohashi et al., 2007), hence, crowding is encouraged in protein-protein interaction studies.

7.3.5 Co-factor stabilizes FRET pair assembly

Novak and colleagues reported that the full-length STAT3 homodimerizes prior to activation and that the process requires the presence of C-terminus of STAT3 as well as a divalent cation including Mg\(^{2+}\) ion (Novak et al., 1998). Furthermore, described in chapter 6 was the ability of the high affinity c-fos sis inducible oligonucleotide DNA (dsM67) to promoted GFP-uSTAT3βtc/dsM67 DNA complex formation. The specific binding of the GFP-uSTAT3βtc construct to its natural recognition site may provide stability and could therefore induce a structural conformation that would promote efficient energy transfer between the interacting FRET pair, eCFP-uSTAT3βtc/eYFP-uSTAT3βtc.

By incubating the FRET reaction with 5 mM MgCl\(_2\), we observed a significant increase in the FRET efficiency by 77 % and 37 % for the preparations containing 1:1 and 1:3 donor to acceptor molar ratios respectively (Figure 7.10). The enhanced FRET efficiency suggested was an indication of stabilizing effect of the Mg\(^{2+}\) cation. However, the enhanced FRET efficiency was reversed with 10-fold molar excess of the chelating agent, EDTA, suggesting that the observed elevation in the FRET signals was indeed due to the presence of the Mg\(^{2+}\) ions that may be engaged in stabilizing the FRET pair assembly. Though the Mg\(^{2+}\) ions may not be involved in actual bond formation, it was thought to be essential in stabilizing the eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly via the uSTAT3βtc interactions. In a related experiment, however, a similar preparation involving impure bacterial cell extract of the FRET pair resulted in a significantly reduced FRET efficiency, \textit{i.e.}, 13 and 7 % for 1:1 and 1:3 donor to acceptor molar ratio’s respectively, relative to those achieved with the pure fractions. The observation indicated that there may be contaminant proteins molecules in the cell extract that either specifically interfere with the assembly of the FRET pair or may be binding the divalent ions needed to stabilize the FRET pair assembly.

Even though Novak and colleagues reported that the C-terminal domain was essential for the formation of a stable inactive homodimer of STAT3, the STAT3 construct employed in his study was different to that used in our study, \textit{i.e.}, STAT3βtc, amino acid residues 127-722. Novak et al employed the full length wild-type STAT3 as well as a truncated isoform which had lost 85 amino acids from the C-terminus, including the Y^705 but retained the SH2 domain. Therefore, on that basis, it was right to suggest that
the stabilizing domain of divalent ion (e.g. Mg²⁺) may well fall between residues 688 and 722 of the STAT3βtc construct, thus suggesting that the eCFP-uSTAT3βtc/eYFP-uSTAT3βtc FRET pair assembly required a functional SH2 domain to associate. In support, Kretzschmar and colleague had reported that a point mutation to render STAT3 SH2 domain incapable of performing its biological functions including "docking" at the cytokine receptor and the subsequent activation via tyrosine phosphorylation in response to cytokines, resulted in the complete loss of FRET signal by the mutant STAT3 molecule even in the absence of the cytokines (Kretzschmar et al., 2004). These and other observations described above suggested that our findings were consistent with those of Novak and colleagues which suggested that stable and inactive STAT3 dimer formation requires the presence of the C-terminal domain and a divalent ion.

Again, in chapter 6, Figure 6.1, a surprising observation was uncovered which suggested that, like the phosphorylated or activated isoform of GFP-pSTAT3βtc, the unphosphorylated isoform, GFP-uSTAT3βtc, also did bind to and form a stable complex with the dsM67 DNA. The stability of the GFP-uSTAT3βtc/dsM67 DNA complex was demonstrated by relative mobility shift of the GFP-uSTAT3βtc protein in the presence of the high affinity dsM67 DNA. Hence, on the basis of this finding, a preparation containing the FRET pairs was also supplemented with 10-fold molar excess of the high affinity dsM67 DNA. A 50 % rise in the FRET efficiency was recorded for the dsM67 DNA treated preparations (Figure 7.11). The observation remained unchanged when an identical preparation was “spiked” or competed with a 30-fold molar excess of sonicated non-specific salmon sperm DNA fragments. This suggested that the FRET pair/dsM67 DNA interaction was specific and led to the formation of a stable protein/DNA complex that was stable and in an orientation that allows energy transfer between the donor and acceptor fluorophores. Also, based on the crystal structure of activated STAT3 dimer, where the C-termini are in close proximity whilst the N-termini are separated by approximately 80 Å (Becker et al., 1999), it was likely that FRET would occur more efficiently with C-terminally tagged GFP-pSTAT3βtc constructs than with the N-terminally tagged GFP-uSTAT3βtc construct. Further to the distance requirement of FRET, the transition dipole of the donor and acceptor species must be oriented favourably relative to each other. Based on that, it was reasonable that GFP-uSTAT3βtc binding to the dsM67 DNA influences the orientation of the GFP tag and enhances the ability of the DNA-bound FRET pair to elicit efficient energy transfer. Thus, the observed FRET signals may reflect the formation of a higher form of dimers or a favourable conformational change induced by the presence of the dsM67 DNA. Though these observations could not be verified via
FRET, it seems reasonable to suggest that the eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly assumes a significantly different structural conformation when in complex with the DNA.

Furthermore, on the basis of the observation described in chapter 6, where the presence of the gp130 phosphotyrosyl peptide resulted in a dose-dependent reduction in the level of GFP-uSTAT3βtc binding to the dsM67 DNA, the ability of the phosphotyrosyl peptide to inhibit or disrupt the (GFP-uSTAT3βtc)₂/dsM67 DNA complex was investigated via FRET. Following preparation and evaluation of spectra, a near negative FRET efficiency value was recorded for the peptide treated preparation relative to the 26 % recorded for the untreated sample. A similar inhibition profile was observed for a corresponding reaction containing no stabilizing dsM67 DNA probe (Figure 7.12 (C)). These observations did not only agree with the findings in chapters 5 and 6 but with those of other groups that demonstrated the specificity of the pY modified residue (H-pYLPQTV-NH₂) to the STAT3 SH2 domain and their concomitant inhibitory and disruptive potency. Moreover, the observations also demonstrated that like the phosphorylated pSTAT3βtc dimer isoforms, the unphosphorylated uSTAT3β inactive dimers also required an intact SH2 domain for stability. These allowed speculation that perhaps, the uSTAT3 inactive dimer may not necessarily assume a completely different conformation upon activation.

Recognising the inhibitory activity of the phosphotyrosyl peptide, the N-terminal 5-carboxyfluorescein (5-FAM) labelled H-pYLPQTV-NH₂ was used to investigate the pY-SH2 domain interaction and the conformational orientation that the eCFP-uSTAT3βtc/peptide heterocomplex may assume relative to the GFP fluorophore and the 5-FAM dye. Near identical FRET efficiency values were obtained for the test (where 5-FAM labelled phosphotyrosyl peptide was used as the acceptor species) and control (where eYFP-uSTAT3βtc was used as the acceptor species) FRET reactions (Figure 7.12 (A)). The observation further supported the affinity of the phosphotyrosyl peptide sequence as well as providing an insight into the conformational orientation of the eCFP-uSTAT3βtc/5-FAMG-pYLPQTV-NH2 heterocomplex relative to their N-terminus reporters. On the basis of the observed FRET, it was safe to suggest that the protein/peptide heterocomplex assumed a favourable orientation that allowed efficient transfer of energy between the two fluorescent reporters (*i.e.* with 10 nm of each other). In a related study, a corresponding unlabeled peptide was investigated for its ability to inhibit interaction between the donor species and the 5-FAM labelled counterpart. As expected, we recorded near 60 % drop in FRET efficiency upon incubating the FRET pair mixture (eCFP-uSTAT3βtc/5-FAM-peptide) with 6-fold molar excess of the unlabelled peptide sequence. The observation further confirmed that the eCFP-
Edwin Nkansah

uSTAT3βtc/5-FAM-peptide assembly was indeed mediated by the pY-SH2 domain interactions. Further, the observed basal FRET efficiency (~20%) even at 6-fold molar excess of the unlabelled inhibitor sequence could not be explained; however, it was speculated that the signal may have arisen from transient non-specific interactions between the GFP fluorophore and the 5-FAM dye molecule. Moreover, findings in chapter 6 coupled with those of other research groups suggested that the free N-terminus of the unlabelled peptide inhibitor may decrease the affinity of the phosphotyrosyl for the SH2 domain of the donor species (Turkson et al., 2001). No further work was carried out with the acetylated derivative to investigate this observation.

In summary, a homogeneous assay based FRET was developed and used to investigate the structural and conformational status of the N-terminal truncated STAT3β isoform of STAT3α tagged at the N-terminus with spectral variants of GFP, GFP-STAT3βtc. FRET was successfully used to demonstrate the induction of eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly as well as the disruption of the assembled complex with phosphotyrosyl peptide inhibitors specific to the STAT3 SH2 domain. FRET signals from assembled FRET pairs were successfully enhanced and/or stabilized using additives including Mg²⁺ ions, glycerol and dsM67 DNA which could be described as factors modulating the protein/protein assembly. The conformational changes implicated on the basis of the data reported in this study on the efficiency of the energy transferred between the interacting GFP-uSTAT3βtc molecules will require further verification through alternative structural studies.
8 GENERAL DISCUSSION

The principal aims of the work carried out in this thesis were to investigate and understand some of the biological functionalities of STAT3. Fluorescence-based isolated cell-free systems were used in order to learn about the conditions that mediate these functionalities and to facilitate the discovery of small drug-like molecule capable of interfering with the activities of the recombinant protein.

Several research groups have reported the existence of varied amounts of nuclear STAT3 prior to cytokine stimulation in living cells (Haan et al., 2000; Meyer et al., 2002). Whilst both reports were based on co-precipitation of STAT3 protein molecules, Kretzchmar and colleagues investigated the structural and conformational integrity of STAT3 through FRET which allowed visualization of STAT3/STAT3 protein interactions in living cells. Similarly, they reported seeing up to 40% of the fluorescent STAT3 construct in the nucleus prior to cytokine stimulation. Furthermore, they reported that STAT3 species bearing crucial alterations including the Y705F-STATS and the R609Q-STATS mutants demonstrated similar subcellular distributions as the wild-type STAT3, thus suggesting that nuclear translocation of inactive STAT3 may be independent of tyrosine phosphorylation and dimer formation. It is widely accepted that activation by tyrosine phosphorylation is required for the DNA binding activity of STAT3, dimer formation through pY-SH2 interaction is therefore crucial to facilitate STAT3 nuclear translocation and DNA binding activity. Therefore, the revelation of the existence of nuclear dimers considered inactive and lacking tyrosine phosphorylation contradicts the standard STATS signalling paradigm and therefore the point of intervention in relation to drug discovery. Furthermore, it also begs the question, is dimerisation or phosphorylation actually a prerequisite for nuclear translocation and/or DNA binding?

To this end, fluorescent STAT3βtc constructs containing the spectral variant of the GFP protein fused to either the C- or the N-terminus of the N-terminally truncated STAT3β isoform of the full length STAT3 protein were developed to functionally characterize the phosphorylated and unphosphorylated isoforms of STAT3. Although significant levels of expression were observed for the C-terminus GFP tagged STAT3βtc construct, STAT3βtc-GFP, no functional soluble fraction was obtained since near 100% were incorporated into inclusion bodies. In contrast, the N-terminus GFP tagged STAT3βtc construct, GFP-STAT3βtc, designed to substitute the missing N-terminal 126 amino acids of STAT3β with spectral variants of the GFP protein was a functional construct which enabled milligram amounts of the fluorescent GFP-STAT3βtc construct to be purified from E.coli. Phosphorylated (active) isoform of the GFP-STAT3βtc construct,
GFP-pSTAT3βtc, was generated by the co-expression of the protein-tyrosine kinase domain of the Elk receptor in the BL21 (DE3) TKB1 E.coli cells. Both active (dimer; GFP-pSTAT3βtc) and inactive (monomer; GFP-uSTAT3βtc) protein constructs were structurally characterized by size exclusion chromatography following a three-step purification protocol that included ammonium sulphate purification, anion exchange chromatography and size exclusion chromatography. However, almost all the functional assays were performed without the third purification step.

Following from characterization, the stability of the purified protein construct was investigated by Native Triton X-100 polyacrylamide gel electrophoresis. Even though purified protein preparations were properly store either at -80 °C in 10 % glycerol (long term storage) or at 4 °C (short term storage), unlike the GFP-pSTAT3βtc isoforms, the GFP-uSTAT3βtc isoforms was prone to extensive aggregation in response to increasing temperatures, including 37 °C. However, it appeared that the protein aggregates were stabilized by inter-chain disulphide bonds which were consistent with the ability of DTT to reduce, albeit during incubation with DTT at temperatures less than 24 °C. The sharp contrast in conformational stability observed between the GFP-STAT3βtc isoforms (i.e. phosphorylated and unphosphorylated constructs) suggested that activation via phosphorylation at Y705 residue and the subsequent dimerization conferred extensive structural stability even at temperatures of up to 50 °C.

Following from the design, purification and characterization of the GFP-STAT3βtc chimaeric protein, three distinct, but complementary biochemical techniques including ELISA, PEMSA, and FRET were adapted into fluorescence-based techniques to evaluate and functionally characterize the purified, recombinant, GFP-STAT3βtc isoforms.

Both purified, GFP-STAT3βtc isoforms were functionally characterized using biotinylated native targets immobilized on a streptavidin coated surface. The principal functionality of the SH2 domain contained in the GFP-uSTAT3βtc construct was successfully assayed using the phosphotyrosyl peptide encompassing Y905 of the gp130 signal transducer. Similarly, significantly high binding data was recorded when the construct was assayed with the unlabelled and free high affinity dsM67 DNA, though little or no binding interaction was measured when the construct was exposed to the biotin labelled and immobilized version of the probe. The observations did not only confirm the retention of the functionalities of the two crucial domains contained in the STAT3βtc protein, but also suggested that the GFP reporter did not affect the functional conformation of the construct. Though GFP-uSTAT3βtc/phosphotyrosyl
peptide interaction was in good agreement with published data (Haan et al., 1999; Ren et al., 2003; Schust et al., 2004; Lehmann et al., 2006), the GFP-uSTAT3βtc/dsM67 interaction had not been reported in literature, hence, the observation was further explore.

In the case of the active isoform, GFP-pSTAT3βtc, the DNA binding was successfully assayed with both labelled and unlabelled high affinity dsM67 DNA. Significant binding data was obtained for both versions of the DNA probe, thus suggesting that the bacterially expressed purified active construct retained DNA binding functionality. Very little or no binding was observed between the SH2 domain contained in GFP-pSTAT3βtc and the biotin labelled gp130 derived phosphotyrosyl peptide. The observation suggested a stable reciprocal phosphotyrosyl-SH2 domain interaction, thus a stable STAT3β dimer.

The binding data acquired from this functional assay was principally generated by measuring the fluorescence of the GFP reporter fused to the STAT that was the captured molecule. The assay is quick, less laborious, efficient and sensitive, though sensitivity was very much dependent on choosing excitation and emission filters appropriate to the excitation and emission wavelengths of the GFP reporter contained in the protein construct.

Furthermore, by adapting the classical EMSA (Fried and Crothers, 1981) into a technique we term PEMSA, the differential dsM67 DNA binding between GFP-pSTAT3βtc and GFP-uSTAT3βtc was demonstrated as a protein/DNA complex mobility shift relative to the unbound GFP-STAT3βtc protein. Although the GFP-pSTAT3βtc/dsM67 DNA complex formation was expected (Shuai et al., 1994), the GFP-uSTAT3βtc/dsM67 DNA complex was not expected and has not been reported in literature. The shift observed for the GFP-uSTAT3βtc/dsM67 DNA complex was reversible with 10-fold molar excess of the gp130 phosphotyrosyl peptide derivative which competes for the dimer interface. No reversal of the mobility shift was observed for the GFP-pSTAT3βtc/dsM67 DNA complex with up to 333-fold molar excess of the gp130 phosphotyrosyl peptide derivative. The differential mobility observed between DNA bound and unbound GFP-STAT3βtc isoforms suggested that the structural conformation of both isoforms may be varied (as judges by the crystal structures of both) and may have accounted for the differences in affinity for the DNA probe, the phosphotyrosyl peptide probe and the stability of the DNA bound protein complex.

DNA binding and inhibition reactions were analysed on non-denaturating native polyacrylamide or agarose gels. The separated samples were visualized under blue
light 470 nm to reveal the relative mobility shift between DNA bound and unbound fluorescent STAT3βtc proteins as opposed to the case of classical EMSAs where the relative mobility shift is observed between bound and unbound ³²P radio-labelled double stranded oligonucleotide DNA probe.

Finally, GFP-based FRET was used to demonstrate protein/protein assembly between unphosphorylated uSTAT3βtc FRET pair molecules, *i.e.*, the donor (eCFP-uSTAT3βtc) and the acceptor (eYFP-uSTAT3βtc) species, in a homogeneous assay. The eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly was principally achieved by changing the pH from 8.5 to 7.2 which may have resulted in a favourable orientational assembly and the concomitant efficient energy transfer between the FRET fluorophores. The observation that the change in pH to the physiological pH (in our case slightly below) encouraged protein/protein assembly correlates with previous reports that the full-length STAT3α as well as its isolated SH2 domain forms dimers at physiological pH (7.4) (Ren *et al.*, 2008; Haan *et al.*, 1999). Buffer additives including Mg²⁺ cation and glycerol were identified as stabilizers of eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly. Furthermore, the high affinity dsM67 DNA demonstrated a significant high stabilizing effect on the FRET pair assembly which eventually resulted in a 3-fold increment in the recorded FRET efficiency signals. Although the stabilizing effect of the dsM67 DNA on protein assemblies in a homogeneous assay have not previously been reported in literature, that of the Mg²⁺ ion on the uSTAT3 dimer was documented by Novak and colleagues and adapted for our study. Moreover, we detected a specific interaction between the SH2 domain of the donor GFP labelled species and the pY motif of the 5-carboxyfluorescein labelled gp130 phosphotyrosyl peptide probe through the efficiency of energy transfer between the two fluorophores. Also, an unlabelled corresponding peptide was competed to demonstrate the loss of FRET signal either between eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly or eCFP-uSTAT3βtc/5-carboxyfluorescein-G-pYLPQTV-NH₂ heterocomplex, thus suggesting that a specific SH2 domain interaction was essential for the successful eCFP-uSTAT3βtc/eYFP-uSTAT3βtc assembly.

In addition to the FRET signals generated with the GFP-uSTAT3βtc complexes, a much more sustained and efficient energy transfer was recorded for the positive control GFP-chimaeric protein construct containing the TEV protease recognition sequence tethering the C- and N-terminal of the donor (eCFP) and acceptor (eYFP) fluorophores. The eCFP-TEV-eYFP construct was basically applied to demonstrate FRET as well as for comparison with the FRET efficiency between GFP-uSTAT3βtc assemblies.
Thus a FRET-based homogeneous assay has been developed that can be used for screening chemical libraries in a high-throughput format, and has herein been used to study protein-protein interactions.

8.1 FUTURE WORK

Due to time constraints, it was not possible to fully exploit the potentials of the fluorescent STAT3βtc constructs. Thus the future work plan would more fully elucidate the structure/function relationship of the GFP-STAT3βtc construct.

- Together with the amount of structural and functional information gathered in this study, it would be appropriate to obtain sufficient material for crystallographic studies including co-crystallization with DNA oligonucleotides, phosphotyrosyl peptides and drug-like small molecule inhibitors which would facilitate our understanding of the GFP-STAT3βtc/GFP-STAT3βtc assembly, DNA-binding, receptor-binding and GFP-STAT3βtc/drug interactions. Furthermore, quick and easy access to large quantities of pure and functional GFP-STAT3βtc would be beneficial for large scale drug screen in a high-throughput format.

- Mutational analysis of the residues involved in the classic pY-SH2 domain interaction (Lys 591, Arg 609, Ser 611, Ser 613 and Tyr 705) as well as those implicated in the stabilization of inactive STAT3 dimer.

- Re-engineering of the C-terminal fused STAT3βtc-GFP to obtain soluble protein constructs that may be used to further investigate the extent to which the biological functionalities of the fluorescent STAT3βtc construct has been influenced by the location of the GFP.
REFERENCES


Edwin Nkansah


Ferguson, K.A. (1964) Starch-gel electrophoresis-application to the classification of pituitary proteins and polypeptides. Metabolism, 13(10), 985-1002.


Lillemoier, B.F., Köster, M. and Kerr, I.M. STAT1 from the cell membrane to the DNA. *EMBO J*, 20, 2508-2517.


O'Shea, J.J. (1997) JAKs, STATs cytokine signal transduction, and immunoregulation: are we there yet? Immunity, 7(1), 1-11.


