

University of Nottingham
UK | CHINA | MALAYSIA

Unphosphorylated dimers are not required for the assembly of active STAT1 and STAT3 homo- and heterodimers

Sebastian Christoph Krause

Student ID: 4342277

Thesis submitted to the University of Nottingham

for the degree of

Doctor of Philosophy

March 2023

Abstract

STATs are latent signalling proteins and transcription factors that are activated by phosphorylation upon ligand binding to cytokine receptors. All seven mammalian STATs form dimers upon activation. So far, research focused on homodimers, but it is now recognised that some STATs can also form heterodimers. Yet, it is unclear what determines the formation of individual dimer types, especially when multiple STATs get activated simultaneously. Several STATs assemble inactive, unphosphorylated dimers. However, their structure is different from the active signalling complex. Therefore, it is conceivable that unphosphorylated STAT dimers may regulate the composition of active signalling complexes. This would mean that activated complexes arise from a pre-defined pool of unphosphorylated STAT homo- and heterodimers via a cytokine-induced change in dimer conformation. Alternatively, active signalling complexes may form independent of their unphosphorylated counterparts. This work describes two new cell-based assays and their application with a focus on STAT1 and STAT3, whose capability to form active heterodimers is relatively well characterised, to test these alternatives. Firstly, the Nuclear Translocation Assay is based on the ability of STATs to accumulate inside the nucleus after cytokine treatment. This phenotype is mediated by a nuclear localisation signal (NLS) sequence and can be disrupted by mutation. It was found here that said mutation is dominant-negative for STAT1, STAT3 and STAT5B nuclear import. Therefore, it is possible to assess active dimer formation between an accumulation mutant and another STAT by observing their cytokine-induced nuclear accumulation pattern. If the two proteins form active dimers, cytokine-induced nuclear accumulation of both is diminished. In turn, if they do not form active dimers, cytokine-induced nuclear accumulation is diminished for the mutant but not the other STAT. In the Nuclear Translocation Assay, this effect is made visible in live-cell microscopy by fusing the putative interaction partners with different fluorescent proteins. Testing several STAT1 and STAT3 wild-types and mutants, that disrupt the unphosphorylated but not active dimers, using this assay revealed that active homo- and heterodimers formed independent of their unphosphorylated counterparts. Secondly, the Co-localisation Assay was used to assess the formation of unphosphorylated dimers between different STATs. The assay

takes advantage of the observation that unphosphorylated STATs shuttle to and from the nucleus constitutively and that the balance of this can be altered in favour of either the nucleus or cytoplasm by fusion with an external signal sequence. If a certain bait-STAT fusion protein can form latent dimers with a co-expressed test-STAT, then both proteins co-localise in the targeted compartment as visualised by different fluorescent protein tags. If not, then co-localisation is diminished. Applying this to STAT1 and STAT3 revealed that both paralogues readily formed unphosphorylated homodimers but not heterodimers. Overall, the findings from both assays suggest that active STAT1 and STAT3 homo- and heterodimer formation is independent of prior assembly as unphosphorylated dimers. Furthermore, the evidence presented here suggests that their homo- and heterodimers may form at equal rates.

Acknowledgements

I would like to thank the great host of friends and family who have patiently backed me while generating this work. Scientific research can be daunting at the best of times, and even more so when disrupted by a global pandemic such as Covid-19. Yet, a caring support network and a great team will undoubtedly guarantee success and this work is the best example of this.

Firstly, I would like to thank Professor Dr Uwe Vinkemeier for his continued support during the 4.5 years in his laboratory and for the opportunity to work alongside him. He has taught me the importance of robust, high-quality science to challenge the status quo and further our understanding of biological mechanisms. Many thanks also to Dr. James Cavey and Dr. Andreas Begitt for continuously sharing their immense experience in laboratory research with me. I would also like to thank them for the fun evenings in the pub and elsewhere we shared over the course of my stay and will hopefully continue to do so in the future. Further thanks to Dr Liu He for her moral support, the most amazing cakes and for her scientific advice.

Special thanks go out to members of the D61 office for their company. I will miss the fun times we had at lunch and outside of work and the many stories that came about or were shared during our time together.

Finally, I want to thank my girlfriend Emma Snow who was always supportive and has brought a lot of joy into my life and my lovely mum Angelika Krause without whom I would have never made it this far and who has always given her all to make sure that I have a bright future.

Declaration

I hereby declare that this thesis and the work reported herein were composed by and originated entirely from me, except for experiments depicted in figures 62-66 of chapter 4.5 which were conducted by Dr. Andreas Begitt. Information derived from the published and unpublished work of others has been labelled in the text and references are listed at the end of this work.

Signed,

_____

Sebastian Christoph Krause

Experiments discussed in this thesis were conducted in the laboratory of Prof Dr Uwe Vinkemeier at the School of Life Sciences, University of Nottingham, Nottingham, UK.

Primary Supervisor: Prof. Dr. Uwe Vinkemeier

Secondary Supervisor: Dr. David Scott

Publications

Begitt, A., **Krause, S**., Cavey, J. R., Vinkemeier, D. E., & Vinkemeier, U. (2023). A family-wide assessment of latent STAT transcription factor interactions reveals divergent dimer repertoires. Journal of Biological Chemistry, 299(5).

Table of Contents

Abstract I
Acknowledgements III
Declaration IV
Publications V
Table of Contents VI
Abbreviations X
1. Introduction 1
1.1 Cytokine biology: Mediators of immune cell communication 1
1.2 Cytokine signalling: The Jak-STAT pathway. 4
1.3 STAT structure and dimerization 8
1.3.1 STAT domain structure 9
1.3.2 STAT dimerization 12
1.4 Nuclear translocation of STATs 14
1.5 Areas of limited understanding relevant to this work 18
1.6 Hypothesis and aim of this thesis. 22
2. Materials 24
2.1 Media 24
2.1.1 Media for bacterial work 24
2.1.2 Media for mammalian cell work 25
2.2 Buffers 25
2.2.1 Buffers for DNA work 25
2.2.2 Buffers for protein work 26
2.3 Organisms 29
2.4 Software 30
3. Experimental Procedure 32
3.1 Working with DNA 32
3.1.1 Polymerase chain reaction (PCR) 32
3.1.2 Gel electrophoretic separation of DNA molecules 32
3.1.3 Restriction enzyme digestion 33
3.1.4 Molecular cloning 33
3.1.5 Colony PCR 35
3.1.6 Sanger sequencing 36

3.2 Working with Bacteria 37
3.2.1 Generation of Chemically Competent Cells: Inoue Method 37
3.2.2 Transformation into Chemically Competent *E. coli* Cells 38
3.2.3 Plasmid isolation from *E. coli* 38
3.3 Mammalian cell culture 39
3.3.1 Initiation of mammalian cell culture 39
3.3.2 Cell Culture Maintenance 39
3.3.3 Generation of frozen cell stocks 39
3.3.4 Transient transfection of cell lines using Lipofectamine LTX and Plus reagent 40
3.3.5 Cytokine treatment of cells 41
3.3.6 Culturing cells for live-cell imaging 41
3.4 Working with Proteins 42
3.4.1 Whole cell extraction for Western blot analysis 42
3.4.2 Whole cell extraction for EMSA 42
3.4.3 Bradford Assay 43
3.4.4 SDS Polyacrylamide gel electrophoresis (SDS-PAGE) 43
3.4.5 Western blotting 44
3.4.6 Electrophoretic mobility shift assay - EMSA 45
3.5 Fluorescence Microscopy 46
3.5.1 Widefield Microscopy 46
3.5.2 Live-Cell Imaging and Confocal Microscopy 46
3.5.3 Foerster Resonance Energy Transfer (FRET) 47
3.5.4 Nuclear Translocation Assay Image Analysis Workflow 48
3.5.5 Image Analysis Workflow for the Co-localisation Assay 52
4. Results 54
4.1 Assay Development: The Nuclear Translocation Assay 54
4.1.1 Theoretical Considerations ..... **Error! Bookmark not defined.**
4.1.2 Measuring relative Protein Amounts based on fluorescence intensity. 59
4.1.3 Dynamics of STAT1, STAT3 and STAT5B nuclear accumulation 73
4.1.4 A conserved stretch of basic amino acids in the DBD of STATs determines cytokine-induced nuclear accumulation. 85
4.1.5 Discussion of Chapter 4.1 92
4.1.6 Nuclear Translocation Assay protocol 99

4.2 Assay Validation: Testing the Nuclear Translocation Assay using known interactors and known non-interactors. 102
4.2.1 The STAT1ΔNLS mutation is dominant negative. 102
4.2.2 The STAT3ΔNLS mutation is dominant negative. 111
4.2.3 The STAT5BΔNLS mutation is dominant negative. 118
4.2.4 STAT5BΔNLS does not inhibit cytokine-induced nuclear import of co-expressed STAT1wt. 125
4.2.5 Discussion of Chapter 4.2 134
4.3 Active STAT1 and STAT3 homodimers form independent of unphosphorylated dimer integrity. 136
4.3.1 Theoretical considerations **Error! Bookmark not defined.**
4.3.2 Characterising point mutations of STAT1 residue F172 140
4.3.3 The monomerisation of unphosphorylated dimers does not affect the formation of activated STAT1 homodimers. 145
4.3.4 Characterising STAT3-L78R 151
4.3.5 The monomerisation of unphosphorylated dimers does not affect the dimerisation of activated STAT3 constructs. 155
4.3.6 Discussion of Chapter 4.3 162
4.4 Unphosphorylated dimers do not pre-determine the formation of active STAT1/STAT3 heterodimers. 165
4.4.1 Theoretical Considerations **Error! Bookmark not defined.**
4.4.2 Active STAT1 and STAT3 readily interact despite monomerization of their latent dimers. 168
4.4.3 Monomerization of latent dimers does not lead to the creation of artificial active STAT1/STAT5B heterodimers. 186
4.4.4 Discussion of Chapter 4.4 200
4.5 Detecting unphosphorylated STAT dimers using the Co-localisation Assay 203
4.5.1 Theoretical considerations **Error! Bookmark not defined.**
4.5.2 Assay Validation 206
4.5.3 STATs form six latent homodimers and two latent heterodimers. 208
4.5.4 STAT1, STAT3 and STAT5B latent homodimers are split by mutating a conserved hydrophobic residue in the NTD. 209
4.5.5 Discussion of Chapter 4.5 213
5. Concluding remarks 217
References VIII
Appendix XXVI

A. Raw Western blot and EMSA data XXVI
(i) Single transfections..... XXVI
(ii) Western blots of cells co-expression different constructs of the same STAT species..... XXIX
(iii) Western blots of cells co-expressing different STAT1 and STAT3 constructs XXXII
B. Equipment settings used in this work. XXXVI
(i) LICOR Odyssey settings for Western blot and EMSA analysis..... XXXVIII
(ii) Live cell imaging settings for Zeiss LSM710 XXXIX
C. Computer programming code used in this work. XLIII
(i) ImageJ macros coded in Java. XLIII
(ii) Data analysis and graphical plotting in R XLIII

Abbreviations

APC ..... Antigen Presenting Cell

APS ..... Ammonium Persulphate

ARDS ..... Acute Respiratory Distress Syndrome

AU ..... Artificial Units

BFP ..... Blue Fluorescent Protein

BSA ..... Bovine Serum Albumin

CCD ..... Coiled-Coil Domain

CF ..... Core Fragment

CI ..... Confidence Interval

CMV ..... Cytomegaly Virus

DBD ..... DNA Binding Domain

DC ..... Dendritic Cell

DMSO ..... Dimethyl Sulfoxide

DNA ..... Desoxyribonucleic Acid

DTT ..... Dithiothreitol

EGF ..... Epidermal Growth Factor

EMSA ..... Electrophoretic Mobility Shift Assay

FBS ..... Foetal Bovine Serum

FRET ..... Förster Resonance Energy Transfer

GAF ..... IFN-γ Activated Factor

GAS ..... IFN-γ Activation Site

GDP ..... Guanosine-Diphosphate

GFP ..... Green Fluorescent Protein

GH ..... Growth Hormone

GHR ..... GH Receptor

GOF ..... Gain-Of-Function

GRR ..... IFN-γ Response Region

GTP ..... Guanosine-Triphosphate

IFN ..... Interferon

IFNAR ..... IFN-α Receptor

IFNGR ..... IFN-γ Receptor

IL ..... Interleukin

IRF9 ..... Interferon Response Factor 9

ISGF3 ..... Interferon-Stimulated Gene Factor 3

kDa ..... Kilo Dalton

LD ..... Linker Domain

MAPK ..... Mitogen-Activated Protein Kinase

M-CSF ..... Macrophage-Colony Stimulating Factor

mIL6R ..... Membrane-Bound IL-6 Receptor

MW ..... Molecular Weight

NES ..... Nuclear Export Signal

NF-kb ..... Nuclear Factor κB

NLS ..... Nuclear Localisation Signal

NPC ..... Nuclear Pore Complex

NTD ..... NH_2-Terminal Domain

PAGE ..... Polyacrylamide Gel Electrophoresis

PBS ..... Phosphate Buffered Saline

PCR ..... Polymerase Chain Reaction

PDGF ..... Platelet-Derived Growth Factor

PFA ..... Paraformaldehyde

PIE ..... Prolactin-Inducible Element

PKI ..... Protein Kinase A Inhibitor Protein

PMSF ..... Phenylmethylsulfonylfluoride

RNA ..... Ribonucleic Acid

ROI ..... Region of Interest

SD ..... Standard Deviation

SDS ..... Sodium Dodecyl Sulphate

SE ..... Standard error

SH2 ..... Src Homology 2

sIL6R ..... Soluble IL-6 Receptor

STAT ..... Signal Transducer and Activator of Transcription

SV40 ..... Simian Virus 40

TAD ..... Transactivation Domain

WCE ..... Whole-Cell Extract

WT ..... Wild-Type

1. Introduction

The work presented in this thesis focuses on the dimerization of proteins of the signal transducer and activator of transcription (STAT) family. These proteins are an essential part of the Jak-STAT pathway, a mechanism that is utilised by most cytokines and is therefore of immense importance for cell-to-cell communication during immune responses. As such, STATs, which act as signalling components of the Jak-STAT pathway, can form two different dimer conformations depending on the activation status of the pathway. Furthermore, it is now clear that STAT family members can not only form dimers with themselves (homodimers) but also with other members of the family (heterodimers).

This study evaluates whether the formation of active heterodimers is predetermined by the composition of the inactive dimer form. Furthermore, the data presented here improves the understanding of the mechanism behind the switch from inactive to active STAT dimers. These conclusions were drawn from experiments on different STAT family members, namely STAT1, STAT3 and STAT5b. The following chapter will outline the current scientific knowledge surrounding cytokine biology, the Jak-STAT pathway and STAT biology while specifically focusing on the three STATs and accompanying cytokines relevant to this work. The last subchapter is dedicated to the areas of limited understanding within the literature and outlines the aspects that shall be answered within this thesis.

1.1 Cytokine biology: Mediators of immune cell communication

Our body is constantly under attack. Different pathogens like viruses, bacteria or eukaryotic parasites are trying to take advantage of us to facilitate their growth and spread to new hosts. At the same time, our cells can become cancerous, growing uncontrollably and spreading to other parts of the body. If unrestrained, these hazards from the outside and within may eventually threaten the existence of the whole organism. Fortunately, evolution has come up with a very effective mechanism for dealing with these threats: the mammalian immune system.

The mammalian immune consists of a complicated network of cells and mechanism that are coordinated by specialised soluble mediators termed “cytokines” after their

ability to mobilise cells to sites of infections (Morris, Kershaw and Babon, 2018). Cell-to-cell communication via these secreted factors ensures the efficient clearance of pathogens like viruses, bacteria and parasites while avoiding hyperimmune responses that can damage the host own and symbiotic cells. Cytokines can act both paracrine in the direct vicinity of their source and endocrine throughout the mammalian organism. Their specific effects depend on the target cell and can be redundant between different cytokines. Furthermore, distinct cell types may secrete the same cytokine (Kany, Vollrath and Relja, 2019).

Cytokines can be broadly divided into the following groups based on similar structures: TNF-αs, Il-1-like cytokines, TNF-βs, chemokines, cytokines that signal through receptor tyrosine kinases and cytokines that signal through the Janus kinase- Signal transducer and activator of transcription (Jak-STAT) pathway. The signalling pathway utilised by the latter group of cytokines shall be the focus of this work and is therefore explored in more detail below by discussing two relevant examples.

Cytokine signalling through the Jak-STAT pathway can be divided into two classes based on the structure of their receptors. Class I cytokines comprise most of the interleukins (ILs), like IL-6 family and IL-2 family cytokines, and some growth factors like erythropoietin (EPO) and growth hormone (GH). Interferons (IFN) and the IL-10 family of cytokines make up class II (Morris, Kershaw and Babon, 2018).

The key effector cytokine in cell-mediated immunity is IFN-γ. It is produced by NK cells and antigen-presenting cells (APC) such as macrophages and dendritic cells (DCs) during the early, acute phase of the immune response. Later, during adaptive immunity, $CD8^+$ T-cells and $CD4^+$ cells of the T helper 1 (Th1) lineage become the main source of IFN-γ (Lippitz, 2013; Green, Young and Valencia, 2017; Kak, Raza and Tiwari, 2018). Effects elicited by this cytokine include improved antigen presentation by dendritic cells, macrophages and other APCs and the preservation of the pro-inflammatory state through Th1 cell response promotion and inhibition of anti-inflammatory regulatory T (Treg) and T helper 2 (Th2) cells. IFN-γ is also the main macrophage-activating cytokine leading to differentiation into the pro-inflammatory "M1" phenotype (Green, Young and Valencia, 2017). Overall, IFN-γ has been associated with favourable outcomes during infection with bacterial pathogens such

as *Chlamydia, Mycobacterium* or *Staphylococcus aureus* and viral infections with human immunodeficiency virus (HIV), Influenza or varicella zoster virus (Kak, Raza and Tiwari, 2018). It was the ability of IFN-γ to inhibit viral replication that gave it its name (Frederick Wheelock, 1965). The IFN-γ receptor subunits are expressed in virtually all cells (Bach, Aguet and Schreiber, 1997; Chen *et al.*, 2012). Consequently, the direct inhibitory effects on viral replication are not limited to immune cells (Schroder *et al.*, 2004).

IL-6 is another pro-inflammatory cytokine that has recently gained significant attention as one of the drivers of severe acute respiratory distress syndrome (ARDS) following SARS-Cov-2 infection. More specifically, ARDS was shown to be a consequence of cytokine storm – an out-of-control release of pro-inflammatory cytokines that includes IL-6. IL-6 levels in the blood are a strong predictor of disease severity and an indication of mechanical ventilation (Mahmudpour *et al.*, 2020; Hirano, 2021). Furthermore, IL-6 is a driving force in cytokine storms which are caused by other viral and bacterial infections and non-infectious inflammation of tissues and can be a side effect of immune cell therapy (Tisoncik *et al.*, 2012; Hirano, 2021). During a "normal" immune response, IL-6 is considered to be involved in both the acute phase of inflammation and the transition towards the chronic inflammatory state (Kaur *et al.*, 2020). IL-6 expression can be induced by different stimuli like bacterial lipopolysaccharide (LPS) or other cytokines including TNFα or IL-1. The transcription factor NF-κB seems to play a major role in this but other transcription factors may contribute (Wolf, Rose-John and Garbers, 2014). The newly synthesised IL-6 leads to the induction of acute phase proteins such as C-reactive protein (CRP) and fibrinogen in the liver. In the bone marrow, IL-6 leads to changes in blood count and increases platelet release (Tanaka, Narazaki and Kishimoto, 2014). Regarding later stages of immunity, IL-6 has important functions on cells of the adaptive response. B-cells are promoted to antibody (Ab) producing plasma cells by IL-6 through direct and indirect mechanisms via T-follicular helper cell differentiation. IL-6 is also involved in shifting the T-helper 17 (Th17)/ regulatory T-cell (Treg) balance towards the former cell population and promotes CD8+ T-cell differentiation into cytotoxic effector T-cells (Tanaka, Narazaki and Kishimoto, 2014; Kishimoto and Kang, 2022). It is also important to note that IL-6 is associated with several immunological pathologies such as autoimmune disease and cancer. Its

effects on the Th17/Treg balance and direct effects on disease-specific tissues influence the progression of autoimmune diseases such as arthritis, inflammatory bowel disease and multiple sclerosis (Kimura and Kishimoto, 2010; Tanaka, Narazaki and Kishimoto, 2014). Moreover, it is now well appreciated that IL-6 is one of the major cytokines driving cancer progression through activating genes important for cell proliferation, differentiation, apoptosis, angiogenesis, and metastasis (Kaur *et al.*, 2020).

As mentioned above, both IFN- γ and IL-6 signals facilitate intracellular signalling mainly through the Jak-STAT pathway. The details of this will be described in the following chapter.

1.2 Cytokine signalling: The Jak-STAT pathway.

More than 60 cytokines encoded by the human genome use the Jak-STAT pathway for intracellular signalling (Morris, Kershaw and Babon, 2018). The pathway is remarkably simple, does not involve any secondary messengers, and only requires three components for signal transduction from the cell surface to the nucleus. These components are the cytokine receptors, the Jaks, and the STATs (OShea, Gadina and Schreiber, 2002). Broadly, the cytokine receptor and Jaks can be viewed as on/off switches for the pathway, while the STATs act as intracellular signalling proteins and travel from the membrane-bound receptors to the cell nucleus where they activate specific genes. Roughly equally as many different receptors as individual cytokines have been identified in humans. In contrast, only four different Jak species and seven distinct STAT paralogues have been identified. Discrete combinations of the three components determine the specificity of individual cytokine signals. Therefore, while details are distinct between different cytokine pathways, the general principle of Jak-STAT signalling is well conserved and will be exemplified by the cytokines most relevant to this work, IFN-γ and IL-6.

The IFN-γ receptor is comprised of two subunits termed IFNGR1 and IFNGR2. It is the former subunit that initially binds to IFN-γ with sub-nanomolar affinity before the other chain is recruited to the receptor-ligand complex to initiate signalling (Bach,

Aguet and Schreiber, 1997). It is worth noting that IFN-γ generally acts as a homodimer and therefore binds two sets of its receptor, making the complete signalling complex hexameric (figure 1) (Morris, Kershaw and Babon, 2018). The IFN-γ receptor is expressed at moderate levels in all cells of the human body (Bach, Aguet and Schreiber, 1997).

The IL-6 receptor is comprised of two subunits and ligand binding follows a two-step mechanism similar to the one described for IFN-γ. First, IL-6 binds to one receptor subunit with high affinity before recruiting the second subunit to form the active receptor complex. The latter subunit is called gp130 and is shared with other receptor complexes of the IL-6 cytokine receptor family such as IL-11 and IL-27. Consequently, it comes as no surprise that gp130 is ubiquitously expressed in most cells of the human body. In contrast, expression of the subunit that binds IL-6 with high affinity is limited to only a few cell types including macrophages, T-cells and hepatocytes. It is called the IL-6 receptor alpha chain and is generally membrane-bound (mIL6Rα) (Rose-John, 2020). However, the cleavage of the membrane-bound receptor chain by the disintegrin and metalloprotease 17 (ADAM17) during inflammation and, to a minor extent, alternative splicing during translation means that it can also exist as soluble (s)IL6Rα chain (Kishimoto and Kang, 2022). Importantly, the α-chains only facilitate cytokine binding whereas it is the gp130 subunit that confers intracellular signalling through mechanisms explained later in this chapter. Therefore, IL-6 binding to both the membrane-bound and the soluble IL-6 receptor alpha chain can elicit cytokine signalling. Since gp130 is found on most cells and sIL6Rα is circulating through the bloodstream, IL-6 signalling is observed on cells that traditionally lack expression of mIL6Rα, a process called trans-signalling. Trans-signalling is involved in the pro-inflammatory effects of IL-6 while classic signalling through the mIL6R is associated with antibacterial and positive metabolic effects (Garbers *et al.*, 2018; Kishimoto and Kang, 2022). Recently, a third mode of IL-6 signalling was discovered in mice that was termed trans-presentation because it involved the presentation of an IL-6/mIL6R complex bound to the membrane of DCs to gp130 bound to the membrane of T-cells. The T-cell then “receives” the IL-6 signal, leading to the priming of a pathogenic Th17 phenotype (Heink *et al.*, 2017).

Cytokines binding to their cognate receptors is not sufficient to activate signalling. This is because the receptors themselves lack intrinsic tyrosine kinase activity and thus must associate with one of the four Jak kinases found in vertebrates (Spangler *et al.*, 2015; Morris, Kershaw and Babon, 2018). In general, one complete cytokine/cytokine receptor signalling complex associates with exactly two Jaks and different receptors have binding sites for different Jaks. The IFNGR is associated with Jak1 and Jak2, whereby the IFNGR2 subunit binds the former and IFNGR1 to the latter (Bach, Aguet and Schreiber, 1997; Morris, Kershaw and Babon, 2018). The gp130 subunit of the IL-6 receptor is primarily associated with Jak1 but can also bind Jak2 and Tyk2 to a lesser extent. The IL-6 receptor alpha subunits do not have an intracellular binding site for Jaks which explains their aforementioned inability to initiate signalling. Furthermore, this might also explain why the IL-6 receptor signalling complex adopts the stoichiometry of 2:2:2 cytokine:IL-6Rα:gp130, fulfilling the requirement of having two Jaks associated with each receptor signalling complex (figure 1) (Morris, Kershaw and Babon, 2018). Jak3 is the only other Jak that is encoded in the mammalian genome. This Jak is exclusively expressed in leukocytes where it binds the common γ receptor chain (γc) that is a component of multiple cytokine receptors (e.g., IL-2, IL-4 and IL-21) (Schindler, Levy and Decker, 2007).

Jaks initiate activation of the Jak-STAT pathway upon cytokine binding through a series of phosphorylation steps. First, the Jaks phosphorylate each other at a tyrosine residue in the kinase activation loop, a process required because Jaks are generally inactive when no cytokine is present (figure 1) (Feng *et al.*, 1997). Of note, this further highlights the need for two Jaks to be present per functional cytokine/cytokine receptor complex. Next, the Jaks phosphorylate conserved tyrosines on the intracellular ends of the cytokine receptors that in turn act as binding sites for STAT dimers. Again, different receptors only bind specific STATs.

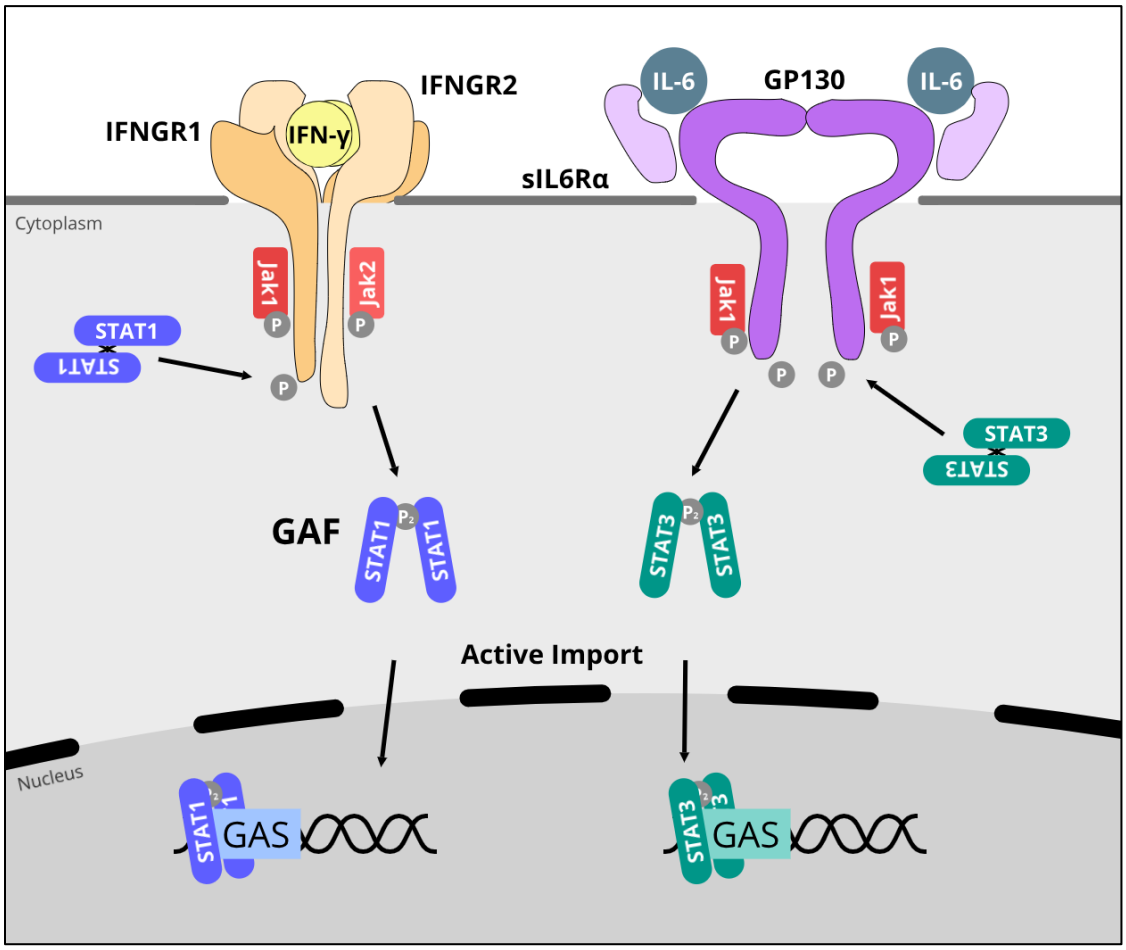


Figure 1 | Jak-STAT signalling by IFN-γ and IL-6: *Each monomer of an IFN-γ dimer bind the IFNGR comprised of IFNGR1 and IFNGR2. The two receptor chains associate with Jak1 and Jak2 respectively. Upon cytokine binding the Jaks phosphorylate each other and then the IFNGR1 subunit. A STAT1 dimer is recruited to the newly phosphorylated tyrosine of the IFNGR1 which leads to both STAT monomers being phosphorylated by the Jaks. The STAT1 dimer changes its conformation to form the GAF which is actively imported into the nucleus where it binds GAS consensus sequences. IL-6 signalling works in a similar way, though the IL-6 receptor is comprised of the sIL6Rα and gp130 subunits. Furthermore, the receptor associates with two Jak1 molecules and recruits STAT3 to the gp130 subunit.*

For example, the IFN-γ receptor predominantly binds STAT1 but can also induce signalling via STAT3 at low levels (or even at high levels in cells lacking STAT1) (figure 1). STAT1 and STAT3 both bind to the same binding site on the IFNGR1 subunit (Qing and Stark, 2004). On the other hand, IL-6 activates both STAT3 and STAT1, whereby binding is maintained by two distinct motifs on the gp130 receptor chain (figure 1) (Lütticken *et al.*, 1994; Heinrich *et al.*, 1998; Kishimoto and Kang, 2022). However, some evidence suggested that the STAT1 signal conferred by the IL-6 receptor is short-lived and perhaps even downmodulated by activated STAT3 (Costa-Pereira *et al.*, 2002). STAT activation is facilitated by the phosphorylation of a conserved tyrosine by the Jaks. Phosphorylation triggers STAT dimers to dissociate from the receptor and change their conformation in a process that is ill-understood (more details on the

individual STAT dimer conformations are given in chapter 1.3.2). Active STATs are capable of binding to proteins of the nuclear import machinery that facilitate active nuclear import (figure 1). This leads to a drastic shift in STAT subcellular distribution in favour of the nucleus. Furthermore, the active conformation allows STATs to bind to specific consensus sequences in the promotors of interferon-stimulated genes. Again, different STATs can bind to different consensus sequences. For example, STAT1 binds the consensus TTCNNNNAA which was termed IFN- γ activation site (GAS, figure 1) (Decker, Kovarik and Meinke, 1997), whereas the TTNNNNAA binding site is specific to STAT3 (Seidel *et al.*, 1995). Eventually, the STATs get dephosphorylated which allows them to return to the cytoplasm and complete the signalling cycle. The intricacies of the intracellular movement of STATs in the latent and active states are given further below (chapter 1.4).

Lastly, it is worth noting that some cytokine receptors that predominantly signal through the Jak-STAT pathway can also activate other pathways that are independent of STATs. For example, both IFN-γ and IL-6 activate mitogen-activated protein kinase (MAPK) and phosphatidyl-inositol-3-kinase (PI3K) pathways (Green, Young and Valencia, 2017; Garbers *et al.*, 2018). The transcription factors activated by these pathways bind to response elements different to GAS and therefore activate different genes than the STATs. One example of this is the activation of the MAPK extracellular signal-regulated kinase (ERK) in response to IFN-γ leading to CCAAT/enhancer-binding protein-β (C/EBPβ) binding to the response element IFN-gamma-activated transcriptional element (GATE) (Green, Young and Valencia, 2017).

1.3 STAT structure and dimerization

The STAT family of proteins comprises seven members, namely, STAT1, STAT2, STAT3, STAT4, STAT5a, STAT5b and STAT6 (Stark and Darnell, 2012). Within the Jak-STAT pathway, it is the STATs that transduce the cytokine signal that is perceived on the cell surface to a cellular response by binding to promoter regions of specific genes in the cell nucleus. Therefore, STATs can be viewed as both signalling proteins and transcription factors. The different functions of these proteins are facilitated by different functional domains within their amino acid sequence that will be described

below. As explained earlier, STATs can form two distinct sets of dimers. These dimers have important functions in that they regulate the activity of some of the STAT's functions. The process behind this will also be explained further below in this chapter.

1.3.1 STAT domain structure

The STAT amino acid sequence is well-conserved among the seven family members (Schindler, 1995). They share three structural domains which can be proteolytically cleaved (Vinkemeier *et al.*, 1996), and six functional domains that are arranged in the same order in all STATs (Figure 2) (Mohr *et al.*, 2012).

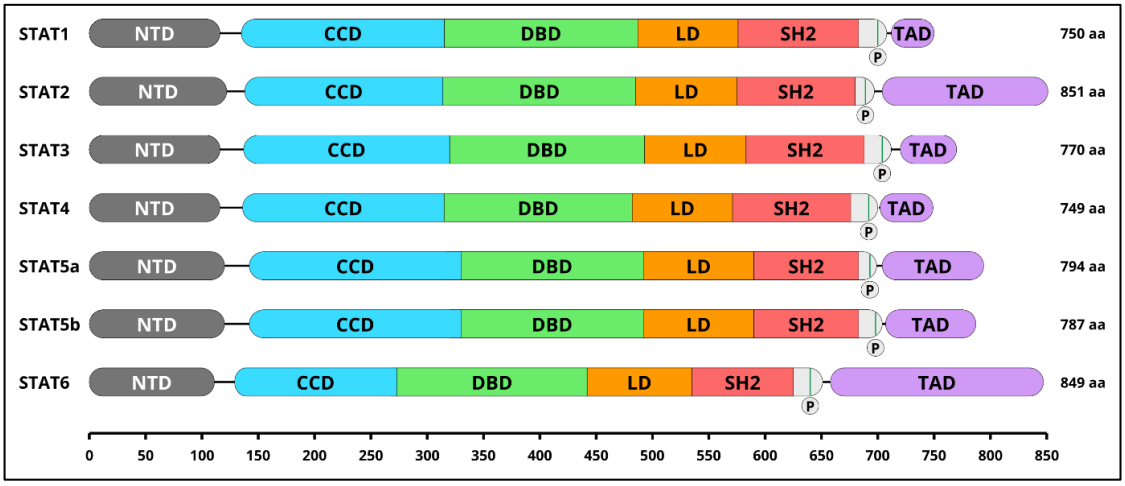


Figure 2 | STAT domain structure: *The STAT family of proteins contains 7 members: STAT1-4, STAT5a+b and STAT6. The organisation of structural and functional domains is highly conserved. The three structural domains are the NH_2-terminal domain (NTD), the core fragment and the transactivation domain (TAD). Both, the NTD and TAD are also functional domain. The core fragment contains five functional domains namely the coiled-coil domain (CCD), DNA-binding domain (DBD), linker domain (LD) and SH2 domain. STATs are activated at a conserved residue in the c-terminal part of the protein (dark green line).*

The N-terminal domain (NTD) is both a structural and a functional domain of STATs and it is the second-best conserved domain within the STAT protein family (Vinkemeier *et al.*, 1998). The NTD can be released from the STAT core fragment with proteolytic cleavage and is connected to it via a flexible linker (Vinkemeier *et al.*, 1996; Mertens *et al.*, 2006). The first crystal structure that was solved for any STAT component was that of a STAT4 NTD dimer which revealed two hook-like shaped monomers folded into each other by their concave sides (Vinkemeier *et al.*, 1998; Chen, 2003). Even before solving the crystal structure of unphosphorylated STAT1, it became clear from biochemical experiments that NTD-NTD interactions are required for efficient dimerization of unphosphorylated STAT1, STAT3, STAT4, STAT5 and STAT6 (more in chapter 1.3.2) (Ota *et al.*, 2004). The crystal structure of an unphosphorylated STAT1

dimer then confirmed that the dimer formation of isolated STAT4 NTDs is physiologically relevant (Mao *et al.*, 2005). It is believed that, because of the relatively short linker, the NTD-NTD interactions within a given STAT dimer are disrupted upon DNA binding. Still, the NTD has an important function in activated STAT dimers as it facilitates the tetramerization of two tyrosine phosphorylated STAT dimers and cooperative binding to tandem GAS sites (Vinkemeier *et al.*, 1996; Xu, Sun and Hoey, 1996). STAT1 carrying the amino acid substitution F77A showed disrupted NTD-NTD interactions in unphosphorylated dimers and was unable to tetramerise on DNA sites (Chen, 2003; Meyer *et al.*, 2004). Yet, mutated STAT1 was still able to translocate into the nucleus and bind DNA, showing that the NTD is dispensable for these actions (Meyer *et al.*, 2004).

The STAT core fragment is the second structural domain and contains five functional domains. The first functional domain within the core fragment is called the coiled-coil domain (CCD, figure 2), earning its name from its appearance of four anti-parallel α-helices forming a coiled-coil structure. The CCD is also important for the formation of some unphosphorylated STAT dimers through interaction with the other monomer's DNA-binding domain (DBD, see details in chapter 1.3.2) (Mao *et al.*, 2005). Furthermore, its structure enables the interaction with several non-STAT proteins, e.g. STAT3 interacting with c-Jun, StIP1 and GRIM-19, and STAT5a and STAT5b interacting with SMRT (Levy and Darnell, 2002; Lim and Cao, 2006; Morris, Kershaw and Babon, 2018). The CCD is further implicated in nucleocytoplasmic shuttling (see details in chapter 1.4). Lastly, the CCD of STAT3 was shown to be essential for its activation in response to IL-6 and EGFR (Zhang *et al.*, 2000).

Directly following the CCD, the DBD is the second functional domain within the STAT core fragment (figure 2). As the name suggests, it is this domain that recognises the consensus DNA sequence $TTCN_{2\text{-}4}GAA$ in promotors of interferon-stimulated genes (ISGs). Different STATs thereby prefer different sequences and sequence lengths within this palindromic sequence (Morris, Kershaw and Babon, 2018). The crystal structure of activated STAT1 bound to DNA revealed that binding to the major and minor grooves of DNA is conferred by several β-sheets contained within the DBD and that this mode of action was similar to the transcription factor nuclear factor κB (NF-

κB) (Chen *et al.*, 1998). Furthermore, STAT1 binding to importin-α5 is mediated by residues located in the DBD (Meyer and Vinkemeier, 2004). The exact mechanism for this will be explained later (Chapter 1.4). Importantly, only activated STAT dimers bind DNA and importin-α because residues maintaining these interactions are buried within the unphosphorylated dimer (see chapter 1.3.2).

The linker domain (LD) connects the DBD with the Src Homology 2 (SH2)-domain (figure 2). Its function is largely unknown, but it was shown that mutations within this domain influence DNA binding and transcriptional activity (Yang *et al.*, 2002; Hüntelmann *et al.*, 2014; Mertens *et al.*, 2015). Moreover, a cysteine residue located in the LD was shown to influence the nuclear export of unphosphorylated STAT1 (Marg *et al.*, 2004).

The last functional domain within the STAT core fragment is the SH2-domain (figure 2). It is the most highly conserved domain among the different STAT protein family members but rather divergent to SH2-domains of other proteins (Schindler, 1995). The SH2-domain determines cytokine receptor specificity by binding to a conserved tyrosine on the intracellular end of the receptor, only when it was previously phosphorylated by Jaks (Bach, Aguet and Schreiber, 1997). As a consequence, the exchange of virtually every other domain with sequences from other STAT family members does not change cytokine sensitivity (Schindler, 1995), and this has proven to be an excellent tool for functional investigations of the other domains (Stark and Darnell, 2012). The SH2-domain has another important function, that is, the formation of activated STAT dimers through interaction with a conserved tyrosine in the C-terminal part of STATs that gets phosphorylated after receptor binding (for STAT1 this is residue Y701). The SH2 of one STAT monomer showed a higher affinity for the phosphotyrosine of the other STAT monomer within the dimer than the phosphotyrosine of the receptor, and this difference in affinity was found to drive receptor dissociation and formation of active STAT dimers (Greenlund *et al.*, 1995). Self-association of STAT monomers is prevented since the phosphorylation site and SH2 domain are directly adjacent to each other within the STAT amino acid sequence (figure 2) (Morris, Kershaw and Babon, 2018).

The third structural and final functional domain of STATs is the transcription-activation domain (TAD). It has not been revealed in the crystal structures of activated STATs

since it is intrinsically disordered (Chen *et al.*, 1998). Alternative splicing results in STAT1, STAT3 and STAT4 isoforms lacking this domain (Horvath, 2000). It is this domain that facilitates STAT interactions with transcriptional co-activator CREB-binding protein (CBP), whereby an especially strong interaction was found for STAT2 (Wojciak *et al.*, 2009). The TADs of STAT1, STAT3, STAT4, STAT5a and STAT5b contain a secondary phosphorylation site: e.g. S727 in STAT1 (Darnell, 1997). Mutational and biochemical analysis showed that phosphorylation of this residue is required for the full transcriptional activity of these STATs (Decker and Kovarik, 2000). The process of C-terminal serine phosphorylation is thereby independent of tyrosine phosphorylation and facilitated by growth factor-regulated MAPK as well as stress-regulated MAPK (Kovarik *et al.*, 1999; Decker and Kovarik, 2000). Additional evidence suggested that STAT3 S727 phosphorylation is mediated by protein kinase C delta (PKCδ) and mTOR kinase (Decker and Kovarik, 2000).

1.3.2 STAT dimerization

STATs form two distinct sets of dimers, and the formation of these dimers is largely dependent on their activation status.

In resting cells, most STATs reside in the cytoplasm as unphosphorylated dimers. These dimers adopt an “anti-parallel” shape with the SH2 domains on opposite ends and are held together by (1) NTD-NTD interactions and (2) reciprocal CCD/DBD interactions (figure 3) (Mao *et al.*, 2005; Mertens *et al.*, 2006). The former interaction site was discovered when the crystal structure of isolated STAT4 NTD was resolved (Vinkemeier *et al.*, 1998; Chen, 2003). This study also identified the residue F77 of STAT1 as the key residue for this interaction (Chen, 2003). Later, it was found that this is also true for homologous residues in STAT3 and STAT4 (Chen, 2003; Ota *et al.*, 2004; Domoszlai *et al.*, 2014). Through X-ray crystallography, the residue F172 of STAT1 was determined to be a major driver of the reciprocal interaction between the CCD and DBD in STAT1 unphosphorylated dimers. This is due to F172 protruding towards the other monomer where it is buried in a hydrophobic binding pocket formed by residues of the complementary DBD. The significance of this interaction surface for the assembly of the unphosphorylated STAT1 dimer was then also confirmed physico-chemically and immuno-chemically (Mao *et al.*, 2005). Furthermore, when resolving the crystal

structure of an unphosphorylated STAT5a dimer, the authors found a similar interaction between the I172 of one monomer and the DBD of the other holding this dimer together (Neculai *et al.*, 2005). However, mutational analysis in STAT3 suggested that the CCD-DBD is of less importance for the formation of its homodimer (Ren *et al.*, 2008).

As mentioned before, STAT phosphorylation by the Jaks leads to the formation of a STAT dimer different from the one that is apparent in resting cells. The active STAT dimer is held together by reciprocal SH2-domain/phosphotyrosine interactions (figure 3) and thus, phosphorylation of the conserved tyrosine in the C-terminus of STATs favours the formation of this dimer (Darnell, 1997). Furthermore, this new arrangement means domains that were involved in the dimerization of the unphosphorylated dimer, such as the DBD, are now available for other functions such as binding to proteins of the nuclear import machinery and specific response elements in genomic DNA. This is further supported by the architecture of the active STAT dimer. Upon forming, the dimer interaction surface is shifted towards the C-terminal ends of each monomer and the SH2 domains are now in a parallel conformation, that is, at the same end of the dimer (figure 3). The CCDs project outwards from the core fragments, meaning that residues within the DBD involved in the recognition of GAS elements are now in a structurally favourable position for efficient DNA binding (Becker, Groner and Müller, 1998; Chen *et al.*, 1998; Li *et al.*, 2016).

While the canonical model of the Jak-STAT pathway states that STATs mostly act as homodimers, it is now clear that STAT heterodimers play important roles in immunity. The prototypical heterodimer is formed between STAT1 and STAT2 which is further associated with a third transcription factor called interferon regulatory factor 9 (IRF9). This ternary complex also termed interferon-stimulated gene factor 3 (ISGF3) was the first STAT complex to be identified and is activated by type I interferons (Fu *et al.*, 1990). ISGF3 is vital for establishing an antiviral cell state during viral infections but is also important in the defence against other pathogens (Decker, Müller and Stockinger, 2005; McNab *et al.*, 2015). STAT1 can also form heterodimers with STAT3 downstream of IL-6 signalling and it was shown that these heterodimers can bind to DNA sequences that are not recognised by STAT1 homodimers (Zhong, Wen and Darnell, 1994a,

1994b). The function of STAT1/3 heterodimers in the context of IL-6 signalling is largely unknown (Delgoffe and Vignali, 2013). However, a study on STAT1/3 heterodimer formation following IL-21 stimulation of CD8+ T-cells found based on immunoprecipitation data, that this dimer together with the STAT3/5 heterodimer is important for elevating IL-1Rα expression to increase the responsiveness of these cells to IL-1β (Kim, Kim and Lee, 2021). Other cytokines that were shown to induce the formation of STAT1/3 heterodimers include IL-27, epidermal growth factor (EGF) and IFNs. Lastly, STAT1 can form heterodimers with STAT4 in response to IL-35 stimulation of T-cells which is required for conversion to Tregs (Collison *et al.*, 2012). Besides STAT1/3 heterodimers, STAT3 can also dimerize with STAT4 and STAT5. The former is only formed downstream of IL-23 signalling, whereas the latter is induced by multiple cytokines including IL-2, IL-7 monocyte colony stimulation factor (M-CSF) and platelet-derived growth factor (PDGF) (Delgoffe and Vignali, 2013; Harrison and Moseley, 2020).

Evidence for heterodimer formation of unphosphorylated STATs is limited to such reports only existing for STAT1/2 and STAT1/3. Biochemically, the pre-association of STAT1 with STAT2 was shown early on (Stancato *et al.*, 1996). This was recently confirmed by a low-resolution cryo-electron microscopy structure of an unphosphorylated STAT1/2 heterodimer. This study also confirmed that the conformation of this dimer is similar to the unphosphorylated homodimers of STAT1, STAT3 and STAT5a (Wang *et al.*, 2021). The existence of unphosphorylated STAT1/3 heterodimers was inferred from immunoprecipitation experiments performed in several independent studies (Stancato *et al.*, 1996; Haan *et al.*, 2000; Blumert *et al.*, 2013; Nivarthi *et al.*, 2016) and a weak interaction between these two in living cells was detected in another (Wedeking *et al.*, 2015).

1.4 Nuclear translocation of STATs

STAT must enter the cell nucleus to convert the extracellular cytokine signal into a cellular response. However, the nucleus is separated from the rest of the cell by its membrane, creating a diffusion barrier to molecules larger than ~ 40 kDa. Still, larger molecules can cross the nuclear membrane by interacting with a component of the nuclear pore complex (NPC). This component is termed nucleoporin and the

interaction is mediated by phenylalanine/glycine-rich domains on the surface of the pore. Some proteins can bind these F/G-rich regions or rely on assistance from carrier proteins from the karyopherin superfamily. Direct interaction with nucleoporins is passive but carrier-mediated translocation requires active energy expenditure and as such is mediated by a RanGTP/RanGDP gradient across the nuclear membrane. The gradient provides directionality and is maintained by two proteins, the RanGTPase-activating protein (RanGAP) which activates GTP hydrolysation and resides in the cytosol and the guanine nuclear exchange factor (GEF) RCC1 in the nucleus which helps Ran to exchange GDP with GTP (figure 4). Carrier proteins that help with nuclear import are called importins. To bind to this transporter, proteins require a special signal sequence, the nuclear localisation signal (NLS). Importin-β is the carrier that finally targets nuclear pores and interacts with F/G repeats. Binding to this protein is either direct or indirect via an isoform of the adapter molecule importin-α. Loading in the cytoplasm is thereby mediated by RanGDP and the release of cargo in the nucleus by RanGTP (Paine and Feldherr, 1972; Macara, 2001; Fahrenkrog, Köser and Aebi, 2004).

STATs are generally capable of both active and passive nuclear translocation, even before activation downstream of the cytokine receptor. Most unphosphorylated STATs are distributed equally between the nucleus and cytoplasm, a phenotype that is governed by a steady state of constant nuclear import and export (Vinkemeier, 2004). Still, individual patterns of STAT subcellular distribution in resting cells and the mode by which this is maintained differ among the seven protein family members (Mohr *et al.*, 2012). Unphosphorylated STAT1, STAT3 and STAT5 dimers are able to cross the nuclear membrane passively and without assistance from other proteins (figure 4) (Meyer, 2002; Marg *et al.*, 2004). In the case of STAT1, passive translocation is assisted by active nuclear export mediated by the protein chromosomal maintenance 1 (CRM1) and this leads to a predominantly cytoplasmic localisation (figure 4) (Begitt *et al.*, 2000). CRM1, also termed exportin-1, facilitates carrier-dependent nuclear export. It is able to bind to importin-β and utilises the RanGTP/RanGDP gradient for active export similarly to the importin but in the opposite way. Proteins can bind to CRM1 via a short leucine-rich nuclear export signal (NES) (Reich, 2013). Such signals have been identified in STAT2, STAT3 and STAT5 (Vinkemeier, 2004). STAT2 shows a very strong cytoplasmic distribution, but STAT3 and STAT5 are more pancellular (Reich, 2013; Steen and

Gamero, 2013). This is likely because the latter two proteins are capable of a third mode of transport while unphosphorylated, namely active nuclear import mediated by an NLS located in the CCD of both STAT3 and STAT5 (Reich, 2013). Furthermore, STAT2 carries a strong NES in its TAD that is not conserved in other STATs (Frahm, Hauser and Köster, 2006).

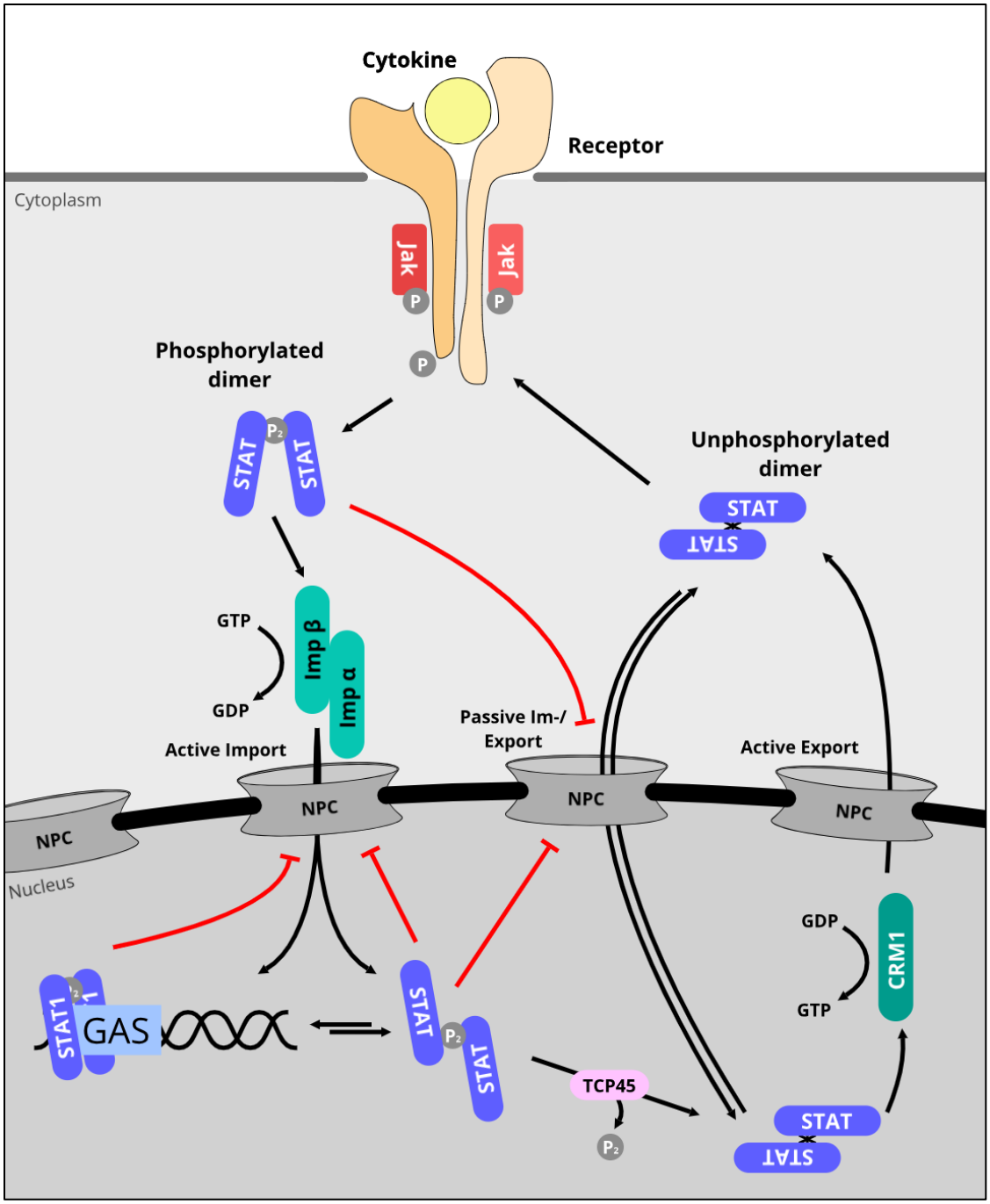


Figure 3 | Nuclear translocation of inactive and activated STATs: *Unphosphorylated inactive dimers adopt an antiparallel conformation which can enter and exit the nucleus via direct interaction with the NPC. Additionally, export of inactive dimers is supported by the carrier protein CRM1 in an energy dependent manner. STATs are phosphorylated at their surface receptor which leads to formation of parallel dimers. These dimers exclusively enter the nucleus via carrier dependent nuclear import facilitated by importin α and importin ß. Activated dimers are retained inside the nucleus and require dephosphorylation by phosphatases such as TCP45 before they can get exported. STATs bound to DNA are protected from dephosphorylation and the formation of antiparallel dimers is required to give TCP45 access to the phosphotyrosine.*

The subcellular distribution of STATs is shifted drastically towards almost complete nuclear accumulation upon activation downstream of ligand binding to the cytokine receptors (Köster and Hauser, 1999). As mentioned earlier, STATs contain another NLS that has been mapped to the DBD but is unique compared to other NLS in many ways as (1) the NLS is only active in STATs that are tyrosine phosphorylated and adopt the parallel dimer form, (2) the sequence of the STAT NLS is different from most other NLS found in other proteins and (3) the NLS is non-transferable (Vinkemeier, 2004). Importantly, parallel phosphorylated STAT dimers are no longer able to enter the nucleus through direct interaction with the NPCs but instead rely fully on carrier-dependent nuclear import via the previously buried NLS (Figure 4). Yet, this switch to active nuclear import is not the driving force for the nuclear accumulation of activated STATs. In fact, it is the nuclear retention of the parallel dimer that is causing this. Consequently, STATs must be dephosphorylated by nuclear phosphatases such as TCP45 before they can re-enter the cytoplasm (Figure 4) (ten Hoeve *et al.*, 2002; Marg *et al.*, 2004). Interestingly, STAT1 binding to DNA was shown to protect the activated dimer from dephosphorylation (figure 4), which explains the apparent correlation between DNA binding affinity and the duration of nuclear accumulation (Meyer *et al.*, 2003; Vinkemeier, 2004). Furthermore, for phosphatases to gain access to the tyrosine residue of STAT1 the formation of an antiparallel dimer is required (figure 4) (Mertens *et al.*, 2006).

1.5 Areas of limited understanding relevant to this work

Dimerization is central to STAT function. Despite several crystal structures of STAT dimers being available, many aspects of dimer formation remain ill-understood, especially surrounding heterodimer formation and unphosphorylated STATs.

First and foremost, most advances in the STAT field were made based on investigations into STAT homodimers. To this day, only a few studies have addressed the relevance of STAT heterodimers other than the ISGF3 complex comprised of STAT1, STAT2 and IRF9. Still, current evidence suggests specific roles for some active STAT heterodimers like STAT3/STAT5 and STAT1/STAT3 downstream of different cytokines. Yet, the molecular mechanisms that determine the formation of individual STAT heterodimers are poorly understood, though several mechanisms have been proposed. On one hand,

the formation of STAT heterodimers might be enabled by post-translational modifications (e.g. serine/tyrosine phosphorylation, SUMOylation or others). Such an effect has been suggested for the STAT1/2 heterodimer. A recent study showed that phosphorylation of STAT2 T404 disrupts the unphosphorylated STAT1/2 dimer (Wang *et al.*, 2021). The authors pointed out that this dimer is exceptionally strong and that the pT404 modification may alleviate the inhibitory effect that the unphosphorylated STAT1/2 dimer has on signalling through STAT1 homodimers (Ho *et al.*, 2016). On the other hand, STAT heterodimerization may be driven by the structure of different cytokine receptors. Though, likely, this theory can only be proven through direct structural insight which is currently unavailable (Delgoffe and Vignali, 2013). Finally, the mere proximity of different activated cytokine receptors may promote the formation of heterodimers. This theory is supported by evidence that showed STAT2/6 heterodimer formation in B-cells only after activation downstream of the IFN-α receptor. Hereby, STAT6 is interacting with the IFN-α receptor 1 (IFNAR1) subunit and STAT2 with the IFNAR2 subunit (Gupta, Jiang and Pernis, 1999; Delgoffe and Vignali, 2013). Furthermore, STAT2/6 heterodimers have also been found in T-cells. Here, it was found that co-stimulation with IFN-α enhanced STAT6 phosphorylation upon treatment with IL-4. Moreover, the formation of STAT2/6 dimers was linked to the proximity of the common γ-chain (γc) subunit of the IL-4 receptor and the IFNAR1 subunit (Eriksen *et al.*, 2004; Delgoffe and Vignali, 2013).

Another issue in STAT research that is ill-understood is the formation of unphosphorylated STAT dimers in latent cells and their functions. Crystal and Cryo-EM structures of unphosphorylated STATS have been resolved for the STAT1, STAT3 and STAT5a homodimer and the STAT1/2 heterodimer (Mao *et al.*, 2005; Neculai *et al.*, 2005; Ren *et al.*, 2008; Wang *et al.*, 2021). Additionally, immunoprecipitation analysis indicated that STAT4 is dimeric before phosphorylation (Ota *et al.*, 2004). In contrast, the purified core fragment of STAT6 lacking the NTD and TAD was found to be monomeric in analytical ultracentrifugation experiments (Li *et al.*, 2016). Considering that the NTD is the main driver for the formation of unphosphorylated STAT dimers and that the contribution of the DBD-CCD interaction to antiparallel dimer stability is limited in some STATs (Neculai *et al.*, 2005; Ren *et al.*, 2008; Wenta *et al.*, 2008), it is questionable whether these results are applicable for full-length STAT6. No

biochemical or structural data is available for STAT5b homodimers. Despite the aforementioned low-resolution structure of STAT1/2 the only other experimental evidence available for the existence of unphosphorylated STAT heterodimers are a few studies that have identified a STAT1/3 heterodimer via immunoprecipitation when specific experimental conditions were met (Stancato *et al.*, 1996; Haan *et al.*, 2000; Blumert *et al.*, 2013; Nivarthi *et al.*, 2016) or at low affinity *in situ* studies (Wedeking *et al.*, 2015). As described earlier, the existence of activated STAT1/4, STAT3/4, STAT3/5 and STAT5a/5b heterodimers have been confirmed (Delgoffe and Vignali, 2013), yet it is unknown whether their unphosphorylated counterparts exist. It is also worth noting that most of the findings identifying latent STAT dimers are based on *in-vitro* experiments using purified protein or cell extracts. In-depth studies looking at the prevalence of unphosphorylated dimers within cells are largely absent, except for a couple of studies that confirmed unphosphorylated STAT3 dimers in living cells using Förster resonance energy transfer (FRET) measurements (Kretzschmar *et al.*, 2004; Domoszlai *et al.*, 2014).

A further unresolved issue is the transition from unphosphorylated to activated STAT dimers. Indeed, this aspect may explain the other two large open questions in STAT dimerisation research stated above. While no direct evidence is available on the mode of transition, two likely scenarios come to mind. Firstly, a conformational change model, where the unphosphorylated dimer merely rearranges after activation to form the activated parallel dimer, and secondly, a dissociation/re-association model where the unphosphorylated dimer splits and reforms as activated dimer following tyrosine phosphorylation. The former scenario was extensively discussed in the context of the crystal structure of an unphosphorylated STAT1 dimer lacking the phosphotyrosine tail segment and TAD (Mao *et al.*, 2005). Importantly, several disordered regions remained unresolved, including the flexible tether that connects the structurally separate NTD with the core fragment. Consequently, the authors could not confirm which of the NTDs and core fragments within the tetrameric crystalline arrangement stemmed from the same molecule. Thus, two dimer conformations were proposed for unphosphorylated STAT1, one parallel and one antiparallel. The authors reasoned that the latter is the predominant form when STAT1 resides in the cytoplasm. However, they theorised that the parallel form might play a role in receptor binding during STAT

activation. They argue that the dimers stay together rather than dissociate and re-assemble during the transition from antiparallel to parallel. This is made possible by the strong NTD-NTD interaction, the relatively weak CCD-DBD interaction, and the flexible tether length (Mao *et al.*, 2005). Importantly, if this conformational change model is confirmed for STAT1 and extended to other STATs, it would also stipulate that active homo and heterodimers arise from a pre-defined cell pool of unphosphorylated STATs since no monomer exchange is taking place during activation at the cytokine receptor.

In another study, the authors presented data in favour of the latter transition model, dissociation/re-association (Wenta *et al.*, 2008). They showed that unphosphorylated STAT1 dimers can bind to GAS sequences with weak affinity and that a parallel conformation is required for this. Furthermore, they observed that deletion of the NTD and disruption of the CCD-DBD interaction through mutations did not adversely affect the DNA binding of the unphosphorylated dimer. Deletion of the NTD increased the affinity of unphosphorylated STAT1 to DNA and thus the authors concluded that the antiparallel dimer inhibits DNA binding. Furthermore, they argue that this makes it unlikely that NTD-NTD interactions facilitate the conformational change between the antiparallel and parallel dimers but instead, the dimers must dissociate and re-associate to transition (Wenta *et al.*, 2008). It is important to emphasise that this study only focused on the transition between dimer conformations within the unphosphorylated STAT1 dimer during the latent. Yet, the theory may also apply to the transition from unphosphorylated to activated STAT dimers during activation at the cytokine receptor. In any case, should this transition model be confirmed, it would be unlikely that unphosphorylated STAT dimers pre-determine the composition of active STAT homo- and heterodimers. Instead, active dimers may be able to exchange monomers following the split of latent dimers at the cytokine receptor.

Lastly, it is worth noting that the transition from phosphorylated to unphosphorylated STATs during deactivation is relatively well understood for STAT1 and that this may inform the transition model during activation at the cytokine receptors. It was recognised soon after solving the crystal structure of unphosphorylated STAT1, that mutations that disrupt the dimerization interface (STAT1 F77A, L78A and F172W) cause

persistent STAT1 phosphorylation in cells. Further analysis uncovered that this is due to STAT1 mutants resisting dephosphorylation by the phosphatase TC45. Based on these observations it was hypothesised that the formation of antiparallel dimers is required for dephosphorylation by nuclear phosphatases (Zhong *et al.*, 2005). In a follow-up study in which U3R cells that do not express endogenous STAT1 where reconstituted with transfected STAT1-WT or NTD and CCD mutants (F77A/L78A and F172W, respectively) and treated with Staurosporine, a kinase inhibitor that stops further phosphorylation of STATs, this dephosphorylation phenotype was confirmed solidifying evidence that both the NTD and the DBD binding pocket facilitate antiparallel dimer formation and that this is essential for efficient dephosphorylation. More importantly, the authors, using pulldown experiments, also showed that dimers stay together and do not dissociate during dephosphorylation, suggesting that the switch from active dimers in the parallel form to unphosphorylated dimers in the antiparallel form is merely mediated by a conformational change (Mertens *et al.*, 2006).

1.6 Hypothesis and aim of this thesis.

The Jak-STAT pathway maintains intracellular signalling of more than 60 cytokines. The paradigm for signal transduction via this pathway is that the biological response of a given cytokine is mainly driven by a homodimer of one of the seven STAT paralogues. However, recent evidence suggests that the concurrent activation of STAT heterodimers may contribute to the specificity of cellular responses to cytokines. The mechanism that determines whether a STAT monomer is incorporated into a homo- or heterodimer after activation is yet to be resolved and shall be the main focus of this work.

Unlike their active counterparts, unphosphorylated STAT dimers are poorly researched and the existence of only a handful of them, mostly homodimers, has been confirmed experimentally. In general, the biological function of unphosphorylated STAT dimers is unknown. However, it is conceivable that they act as regulators of active dimer formations. Two hypotheses come to mind:

Firstly, should the formation of active dimers be regulated by their unphosphorylated counterparts, then the cellular pool of active dimers is predetermined long before exposure to cytokines. This would necessitate that the dimers transition from unphosphorylated to active without the exchange of monomers at the cytokine receptor.

Alternatively, if active dimer formation is independent of unphosphorylated STAT dimers, then unphosphorylated dimers are expected to split upon activation and reform in their active conformation. This process may involve the exchange of monomers and potential recombination into other dimer species.

The hypothesis and its alternative shall be investigated in this work from two angles. First, the effects of latent dimer splitting mutations on the formation of active STAT1 and STAT3 homo and heterodimer formation will be assessed. Secondly, a systematic analysis of unphosphorylated STAT dimer formation in living cells across the whole family will confirm or disprove the existence of latent counterparts to the active STAT homo- and heterodimers that have been reported before.

Thus far, the tools to investigate these aspects of cytokine signalling have been limited. Therefore, significant efforts were put into the development of two cell-based imaging assays that enable the assessment of the two alternative mechanisms mentioned above. Further methodologies used in this work include Western blotting, electrophoretic mobility shift assays (EMSAs) and other biochemical and molecular biology techniques. These will be described in detail in the following chapters.

2. Materials

2.1 Media

2.1.1 Media for bacterial work

LB medium was used for the general cultivation of *Escherichia coli* following DNA transformation. The composition of LB is described in table 1. For the preparation of agar plates, 1.4 % (w/v) agar was added to the broth before autoclaving. Antibiotics were added to the media after autoclaving.

Table 1 | Composition of LB media used for the cultivation of E. coli.

LB medium	
NaCl	5 g/L
Tryptone	10 g/L
Yeast - extract	5 g/L

SOB and SOC media were used for the preparation of chemically competent *E. coli* (below). The contents of the SOB medium are listed in table 2.

Table 2 | Composition of the SOB medium used for the preparation of chemically competent E. coli.

SOB medium	
NaCl	0.5 g/L
Tryptone	10 g/L
Yeast - extract	5 g/L
KCl	2.5 mM

The media was autoclaved and sterile $MgCl_2$ was added to a final concentration of 10 mM just before use.

SOC medium is identical to SOB medium, except for the addition of sterile glucose to a final concentration of 20 mM after autoclaving.

2.1.2 Media for mammalian cell work

In general, mammalian cells were grown in fully supplemented Dulbecco's Modified Eagle's Medium (DMEM) and the composition is outlined in table 3.

Table 3 | Composition of supplemented DMEM

DMEM (supplemented)		
DMEM	500 mL	100 %
Heat Inactivated Foetal Calf Serum (FCS)	50 mL	10 %
Penicillin/Streptomycin	5 mL	1 %

For live-cell imaging, cells were grown in imaging medium, that is, full medium lacking the pH indicator phenol red (table 4), to avoid the introduction of artefacts during image acquisition.

Table 4 | Composition of supplemented imaging media

Imaging media (supplemented)		
DMEM phenol red free	500 mL	100 %
Heat Inactivated Foetal Calf Serum (FCS)	50 mL	10 %
Penicillin/Streptomycin	5 mL	1 %
L - Glutamine	10 mL	2 %
Sodium Pyruvate	5 mL	1 %
HEPES	12.5 mL	2.5 %

All media supplements were added using aseptic technique.

2.2 Buffers

2.2.1 Buffers for DNA work

The following buffers were used for the work with DNA:

50x TAE

Tris	2M
Concentrated Acetic Acid	5.7 % (v/v)

EDTA	0.05 M
pH = 8.3	

Inoue transformation buffer

$MnCl_2$	55 mM
$CaCl_2$	15 mM
KCl	250 mM
PIPES pH = 6.7	10 mM

2.2.2 Buffers for protein work

The following buffers were used to work with proteins:

10x TBS

NaCl	1.4 M
Tris (100 mM)	100 mM
pH = 7.3	

TBS-T

500 µl Tween-20 in 1 L 1x TAE

10x SDS running buffer

Glycine	2 M
Tris	250 mM
SDS	1 % (w/v)

2x SDS sample buffer

Tris	42 mM
Glycerol	8 % (v/v)
SDS	1.3 % (w/v)
DTT	0.25 M
Bromphenylblue	1 tip of a spatula

4x SDS PAGE Resolving gel buffer

Tris	1.5 M

SDS	0.4 % (w/v)
pH = 8.8	

4x SDS PAGE stacking gel buffer

Tris	0.5 M
SDS	0.4 % (w/v)
pH = 6.8	

10 % SDS PAGE resolving gel

H_2O	4.1 mL
4x Resolving gel buffer	2.5 mL
Acrylamide/Bisacrylamide (30 %)	3.15 mL
APS (10 %)	100 µL
TEMED	10 µL

SDS PAGE stacking gel

H_2O	3.15 mL
4x Stacking gel buffer	1.25 mL
Acrylamide/Bisacrylamide (30 %)	550 µL
APS (10 %)	100 µL
TEMED	10 µL

10x Blotting buffer

Tris	0.48 M
SDS	13 mM
Glycine	0.39 M

1x Blot buffer

10x Blotting buffer	10 % (v/v)
Methanol	20 % (v/v)
Water	70 % (v/v)

WCE for Western blotting

Tris	50 mM
EGTA	2 mM
EDTA	0.2 mM

Na_3VO_3	1 mM
NaF	50 mM
NaCl	280 mM
Glycerol	10 %
pH	7.5
Add just before use:	
cOmplete™ protease inhibitors	1x
IGEPAL	0.5 % (v/v)
DTT	1 mM
PMSF	0.1 mM

WCE for EMSA

Tris	50 mM
EDTA	5 mM
Na_3VO_3	1 mM
NaCl	150 mM
Glycerol	10 % (v/v)
pH	7.5
Add just before use:	
cOmplete™ protease inhibitors	1x
DTT	1 mM
PMSF	1 mM

5x TBE

Tris	0.7 M
Boric Acid	0.44 M
EDTA	10 mM

EMSA gel (4.5 %)

5x TBE	2.4 mL
Acrylamide/Bisacrylamide (40 %)	6 mL
H_2O	40.6 mL
APS (10 %)	1 mL
TEMED	50 µL

5 x Shift buffer

HEPES	100 mM
$MgCl_2$	5 mM
KCl	200 nM
EGTA	0.5 mM
EDTA	2.5 mM
Ficoll	20 % (v/v)
pH	7.9

Shift reaction

EMSA protein extracts (1 µg/µL)	8.5 µL
5x Shift buffer	4 µL
dIdC (11 U/mL)	3.5 µL
Probe	1.8 µL
DTT (1 M)	0.2 µL

6x SDS running buffer

Tris	125 mM
SDS	4 % (w/v)
DTT	0.7 M
Glycerol	25 % (v/v)
Bromophenol Blue	Tip of a spatula

2.3 Organisms

The following bacteria were used in this work (table 5):

Table 5 | ***List of bacterial organisms used in this work.***

Name	Genotype	Usage
***E. coli* NEB5α** (NEB, USA)	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	Cloning

The following mammalian cell lines were used in this work (table 6):

Table 6 | List of mammalian cell lines used in this work.

Name	Cell type	Usage
HeLa	Human Cervix carcinoma	Microscopy
HEK 293 + GHR	Human embryonic kidney	Western Blot, EMSA
HEK 293T	Human embryonic kidney	Western Blot, EMSA
U3R	Human fibrosarcoma	Western Blot, EMSA

2.4 Software

Different software packages were used for experiment design and data analysis.

The following software was used to work with DNA and plasmids. SnapGene Viewer version 6.1.2 (Dotmatics, USA) was used to visualise plasmid maps, nucleotide, and amino acid sequences of different genes. NEBaseChanger version 1.3.3 (NEB, USA) was used to design site-directed mutagenesis experiments. NEBioCalculator 1.15.0 (NEB, USA) was used to calculate the mass of insert required for a 1:3 vector-insert ratio for DNA ligation reactions. Agarose gels were imaged using a GeneGenius (Syngene, USA) gel imaging system with accompanying software.

Image Studio™ Lite (LI-COR, USA) was used for the quantification of Western blots and EMSA gels.

Gene sequence alignments were created using Clustal Omega (Sievers *et al.*, 2011) and viewed in Jalview version 2.11.2.6 (Waterhouse *et al.*, 2009).

Protein structures were viewed and processed using UCSF Chimera version 1.16 (Pettersen *et al.*, 2004).

Microscopy images were visualised with Zen blue edition version 3.4 (Zeiss, Germany). Images taken for use in the Nuclear Translocation Assay were quantified using stock and in-house macros (see Appendix D) for Fiji/ImageJ v1.53t (Schindelin *et al.*, 2012). All graphs were generated in RStudio 2022.07.2+576 using the tidyverse and ggpubr libraries version 1.3.0 for R 4.2.2. Statistical analysis was performed in RStudio 2022.07.2+576 using R 4.2.2 and show the mean and standard error f not otherwise stated. The student t-test was used for hypothesis testing unless indicated otherwise. Correlation coefficients were calculated using Pearson's method. A full list of code used

for data analysis in R is available in Appendix D. Figures were assembled using Corel Vector version i013.1.

3. Experimental Procedure

3.1 Working with DNA

3.1.1 Polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) was used to amplify fragments of genomic or plasmid DNA. The boundaries of the fragments are specified by a forward (5'→3') and a reverse (3'→5') primer oligonucleotide. Overhangs containing restriction enzyme recognition sites were included in primers for molecular cloning. A Q5® High-Fidelity DNA Polymerase (NEB, USA) was used for all PCR reactions as described in the manufacturer's protocol (table 7).

Table 7 | Reaction mixes for PCR using the Q5® High-Fidelity DNA Polymerase.

Components	25 µL Reaction	Final concentration
5x Q5 reaction buffer (NEB, USA)	5 µL	1x
10 mM dNTPs (NEB, USA)	0.5 µL	200 µM
10 µM Forward primer	1.25 µL	0.5 µM
10 µM Reverse primer	1.25 µL	0.5 µM
Template DNA	Variable	5-50 ng
Q5® High-Fidelity DNA Polymerase (NEB, USA)	0.25 µL	0.02 U/µL
Nuclease-free H_2O (Qiagen, Netherlands)	To 25 µL	

The standard reaction protocol was performed in 14-25 cycles after initial denaturation at 98 °C for 30 s. Each cycle included another denaturation step at 98°C for 10 s, a primer annealing step at 50-72 °C for 30 s and a DNA elongation step at 72 °C for 30s/kbp. After all cycles were finished, the elongation of all fragments was completed at 72 °C for another 2 min.

3.1.2 Gel electrophoretic separation of DNA molecules

Agarose gel electrophoresis was utilised to separate DNA molecules according to their size using an electric current. The negatively charged DNA will hereby migrate through the agarose gel, typically containing 1 % agarose in TAE buffer, to the positively charged anode. Separation by size is elicited because small nucleic acid fragments can travel faster through the agarose network than large constructs. To make DNA molecules visible under UV light, agarose gels contained 0.5 µg/mL ethidium bromide. Gels were run at 5-8 V per 1 cm between electrodes for 30-60 min.

3.1.3 Restriction enzyme digestion

Restriction enzymes were used to cut DNA fragments. Depending on the respective enzyme(s) used (table 8), different sticky ends are produced during digestion. The second step in molecular cloning experiments comprised cutting a vector and an insert with two different enzymes to generate matching sticky ends (double digest). Additionally, double digests were used for analytical purposes including quality control of molecular cloning experiments.

Table 8 | Restriction enzymes used in this work. *All chemicals were purchased from NEB, USA.*

Restriction enzyme	Recognition site	Reaction Buffer	Optimal temperature
AgeI-HF®	ACCGGT	CutSmart®	37 °C
BamHI-HF®	GGATCC	CutSmart®	37 °C
EcoRI-HF®	GAATTC	CutSmart®	37 °C
SalI-HF®	GTCGAC	CutSmart®	37 °C
NheI-HF®	GCTAGC	CutSmart®	37 °C

Reactions for digestion of up to 0.5 µg template DNA were performed on a 25 µL scale, containing 1 µL (20 U) of each restriction enzyme and 2.5 µL of 10x CutSmart® buffer (NEB, USA). The reaction mix was incubated at 37 °C for 1 h.

3.1.4 Molecular cloning

The method described in this section was used to insert a specific gene into a plasmid vector (table 9) for use in different types of molecular biology assays. Molecular cloning was performed in four steps.

Firstly, the gene of interest was amplified from a DNA template using PCR (3.1.1). The primers used for the PCR reaction thereby determined the fragment amplified and included restriction enzyme recognition sites as overhangs at both ends of the gene of interest. To facilitate the insertion of the gene of interest in the right orientation, the restriction enzyme recognition site at the 5' overhang was different from the recognition site at the 3' end. The PCR product was separated from other components of the reaction using the NucleoSpin™ Gel and PCR Clean-up Kit (Marcherey-Nagel™, Germany) according to the manufacturer's instructions.

Table 9 | Plasmid vectors used in this work.

Construct	Insert	Origin
pcDNA 3.1 zeo (-) + STAT1 F172W	STAT1 F172W	Vinkemeier lab
pEGFP-N1 + STAT1wt Flag Stop	STAT1wt-Flag	Vinkemeier lab
pmCherry-N1 lab	Empty Vector	Vinkemeier lab
pmGFP-N1	Empty Vector	Vinkemeier lab
pmGFP-N1-pkiNES	Empty Vector	Vinkemeier lab
pmGFP-N1-sv40NLS	Empty Vector	Vinkemeier lab
pEGFP-N3	Empty Vector	Vinkemeier lab
pASK-Iba3	Empty Vector	Vinkemeier lab
pmCherry-GFP flexible linker	fGFP-Cherry	Vinkemeier lab
pmCherry-GFP rigid linker	rGFP-Cherry	Vinkemeier lab
pCMV-GHR	GHR	Dr RJ Ross, Uni of Sheffield
pSport6-Jak2	Jak2	Uni of Cambrifge

Secondly, the amplified PCR fragment and the vector were double-digested with matching restriction enzymes (3.1.3). The reaction was stopped by supplementation with Gel Loading Dye, Purple (NEB, USA) and loaded onto an agarose gel. The DNA reaction products were separated (3.1.2) and the bands representing the cut vector and insert were isolated from the agarose gel using NucleoSpin™ Gel and PCR Clean-up Kit (Marcherey-Nagel™, Germany) as described in the manufacturer's manual.

Next, the cut insert was ligated into the cut vector in a ligase reaction. For this, the DNA isolated from agarose gels was mixed so that the insert DNA was at a 3x molar excess compared to vector DNA and supplemented with T4 DNA Ligase (NEB, USA) and a buffer as described in table 10.

Table 10 | Ligation reaction mix.

Components	20 µL Reaction mix
T4 DNA-Ligase Buffer (10x) (NEB, USA)	2 µL
Vector DNA	50 ng
Insert DNA	3x molar excess to vector DNA
T4 DNA-Ligase (NEB, USA)	1 µL
Nuclease-free H_2O (Qiagen, Netherlands)	To 20 µL

The reaction was incubated at 16 °C overnight. On the next morning, T4 DNA Ligase (NEB, USA) was deactivated by heating the mixture to 65 °C for 10 min.

Finally, 5 µL of the ligation mix was transformed into chemically competent *E. coli* NEB5α cells for high-quality plasmid preparations (3.2.2).

This work relied heavily on various GFP or mCherry-tagged STAT proteins. To ensure reproducibility and ease of use, the cloning strategy for these constructs was standardised. This involved changing the multiple cloning site of pmCherry-N1 to resemble that of pmGFP-N1. Thus, all STAT constructs can be cloned into both vectors using a NheI/AgeI restriction digest. Some experiments required additional tagging with NES or NLS signals. To create pmGFP-N1-pkiNES, the NES signal was introduced between the STAT and fluorescent protein using an AgeI single digest in a way that maintained the ability for easy introduction of STAT constructs by the above-mentioned double digest. The NLS was introduced by site-directed insertion as per the Q5® Site-Directed Mutagenesis Kit (NEB, USA) instructions.

3.1.5 Colony PCR

Colony PCR is a rapid way to screen for the successful transformation of an artificial plasmid containing a specific gene into bacterial cells. Subjecting the DNA contents of a bacterial colony to a PCR reaction using primers flanking the gene of interest will lead to a PCR product of a specific size if the gene was successfully inserted.

Therefore, following molecular cloning, 10-14 colonies were picked from LB agar plates using an inoculation loop and struck onto a fresh gridded LB agar plate containing appropriate antibiotics. Directly afterwards, the residual bacteria stuck to the inoculation loop were transferred to 50 µL dH_2O. The gridded plate was incubated at 37 °C for up to 6 h and the bacterial suspensions were heated at 95 °C for 5 min and vortexed vigorously.

A PCR reaction mix was prepared (table 11) and the reaction was performed as described in table 12.

Table 11 | Reaction mixes for a 25 µL colony PCR reaction.

Reagent	Final Concentration	Volume
10x Standard Taq Reaction Buffer (NEB, USA)	1x	2.5 µL
dNTPs (NEB, USA)	200 µM	4 µL
Forward Primer	0.2 µM	0.5 µL
Reverse Primer	0.2 µM	0.5 µL
Template DNA		4 µL
Taq DNA-Polymerase (NEB, USA)	250 U/mL	0.125 µL
Nuclease-free H_2O (Qiagen, Netherlands)		to 25 µL

The primers used for the colony PCR were typically identical to the primers used for molecular cloning. Therefore, if the gene of interest is not present in the respective bacterial colony, no fragment will be amplified in the PCR reaction.

Table 12 | PCR programme for colony PCR.

Step	Temperature	Time	
Initial denaturation	95 °C	30 s	
Denaturation	95 °C	30 s	30 cycles
Annealing	61 °C	60 s	
Elongation	68 °C	60 s/kb	
Final elongation	68 °C	5 min	

The PCR products were run on an agarose gel and visualised under UV light (3.1.2). Positive clones will show a fragment of the size of the insert. Two to four positive clones were picked from the gridded LB agar plate to inoculate a 5 mL LB culture containing suitable antibiotics for plasmid mini prep (3.2.3).

3.1.6 Sanger sequencing

DNA sequencing was used to validate the integrity of genes, mainly after molecular cloning experiments (3.1.4).

A 20 µL BigDye™ sequencing PCR was set up which included 4 µL of 5x sequencing buffer (obtained from DeepSeq, UoN), 0.5 µM of a specific sequencing primer (see table 13), 150 ng of sequencing template and 1 µL of BigDye™ (DeepSeq, UoN).

Table 13 | Oligonucleotides used for Sanger sequencing.

Name	Sequence (5'-3')	Position	Direction
C275	GCAGAGCTGGTTTAGTGAAC	pEGFP-N1 MCS	5'-3'
IL5	CAGGCCAAAGGAAGCACCAGAG	STAT1 aa 690	5'-3'
C871	CTACAGCTCTCAGCCTTGCG	STAT3 aa 690	5'-3'
C872	GGCCGAGAAGCACCAGAAG	STAT5 aa 230	5'-3'
C874	CAAGCAAGTGGTCCCTGAG	STAT5B aa704	5'-3'
318	GACAATAAGAGAAAGGAAGTAGTTC	STAT1 aa 207	5'-3'
C621	CTAGATCGGCTAGAAAAC	STAT3 aa 207	5'-3'
467	CGCAAATGGGCGGTAGGCGTG	CMV Promotor	5'-3'
517	CGTCGCCGTCCAGCTCGACCAG	mEGFP aa 15	3'-5'
C817	CCTTGGAGCCGTACATGAACTG	mCherry aa 69	3'-5'
C563	GAAATTTGTGATGCTATTGC	SV40 poly (A)	3'-5'
C570	GACAAGCAGAAGAACGGCATCAAG	mEGFP aa C570	5'-3'
C244	GTCAATGGGAGTTTGTTTTGGCACC	CMV Promotor	5'-3'

The Sanger sequencing PCR programme included an initial denaturation step at 94 °C for 1 min, followed by 30 cycles of denaturation at 94 °C for 16 s, primer annealing at 52 °C for 16s and elongation at 60 °C for 2 min.

The PCR products were purified by precipitation. For this, the 20 µL reaction mixes were supplemented with 5 µL 75 mM sodium acetate, 12.5 mM EDTA solution and 80 µL 95 % ethanol. The mixture was incubated on ice for 15 min. Thereafter, the precipitated DNA was pelleted with centrifugation at 16100 x g for 15 min. The pellet was washed with 200 µL of a 70 % ethanol solution and centrifuged again for 15 min at 16100 x g. The supernatant was removed, and the pellet dried in a thermomixer (Eppendorf ThermoMixer® 5350). The Sanger sequencing service at DeepSeq, School of Life Sciences, University of Nottingham was used for sequencing sample electrophoresis and evaluation of the DNA sequence.

3.2 Working with Bacteria

3.2.1 Generation of Chemically Competent Cells: Inoue Method

E. coli NEB5α cells were made competent using the Inoue method (Im, 2011).

The cells were scraped from a laboratory glycerol stock and streaked onto LB agar plates. The plates were incubated at 37 °C overnight. On the next morning, a single colony was picked, transferred to 25 mL SOB medium starter culture, and incubated at 37 °C, 300 rpm for 8 h. Afterwards, the starter culture was used to inoculate 250 mL SOB medium with 10 mL starter culture and shaken at 18 °C, 180 rpm overnight. Once the OD_{600} had reached 0.55 the following morning, the bacterial culture was cooled in an ice-water bath for 10 min. Thereafter, the cells were harvested by centrifugation at 2500 x g for 10 min, 4 °C. The cell pellet was washed in 80 mL ice-cold Inoue transformation buffer followed by centrifugation at 2500 x g for 10 min, 4 °C. The cells were resuspended in 20 mL of ice-cold Inoue transformation buffer to which 1.5 mL of DMSO was added. After incubating on ice for 10 min the competent *E. coli* were aliquoted in 200-400 µL aliquot, snap-frozen and stored at -80 °C until use.

3.2.2 Transformation into Chemically Competent *E. coli* Cells

This method was used to transform artificial plasmids into chemically competent bacteria. The competent *E. coli* cells were thawed on ice for transformation. 5 µL of a ligation mix or 30-50 ng of purified plasmid was added to the thawed cells and incubated on ice for 30 min. After a heat shock at 42 °C for 30 s which enables the uptake of DNA, the cells were cooled on ice for 5 min. Subsequently, 900 µL of SOC media was added to the cell suspension and the cells were shaken for 1 h at 37 °C to allow expression of the antibiotic resistance genes encoded by the plasmid. Agar plates containing antibiotics for selection were incubated with 50-100 µL of the cell suspension, the rest of the cell suspension was centrifuged at 4000 x g for 5 min, 850 µL of the supernatant was discarded and the cell pellet was resuspended in the leftover supernatant. Afterwards, the cell suspension was plated on another agar plate containing appropriate antibiotics. The agar plates were incubated overnight at 37°C.

3.2.3 Plasmid isolation from *E. coli*

For small-scale isolation, *E. coli* carrying the desired plasmid were grown in 5 ml LB and appropriate antibiotics at 37 °C, and shaken overnight. The plasmid was isolated using the NucleoSpin Plasmid – Plasmid Mini-Prep Kit (Marcherey-Nagel™, Germany) according to the manufacturer's instructions. Plasmid DNA isolated with this method amongst others was used for quality control by sequencing and analytical restriction enzyme digestion in molecular cloning experiments.

Transient transfection of artificial plasmids into mammalian cells required larger amounts and higher purity. Thus, *E. coli* transformed with the plasmid of interest were grown in 100 mL LB and appropriate antibiotics at 37 °C, shaking overnight. The Marcherey-Nagel™ NucleoBond™ Xtra Midi Prep Kit was used to purify the recombinant plasmid from *E. coli* cultures according to the manufacturer's instructions.

3.3 Mammalian cell culture

3.3.1 Initiation of mammalian cell culture

Mammalian cell lines were stored in the liquid phase of liquid nitrogen cell bank storage in a solution containing 90 % (v/v) FBS and 10 % (v/v) dimethyl sulfoxide (DMSO). To initiate cell culture, the vial containing 1 mL cell suspension was thawed rapidly in lukewarm water until slightly frozen. Subsequently, the cells were added to 9 mL full medium pre-warmed to 37 °C and pelleted at 200 x g for 5 min to remove any remaining DMSO. The cell pellet was resuspended in 18 mL full medium and divided into a T-25 cell culture flask and a 10 cm cell culture treated dish. Cells were grown in a humidified atmosphere containing 5 % CO_2 (v/v) at 37 °C.

3.3.2 Cell Culture Maintenance

All mammalian cells were cultured in full medium in a humidified atmosphere containing 5 % CO_2 (v/v) at 37 °C. Generally, cells were plated in a T-25 cell culture flask or 10 cm diameter cell culture dish. Confluency was assessed using a Knauer light microscope. Upon reaching 70-90 % confluency, cells were washed in phosphate-buffered saline (PBS) and dislodged from culture vessels using Trypsin/EDTA and diluted 1:5 or 1:7 in fresh culture medium. Cells were reseeded using 1 mL cell suspension and 5 mL fresh full medium in T-25 flasks and using 1.5 mL cell suspension and 10.5 mL full medium in 10 cm cell culture dishes.

3.3.3 Generation of frozen cell stocks

Cells were allowed to reach 70-90 % confluency before processing into frozen cell stocks. A single-cell suspension was generated by washing the cell-containing culture flasks with PBS, followed by incubation with Trypsin/EDTA for several minutes at 37 °C before deactivating Trypsin/EDTA through the addition of 5 mL full media. Then, the total cell number was determined by cell counting using a Luna-II™ automated cell counter (logos biosystems, South Korea). The cells were pelleted by centrifugation at 200 x g for 5 min and the cell pellet was resuspended in freezing medium (90% (v/v) FBS and 10 % (v/v) DMSO) to a concentration of 1-5 x 10^6 cells/mL. Cryovials were filled with 1 mL of the cell suspension and cooled in a Mr Frosty™ cryo-freezing container

(Nalgene, USA) in a -80 °C freezer. The next day, the cryovials were taken from the Mr Frosty™ and transferred to a liquid nitrogen cell bank for long-term storage.

3.3.4 Transient transfection of cell lines using Lipofectamine LTX and Plus reagent

Mammalian cells were transiently transfected using Lipofectamine® LTX & PLUS™ Reagent (Life Technologies, USA). The transfection reagents form lipid complexes with DNA that are then taken up by eukaryotic cells via endocytosis.

One day before transfection, cells were seeded at different concentrations on 24- or 6-well dishes or 35 mm µ-Dish (ibidi, Germany), depending on the downstream application and cell type. HeLa cells were seeded at 5 x 10^4 cells/well on coverslips inside 24-well plates and 7.2 x 10^4 cells per plate on the inside well of 35 mm µ-Dishes. HEK 293T cells were plated at 8 x 10^5 cells/well on 6-well plates. U3R cells were plated at 5.6 x 10^5 cells/well on 6-well plates. This resulted in 50-70 % cell confluency on the day of transfection.

Table 14 | Transfection reagent mixes A & B

		6-Well	24-Well	35 mm µ-Dish
Mix A	Opti-MEM® Medium	100 µL/well	25 µL/well	50 µL/well
	DNA	1 µg/well	0.3 µg/well	0.5 µg/well
	PLUS™ Reagent	5 µL/well	1.5 µL/well	3 µL/well
Mix B	Opti-MEM® Medium	100 µL/well	25 µL/well	50 µL/well
	Lipofectamine® LTX	5 µL/well	1.5 µL/well	3 µL/well

Transfection reaction mixes were prepared as outlined in table 14. Both mixes were incubated separately for 20 min at room temperature and then mixed and incubated for another 20 min at RT. In the meantime, the cells were subjected to a media change with 6-well plates filled with 1.8 mL/well, 24-well plates with 450 µL/well and 35 mm µ-Dishes with 1.9 mL/well. After incubation, 200 µL transfection reagent was added to each well of 6-well plates, 50 µL transfection reagent to each well of a 24-well plate and 100 µL transfection reagent to each 35 mm µ-Dish. Transiently transfected cells were incubated for at least 24h before further processing.

3.3.5 Cytokine treatment of cells

To trigger the phosphorylation of STATs, transfected and untransfected cells were treated with different cytokines (table 15). For this, the culture medium was replaced with fresh medium containing one or more cytokines. Treatment with IL-6 required co-incubation with the sIL6Rα and treatment with GH required prior transfection of the growth hormone receptor (GHR). Stimulation lasted for 45-120 min depending on individual experiments.

Table 15 | Recombinant cytokines used in this work.

Cytokine	Concentration	Source
Human IFN-γ	50 U/ml	Gibco, USA
Human IL-6 & sIL6Rα	200 ng/mL & 250 ng/mL	BioLegend, USA
Human GH	100 ng/mL	BioLegend, USA

At the end of the incubation time, cells were washed with PBS and either snap-frozen on dry ice (6-well plates for Western blotting and EMSA) or fixed using 2 % paraformaldehyde (PFA) for 10 min (fluorescence microscopy).

3.3.6 Culturing cells for live-cell imaging

As described previously, cells were plated on the inside well of a 35 mm µ-Dish before transfection. Sixteen hours after transfection, cells were prepared for live-cell imaging experiments. For this, the full growth medium was removed and replaced with 400 µL imaging media containing 5 µg/mL Hoechst 33342 (Merck, Germany) for staining of the cell nucleus. After incubation for 5 min at 37°C, cells were washed three times with 2 mL PBS for 3 min. Subsequently, dishes containing cells transiently transfected with STAT DNA were filled with 1.5 mL imaging medium and stored in a humidified atmosphere containing 5 % CO_2 (v/v) at 37 °C until use. All other dishes were filled with 2 mL medium. Live-cell imaging was performed using an LSM710 laser scanning confocal microscope (Zeiss, Germany). During this process, STAT-containing cells were either stimulated by adding 0.5 mL imaging medium containing appropriate quantities of cytokines for a final concentration as indicated in table 15 or left untreated.

3.4 Working with Proteins

3.4.1 Whole cell extraction for Western blot analysis

Human cell lines were used for transient expression of wild-type (wt) or mutant STAT1, STAT3 and STAT5B (3.3.4) Cells were lysed using the following method to analyse expression and phosphorylation status by Western blot (3.4.5).

Frozen cells on 6-well plates were rapidly thawed and immediately supplemented with 120 µL/well Western blot whole cell extraction buffer (2.2.2). The cells were scraped from culture plates, transferred to microcentrifuge tubes, and incubated on ice for 30 min to allow for effective lysis. Subsequently, cell fragments were separated from protein-containing supernatants by centrifugation at 16,100 x g for 10 mins, 4 °C. The supernatant was transferred to a fresh microcentrifuge tube and the pellet containing cell debris was discarded. A 5 µL sample was taken to determine protein concentration via Bradford assay (3.4.3). The remaining supernatant was mixed with appropriate volumes of 6x SDS running buffer (2.2.2), boiled at 95 °C for 5 min, vortexed for 1 min and rapidly frozen on dry ice. The SDS protein samples were aliquoted and stored at -80 °C until use.

3.4.2 Whole cell extraction for EMSA

The whole cell extraction methodology was modified slightly for EMSA experiments to avoid the introduction of artefacts when detecting DNA-bound STAT dimers (Haan *et al.*, 2005; Haan and Haan, 2013). Frozen cells on 6-well plates were rapidly thawed and immediately supplemented with 150 µL/well EMSA whole-cell extraction buffer (2.2.2). The cells were scraped from culture plates and transferred into microcentrifuge tubes. Subsequently, cells were lysed by 20 rounds of Dounce homogenisation. The soluble cell fraction was separated from cell debris through centrifugation at 16,100 x g for 10 mins, 4 °C. A 5 µL sample was taken to determine the total protein concentration via Bradford assay (3.4.3) and the whole cell extract samples were diluted with EMSA whole cell extraction buffer to a final protein concentration of 1 µg/µL. The sample was divided into 20 µL aliquots and rapidly frozen on dry ice before storage at -20 °C until further use.

3.4.3 Bradford Assay

The total protein concentration was determined via Bradford assay. For this, a 5 µL whole cell extract sample was diluted in 95 µL deionised water (dH_2O) and then mixed with 100 µL Bradford reagent (a 2:3-part mixture of Protein Assay Dye Reagent Concentrate (BioRad, USA) and dH_2O). A BSA protein standard was run alongside in duplicates and treated in the same way as protein extract samples. Absorbance was measured at 450 nm and 595 nm using a SPECTROstar Nano (BMG Labtech, Germany). Linear regression between BSA standard protein concentrations and the respective 595/450 nm determined the linear relationship that was used to calculate protein sample concentrations from spectroscopic measurements.

3.4.4 SDS Polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were separated by size using vertical discontinuous electrophoresis.

SDS polyacrylamide gels were poured as follows. First, the freshly prepared resolving gel was poured into a Mini-PROTEAN® Tetra Cell casting frame (Bio-Rad, USA), coated with isopropanol and left to polymerase. Afterwards, the isopropanol was removed, and the stacking gel was poured on top of the resolving gel. Combs were inserted and the gel was left to polymerise. Meanwhile, previously frozen SDS protein samples were heated for 5 min at 95 °C and vortexed for 1 min. The mixtures were allowed to cool and subsequently loaded onto the freshly poured 10 % polyacrylamide gels. A protein standard (Blue Prestained Protein Standard, Broad range (11-250 kDa); NEB, USA) was run alongside protein samples to approximate the molecular weight of proteins. Gel electrophoresis was performed in two steps in a Mini-PROTEAN® Tetra Cell gel system (Bio-Rad, USA). Firstly, running the gel at 120 V for 10 min ensured protein stacking at the stacking gel border. After this, proteins were separated at 160 V for 1 h.

After electrophoresis, gels were used in Western blot experiments to detect a specific protein (3.4.5).

3.4.5 Western blotting

Western blotting was used to detect specific proteins in a sample of a protein mixture using immunoglobulins. This requires the transfer of electrophoretically separated proteins onto a carrier membrane.

Thus, SDS PAGE gels (3.4.4) were layered between three filter membranes on the bottom and a nitrocellulose membrane and two filter membranes on top which were pre-soaked in 1x blotting buffer. A semi-dry transfer was performed for 1 h at 16 V in a Trans-Blot® SD Semi-Dry Transfer Cell (Bio-Rad, USA). To confirm the transfer, proteins were made visible by staining the nitrocellulose membrane with a Ponceau S stain (0.1 % w/v Ponceau S, 5 % acetic acid). Background stains were removed by rinsing in dH_2O and a picture of the band pattern. Subsequently, the nitrocellulose membrane was blocked with 4 % semi-skimmed milk powder (Marvel, UK) in TBS-T for 1 h to prevent the non-specific binding of antibodies. Additionally, incubation in milk removed the Ponceau S stain from the membrane completely. The membrane was rinsed in TBS and then incubated with one or two primary antibodies of different species (table 16) in 3 mL TBS-T at 4 °C overnight.

Table 16 | Primary and secondary antibodies used for immunoblotting.

Name	Species	Concentration	Source
Anti-STAT1 Clone 42	Mouse	1:1000	BD Bioscience
Anti-STAT1 pY701 (D4A7)	Rabbit	1:1000	Cell Signaling Technology
Anti-STAT3 Clone 84	Mouse	1:5000	BD Bioscience
Anti-STAT3 pY705 (D3A7) XP	Rabbit	1:2000	Cell Signaling Technology
Anti-STAT5B (G-2)	Mouse	1:1000	Santa Cruz
Anti-STAT5 pY694 (D47E7) XP	Rabbit	1:1000	Cell Signaling Technology
Anti-ß-Actin	Mouse	1:1500	Merck
Anti-GFP	Rabbit	1:1000	Cell Signaling Technology
Anti-Mouse IRDye 680LT	Goat	1:5000	LICOR Bioscience
Anti-Rabbit IRDye 800CW	Goat	1:5000	LICOR Bioscience

On the next morning, the membrane was washed 3 x 5 min with TBS-T and incubated with either 1 µL goat α-mouse IR-Dye®RD 680 (Licor, USA), 1 µL goat α-rabbit IR-Dye®CW 800 or both antibodies in 7 mL TBS-T for 45 min at RT. Next, the membrane was washed again 3 x 5 min with TBS-T and imaged on a LiCor Odyssey® imaging system.

3.4.6 Electrophoretic mobility shift assay - EMSA

EMSA was used to test the binding of different activated STAT dimers to specific DNA sequences. This required incubation of EMSA whole cell extracts with IRDye700-labelled oligonucleotide probes and, optionally, specific antibodies in a process called the shift reaction. This was then followed by electrophoretic separation on a native PAGE gel.

Before the shift reaction, native EMSA PAGE gels (2.2.2) were equilibrated in 0.25 % TBE running buffer at 400 V until the current had fallen below 35 mA. For each shift reaction, 8.5 µL EMSA whole cell extracts (concentration 1µg/µL, 3.4.2) were mixed with 0.2 µL DTT (1M), 3.5 µL poly-deoxyinosinic-deoxycytidylic (dIdC; 11 U/mL), 4 µL Shift buffer (5x, 2.2.2) and 1.8 µL DNA probe at various concentrations (table 17; Biolegio BV, Netherlands).

Table 17 | Oligonucleotide probes used for EMSA experiments in this work. *All probes contained a covalent IRDye 700 label at the 5' end.*

DNA Probe	Sequence
M67 fwd	5‘ – GCCGACATTTCCCGTAAATCTG – 3‘
M67 rev	5‘ – CATGTATTTCCCAGAAAAGGAAC – 3‘
GRR fwd	5‘ – GTTCCTTTTCTGGGAAATACATG – 3‘
GRR rev	5‘ – GCCGACATTTCCCGTAAATCTG – 3‘
PIE fwd	5‘ – ATCAGATTTCTAGGAATTCAATG – 3‘
PIE rev	5‘ – CATTGAATTCCTAGAAATCTGAT – 3‘

Immediately after, 2 µL of specific antibody (table 18) or 0.1 % BSA solution was added to the shift reactions and incubated for 10 min at room temperature.

Table 18 | Antibodies used for EMSA supershifts in this work.

Name	Species	Concentration	Source
Anti-STAT1 (C-24)	Mouse	1:10	Santa Cruz
Anti-STAT3 (C-20X)	Mouse	1:10	Santa Cruz
Anti-STAT5B (G-2)	Mouse	1:10	Santa Cruz

Following the incubation, 18 µL of the shift reaction were loaded onto the equilibrated native PAGE gels and run at 400 V for approximately 95 min. STAT proteins bound to the oligonucleotide probes were visualised using a LiCor Odyssey® imaging system.

3.5 Fluorescence Microscopy

3.5.1 Widefield Microscopy

For widefield microscopy, HeLa cells were grown on suitable coverslips in 24-well plates (see 3.3.4 for details). Sixteen hours after transient transfection, cells were further processed for widefield microscopy. This entailed washing the coverslips containing cells with 500 µL cell culture grade PBS followed by fixation with 2 % paraformaldehyde (PFA) for 10 min at room temperature. For the nuclear staining, the cells were incubated with PBS containing 5 µg/mL Hoechst 33342 for 5 min at room temperature. The nuclear dye was washed off by incubating with PBS three times for 5 min. Finally, the coverslips were mounted on microscopy slides using fluorescence mounting media (Agilent Dako, USA) and sealed with nail varnish. Widefield microscopy was performed on a Zeiss Axioplan 2 microscope using a 63x/1.4 Plan-Apochromat oil lens, DAPI (exciter filter BP 365/12, beam splitter FT395, barrier filter LP397), FITC (exciter filter BP 470/20, beam splitter FT510, barrier filter BP540/25) and TexasRed (exciter filter BP 575.5/27.5, beam splitter FT600, barrier filter LP615) filter sets, and an Axiocam CCD camera (all Zeiss, Germany).

3.5.2 Live-Cell Imaging and Confocal Microscopy

All live-cell imaging experiments were conducted using a Zeiss Fluar 40x/1.4 oil Ph3 lens on a Zeiss LSM710 laser scanning confocal microscope and Zeiss Zen software. Imaging was done using three tracks in line sequential scanning mode. The first track was used to detect nuclear Hoechst stain and to take phase contrast images. Thus, the excitation wavelength was set to 405 nm and fluorescence was detected between 415-465 nm. The second track was used to detect GFP fluorescence using an excitation wavelength of 488 nm and detection wavelengths of 496-545 nm. The third track was used to detect mCherry fluorescence with the excitation wavelength set to 561 nm and detection conducted between 583-631 nm.

This work describes a novel live-cell imaging assay termed the Nuclear Translocation Assay that uses the change in subcellular localisation of STATs as readout to detect an interaction between different family members (chapter 4.1). Imaging was done with HeLa cells transiently transfected with two fluorescently tagged STAT DNA constructs at one time. Transfection efficiencies varied but generally ranged from 10-40%.

Therefore, to ensure that only cells expressing the STAT protein are imaged for the nuclear translocation assay, live cell imaging was conducted in three steps utilising different imaging setups. Live cell imaging settings are also available in Appendix B.

First, a large area (990.78 µm x 990.78 µm) in form of a 3x3 tile scan was imaged at a low resolution (fast pixel speed) and low magnification (scan zoom set to 0.6) to "map" the cells expressing GFP and mCherry tagged protein. Based on these images, 30 small area frames (scan zoom set to 2.0; 106.27 µm x 106.27 µm) were picked and positions saved using the positions tool in the Zeiss Zen software. Next, the 30 small areas were imaged at high resolution using a 7-frame z-stack spanning 24 µm to capture the fluorescence of the whole cell. The cells were stimulated with cytokines by adding 500 µL imaging medium containing appropriate amounts of the cytokine to reach final concentrations as outlined in chapter 3.3.5. Lastly, the same 30 small areas were reimaged 90 min after cytokine stimulation, using the same settings as previously. Importantly, to reduce bias, all Nuclear Translocation Assay experiments were conducted blind to genotype if not otherwise stated. In experiments described in this work, several STAT-GFP/STAT-mCherry combinations were examined using the Nuclear Translocation Assay (see detailed protocol in chapter 4.1.6). The contents of individual experiments were blinded by assigning them a random number which was then communicated to the experimenter. To uphold the blinding during experiments, DNA encoding for the STAT-GFP/STAT-mCherry constructs was combined in the transfection reaction mix (see chapter 3.3.4) by a colleague. Experiments were unblinded only after all imaging and image analysis was completed.

3.5.3 Foerster Resonance Energy Transfer (FRET)

FRET experiments were conducted to rule out fluorescence intensity measurement artefacts introduced by resonance energy transfer in fusion proteins containing both a GFP and mCherry fluorophore (for more details see 4.1.2). All experiments were performed using a Zeiss Fluar 40x/1.4 oil Ph3 lens on a Zeiss LSM710 laser scanning confocal microscope and Zeiss Zen software. Generally, these experiments were conducted with fixed cell samples on coverslips that were prepared as outlined in chapter 3.5.1. Furthermore, FRET imaging was carried out employing the same scanning settings as described in 3.5.2. The FRET analysis was initiated by taking a

354.25 x 354.25 µm snapshot. Then, ten cells that showed observable GFP and mCherry fluorescence were selected based on the snapshot and marked for bleaching using the regions tool of Zeiss Zen. The FRET imaging workflow was comprised of taking five images in 5 s intervals, followed by acceptor bleaching for 30 iterations using 100% power of the 561 nm laser. Then, another set of five images was taken in 5 s intervals. GFP and mCherry mean fluorescence intensity at different time points within the 10 regions was exported from the FRET tool of Zeiss Zen and further analysed using Excel and R (see code in appendix D).

3.5.4 Nuclear Translocation Assay Image Analysis Workflow

Quantification of live-cell imaging micrographs was necessary to determine key parameters for use in the Nuclear Translocation Assay (chapter 4.1.6). These include total GFP and mCherry fluorescence and average GFP and mCherry intensity within the nucleus before and after cytokine treatment. Image analysis was done in Fiji/ImageJ and workflows used to determine the two required parameters are outlined in the following. As mentioned above, Nuclear Translocation Assay experiments were conducted blind to genotype, and the contents investigated were only revealed when image analysis was completed in full.

Cell segmentation and measurement of total GFP and mCherry fluorescence

The nuclear translocation assay requires knowledge of the total amount of GFP and mCherry fluorescence within individual cells to determine if these cells were eligible for further analysis (see 0). This was determined in a semi-automated fashion based on principles that were described previously (Kelley and Paschal, 2019).

Essentially, a Fiji macro termed Whole Cell Segmentation, developed specifically for this work, was applied that allowed for efficient segmentation of GFP-tagged protein containing cell clusters from non-expressing cells and background. Because the macro could not distinguish cell borders of neighbouring cells, individual cells were then segmented manually (figure 5).

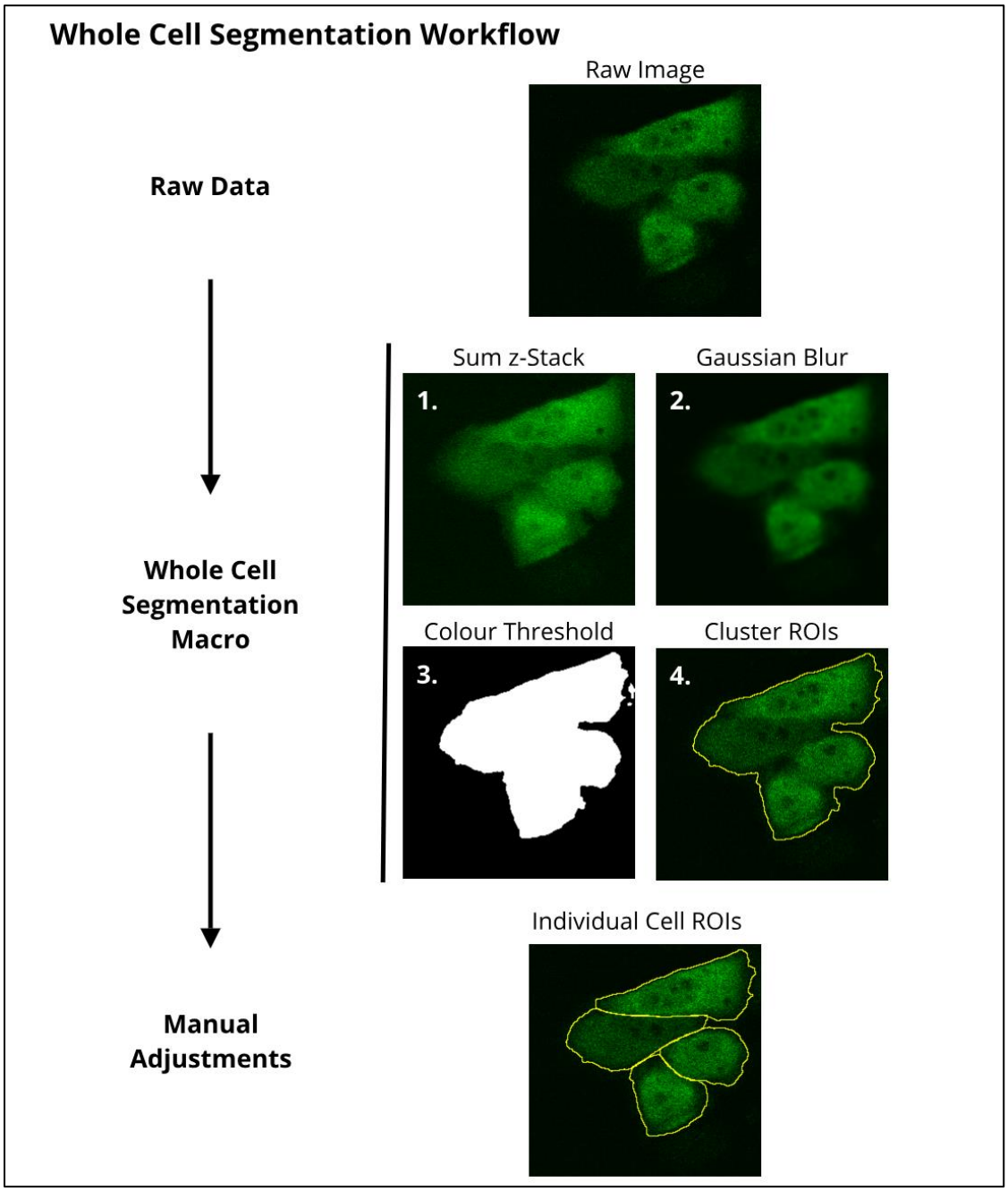


Figure 4 | Whole cell segmentation workflow: *Microscopy images were segmented for whole cell ROIs in two steps. First the Whole Cell Segmentation Macro was applied. This macro automatically, sums up intensities of all z-stack images, duplicates the GFP channel, applies Gaussian blur and uses a colour threshold to create a binary image, segmenting GFP-tagged protein expressing cells from background and non-expressing cells that is then used to create a cell cluster ROI. Secondly, based on this, individual cell ROIs were created manually.*

Details on the Whole Cell Segmentation macro code are described in Appendix D but a summary of its mechanism is given in the following. First, the macro created a new image by summing up the intensities of all z-stack images in individual channels, utilising the Z Project function. Next, the GFP channel image was duplicated, and a

Gaussian blur was applied (Sigma = 2.0) to smoothen the image. Using the colour threshold function, a binary image was created with all pixels below the threshold being assigned a pixel value of 0 (black) and all pixels above the threshold being assigned a pixel value of 255 (white). The black areas thereby represent areas containing no cells or cells that do not express GFP-tagged protein. Lastly, the macro identified outlines of the high pixel value areas using the Analyse Particles function and saved them to the ROI manager as regions of interest (ROIs). Importantly, the macro could not distinguish between the cell borders of neighbouring cells. Therefore, the macro's ROI outputs may represent the outer borders of GFP-tagged protein-expressing cell clusters rather than individual cells. To accommodate this, the cell cluster ROI put out by the Whole Cell Segmentation macro was divided manually into individual whole cell ROIs and saved to the ROI manager. Total GFP and mCherry fluorescence of individual cells was measured in summed-up images put out by the Whole Cell Segmentation macro using the previously saved individual cell ROIs. Measured values were transferred to an Excel sheet for further analysis.

Segmentation of cell nuclei and measurement of average nuclear fluorescence

The Nuclear Translocation assay compares the average nuclear fluorescence before and after cytokine stimulation in individual fluorescence channels to score the interaction between the GFP and mCherry-tagged proteins.

Nuclear ROIs for intensity measurements were created based on Hoechst staining using a Nucleus Segmentation macro that was created in Fiji specifically for this work (see code in Appendix D, figure 6). Again, the macro essentially follows an image analysis workflow that has been previously described (Kelley and Paschal, 2019).

In short, the macro first duplicates the Hoechst channel image of the median z-slice displaying the middle of the cell nucleus. The median z-slice had to be defined visually and was fed into the code for individual sample image sets (containing 30 individual images, see chapter 3.5.2) before the use of the macro. Next, a colour threshold was applied to create a binary image that segments cell nuclei from the background. This was then used to create nuclear ROIs using the Analyse Particles function, also saving

them to the ROI manager. Importantly, the Hoechst 33342 dye stained the nuclei of all cells, irrespective of their transfection status.

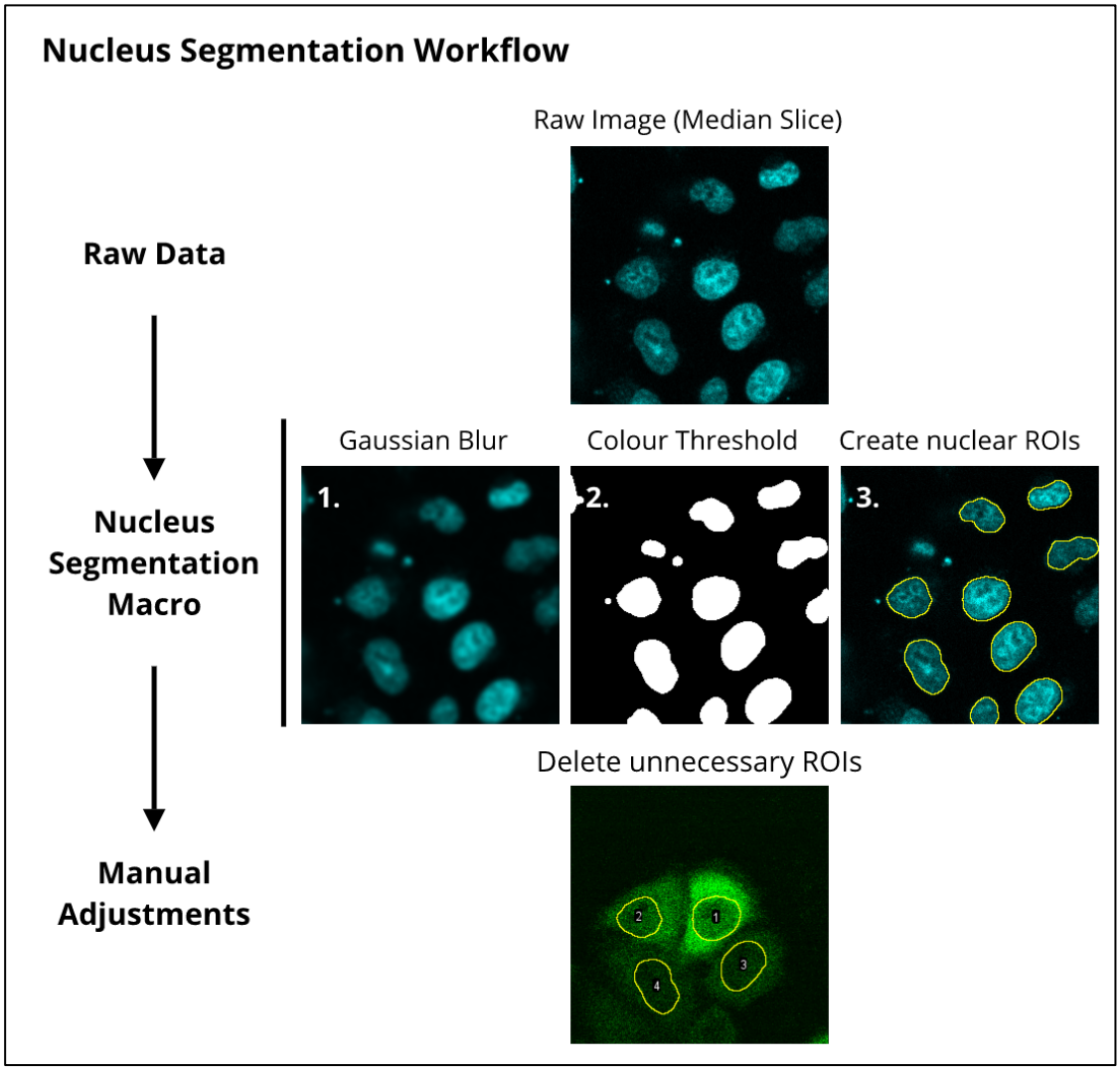


Figure 5 | Nucleus segmentation workflow: *Cell nucleus ROIs were generated in Fiji in a three-step process. First, the median slice running through the middle of the cell nucleus was picked. Then, the Nucleus Segmentation macro was applied that automatically applies a Gaussian blur filter to the Hoechst channel image of the median slice, uses a colour threshold to create a binary image of nuclei and then uses this image to create nuclear ROIs. Lastly, ROIs that corresponded to nuclei of cells that do not express GFP-tagged protein were deleted manually.*

Therefore, nuclear ROIs of non-transfected cells had to be deleted manually. Average GFP and mCherry intensity in the median z-slice image was measured using the ROIs put out by the Nucleus Segmentation macro and transferred to an Excel sheet for further analysis. This process was applied to both images of cells taken before cytokine stimulation and again to images of the same cells captured after treatment for comparison of the two cell states.

3.5.5 Image Analysis Workflow for the Co-localisation Assay

The Nuclear Translocation Assay was used to investigate active STAT dimer formation in living cells based on altered cytokine-induced nuclear accumulation patterns. Inherently, cytokine-induced phenomena cannot be used to assess latent dimer formation because cytokine treatment leads to the transition of latent to active dimers. Therefore, the Co-localisation Assay was developed (see chapter 4.1 for further details). Compared to the Nuclear Translocation Assay, this assay involved an altered image acquisition process using fixed cell samples and widefield microscopy (chapter 3.5.1) instead of live-cell confocal imaging as outlined below.

Measuring total GFP and mCherry fluorescence per cell

To avoid measurement artefacts, only cells containing a relative molecular excess of the GFP-tagged STAT were eligible for further analysis in the two-step image analysis workflow. Relative protein amounts were derived from total GFP and mCherry fluorescence ratios. To determine total GFP and mCherry fluorescence for individual cells, images were segmented using manual drawings of whole cell ROIs based on GFP fluorescence using the Freehand Selection tool in Fiji. ROIs were loaded into the ROI manager and saved for later use. Total GFP and mCherry fluorescence was measured in both channels for individual ROIs using the Analyse Particles function of Fiji. Values were loaded into an excel spreadsheet and relative protein amounts were calculated by comparison to the fluorescence standard (outlined in chapter 4.1.2). Only cells containing equimolar or up to a 4-fold molecular excess of the GFP-tagged STAT were progressed to the second image analysis step.

Quantifying GFP and mCherry co-localisation in individual cells

The Co-localisation Assay scores the degree of colocalization between GFP and mCherry-tagged STAT proteins to determine if the two can form latent dimers. Several plugins that automatically calculate the co-localisation of proteins within individual cells are available in the Fiji/ImageJ software package. In this work, the PSC Colocalization plugin was used for this purpose (French *et al.*, 2008). Here, co-localisation was scored by calculating the average correlation of GFP and mCherry intensity per image pixel, represented by the Pearson correlation coefficient. Essentially, a high negative correlation of r_P=-1 indicated that the fluorescence-tagged

proteins distribute to different compartments, a correlation of r_P=0 showed that distributions of individual proteins are not correlated and a high positive correlation of r_P=1 indicated that both proteins co-localise. STATs were considered to form latent dimers when the correlation of GFP and mCherry distribution was r_P>0.9. For use in the PSC Colocalization plugin, raw 14-bit image files had to be converted into RGB images using ImageJ's image type converter function. The default intensity setting of 40 was used as the image noise threshold within the PSC Colocalization plugin. Pearson's correlation coefficients were calculated for individual cells that met the eligibility criteria outlined above, based on the ROIs generated in the first image analysis step.

4. Results

4.1 Assay Development: The Nuclear Translocation Assay

STATs can form active homo- and heterodimers but it is unclear what governs the formation of one over the other. In this work, this shall be assessed using a novel cell-based assay, termed the Nuclear Translocation Assay. This chapter describes the theory behind the assay and walks the reader through its development.

4.1.1 Introduction

Activation at the cytokine receptor and the formation of the active signalling complex is followed by the accumulation of STATs in the cell nucleus (figure 7A). For several STATs, including STAT1, STAT3 and STAT5, the residues mediating this phenotype have been mapped to an NLS sequence in the DBD. Mutating these residues resulted in abrogated nuclear accumulation following cytokine treatment (figure 7B) (Melén, Kinnunen and Julkunen, 2001; Ma *et al.*, 2003; Iyer and Reich, 2008). Additionally, for STAT1 the phenotype was shown to be dominant negative in STAT1/STAT2 heterodimers, meaning that mutating the NLS (a ΔNLS mutation) in one monomer rendered the entire dimer unable to accumulate inside the nucleus in response to cytokine stimulation (Melén, Kinnunen and Julkunen, 2001; Fagerlund *et al.*, 2002).

Currently, evidence is lacking that the STAT1ΔNLS is dominant negative in STAT1 homodimers or that ΔNLS mutations of other STATs exert a dominant negative effect. However, the presumption that this may be the case inspired the creation of the Nuclear Translocation Assay. Accordingly, should a given STATΔ show a dominant negative effect, then the active dimer formation with another STAT could be assessed by observing their nuclear accumulation pattern when co-expressed in the same cell. If the two proteins form an active signalling complex, then the cytokine-induced nuclear accumulation of both would be diminished (figure 7C). This is because the dominant negative effect of the ΔNLS contained in one monomer would cross-inhibit the cytokine-induced nuclear accumulation of the other STAT if the two are contained in a dimer. If the two do not interact, the cytokine-induced nuclear accumulation of the STATΔNLS but not the other STAT would be diminished (figure 7D). Importantly, should the ΔNLS exert a dominant negative effect on both homo- and heterodimers,

then Nuclear Translocation Assay experiments would not be limited to the assessment of active STAT homodimers but could also be used to detect active STAT heterodimers. The accumulation phenotypes of putative interaction partners were made visible in live-cell imaging by fusing the two STAT constructs with different fluorescent proteins (mCherry and GFP for the STATΔNLS for the other STAT, respectively).

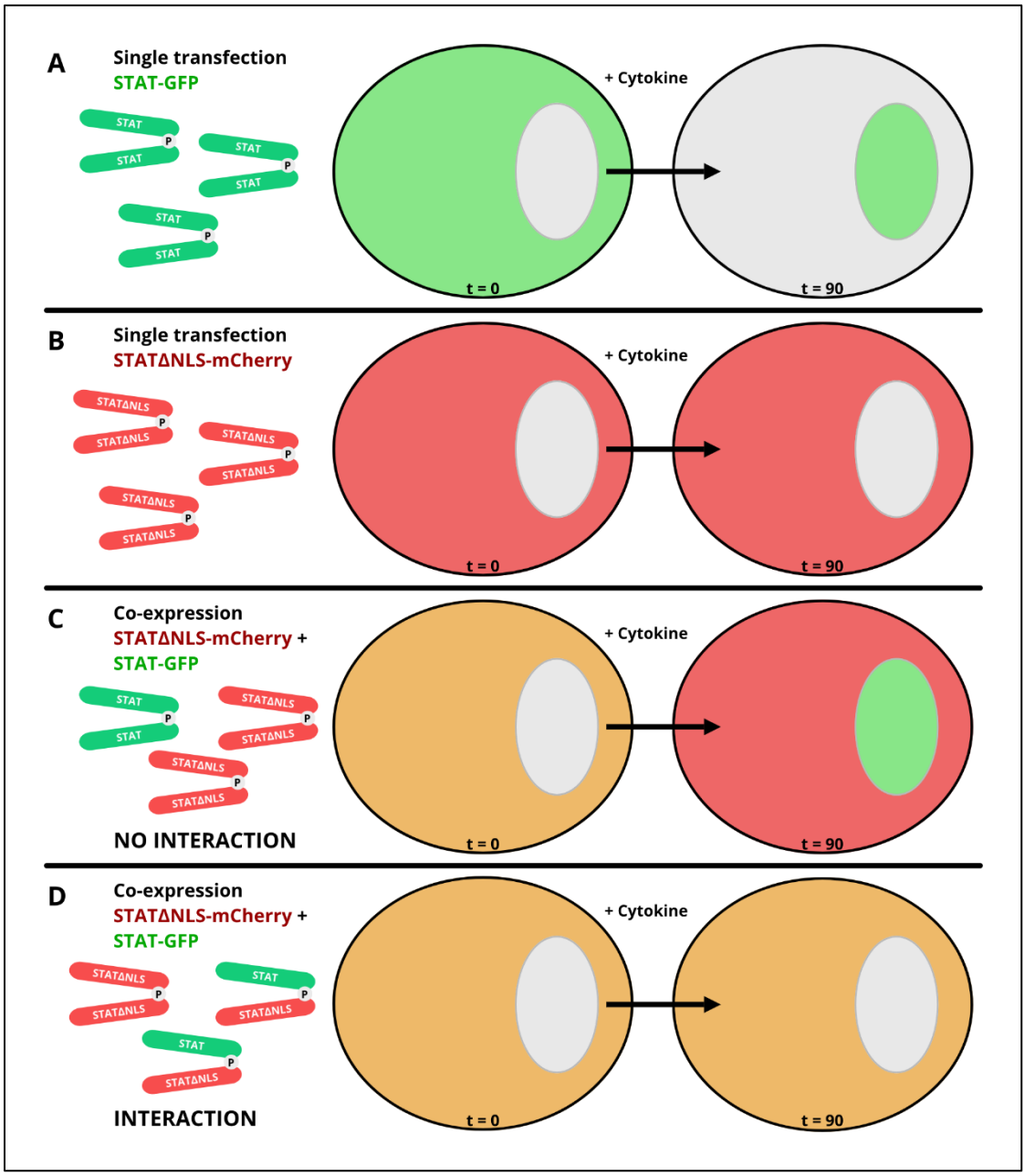


Figure 6 | The Nuclear Translocation Assay can detect STAT homo and heterodimers in live-cells: ***[A]*** *A GFP-tagged STAT is transiently expressed in HeLa cells. Nuclear accumulation is triggered when cells are stimulated with cytokine. This is made visible by a shift in GFP fluorescence from the cytoplasm to the nucleus.* ***[B]*** *When mCherry-tagged STATΔNLS expressing cells are stimulated with cytokine, the subcellular distribution of mCherry fluorescence does not change.* ***[C]*** *If a STAT-GFP and a STATΔNLS-mCherry do not interact cytokine-induced nuclear accumulation of the latter but not the former is diminished.* ***[D]*** *If the STAT/STATΔNLS pair interact, both proteins are retained in the cytoplasm, staining it orange when GFP and mCherry channels are overlaid in microscopy images.*

Accumulation was scored by measuring the cytokine-induced increase in nuclear fluorescence. The formation of active complexes was declared if this score was diminished for the STAT-GFP in the presence of STATΔNLS-mCherry compared to when the STAT-GFP was co-expressed with the corresponding wild-type (wt) of STATΔNLS.
It is important to note that the ΔNLS mutation does not affect the assembly of active STAT1 or STAT3 homodimers (Melén, Kinnunen and Julkunen, 2001; Ma *et al.*, 2003). Therefore, when combining a STAT-GFP and a STATΔNLS-mCherry of the same paralogue, three different dimer species can be formed:

- STAT-GFP/STAT-GFP
- STATΔNLS-mCherry/STATΔNLS-mCherry
- STATΔNLS-mCherry/STAT-GFP

Based on the theory outlined above, of these three, cytokine-induced nuclear accumulation would only be diminished for the latter two. The STAT-GFP/STAT-GFP does not carry the ΔNLS mutation in either monomer. Thus, its accumulation after cytokine treatment should not be inhibited. Since the cytokine-induced change in nuclear GFP fluorescence was used as readout in the Nuclear Accumulation Assay, uninhibited accumulation of residual STAT-GFP/STAT-GFP dimers would lead to measurement artefacts. Consequently, the formation of these dimers had to be minimised, while, at the same time, maximising the assembly of STATΔNLS-mCherry/STAT-GFP dimers. This was achieved by expressing the STATΔNLS-mCherry in molar excess over the STAT-GFP. In this case, the STAT-GFP monomer would be most likely to assemble a dimer with STATΔNLS-mCherry, should the two interact, instead of self-assembly. Of course, having a molar excess of STATΔNLS-mCherry present should also increase the fraction of STATΔNLS-mCherry/STATΔNLS-mCherry dimers formed. However, since dimer formation between STATΔNLS and the other STAT was assessed based on GFP fluorescence, the increased number of STATΔNLS-mCherry/STATΔNLS-mCherry should not distort the measurements.

It is important to note that it is unclear if individual STATs form homo- and heterodimers at equal rates or if they preferably assemble into homo- over heterodimers or vice-versa. Should the latter be true, then the predictions made for

dimer compositions in the last paragraph would not be applicable when investigating STAT heterodimers in the Nuclear Translocation Assay.

Recent reports have shown that co-transfected foreign DNA can compete with each other and intrinsic components for cellular resources such as RNA polymerases, ribosomes, and energy. Furthermore, it was found that this effect is dependent on the DNA concentration ratios used during transfection (Huliák *et al.*, 2012; Di Blasi *et al.*, 2021). Therefore, excess STATΔNLS-mCherry over STAT-GFP expression should be achievable by adapting relative plasmid DNA amounts used during the transfection stage. However, transfection is often accompanied by significant variability in transfection efficiencies within a given cell pool, even if the cells originated from the same batch (Vitor *et al.*, 2018; Di Blasi *et al.*, 2021). For this reason, it was vital to confirm relative protein amounts in individual cells after transfection.

Determining absolute protein amounts based on fluorescence can be a difficult task. This work describes a novel approach for this based on comparison to a fluorescence standard. The fluorescence standard was comprised of a GFP and mCherry molecule fused using an amino acid linker, which should ensure that the two fluorophores are expressed in a known and fixed ratio of one-to-one. Consequently, the ratio of mCherry-to-GFP fluorescence should follow a linear dependency and signify the fluorescence ratio when equal amounts of GFP-tagged, and mCherry-tagged proteins are present within the cell (figure 8, blue line). In turn, extrapolation from this should allow for the determination of relative amounts of individual GFP- and mCherry-tagged proteins present, on a cell-to-cell basis. To control for measurement artefacts evoked by a low STATΔNLS-mCherry to STAT-GFP ratio, a minimum excess of 1.5 times and a maximum excess of 10 times was stipulated (figure 8, green box).

In the following, the process of developing the Nuclear Translocation Assay is outlined. This included verifying the functionality of the GFP-MCherry fusion protein as a fluorescence standard, measuring the dynamics of STAT1, STAT3 and STAT5B cytokine-induced nuclear accumulation to determine the ideal duration of cytokine treatment for the Nuclear Translocation Assay and confirming the ΔNLS phenotype for specific mutations in STAT1, STAT3 and STAT5B. Whether these ΔNLS mutations exert a dominant negative effect on cytokine-induced nuclear accumulation will be assessed in chapter 4.2.

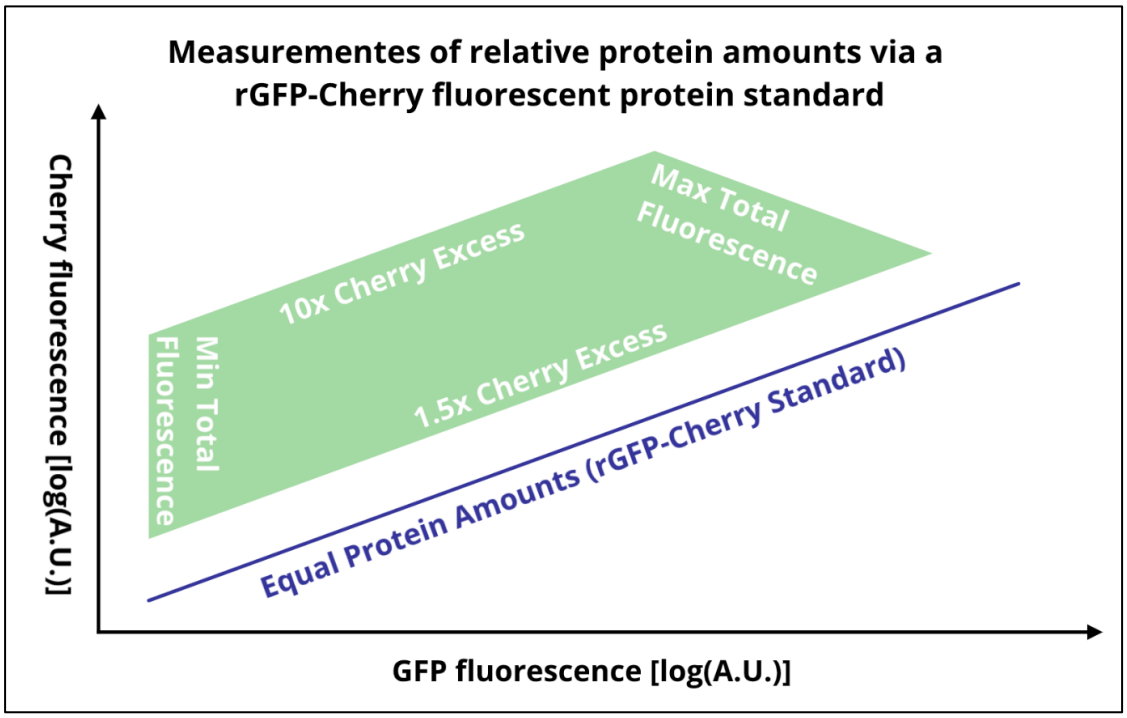


Figure 7 | Relative protein amounts are measured by using a bifunctional fluorescence standard: *Schematic sketch showing the rationale behind measurements of relative protein amounts using a fluorescent standard. Measurements for cells expressing a GFP-mCherry fluorescent standard show a linear dependency when total GFP fluorescence is plotted against total mCherry fluorescence. Since the GFP-mCherry fusion construct ensures that cells express the two fluorophores at a fixed and known ratio of one-to-one, this straight line represents the GFP/mCherry fluorescent ratio at which equal amounts of protein are present in the cell. From this, molar ratios of GFP vs mCherry-tagged proteins can be extrapolated. The required minimum and maximum excess of mCherry-tagged protein to be eligible for assessment in the Nuclear Translocation Assay is indicated as a green box.*

4.1.2 Measuring relative Protein Amounts based on fluorescence intensity.

Relative protein amounts contained in cells co-expressing different GFP- and mCherry-tagged STATs, were extrapolated from fluorescence intensities. For this, the intensities were normalised to the ones measured for a GFP-mCherry fluorescence standard that expressed the fluorophores in a known and fixed ratio of one-to-one. The process of generating and validating the GFP-mCherry fluorescence standard is outlined in the following.

Creating a GFP-mCherry fusion protein

The GFP-mCherry fusion cDNA was generated by excising a GFP cDNA fragment from a pEGFP-C3 vector and cloning it into the multiple cloning site of the pmCherry-N1(lab) vector. The fluorophores were separated by a flexible linker spanning 25 amino acids and thus the generated construct was termed flexible GFP-mCherry (fGFP-mCherry).

To confirm its expression and phenotype, the fGFP-mCherry fusion protein was transiently expressed in HeLa cells for fluorescence microscopy and Western blot analysis. Fluorescent microscopy images were taken using a widefield microscope. GFP and mCherry fluorescence was observed in HeLa cells transfected with the fGFP-mCherry fusion protein confirming the structural integrity of individual fluorophores (figure 9A). For Western blotting, whole-cell extracts of HeLa cells expressing either GFP alone, the fGFP-mCherry fusion protein or a mCherry-tagged STAT1 construct were generated. Individual proteins were detected using an antibody targeted to the GFP epitope. The lane containing extracts from cells transfected with pmGFP-N1 empty vector showed a strong band running at approximately 26 kDa which is consistent with the theoretical molecular weight of GFP (figure 9B, lane 2). No bands were observed in the lane where extracts from STAT1-mCherry expressing cells were loaded. This confirmed that the antibody is specific to GFP and did not detect mCherry (figure 9B, lane 3). The lane containing extracts from cells transfected with the fGFP-mCherry construct showed a strong band running at around 63 kDa which is slightly higher than the theoretical mass of the fGFP-mCherry protein at 56 kDa (figure 9B, lane 1). Importantly, no additional bands were observed in this lane which indicated that the fGFP-mCherry fusion protein was not split into separate GFP and mCherry molecules.

Taken together, this data confirmed the expected phenotype for the fGFP-mCherry protein.

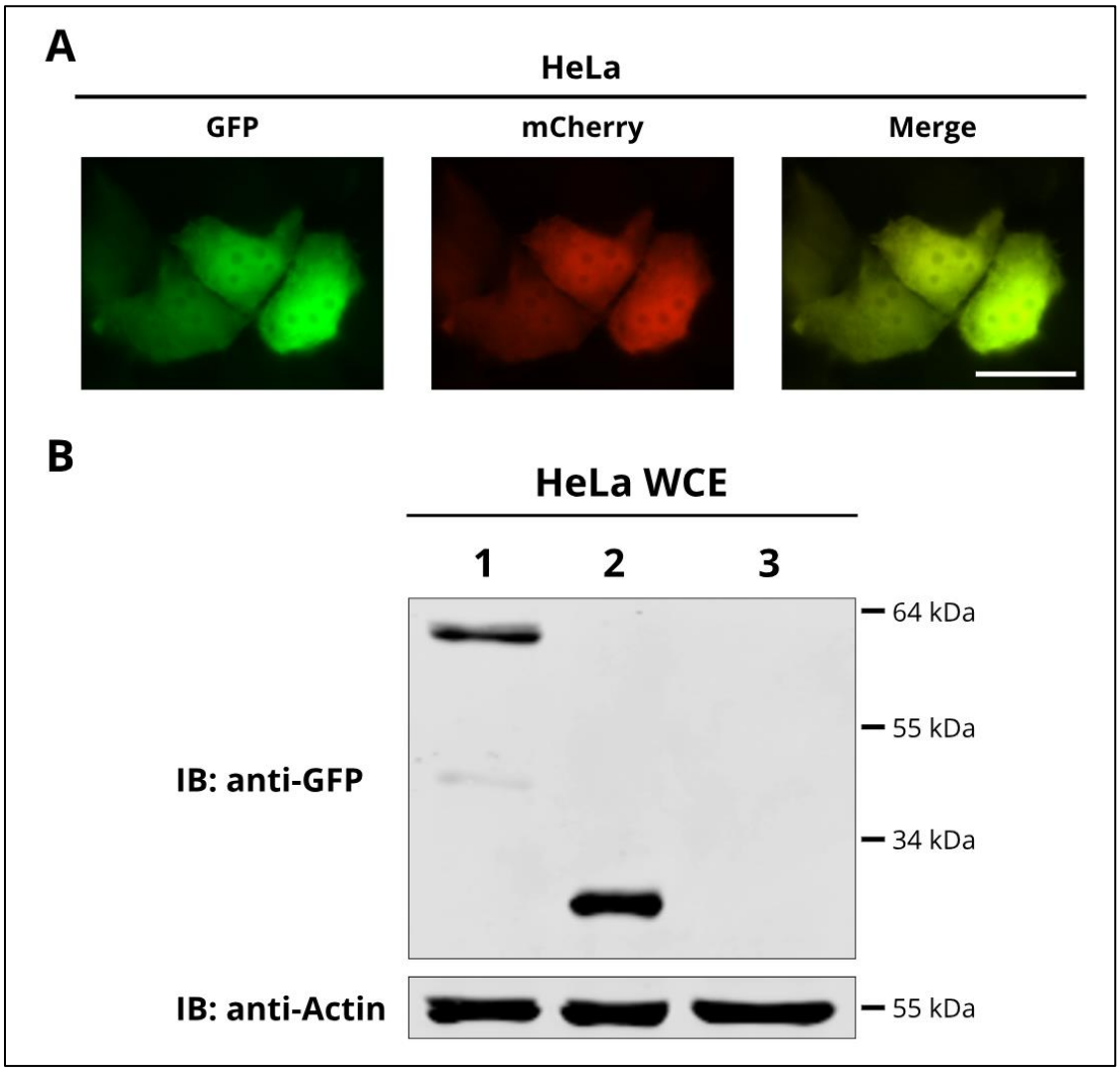


Figure 8 | The fGFP-Cherry protein is intact and shows the expected phenotype: ***[A]*** *Fluorescent micrographs of HeLa cells expressing the fGFP-Cherry fluorescent standard. Images were taken on a widefield microscope.* ***[B]*** *Western blot of HeLa cells expressing the fGFP-Cherry (lane 1), GFP (lane 2) or a Cherry-tagged STAT1 (lane 3). The membrane was probed with anti-GFP and anti-Actin antibodies.*

Cells expressing the fGFP-mCherry fluorescent standard were imaged in parallel to other samples during optimisation and application of the Nuclear Translocation Assay. For reasons explained in detail later (chapter 4.1.3), different settings were used to image cells expressing fluorescent-tagged STAT1, STAT3 and STAT5 proteins, respectively. To determine the relative protein amounts in these samples, the fGFP-mCherry was imaged using the same settings appropriate for individual STATs. When plotting GFP against mCherry intensity measured for fGFP-mCherry expressing cells, the data converged into two straight lines (figure 10). This suggested that the intensities measured for the fGFP-mCherry indeed represented a fixed protein ratio of

one-to-one, GFP-to-mCherry. Data imaged with STAT1 settings aligned to a straight line with a larger slope compared to cells imaged with STAT3 settings (figure 10, red and blue dots, respectively). This indicates that the GFP-to-mCherry fluorescence intensity representing a protein ratio of one-to-one is different when changing imaging settings. Therefore, only samples imaged with the same settings could be compared.

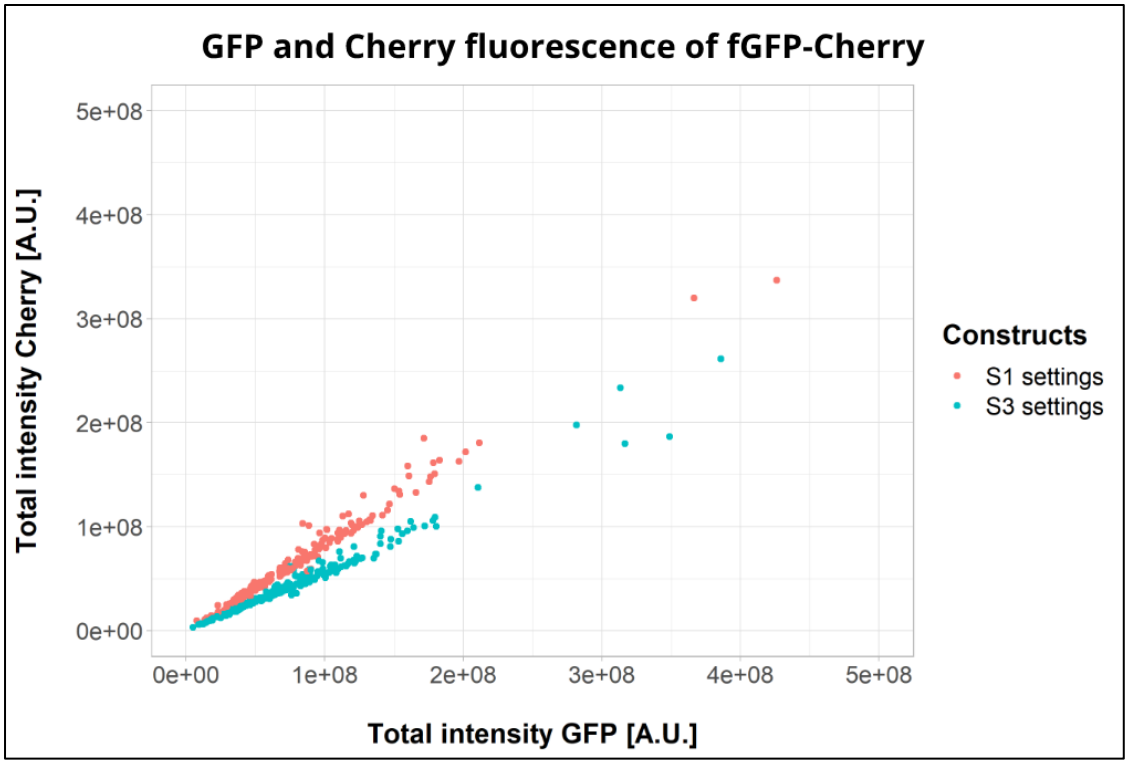


Figure 9 | GFP and mCherry fluorescence of fGFP-Cherry: *Cells expressing the fGFP-Cherry protein were imaged using different imaging settings. Cells imaged using STAT1 settings are indicated in red, while cells imaged using STAT3 settings are indicated in blue.*

Linear regression was used to confirm the linear dependency of GFP/mCherry fluorescence in cells expressing the fGFP-mCherry protein. The algorithm was applied to measurements from individual experiment days and results are summarised in table 19. All measurements showed a significant linear dependency between GFP and mCherry fluorescence ($p < 10^{-13}$) (table 19). Coefficient of determination R^2 and adjusted coefficient of determination adj. R^2 indicated the efficient application of the linear regression algorithm (R^2_{mean} = 0.99 ±0.01, adj. R^2_{mean} = 0.99 ± 0.01). The linear slope of individual measurements varied between 0.8-1 for STAT1 settings (table 19, rows 1-7), meaning that a GFP-to-mCherry intensity ratio of roughly one-to-one represents equal proteins being present in the cell. For STAT3 settings the slope was 0.5-0.7 meaning that an intensity ration of two-to-one/three-to-two indicates the

presence of equal protein amounts. Slopes determined with this method were used to extrapolate the relative protein contents of cells expressing individual GFP-tagged and mCherry-tagged proteins. Taken together, the results from table 19 showed that the fGFP-mCherry fluorescent standard allowed for accurate estimates of the GFP/mCherry fluorescence ratio when fluorophores are present at a fixed ratio of one-to-one. Furthermore, the low variation of slopes determined on different experiment days showed that measurements were consistent. This was especially obvious when linear regression fits for individual experiment days were plotted on a GFP/mCherry fluorescence graph (figure 11).

Table 19 | Linear regression of total GFP-vs-mCherry measurements in cells expressing the fGFP-mCherry protein: *Cells expressing the fGFP-mCherry protein were imaged on several experiment days using different imaging settings. Linear regression was performed using R.*

	No	Experiment Date	Slope	Slope SE	p-value	R^2	adj. R^2
STAT1 settings	**1**	05/11/2021	0.916	0.015	2.7×10^{-20}	0.996	0.995
	2	10/11/2021	1.044	0.029	1.2×10^{-13}	0.991	0.990
	3	17/11/2021	0.896	0.020	6.1×10^{-30}	0.985	0.985
	4	24/11/2021	0.819	0.016	3.2×10^{-32}	0.988	0.987
	5	26/11/2021	0.856	0.017	1.4×10^{-32}	0.987	0.986
	6	01/12/2021	0.784	0.027	1.5×10^{-18}	0.976	0.975
	7	08/12/2021	0.793	0.003	4.6×10^{-42}	1.000	1.000
STAT3 settings	**8**	10/11/2021	0.680	0.020	3.5×10^{-13}	0.989	0.988
	9	12/11/2021	0.645	0.019	1.0×10^{-21}	0.979	0.978
	10	17/11/2021	0.581	0.011	2.6×10^{-32}	0.990	0.989
	11	24/11/2021	0.548	0.010	1.6×10^{-33}	0.990	0.990
	12	26/11/2021	0.594	0.008	5.6×10^{-39}	0.995	0.994
	13	01/12/2021	0.519	0.019	4.3×10^{-13}	0.974	0.973
	14	03/12/2021	0.682	0.018	1.1×10^{-20}	0.985	0.985
	15	08/12/2021	0.532	0.002	3.1×10^{-41}	1.000	0.999
	16	17/12/2021	0.684	0.022	4.2×10^{-25}	0.970	0.969
	17	22/12/2022	0.563	0.015	1.7×10^{-21}	0.985	0.984

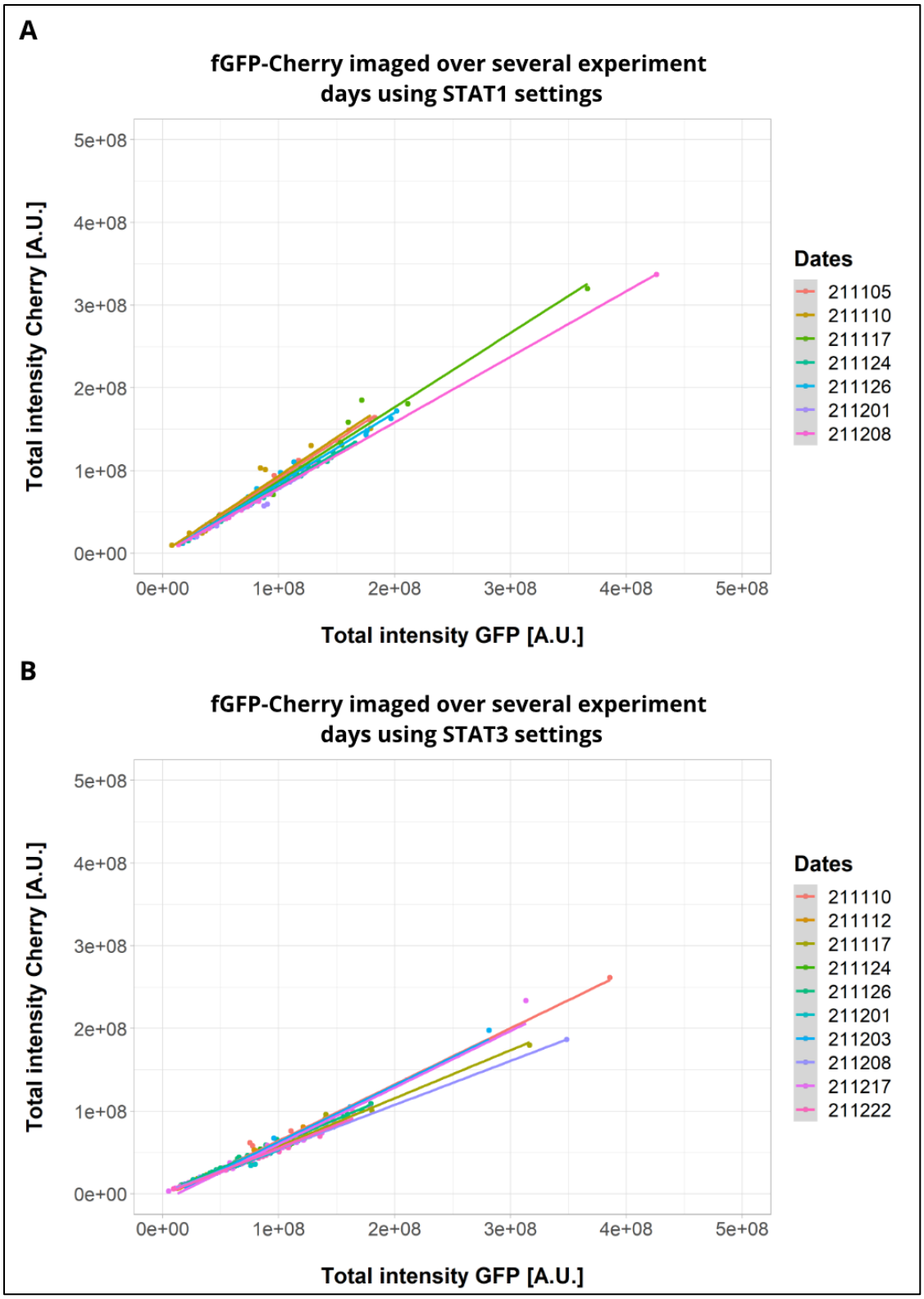


Figure 10 | GFP and mCherry fluorescence of cells expressing fGFP-Cherry show linear dependency: *Cells were imaged using STAT1 (A) or STAT3 settings (B). Linear regression was performed in R and plotted for individual measurements made on different experiment days.*

FRET masks the true GFP-to-mCherry fluorescence ratio of the fGFP-mCherry fusion protein

Förster resonance energy transfer (FRET) occurs when two fluorophores come into proximity and the emission spectrum of the donor molecule overlaps with the excitation spectrum of the acceptor molecule. During FRET, excited state energy from the donor molecule is transferred to the acceptor molecule via long-range dipole-dipole interactions instead of emitting a photon. The acceptor molecule may then emit this energy in the form of light at a higher wavelength (Piston and Kremers, 2007).

The fGFP-mCherry fusion protein was created for the use as fluorescence standard to estimate the GFP-to-mCherry fluorescence ratio that represents the presence of equal molar amounts of GFP-tagged and mCherry-tagged protein within a given cell. However, should FRET occur between the GFP and mCherry moiety of this construct, an accurate determination of the true GFP-to-mCherry fluorescence would no longer be possible. The GFP fluorescence intensity would be underestimated because a part of the excited state energy would be transferred to the acceptor mCherry moiety instead of emitting the energy in form of a photon. The mCherry fluorescence would be overestimated because not only its own excited state energy would be released as light but also the energy it received from the GFP donor moiety via FRET. To test these assumptions, FRET measurements were conducted on HeLa cells transiently expressing the fGFP-mCherry fusion protein.

HeLa cells were fixed using 2% PFA 24h after transfection with fGFP-mCherry cDNA and measurements using a Zeiss LSM710 confocal microscope. The FRET workflow consisted of an initial image taken to confirm fGFP-mCherry expression visually. In this image, 10 cells showing GFP and mCherry fluorescence were picked and marked for photobleaching. The FRET measurements were started by taking five images of the marked cells in the span of 25 seconds. Then, the acceptor fluorophore, mCherry, was bleached using 100% laser power of the 561 nm laser. This was followed by another set of measurements consisting of five images in the span of 25 seconds. This process was repeated for four independent image frames. The average fluorescence intensity within the marked regions was recorded for each time point before and after acceptor bleaching. Figure 12 shows the average mCherry (left) and the average GFP fluorescence (right) per pixel over time. Within the first set of measurements, mCherry

fluorescence remained stable. The same is true for the second set of measurements. This indicated that the fluorophore resisted photobleaching by the 561 nm laser when using standard imaging settings.

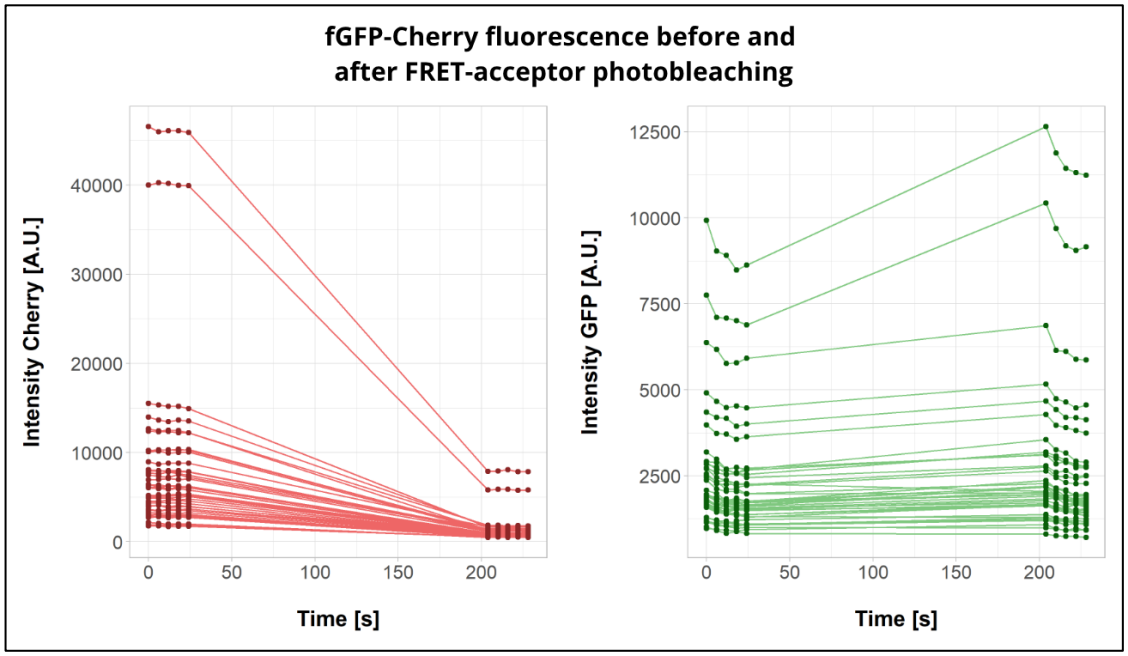


Figure 11 | *fGFP-Cherry fluorescence before and after FRET-acceptor photobleaching:* Cells expressing the fGFP-Cherry fluorescent standard were imaged every five seconds for five repeats. The FRET acceptor (Cherry) was bleached, and the cells reimaged five times every five second. The left plot shows average Cherry intensity per pixel over time for 40 regions of interest. The right plot shows average GFP intensity per pixel over time for the same 40 regions.

Bleaching reduced the mCherry fluorescence to 25 % of the original signal on average (figure 12, left; appendix table 1, compare columns 3 and 5). Considerable variation of the GFP fluorescence was observed within measurement sets before and after bleaching (figure 12, right). As such, the intensity measured in the first image of an image set was on average 12 % higher than in the last image. This suggested that GFP was photobleached by the 488 nm laser when using normal imaging settings. Average GFP fluorescence increased after FRET acceptor photobleaching suggesting that FRET occurred (figure 12, right; appendix table 1). To confirm this suspicion, the FRET efficiency E_{FRET} was calculated using formula I, where I_{DA} is the GFP fluorescence measured in the last image before, and I_D is the GFP fluorescence in the first image after bleaching.

$$E_{FRET} = 1 - \frac{I_{DA}}{I_D} \qquad \text{(I)}$$

The distance r between fluorophores was calculated using formula II, where R_0 is the Förster distance.

$$r = R_0 \left(\frac{1}{E_{FRET}} - 1 \right)^{\frac{1}{6}} \quad \text{(II)}$$

The Förster distance R_0 is defined as the distance between fluorophores where E_{FRET} is 50 % and can be calculated for individual FRET pairs (formula III). The Förster distance is a function of the quantum yield of the FRET donor Q_D (for GFP Q_D = 0.6), the refractive index of the mutual environment n (for water n=1.33), the orientation of dipoles which is commonly assumed to be κ = 2/3 and the spectral overlap integral J(λ). The spectral overlap integral for the GFP/mCherry FRET pair is J(λ) = 1.94 x 10^{15} $M^{-1}cm^{-1}nm^4$ and therefore the Förster distance is 5.288 nm.

$$R_0 = 0.211 * \sqrt[6]{\kappa^2 n^{-4} Q_D J(\lambda)} \quad \text{(III)}$$

Results for FRET efficiencies and distances between fluorophores calculated for 40 cells expressing the fGFP-mCherry protein are summarised in appendix table 1. The average FRET efficiency in cells expressing the fGFP-mCherry was 13.97% with a standard deviation of 8.80%. The average distance between the GFP and mCherry fluorophore was 7.11 nm with a standard deviation of 0.75 nm.

In summary, the experiments described above confirmed the expression of the fGFP-mCherry fluorescence standard as a full-length fusion protein (figure 9). Furthermore, GFP and mCherry fluorescence showed a linear dependency, indicating that it represents the GFP/mCherry fluorescence ratio when a protein ratio of one-to-one was present within cells (table 18; figure 11). However, FRET was detected in cells expressing the fGFP-mCherry construct. This means that the GFP-to-mCherry fluorescence ratio measured for the fGFP-mCherry construct does not represent the GFP-to-mCherry fluorescence ratio apparent when equal molar amounts of individual GFP-tagged and mCherry-tagged proteins (that would most likely not undergo any FRET) are present. Instead, the GFP fluorescence intensity is underestimated and mCherry fluorescence intensity overestimated. Therefore, a considerable error would be expected when extrapolating relative protein amounts of other GFP- and mCherry-tagged proteins based on the fGFP-mCherry fluorescence standard.

A rigid linker does not prevent GFP-mCherry FRET

FRET can only occur when two fluorophores with overlapping donor emission and acceptor excitation spectrums come in close contact with each other. As shown in formula I, the efficiency of the FRET process E_{FRET} is proportional to the inverse sixth power of the distance r between the participating molecules. Thus, it was hypothesised that increasing the distance between the FRET acceptor and the FRET donor would lead to reduced FRET efficiencies. In turn, this would minimise the measurement error when estimating relative protein amounts based on the fluorescence of the GFP-mCherry standard. Arai *et al.* compared different unordered and α-helical linkers of variable lengths on their ability to separate the EGFP and EBFP moieties of an EGFP-EBFP fusion protein via FRET measurements. They found that the linker termed HL5 because it contains five repeats of the EAAAK motive (Marqusee and Baldwin, 1987) (full amino acid sequence: LAEAAAKEAAAKEAAAKEAAAKEAAAKAAA) was most effective in reducing EBFP/EGFP FRET (Arai *et al.*, 2001). This linker was later termed "rigid" based on small-angle X-ray scattering experiments and associated analysis (Arai *et al.*, 2004).

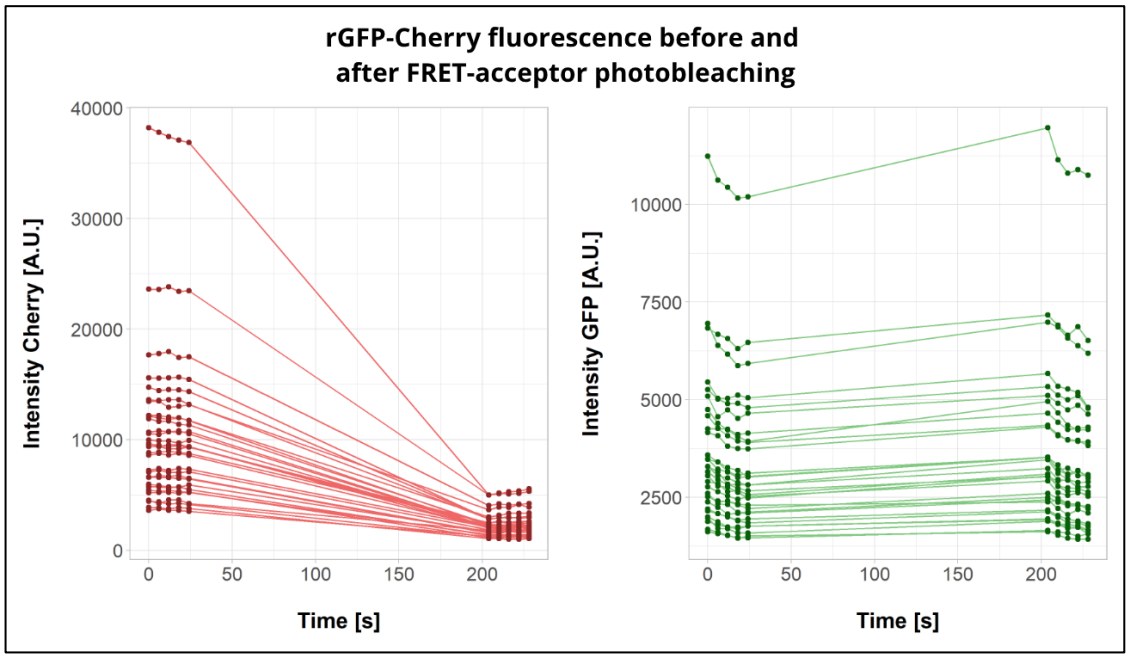


Figure 12 | rGFP-Cherry fluorescence before and after FRET-acceptor photobleaching: *Cells expressing the rGFP-Cherry fluorescent standard were imaged every five seconds for five repeats. The FRET acceptor (Cherry) was bleached, and the cells reimaged five times every five second. The left plot shows average Cherry intensity per pixel over time for 30 regions of interest. The right plot shows average GFP intensity per pixel over time for the same 30 regions.*

To reduce FRET between the fluorophores used in this study, a rigid GFP-mCherry (rGFP-mCherry) construct was generated spanning the HL5 peptide between the GFP

and mCherry moieties. To assess the efficiency of FRET, the rGFP-mCherry cDNA was transfected into HeLa cells and subjected to FRET measurements as described before.

Similar to the results obtained for the fGFP-mCherry construct, acceptor bleaching reduced mCherry fluorescence by 75 % (appendix table 2). The mCherry fluorescence within individual measurement sets comprising five pictures taken in a span of 25 seconds before and after bleaching remained constant (figure 13, left). Average GFP fluorescence increased when comparing images taken just after bleaching with ones taken just before, hinting at the presence of FRET (figure 13, right). Again, GFP fluorescence was negatively affected by imaging with standard settings. Within a set of five images taken before FRET acceptor bleaching, the average GFP fluorescence decreased with every image taken. The same was observed in the five images taken after photobleaching (figure 13, right).

To confirm if FRET was present, the FRET efficiencies were as before using formulas I-III. Calculations were done for 30 individual images. A summary of the results can be found in appendix table 2. On average, a FRET efficiency of 12.68% with a standard deviation of 3.63% was observed. Therefore, the average distance between fluorophores was 7.36 nm with a standard deviation of 0.49 nm. The introduction of a rigid linker reduced the FRET efficiency by 1.3 percentage points compared to when a flexible linker was used. However, care must be taken when interpreting these results. The useful range for exact estimates of FRET efficiencies and distances between fluorophores is generally accepted as 0.5-1.5x R_0. This is due to the inverse 6^{th}-power dependence on the distance r between fluorophores and the signal-to-noise ratio that can be reasonably achieved using standard microscopy techniques (Piston and Kremers, 2007). The low boundary for accurate estimates of FRET efficiencies (1.5x R_0) for the GFP/mCherry donor-acceptor pair is 7.932 nm, so reasonably close to the average distance between GFP and mCherry measured for fGFP-mCherry and rGFP-mCherry (7.11±0.75 nm and 7.36±0.49 nm, respectively). Nevertheless, it was clear that fluorescence standards were capable of low-level FRET. Still, measurements of FRET efficiencies for rGFP-mCherry showed less variance compared to fGFP-mCherry. Perhaps this is due to the rigidity of the α-helical linker keeping the fluorophores at a consistent distance. For this reason, the rGFP-mCherry was used as fluorescence

standard in the Nuclear Translocation Assay hereon after. Still, to compensate for the margin of error expected from FRET, only cells expressing a minimum of 1.5 times more mCherry-tagged over GFP-tagged protein were eligible for analysis in the Nuclear Translocation Assay.

The rGFP-mCherry fluorescence standard was imaged as a control in parallel to experiments for the Nuclear Translocation Assay using separate imaging settings for STAT1, STAT3 or STAT5 expressing cells (for reasons explained in a later chapter). When plotting total GFP fluorescence against total mCherry fluorescence measured for the rGFP-mCherry standard, the data converged to three straight lines based on the imaging settings used (figure 14). The data obtained using STAT1 settings (figure 14, red dots) appeared to converge to a straight line with a greater slope than when STAT3 settings were used (figure 14, green dots). The data obtained from imaging the rGFP-mCherry with STAT5 settings (figure 14, blue dots) also converged to a straight line with a slope similar to that from images taken with STAT3 settings.

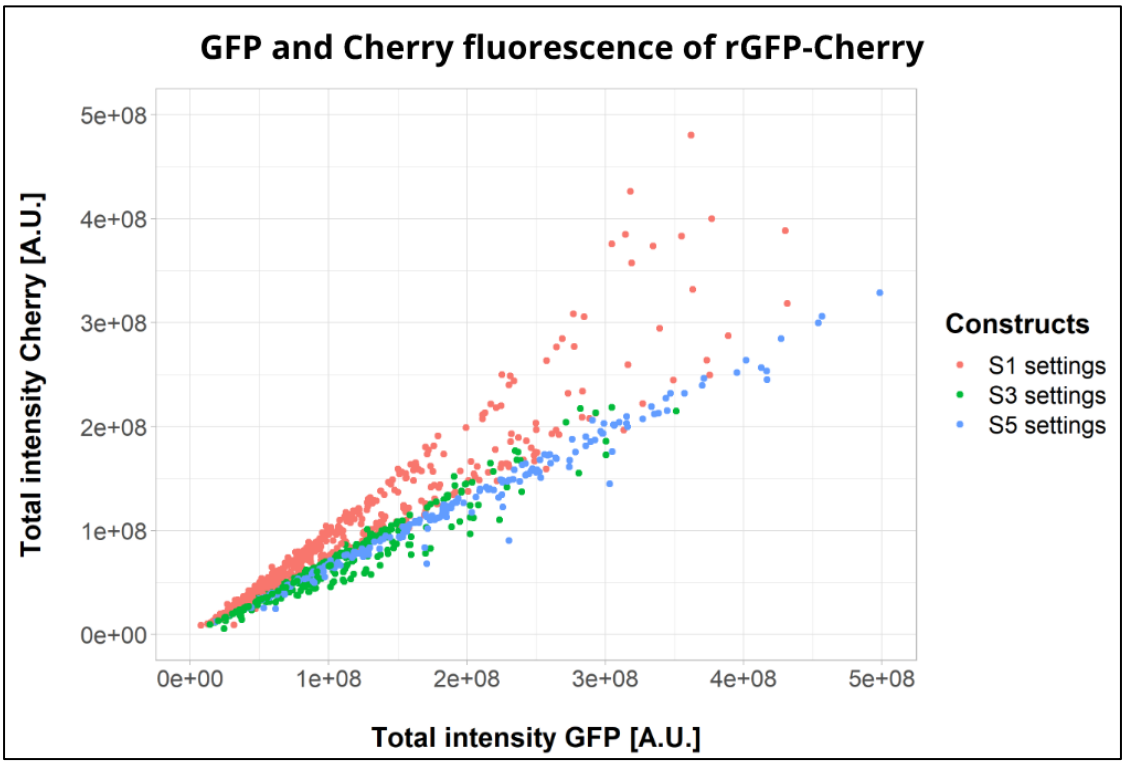


Figure 13 | GFP and mCherry fluorescence of rGFP-Cherry: *Cells expressing the rGFP-Cherry protein were imaged using different imaging settings. Cells imaged using STAT1 settings are indicated in red, while cells imaged using STAT3 settings are indicated in blue.*

Indeed, GFP and mCherry fluorescence intensities showed a linear relationship irrespective of which imaging settings were used, as confirmed when linear regression

was performed on individual data sets for different experiment days (figure 15, table 22). Still, as observed visually, slopes determined for linear regression curves indicated that the degree of the fluorescence relationship was dependent on the imaging setting applied. When using STAT1 settings, slopes were generally the highest (table 22, rows 1-16; figure 15A) with an average slope of 0.96±0.05. The slopes reported for imaging data obtained using STAT3 and STAT5 settings were similar with average slopes of 0.68±0.03 and 0.64±0.01, respectively (figure 15B and C; table 22, row 17-36).

As for fGFP-mCherry measurements, the coefficient of determination, R^2, was used to assess goodness-of-fit for linear regression. On average a coefficient of determination of R^2_{mean} = 0.993±0.004 and an adjusted coefficient of determination of adj. R^2_{mean} = 0.993±0.004 was determined when the linear relationship of GFP to mCherry fluorescence was assessed for rGFP-mCherry expressing cells (table 22, column 6+7). Furthermore, the linear dependence was significant for all measurements ($p < 10^{-18}$). Overall, the regression line slopes determined for the rGFP-mCherry were slightly higher than for fGFP-mCherry. For images taken using STAT1 settings, the average slope was 0.95 for rGFP-mCherry and 0.87 for fGFP-mCherry. Similarly, the average slopes for images taken using STAT3 settings were 0.68 for rGFP-mCherry and 0.60 for fGFP-mCherry. Considering that GFP fluorescence would be underestimated and mCherry fluorescence overestimated due to errors caused by FRET, a higher slope of the GFP/mCherry regression line indicated that FRET was reduced in the rGFP-mCherry when compared to fGFP-mCherry. Moreover, goodness-of-fit was improved for rGFP-mCherry in contrast to fGFP-mCherry measurements, though the difference was small. Nevertheless, this strengthened the decision to use rGFP-mCherry construct instead of fGFP-mCherry as a fluorescence standard in Nuclear Translocation Assay experiments.

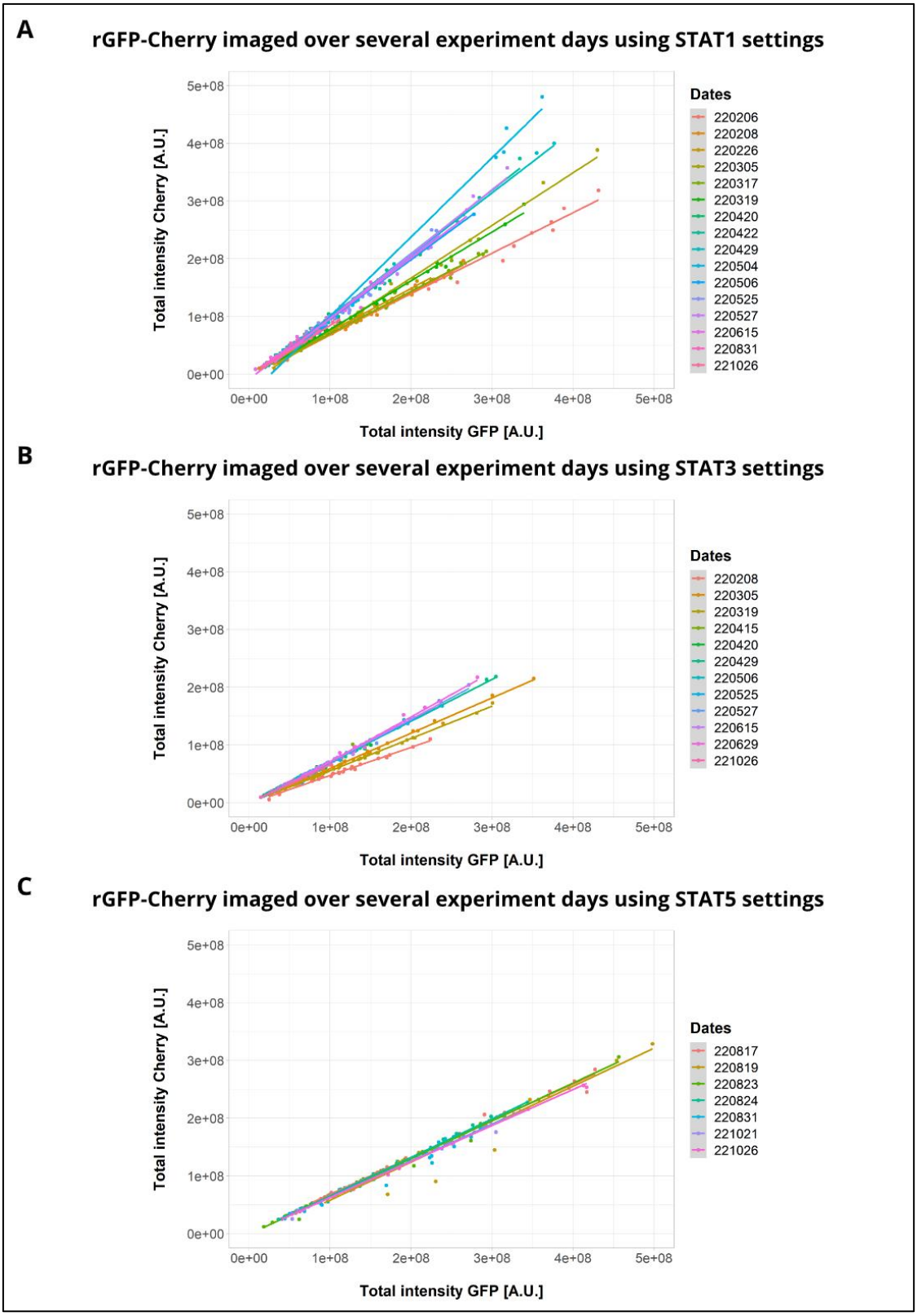


Figure 14 | GFP and mCherry fluorescence of cells expressing rGFP-Cherry show linear dependency: *Cells were imaged using STAT1 (A) or STAT3 settings (B). Linear regression was performed in R and plotted for individual measurements made on different experiment days.*

Table 20 | Linear regression of total GFP-vs-mCherry measurements in cells expressing the rGFP-mCherry protein: *Cells expressing the rGFP-mCherry protein were imaged on several experiment days using different imaging settings. Linear regression was performed using R.*

	No	Experiment Date	Slope	Slope SE	p-value	R^2	adj. R^2
STAT1 settings	1	06/02/2022	0.701	0.017	2.3×10^{-24}	0.985	0.985
	2	08/02/2022	0.737	0.012	1.2×10^{-30}	0.993	0.993
	3	26/02/2022	0.742	0.008	1.1×10^{-36}	0.997	0.997
	4	05/03/2022	0.910	0.017	1.3×10^{-24}	0.992	0.992
	5	17/03/2022	0.732	0.010	9.7×10^{-37}	0.995	0.994
	6	19/03/2022	0.844	0.010	3.7×10^{-38}	0.996	0.996
	7	20/04/2022	1.028	0.012	9.1×10^{-50}	0.994	0.994
	8	22/04/2022	1.100	0.014	2.4×10^{-35}	0.995	0.995
	9	29/04/2022	1.075	0.010	1.6×10^{-46}	0.997	0.997
	10	04/05/2022	1.372	0.030	3.5×10^{-24}	0.989	0.989
	11	06/05/2022	1.029	0.012	2.5×10^{-48}	0.995	0.995
	12	25/05/2022	1.078	0.022	2.1×10^{-32}	0.986	0.986
	13	27/05/2022	1.113	0.017	7.2×10^{-39}	0.991	0.991
	14	15/06/2022	1.048	0.021	5.9×10^{-32}	0.987	0.987
	15	31/08/2022	0.992	0.026	7.4×10^{-34}	0.973	0.973
	16	26/10/2022	0.886	0.018	1.9×10^{-29}	0.988	0.988
STAT3 settings	17	08/02/2022	0.486	0.008	1.9×10^{-30}	0.993	0.993
	18	05/03/2022	0.611	0.007	4.7×10^{-30}	0.997	0.997
	19	19/03/2022	0.559	0.004	6.1×10^{-45}	0.998	0.998
	20	15/04/2022	0.714	0.021	1.3×10^{-18}	0.984	0.984
	21	20/04/2022	0.709	0.006	4.1×10^{-57}	0.997	0.997
	22	29/04/2022	0.719	0.005	8.4×10^{-51}	0.998	0.998
	23	06/05/2022	0.711	0.007	3.8×10^{-52}	0.997	0.996
	24	25/05/2022	0.732	0.009	1.1×10^{-39}	0.995	0.995
	25	27/05/2022	0.745	0.009	1.0×10^{-42}	0.995	0.995
	26	15/06/2022	0.736	0.007	9.9×10^{-42}	0.997	0.997
	27	29/06/2022	0.775	0.016	1.6×10^{-20}	0.992	0.992
	28	26/10/2022	0.662	0.004	8.7×10^{-46}	0.999	0.999
STAT5 settings	29	17/08/2022	0.637	0.014	4.1×10^{-21}	0.991	0.991
	30	19/08/2022	0.640	0.024	3.8×10^{-19}	0.966	0.965
	31	23/08/2022	0.667	0.006	1.4×10^{-41}	0.997	0.997
	32	24/08/2022	0.669	0.005	1.8×10^{-40}	0.998	0.998
	34	31/08/2022	0.632	0.013	5.6×10^{-38}	0.983	0.983
	35	21/10/2022	0.627	0.011	3.0×10^{-36}	0.993	0.993
	36	26/10/2022	0.624	0.004	6.1×10^{-43}	0.999	0.999

4.1.3 Dynamics of STAT1, STAT3 and STAT5B nuclear accumulation

Activation at the cytokine receptor is followed by rapid accumulation of STATs inside the nucleus. As explained earlier, in Nuclear Translocation Assay experiments interaction between the two putative interaction partners was declared when the cytokine-induced nuclear accumulation of a STAT-GFP was cross-inhibited by a co-expressed STATΔNLS-mCherry. To ensure assay robustness, cells needed to be imaged when nuclear accumulation was maximal. Furthermore, the dominant-negative effect of the STATΔNLS would be expected to be maximal at this time point. In the following, the process of quantifying cytokine-induced nuclear accumulation over time for STAT1, STAT3 and STAT5B is described. Generally, cells were stimulated with both IFN-γ and IL-6 to efficiently activate STAT1 and STAT3 they were combined for heterodimer assessment later on in this work. While this may not be necessary when experimenting with either homodimer, cells were still stimulated with both cytokines to ensure comparability and a consistent experimental setup. As outlined below, the cytokine stimulation setup had to be modified when examining STAT5B to ensure efficient activation of this protein.

STAT1 nuclear accumulation plateaus 90 min after cytokine treatment

To investigate the dynamics of STAT1wt nuclear accumulation following treatment with cytokines, HeLa cells were grown on a cell culture dish suitable for live-cell imaging and transfected with cDNA encoding STAT1wt-GFP. Ten image frames containing GFP-fluorescent cells were imaged before and then stimulated with recombinant IFN-γ, IL-6 and sIL6Rα (the latter was added because HeLa cells do not express this receptor chain). The same cells were then re-imaged 45, 60, 75, 90, 105 and 120 minutes after cytokine treatment.

Before stimulation STAT1wt-GFP predominantly resided in the cytoplasm but a weak signal remained detectable inside the nucleus (figure 16A, panel 1). Nuclear accumulation of the GFP signal was apparent as early as 45 min after cytokine treatment (figure 16A, panel 2) as evident by co-localisation with Hoechst 33342 nuclear counterstain (not shown) and GFP fluorescence inside the nucleus remained strong over the rest of the time course (figure 16A, panel 3-7).

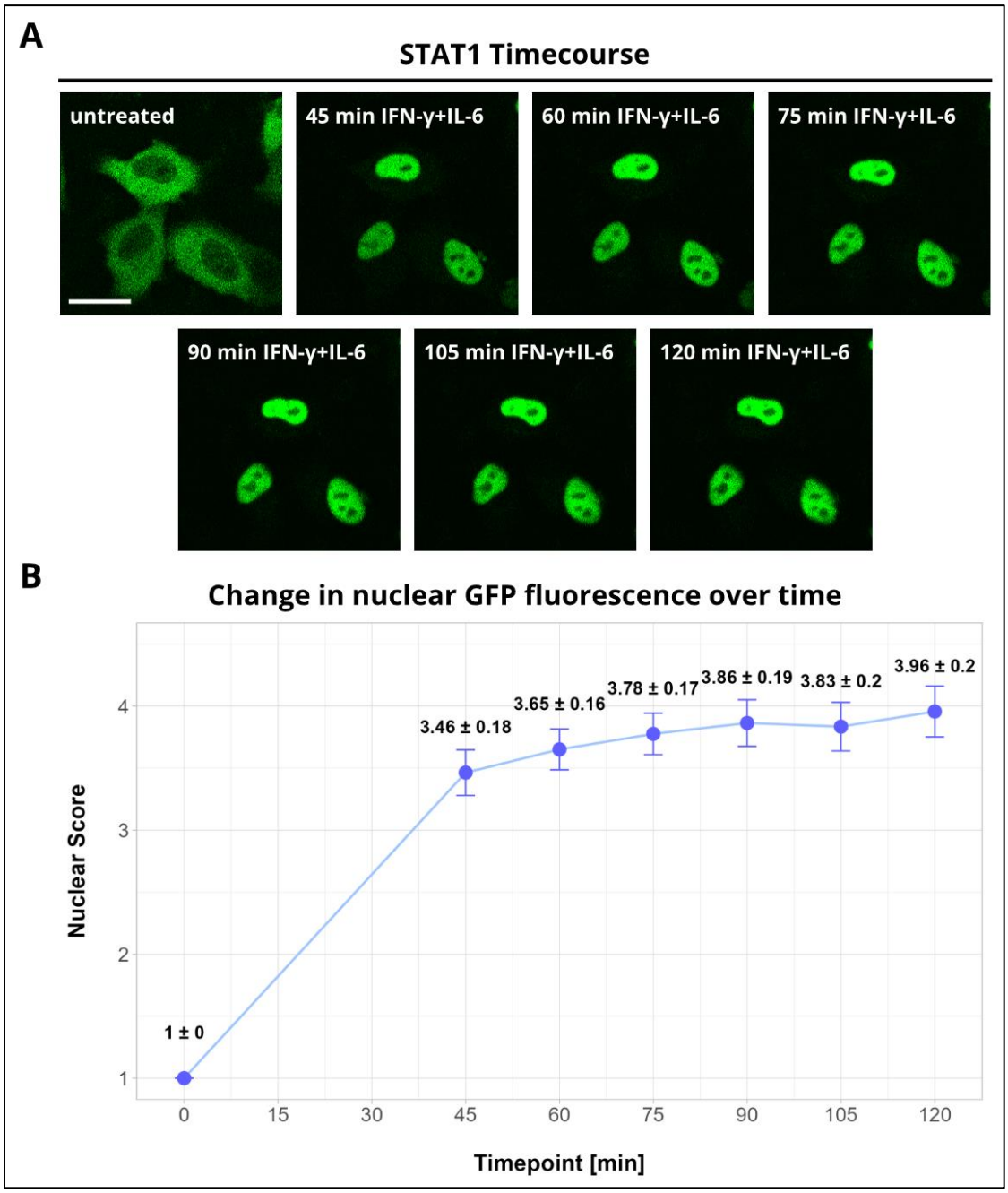


Figure 15 | Nuclear accumulation of STAT1wt over time: *HeLa cells were transfected with STAT1wt-GFP. Ten frames were captured as a z-stack image before stimulation IFN-γ and IL-6. More z-stacks of the same cells were taken in 15 min intervals between 45-120 min post cytokine treatment.* ***[A]*** *Example image of STAT1wt nuclear accumulation over time. The median slice of a 7-image z-stack is shown at different timepoints before and after IFN-γ+IL-6 stimulation.* ***[B]*** *The change of average nuclear GFP fluorescence over time was determined for 20 cells expressing STAT1wt-GFP.*

Cytokine-induced nuclear accumulation was quantified by calculating the nuclear score F_n using formula IV. Essentially, mean fluorescence intensity in the nucleus at time point t $\left(\bar{I}_{n,t}\right)$ was divided by the average nuclear fluorescence intensity before cytokine stimulation $\left(\bar{I}_{n,t_0}\right)$. Thus, the nuclear score represents the fold increase of mean nuclear fluorescence in the period t_0 – t.

$$F_n = \frac{\bar{I}_{n,t}}{\bar{I}_{n,t_0}} \qquad \text{(IV)}$$

Forty-five minutes after cytokine treatment, the nuclear score was 3.5±0.2. Thereafter, the nuclear score increased further before levelling off 90 min after the addition of cytokine to 3.9±0.2. This was followed by nuclear scores of 3.8±0.2 and 4.0±0.2 for 105 min and 120 min, respectively (figure 16B).

In summary, it was found that the nuclear accumulation of STAT1wt plateaus 90 min after cytokine treatment. Based on this, the 90 min time point appeared ideal for application in the Nuclear Translocation Assay.

STAT3 nuclear accumulation also plateaus 90 min after cytokine treatment.

The STAT3 nuclear accumulation time course experiment was essentially conducted as described for STAT1 above. Compared to STAT1wt-GFP which resided mostly in the cytoplasm (figure 16A, panel 1), the STAT3wt-GFP molecules were distributed more equally between the nucleus and cytoplasm before cytokine treatment (figure 17A, panel 1). Treatment with IFN-γ, IL-6 and sIL6Rα triggered the nuclear accumulation of STAT3wt-GFP as confirmed by co-localisation with Hoechst nuclear counterstain (not shown) and was first observed 45 min after stimulation (figure 17A, panel 2). The nuclear score at this time point was 1.9±0.07 and only increased slightly thereafter. Indeed, nuclear scores, determined from images taken 60-120 min after the addition of cytokines, all ranged around 2 (figure 17B). Notably, the nuclear scores determined for STAT3wt were considerably lower than those calculated for STAT1wt (figure 16B vs 17B). This was likely due to the different subcellular distributions of STAT1wt and STAT3wt before the addition of cytokine. As described by formula IV, the nuclear score represents the fold increase in average nuclear fluorescence in period t_0-t. Thus, due to a higher share of STAT3wt-GFP already being present in the nucleus before cytokine treatment, the addition of IFN-γ, IL-6 and sIL6Rα could only trigger a relatively moderate increase in nuclear GFP fluorescence. On the contrary, a relatively high share of STAT1wt-GFP relocated from the cytoplasm to the nucleus upon cytokine treatment, which was reflected in a high nuclear score.

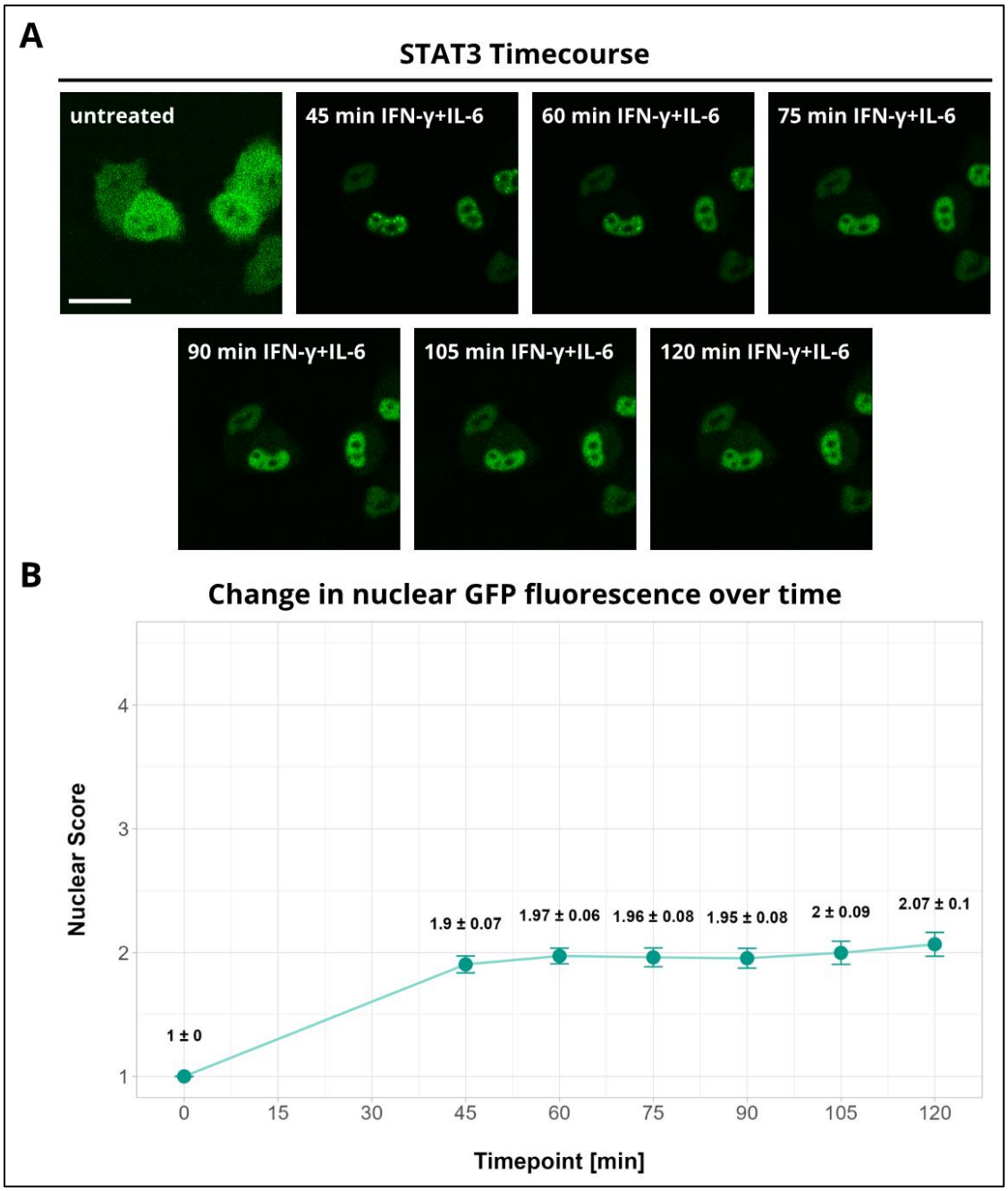


Figure 16 | Nuclear accumulation of STAT3wt over time: *HeLa cells were transfected with STAT3wt-GFP. Ten frames were captured as a z-stack image before stimulation IFN-γ and IL-6. More z-stacks of the same cells were taken in 15 min intervals between 45-120 min post cytokine treatment.* ***[A]*** *Example image of STAT3wt nuclear accumulation over time. The median slice of a 7-image z-stack is shown at different timepoints before and after IFN-γ+IL-6 stimulation.* ***[B]*** *The change of average nuclear GFP fluorescence over time was determined for 20 cells expressing STAT3wt-GFP.*

In several cells analysed for this time course, STAT3wt was observed to crystallise as illustrated by bright GFP spots in and outside the nucleus. In the example shown in figure 17A, these paracrystals disappeared 90 min after cytokine treatment (panel 2 + 3), however, this was not the case for all cells showing these structures. Paracrystal formation has been observed before in multiple cell lines, including primary cells and

mouse acute-phase liver sections (Herrmann *et al.*, 2004; Droescher *et al.*, 2011). Importantly, due to the high GFP intensities measured in these crystals, the imaging settings originally optimised for STAT1wt-GFP expressing cells could not be used to image cells expressing STAT3wt-GFP due to detector oversaturation. Therefore, it was decided that all samples containing STAT3 transfected cells (including this time course) would be imaged with settings specifically optimised for this construct. Details on the optical setup for different experimental constraints can be found in Appendix A.

Overall, the STAT3wt-GFP nuclear accumulation time course confirmed that the nuclear score was maximal 90 min after the addition of cytokine and remained reasonably constant thereafter. Therefore, the 90 min time point was declared ideal for the application in the Nuclear Translocation Assay.

STAT5Bwt shows nuclear accumulation dynamics distinct from STAT1wt and STAT3wt

STAT5B is the main STAT activated downstream of cytokine receptors like GHR, erythropoietin receptor (EPOR), thrombopoietin receptor (TPOR) and several interleukins (Kanai, Jenks and Nadeau, 2012). Additionally, STAT5B activation has been reported downstream of other cytokine receptors that predominantly activate other STATs, including type I and type II IFNs and IL-10 (Grimley, Dong and Rui, 1999).

Examination of protein-protein interactions using the Nuclear Translocation Assay required efficient activation and subsequent nuclear accumulation of STATs. However, cell lines routinely grown during this work like HeLa, HEK 293, and HEK 293T only show low intrinsic expression of the prototypical STAT5B activating receptor GHR (Peritz *et al.*, 1988; Milward *et al.*, 2004; Guesdon *et al.*, 2012). Therefore, STAT5B activation had to be further characterised before time course experiments could be conducted. For this, STAT5Bwt-mCherry was transfected into HEK 293 cells stably expressing the human GHR, a kind gift from Dr Ian Wilkinson, University of Sheffield, UK. After 24 h of incubation, cells were either left untreated or treated with either 50U/mL IFN-γ or 100 ng/mL GH for 1h. IFN-γ stimulation was tested for convenience as it was routinely used for other experiments in this work and because it has been previously shown to activate STAT5B (Meinke *et al.*, 1996). Whole-cell protein extracts were generated and subjected to Western blotting using antibodies specific to full-length STAT5 and the STAT5B pY694 phosphopeptide.

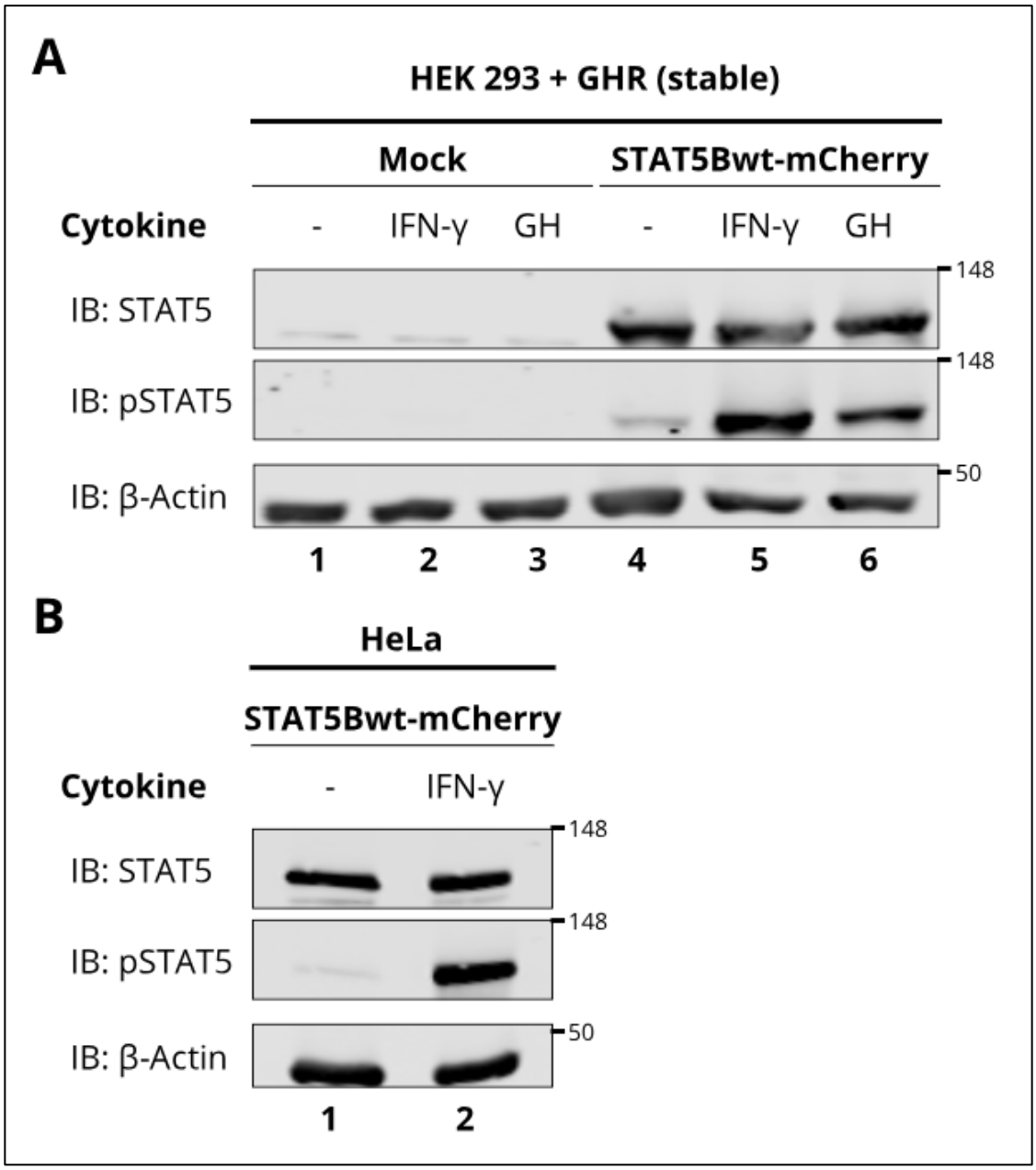


Figure 17 | STAT5B is activated by both GH and IFN-γ: ***[A]*** *HEK 293+GHR (stable) were either left untransfected or transfected with STAT5Bwt-mCherry and subsequently stimulated with IFN-γ, GH or left unstimulated. A Western blot was conducted to specifically detect STAT5 and phosphorylation of STAT5 at Y694 in whole cell extracts. Actin was used as loading control.* ***[B]*** *HeLa cells were transfected with STAT5Bwt-mCherry and treated with IFN-γ or left untreated. STAT5 and phosphorylation of STAT5 at Y694 was detected using immunoblot assays. Actin was used as loading control. MW markers in kDa.*

No full-length STAT5 was detected in non-transfected cells (figure 18A, lanes 1-3) but a strong band appeared when cells were transfected with STAT5Bwt-mCherry cDNA (figure 18A, lanes 4-6). The low levels of phosphorylated STAT5Bwt could already be detected in unstimulated cells but this increased further upon cytokine treatment (figure 18A, lanes 4-6). More phosphorylated STAT5B was found in cells treated with IFN-γ than GH confirming previous studies that showed efficient activation of STAT5B

by the former cytokine (Meinke *et al.*, 1996; Grimley, Dong and Rui, 1999). This result prompted the assumption that IFN-γ treatment may also be sufficient for efficient STAT5B activation and nuclear accumulation in other cell lines. Therefore, the Western blot experiments were repeated with HeLa cells transfected with STAT5Bwt-GFP. Again, cells were left untreated or treated with IFN-γ for 60 min. Detection with an antibody specific for STAT5 showed that transfection resulted in the expression of STAT5B (figure 18B, lane 1+2). Furthermore, a strong signal was detected when probing for phosphorylated STAT5B in whole-cell extracts from IFN-γ stimulated cells (figure 18B, lane 2). Low-level activation of STAT5B was observed in cell extracts from untreated cells (figure 18B, lane 1). These results suggested that IFN-γ was a potent activator of STAT5B in both HEK 293 and HeLa cell lines.

To examine if IFN-γ can also trigger the efficient nuclear accumulation of STAT5B, HeLa cells were grown on microscopy-grade dishes and co-transfected with STAT5Bwt-mCherry and STAT5Bwt-GFP for live-cell imaging. Cells showing GFP and mCherry fluorescence were imaged before, 60min, and 90 min after the addition of IFN-γ. Nuclear accumulation was judged based on co-localisation with the Hoechst 33342 counterstain. Both STAT5Bwt constructs showed a pancellular distribution before cytokine addition, suggesting that the protein's behaviour is not abnormally altered by fusion to different fluorescent proteins (figure 19A, top row). The distribution of neither STAT5B was different 60 min (not shown) or 90 min after treatment with IFN-γ (figure 19A, bottom row). This implied that STAT5B phosphorylation at the IFNGR did not trigger sufficient nuclear accumulation. Earlier results indicated that STAT5B is activated in HEK 293 cells stably expressing the GHR when stimulated with GH, which may translate into nuclear translocation. However, it was decided to stick to the HeLa cell model since quantitative imaging of STAT1wt and STAT3wt nuclear accumulation dynamics was already done in HeLa. It was shown previously that the co-transfection of GHR is a viable technique to induce nuclear accumulation of STAT5B-GFP in response to GH treatment in COS-7 cells (Herrington *et al.*, 1999). Therefore, this was tested for HeLa cells and GHR plasmid DNA co-transfected alongside STAT5B-GFP and STAT5B-mCherry. Again, cells that showed GFP and mCherry fluorescence were imaged before, 60 min and 90 min after treatment with GH. STAT5B showed similar resting distributions as before in both the GFP and mCherry channel, which implied that the

co-transfection of GHR had no adverse effect on this (figure 19B, top row). This time, moderate nuclear accumulation was observed in images taken 90 min after cytokine addition. Yet, considerable fluorescence remained detectable inside the cytoplasm (figure 19B, bottom row). Similar observations were made in images taken 60 min after cytokine treatment (not shown). Receptor availability could be ruled out as the reason for incomplete GH-induced nuclear accumulation because the GHR was expressed under a CMV promoter element.

Like most cytokine receptors, GHR lacks intrinsic tyrosine kinase activity and must associate with one of the four mammalian Jaks to initiate signalling. While GHR receptors have been shown to interact with several other Jaks, Jak2 was necessary for the activation of STAT5 (Argetsinger *et al.*, 1993; Han *et al.*, 1996; Hellgren *et al.*, 1999). Jak2 is also used by other cytokine signalling pathways, some of which are ubiquitously existent in most cell types, including the models used in this work, HEK 293(T) and HeLa cells (Parganas *et al.*, 1998). As a result, it was speculated that the amount of endogenous Jak2 in these cells is not sufficient to enable efficient signalling through overexpressed GHR and other signalling pathways. To test if this can be alleviated through overexpression of Jak2 alongside GHR, STAT5B-GFP and STAT5B-mCherry, HeLa cells growing on microscopy-grade dishes were transfected with the appropriate plasmid DNAs. Again, thirty image frames were captured before, 60 min, and 90 min after the addition of GH. Subcellular localisation of STAT5B-GFP and STAT5B-mCherry before cytokine treatment was similar to previous experiments (figure 19, top rows). When images taken 90 min after GH addition were compared to images taken of GH-treated HeLa cells co-expressing GHR, STAT5B-GFP and STAT5B-mCherry from earlier (figure 19B), a further increase in GH-induced nuclear accumulation was observed in both fluorescent channels. However, STAT5B accumulation still appeared incomplete when compared to images of cytokine-treated cells expressing either STAT1wt or STAT3wt (compare figures 16A, 17A and 19C). Nevertheless, combining GH treatment with the overexpression of GHR and Jak2 was deemed the method of choice for STAT5B activation, since a cytokine-induced effect on nuclear accumulation was distinguishable in microscopy images.

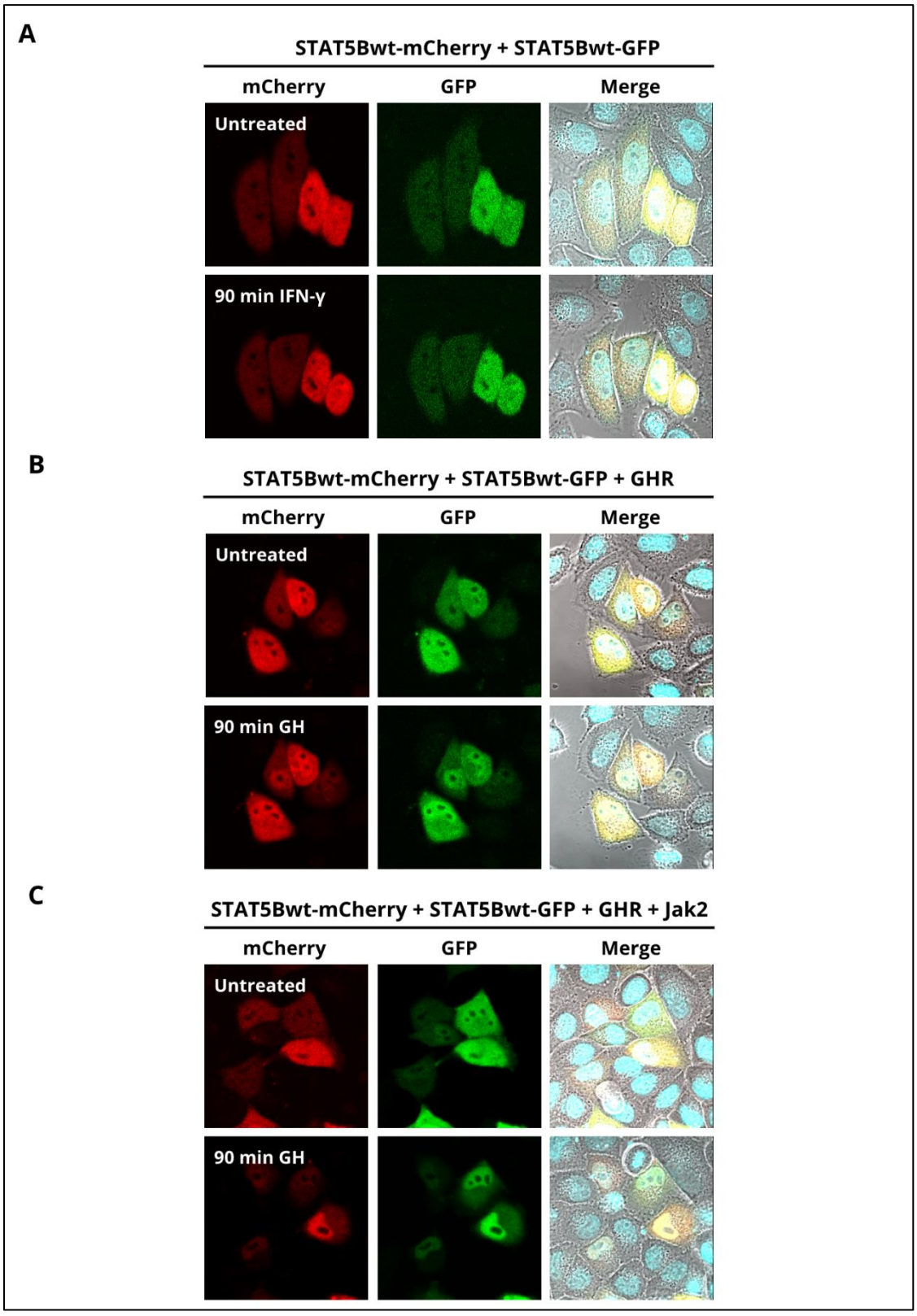


Figure 18 | STAT5B activation at the GHR but not IFNGR results in nuclear accumulation: ***[A]*** *HeLa cells were transfected with STAT5Bwt-GFP and STAT5Bwt-Cherry and imaged before and 90 min after IFN-γ addition.* ***[B]*** *HeLa cells were transfected as above alongside GHR DNA and imaged before and 90 min after treatment with GH.* ***[C]*** *HeLa cells were transfected with STAT5Bwt-GFP, STAT5Bwt-Cherry, GHR and Jak2 and imaged before and 90 min after addition of GH. Merged show the mCherry, GFP, Hoechst and Phase 3 Contrast channels.*

Next, the dynamics of STAT5B cytokine-induced nuclear accumulation over time was examined. For this HeLa cells were transfected with plasmid DNA coding for STAT5Bwt-GFP, GHR and Jak2 and imaged before cytokine addition and at 15 min intervals between 45-120 min thereafter. Importantly, the cytokine cocktail used to stimulate these cells included both GH and the prototypical activator of STAT1 IFN-γ since the two paralogs were combined in further experiments presented in later chapters. As before, the images taken for a total of 18 individual cells were quantified and the nuclear scores $F_{n,S5B}$ were calculated using formula IV. A representative image frame showing the time course of STAT5B nuclear accumulation following treatment with cytokines is depicted in figure 20A. In untreated cells, STAT5Bwt-GFP showed pancellular distribution, which was consistent with previous observations (figure 20A, panel 1). Nuclear accumulation, evidenced by co-localisation with Hoechst stain (not shown), was apparent as early as 45 min after the start of GH stimulation (figure 20A, panel 2). Around the 75 min mark, paracrystals started to appear that primarily localised in the cytoplasm (figure 20A, panel 4). Both the size and the number of these paracrystals increased thereafter but were still mainly localised inside the cytoplasm (figure 20A, panel 5-7). At the 105 min time point, the majority of cells showed aggregates of GFP-tagged STAT5B that seemed to outline the inside of the nuclear membrane (figure 20A, panel 6). The intensity of these aggregate structures increased further over time as seen in images taken 120 min after GH addition (figure 20A, panel 7). The appearance of STAT5B paracrystals matched those of previous accounts from the literature (Herrmann *et al.*, 2004). Furthermore, the STAT5B paracrystals observed in this experiment also looked similar to those observed in STAT3wt-GFP expressing cells, though their subcellular localisation differed (figure 17A). To the best of my knowledge, the STAT5B aggregates outlining the nuclear membranes that appeared at 105 min of GH stimulation were not reported before.

To quantify the cytokine-induced nuclear accumulation of STAT5B over time, nuclear scores were calculated using formula IV. The nuclear score at the time point 45 min was 1.5±0.1. Over the next 45 min, the nuclear score increased further before reaching its maximum at t = 90 min with 1.9±0.2. For the next 15 min, the nuclear score remained fairly constant (1.9±0.2 at t = 105 min) but started to decrease thereafter as indicated by a nuclear score of 1.8±0.1 Eafter 120 min (figure 20B). The absolute values

for the maximum nuclear score obtained for STAT5B were similar to those obtained for STAT3wt-GFP (figure 17B) but considerably lower than those found for STAT1wt-GFP (figure 16B). Again, this difference is likely explained by different subcellular distributions before cytokine treatment observed for the three paralogues. STAT5Bwt and STAT3wt both showed pancellular distributions while STAT1wt was more cytoplasmic.

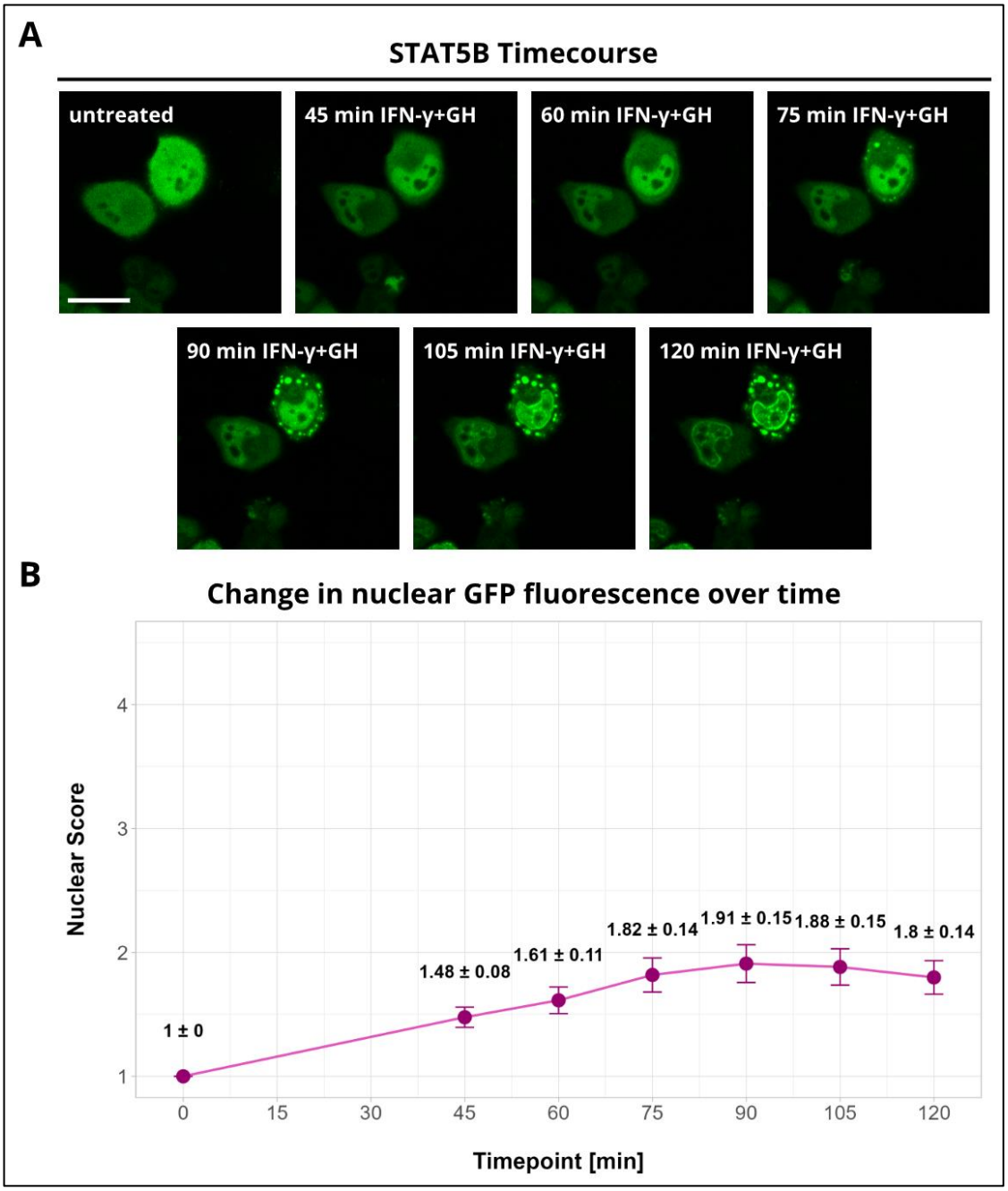


Figure 19 | Nuclear accumulation of STAT5Bwt over time: *HeLa cells were transfected with STAT5Bwt-GFP, GHR and Jak2. Ten frames were captured as a z-stack image before stimulation IFN-γ and GH. More z-stacks of the same cells were taken in 15 min intervals between 45-120 min post cytokine treatment.* ***[A]*** *Example image of STAT5Bwt nuclear accumulation over time. The median slice of a 7-image z-stack is shown at different timepoints before and after IFN-γ+GH stimulation.* ***[B]*** *The change of average nuclear GFP fluorescence over time was determined for 18 cells expressing STAT5Bwt-GFP.*

Taken together, these results indicate that the cytokine-induced nuclear accumulation of STAT5B reaches its maximum 90 min after stimulation onset. This and the observation that accumulation remained stable for at least 15 min afterwards makes the 90 min mark ideal for the application in the Nuclear Translocation Assay.

All in all, the reports presented in this chapter sought to examine the nuclear accumulation dynamics of STAT1wt, STAT3wt and STAT5Bwt following their activation at the relevant cytokine receptors. In the Nuclear Translocation Assay cells were imaged before and after the addition of cytokine to decide whether a STATΔNLS could inhibit the cytokine-induced nuclear accumulation of a putative interaction partner. As explained at the start of this chapter, to detect the maximum inhibitory effect, cells should be imaged at the time point where uninhibited cytokine induced nuclear accumulation is maximal. Furthermore, the nuclear fluorescence should be stable for the next few minutes after this time point. This is because capturing several image frames using confocal laser scanning microscopy can take several minutes. The 90 min time point fulfilled these criteria for STAT1wt, STAT3wt and STAT5Bwt (figure 16, 17 and 20). Consequently, it was deduced that the timeframe t = 90-105 min after cytokine addition could be used to obtain consistent results when imaging multiple frames consecutively. Using the imaging settings optimised for this work means that 30 individual image frames can be captured within this timeframe.

4.1.4 A conserved stretch of basic amino acids in the DBD of STATs determines cytokine-induced nuclear accumulation.

To assess dimerisation in the Nuclear Translocation Assay, a STATΔNLS was combined with another putative STAT interaction partner. The formation of active dimers between the two STATs was declared if the STATΔNLS could cross-inhibit the cytokine-induced nuclear accumulation of the other STAT. This chapter seeks to identify and characterise ΔNLS mutations for STAT1, STAT3 and STAT5B to decide if they were suitable for application in the Nuclear Translocation Assay.

The STAT1 KK410/413AA ΔNLS mutant cannot accumulate inside the nucleus following cytokine treatment.

A stretch of basic amino acids in its DNA binding domain has been identified as essential for carrier-dependent nuclear import of activated STAT1 before. In their study, the authors found that STAT1 mutant KK410/413AA was incapable of nuclear accumulation after IFN-γ treatment and further, also cross-inhibited nuclear import of STAT2 in a dominant negative fashion. Nevertheless, activation by phosphorylation and the formation of STAT1 homo- and heterodimers were not affected (Melén, Kinnunen and Julkunen, 2001). Therefore, STAT1-KK410/413AA should contain all properties required for its use in investigations of STAT1 homo- and heterodimers via the Nuclear Translocation Assay. The STAT1-KK410/413AA mutant will be termed STAT1ΔNLS from now on.

To confirm the nuclear accumulation defect of STAT1ΔNLS, HeLa cells were transfected with either DNA encoding for the STAT1ΔNLS-mCherry or STAT1wt-mCherry. In the previous chapter, the 90 min time point after cytokine addition was identified as ideal for detecting abrogated nuclear accumulation of STAT1ΔNLS. Therefore, in this experiment, cells were imaged before the addition of IFN-γ, IL-6 and sIL6Rα and 90 min thereafter.

Figure 21A shows sample images of STAT1wt-mCherry or STAT1ΔNLS-mCherry expressing cells. Both constructs showed a similar subcellular distribution in latent cells with the majority of mCherry fluorescence located in the cytoplasm (figure 21A, top panels). However, while almost all of the mCherry fluorescence co-localised with the nuclear dye in STAT1wt expressing cells 90 min after cytokine treatment, no difference

in subcellular distribution compared to images taken before cytokine treatment was observed for STAT1ΔNLS-mCherry (figure 21A, bottom panels). These results confirmed accounts from the literature that STAT1ΔNLS carrying the mutation KK410/413AA was incapable of cytokine-induced nuclear accumulation.

To rule out that this phenomenon is caused by defects in cytokine-induced tyrosine phosphorylation, the two STAT1 constructs were transiently expressed in U3R cells. This cell line lacks endogenous STAT1 but stably expresses the IFNGR (Muller *et al.*, 1993) and was used to avoid interference between endogenous and transiently expressed STAT1. The cells were either stimulated with IFN-γ for 60 min or left untreated. Whole-cell lysates were prepared as described (see chapter 3.4.1) and subjected to Western blotting using antibodies against full-length STAT1, STAT1 pY701 and β-Actin. Detection of full-length STAT1 showed that STAT1wt-GFP and STAT1ΔNLS-GFP were similarly well expressed in U3R cells (figure 21B, compare lanes 1-3). No phosphoprotein was detected in extracts from cells expressing STAT1wt-mCherry that were left untreated (figure 21B, lane 1). In contrast, IFN-γ treatment induced phosphorylation of both STAT1wt-mCherry and STAT1ΔNLS-mCherry. Notably, a stronger band was detected for the STAT1ΔNLS construct (figure 21B, lane 2+3). To quantify the relative amounts of activated protein for both STAT1 constructs, band intensities from this and two further Western blots from independent experiments were quantified. The proportion of phosphorylated protein was calculated by dividing the intensity measured for bands detected with the anti-STAT1 pY701 antibody by the intensity measured for the corresponding band when detecting with the anti-STAT1 antibody. This confirmed that the proportion of activated STAT1ΔNLS-mCherry was significantly higher (755%) than the proportion of activated STAT1wt-mCherry (figure 21C). Taken together, this shows that STAT1ΔNLS is significantly better phosphorylated than STAT1wt, ruling out a defect in tyrosine phosphorylation as a cause for the previously observed nuclear accumulation defect.

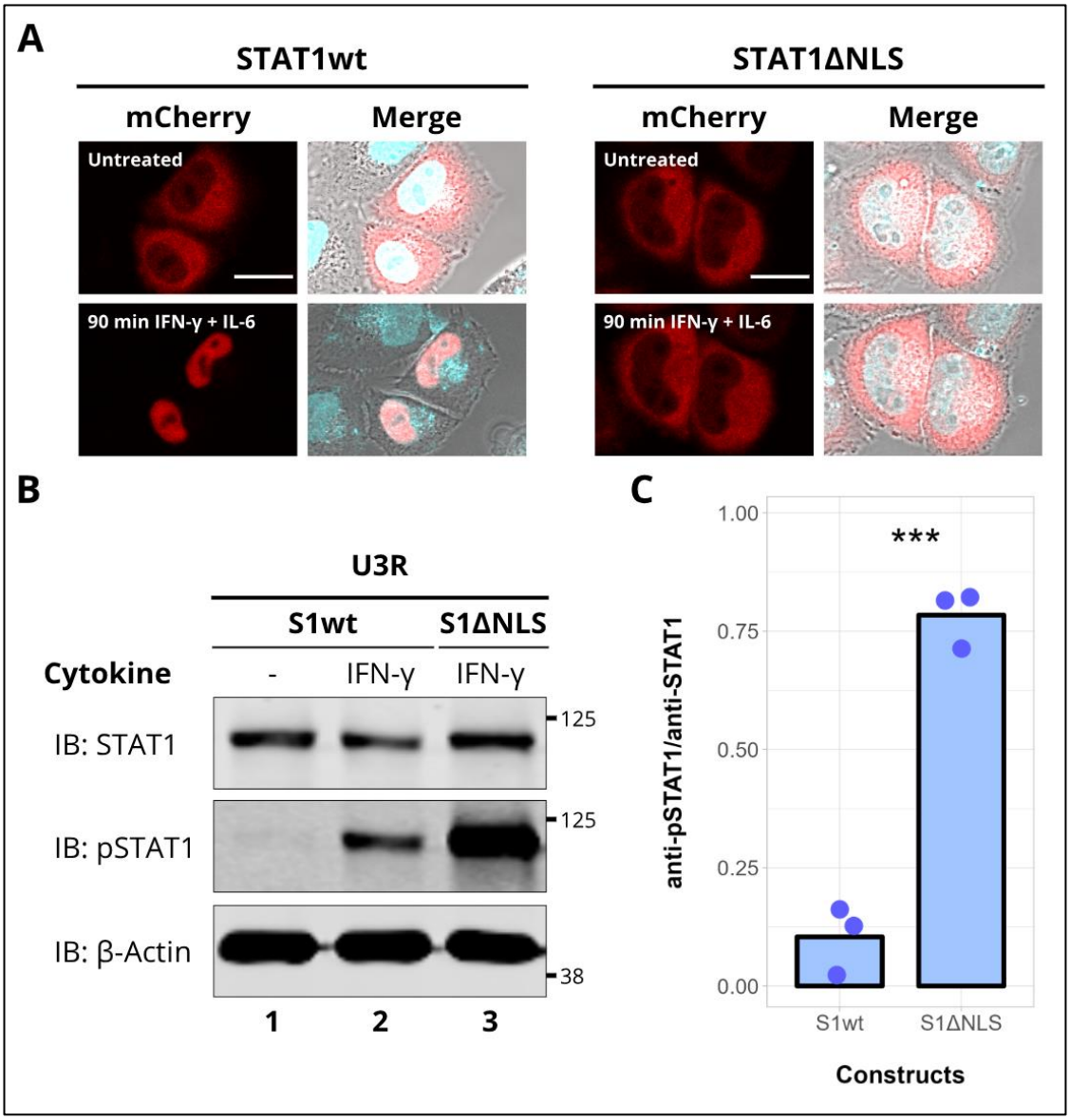


Figure 20 | STAT1ΔNLS subcellular distribution remains unchanged after IFN-γ and IL-6 treatment: ***[A]*** *Fluorescent micrographs of mCherry-tagged STAT1wt or STAT1ΔNLS expressing cells before or 90 min after IFN-γ and IL-6 treatment. The scale bar represents 20 µm. Merge shows the mCherry, Hoechst and Phase 3 Contrast.* ***[B]*** *Whole cell lysates of U3R cells expressing either GFP-tagged STAT1wt or STAT1ΔNLS were subjected to Western blot analysis. Anti-STAT1, anti-pSTAT1 and anti-Actin antibodies were used to detect specific proteins.* ***[C]*** *Quantification of Western blots from [B] and two further independent experiments. Shown is the ratio of fluorescent signal detected for anti-pSTAT1 antibody divided by fluorescence of the anti-STAT1 antibody for IFN-γ treated cells containing either GFP-tagged STAT1wt or STAT1ΔNLS. MW markers in kDa.*

The STAT3 RR414/417AA ΔNLS mutant fails to accumulate inside the nucleus after cytokine treatment.

Ma *et al.* identified two stretches containing surface-exposed arginines as essential for the nuclear accumulation of STAT3 after treatment with IL-6. The first stretch,

R214/R215, is situated in the CCD of STAT3 while the other one comprised of R414/R417 is in the DBD. Interestingly, R414/R417 of STAT3 is in a similar region to the STAT1ΔNLS at residues K410/K413. Indeed, when the authors investigated the properties of a STAT3 RR414/417AA mutant, they found that this mutant did not accumulate in the cell nucleus after treatment with IL-6 or EGF. Furthermore, they were unable to locate STAT3 RR414/417AA in nuclear extracts of EGF-treated cells. Still, Western blot and EMSA experiments confirmed that this mutant was phosphorylated normally, could dimerise, and could bind DNA (Ma *et al.*, 2003). The STAT3 RR414/417AA will be called STAT3ΔNLS from now on.

To confirm the phenotype of STAT3ΔNLS-mCherry, HeLa cells expressing this construct or STAT3wt-mCherry were subjected to live-cell imaging. Several images of cells, expressing either construct, were taken before and 90 min after treatment with IFN-γ and IL-6. Before the addition of cytokine, STAT3 and STAT3ΔNLS showed similar subcellular localisations (figure 22A, top row). Ninety minutes after cytokine treatment both the STAT3wt and STAT3ΔNLS expressing cells showed the formation of paracrystalline structures (figure 22A, bottom row). However, the subcellular distribution of these paracrystals was vastly different. While they were observed mostly in the nucleus for STAT3wt-mCherry expressing cells, STAT3ΔNLS-mCherry paracrystals exclusively localised in the cytoplasm, suggesting that the latter protein, carrying the mutation RR414/417AA could not be imported into the nucleus.

Western blots were performed to rule out that the ΔNLS phenotype was caused by insufficient tyrosine phosphorylation. For this, U3R cells were transfected with STAT3wt-mCherry or STAT3ΔNLS-mCherry and cells were stimulated with IL-6 for 1 h or left unstimulated. Western blots were performed on whole-cell protein extracts. Detection with an antibody targeting STAT3 revealed good expression for both constructs, but the band detected in the STAT3ΔNLS-containing extract was slightly weaker compared to that from STAT3wt-containing extracts (figure 22B, lanes 1-3). No phosphorylated STAT3 was detected in untreated STAT3wt-mCherry-transfected cells (figure 22B, lane 1). In contrast, the phosphorylated protein was found in cytokine-treated extracts containing either STAT3wt-mCherry or STAT3ΔNLS-mCherry (figure 22B, lane 2+3).

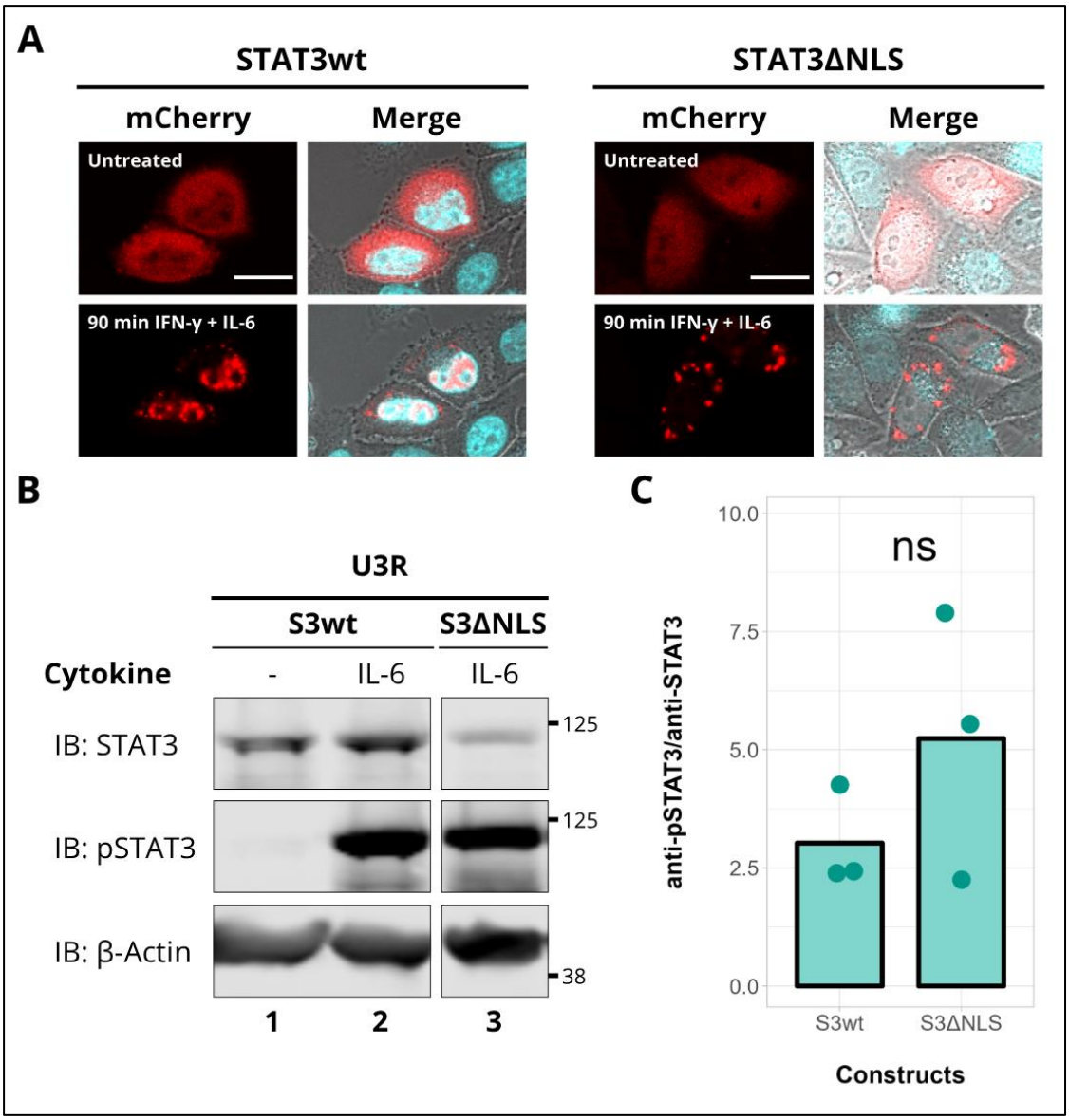


Figure 21 | STAT3ΔNLS subcellular distribution remains unchanged after IFN-γ and IL-6 treatment: ***[A]*** *Fluorescent micrographs of mCherry-tagged STAT3wt or STAT3ΔNLS expressing cells before or 90 min after IFN-γ and IL-6 treatment. The scale bar represents 20 µm. Merge shows the mCherry, Hoechst and Phase 3 Contrast channels.* ***[B]*** *Whole cell lysates of U3R cells expressing either mCherry-tagged STAT3wt or STAT3ΔNLS were subjected to Western blot analysis. Anti-STAT3, anti-pSTAT3 and anti-Actin antibodies were used to detect specific proteins.* ***[C]*** *Quantification of Western blots from [B] and two further independent experiments. Shown is the ratio of the fluorescent signal detected for the anti-pSTAT3 antibody divided by fluorescence of the anti-STAT3 antibody for IL-6 treated cells containing either mCherry-tagged STAT3wt or STAT3ΔNLS. MW markers in kDa.*

The quantification of this and two other independent Western blots showed a non-significant increase in the proportion of phosphorylated STAT3ΔNLS-mCherry compared to STAT3wt-mCherry (173%; p=0.31; figure 22C). Overall, these Western blot experiments confirmed that defective phosphorylation was not the cause for abrogated nuclear accumulation of this construct following cytokine treatment.

Residues RR422/425 mediate the cytokine-induced nuclear accumulation of STAT5B.

The stretch of basic amino acids in the DNA binding domain that is essential for the cytokine-induced nuclear accumulation of STAT1 and STAT3 is conserved in STAT5B (STAT1 K410/K413, STAT3 R414/R417, STAT5B R422/R425). Therefore, it was suspected that these residues may also facilitate STAT5B nuclear accumulation in response to cytokine treatment. Indeed, a previous study reported impaired GH-induced STAT5A nuclear accumulation when all basic amino acids in peptide 422-KRIKRADRR-430 were replaced with alanine. Still, GH efficiently induced activation by phosphorylation of this protein (Iyer and Reich, 2008). Guided by these findings, a mCherry-tagged STAT5B RR422/425AA mutant was generated by site-directed mutagenesis and its nuclear accumulation potential was assessed via live-cell imaging. The STAT5B RR422/425AA mutant will be called STAT5BΔNLS from hereon after.

HeLa cells growing on microscopy-grade cell culture dishes were transfected with plasmid DNA encoding for the GHR, Jak2 and either STAT5Bwt-mCherry or STAT5BΔNLS-mCherry. The transfected cells were then imaged before and 90 min after treatment with IFN-γ and GH. STAT5Bwt-mCherry and STAT5ΔNLS-mCherry showed a similar subcellular distribution before cytokine stimulation that was pancellular (figure 23A, top row). Their distribution also resembled that observed for STAT3wt-mCherry and STAT3ΔNLS-mCherry earlier (figure 22A). Cytokine-stimulated cells expressing the STAT5Bwt-mCherry displayed a noticeable shift towards nuclear fluorescence. Yet, a considerable amount of fluorescence remained detectable inside the cytoplasm. Furthermore, some cells showed the formation of paracrystals inside the cytoplasm (figure 23A, bottom left). Paracrystal formation of STAT5B constructs in response to GH treatment has been observed before (Herrington *et al.*, 1999). In contrast, the GH and IFN-γ treatment of cells expressing STAT5BΔNLS-mCherry appeared to favour nuclear export of the mCherry-tagged protein (figure 23A, right). These findings suggest that STAT5BΔNLS-mCherry carrying the KK422/425AA mutation could not accumulate inside the nucleus following stimulation with GH and IFN-γ.

Once again, to rule out that the ΔNLS phenotype was caused by diminished tyrosine phosphorylation following cytokine treatment, Western blot experiments were performed. For this, U3R cells were transfected with the STAT5B constructs from

before and either treated with IFN-γ for 60 min or left untreated. Whole-cell extracts were generated and subjected to Western blotting using antibodies specific to STAT5B or the STAT5B pY694 phosphopeptide. In both STAT5Bwt-mCherry and STAT5BΔNLS-mCherry containing extracts two bands migrating close together at a speed consistent with the molecular weight of these constructs were observed when detecting with the anti-STAT5 antibody. The slower migrating band was stronger in either sample (figure 23B, lanes 1-3). The additional band might stem from a proteolytically degraded product of STAT5B or may indicate the presence of a STAT5B version containing an additional post-translational modification. Indeed, multi-band patterns in STAT5 Western blots have been observed before and were attributed to distinct serine/threonine phosphorylation in previous studies (Ram *et al.*, 1996; Herrington *et al.*, 1999). A weak band was detected in extracts from untreated STAT5Bwt-mCherry expressing cells when using the anti-STAT5 pY694 antibody. The band co-migrated with the slow-migrating band detected earlier using the anti-STAT5 antibody, indicating that some baseline phosphorylation of this but not the slower-migrating STAT5B entity was present in untreated cells (figure 23B, lane 1). A stronger band migrating at the same speed was detected in extracts from IFN-γ-treated STAT5Bwt-mCherry or STAT5BΔNLS-mCherry expressing cells when the anti-STAT5 pY694 antibody was used (figure 23B, lane 2+3), indicating that both constructs were efficiently phosphorylated. When this and a further two independent Western blot experiments were quantified, no significant difference in the proportion of phosphorylated protein was detected between STAT5Bwt-mCherry or STAT5BΔNLS-mCherry from cytokine-treated cells (figure 23C). Consequently, diminished cytokine-induced phosphorylation of STAT5BΔNLS-mCherry could be ruled out as the cause for the abrogated cytokine-induced nuclear accumulation observed in Nuclear Translocation Experiments.

In summary, the evidence presented in this chapter confirmed that a conserved stretch of surface-exposed basic amino acids in the DBD STAT1, STAT3 and STAT5B mediates their nuclear accumulation following activation at the cytokine receptor. In turn, the mutants STAT1 KK410/413AA, STAT3 RR414/417AA and STAT5B KK422/425AA showed the necessary phenotype for use in Nuclear Translocation Assay experiments.

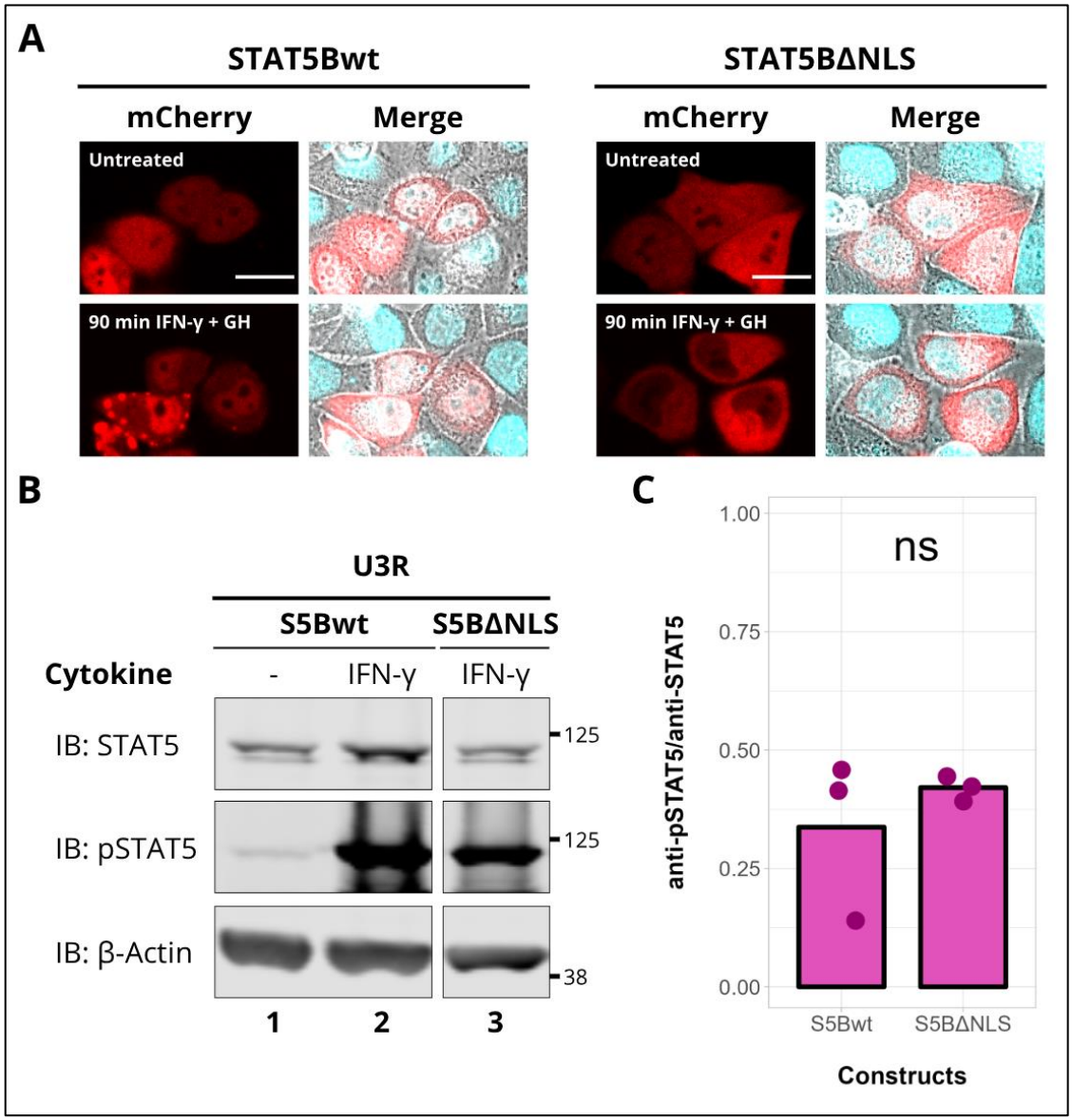


Figure 22 | STAT5BΔNLS subcellular distribution remains unchanged after IFN-γ and GH treatment: ***[A]*** *Fluorescent micrographs of mCherry-tagged STAT5Bwt or STAT5BΔNLS expressing cells before or 90 min after IFN-γ and GH treatment. The scale bar represents 20 µm. Merge shows the mCherry, Hoechst and Phase 3 Contrast.* ***[B]*** *Whole cell lysates of U3R cells expressing either mCherry-tagged STAT5Bwt or STAT5BΔNLS were subjected to Western blot analysis. Anti-STAT5, anti-pSTAT5 and anti-Actin antibodies were used to detect specific proteins.* ***[C]*** *Quantification of Western blots from [B] and two further independent experiments. Shown is the ratio of fluorescent signal detected for anti-pSTAT5 antibody divided by fluorescence of the anti-STAT6 antibody for IFN-γ treated cells containing either mCherry-tagged STAT5Bwt or STAT5BΔNLS. MW markers in kDa.*

4.1.5 Discussion of Chapter 4.1

The molecular mechanisms behind the formation of active STAT homo- and heterodimers are poorly understood. This work describes the development of a novel cell-based assay termed the Nuclear Translocation Assay and its application to the

above research question. In this section, the rationale behind the assay was described and the different components required for the assay were tested.

In the Nuclear Translocation Assay, active dimer formation between a STATΔNLS and another STAT is declared if the former can cross-inhibit the cytokine-induced nuclear accumulation of the latter. To visualise this in live-cell imaging, a STATΔNLS fused to mCherry is co-expressed with the other STAT fused to a GFP in HeLa cells. However, the ΔNLS mutation does not affect dimerisation, meaning that not only active STATΔNLS-mCherry/STAT-GFP dimers can form should the two proteins interact, but STAT-GFP could also self-assemble into active STAT-GFP/STAT-GFP. The cytokine-induced nuclear accumulation of these dimers is not inhibited. Therefore, cytokine-induced nuclear accumulation of STAT-GFP/STAT-GFP dimers may mask the cross-inhibitory effect of STATΔNLS-mCherry on STAT-GFP if the two can interact. It was reasoned that STAT-GFP self-assembly could be minimised and STATΔNLS-mCherry/STAT-GFP dimerisation maximised by expressing the STATΔNLS-mCherry in molar excess to STAT-GFP.

To control for this, it was hypothesised that the ratio of mCherry-tagged-to-GFP-tagged protein present in individual cells could be determined by comparison of total mCherry and GFP fluorescence to that measured in cells expressing a GFP-mCherry fluorescence standard. The GFP-mCherry fluorescence standard was designed to ensure that the fluorophores are expressed in a known and fixed ratio of one-to-one. This was achieved by fusing the two fluorescent moieties via an amino acid linker. Indeed, for both cells expressing the fGFP-mCherry or the rGFP-mCherry variant, a linear relationship between mCherry and GFP fluorescence intensity was found across different expression levels and imaging settings. Therefore, relative protein ratios in cells co-expressing different amounts of individual mCherry-tagged and GFP-tagged protein could be extrapolated from this.

To the best of my knowledge, the specific approach for quantification of relative protein ratios based on extrapolation from a fluorescence standard presented here has not been implemented before. However, a closely related approach has been applied to determine absolute protein concentrations of proteins fused to a GFP within living cells (Cherkas *et al.*, 2018). The authors showed that the ratio of concentrations

between two fluorophores is proportional to the ratio of their fluorescence intensities by a “ratio factor”, an assumption that was also made during the generation of the GFP-mCherry standards in this work. Furthermore, fluorophore concentration can be calculated from fluorescence intensity when the spectral properties of the imaging setup, sample volume and spectral parameters of the fluorophore are known. Therefore, combining these observations, the authors showed that intracellular concentrations of different fluorescent-tagged proteins could be determined from fluorescence intensities by flowing in a reference fluorescence label at known concentration via a patch clamp setup. Interestingly, the authors validated this approach by utilising their version of a fluorescence standard where the fluorescent proteins Cerulean and Venus are separated by a large linker encoding parts of the TRAF2 protein. They found a strong relationship between the fluorescence intensities of the two fluorophores both within individual cells and between different cells (Cherkas *et al.*, 2018). Similar results were also obtained for fGFP-mCherry and rGFP-mCherry in this work, confirming that relative protein ratios can be extrapolated over a broad range of expression levels. Another study used normalisation to the FRET signal of a GFP-mCherry standard to assess the stoichiometry of the rat asialoglycoprotein receptor (Renz *et al.*, 2012). The standard was comprised of the two fluorescent proteins joined by a short linker and thus allowed for substantially more FRET than the fGFP-mCherry and rGFP-mCherry used in this work. They then measured FRET in cells expressing a GFP and a mCherry-tagged version of the asialoglycoprotein receptor subunit rat hepatic lectin 1 (RHL1) and found that the signal was higher than the FRET signal measured for the GFP-mCherry standard, thus confirming that the subunit homo-oligomerise to form the receptor (Renz *et al.*, 2012).

Other methods allowing for the quantification of absolute protein amounts based on GFP fluorescence compared to a standard have been reported. However, they either compare sample fluorescence with standards outside cells (Hirschberg *et al.*, 1998; Chiu *et al.*, 2001, 2002; Khakh *et al.*, 2001; Furtado and Henry, 2002) which may not guarantee consistent experimental conditions or require the production of recombinant protein or virus particle standards (van der Wal *et al.*, 2001; Dundr *et al.*, 2002). The approach described in this work avoided these issues by comparing sample fluorescence with a standard expressed in the same cell model using the same

expression vector and transfection method. Still, this came at the price of not being able to determine absolute protein amounts. Nevertheless, since knowledge of relative protein amounts was deemed sufficient to guarantee the efficient function of the Nuclear Translocation Assay, the relative ease of handling and time efficiency of this approach outweighed its disadvantages.

FRET occurs when two fluorophores come into close contact and the emission spectrum of the donor fluorophore overlaps with the excitation spectrum of the acceptor. Due to the nature of the GFP-mCherry fluorescence standard design presented in this work, it was likely that these prerequisites were fulfilled. If this was the case, then quenching of GFP fluorescence and energy transfer to the mCherry fluorophore may mask the true fluorescence ratio at which equal amounts of protein are present within individual cells. Therefore, two designs of the GFP-mCherry constructs were compared for their ability to minimise FRET. A GFP-mCherry construct containing an unordered flexible linker showed low FRET efficiencies with a high standard deviation (13.97±8.80%). When joining the fluorophores using a rigid linker, the average FRET efficiency was only slightly reduced (12.68%). However, the standard deviation of these values was lower (3.63%). The large spread of FRET efficiencies in the flexible linker design might be due to the GFP-mCherry being able to adopt different conformations with a variety of distances between the two fluorophores. Indeed, a study investigating this notion for a different flexible linker comprising a random coil found that FRET efficiencies predicted by a worm-like chain model agreed well with experimental data, indicating that a multitude of conformations was adopted (Merkx, 2010). In general, linker designs have gained some attention in FRET-based methods. Bimolecular fluorescent proteins are often used in FRET interaction assays as positive controls to gauge the maximum achievable FRET signal for a single interaction (Kaminski, Rees and Schierle, 2014; Algar *et al.*, 2019). Usually, these constructs contain a small linker to allow for close contacts between fluorophores and therefore high FRET efficiencies. Furthermore, FRET-based sensors have been used to detect metal ions or post-translational modifications and these constructs also often require the interconnection of fluorophores with peptide linkers (Merkx, 2010). Several studies have investigated different linker designs. The design of the rGFP-mCherry construct was inspired by work from Arai *et al.* (2001), where different linkers

were compared for their ability to efficiently separate a blue fluorescent protein (BFP) and GFP domain. Using the decrease in FRET efficiency as a readout, the authors found that separation increased with increasing lengths of a helix-forming peptide linker. Furthermore, they found that a flexible linker containing the same number of amino acids as a helical linker showed higher FRET efficiency, indicating that it could not separate the two domains as effectively. Overall, a helical linker containing five repeats of the EAAAK peptide separated the BFP and GFP best which resulted in a low FRET efficiency of 9% (Arai *et al.*, 2001). Adopting this linker for the rGFP-mCherry design, an average FRET efficiency of 12.68% with a standard deviation of 3.63% was found (appendix table 2). Interestingly, a different group found similar results for the GFP-mCherry FRET pair using the same rigid linker design with five EAAAK repeats (Zhuo and Knox, 2022). Furthermore, their study indicates that adding another two EAAAK repeats to the linker could decrease FRET efficiencies below 10%. Still, it was decided to continue using the original rGFP-mCherry construct, especially since low levels of FRET would be expected for latent dimer forming STAT interaction partners that carry a C-terminal fusion to GFP or mCherry (Kretzschmar *et al.*, 2004; Cimica *et al.*, 2011; Domoszlai *et al.*, 2014). Therefore, the relative protein ratios in cells co-expressing a low-FRET STAT-GFP/STAT-mCherry pair should be discernible with reasonable accuracy by comparing fluorescence signal ratios to an rGFP-mCherry standard that equally showed low-FRET efficiencies.

After characterising the fluorescence standards, the focus was shifted to assessing the behaviour of other reagents of the Nuclear Accumulation. It was argued that should the two proteins form dimers, the cross-inhibitory effect on STAT-GFP cytokine-induced nuclear accumulation by a STATΔNLS-mCherry would be most obvious at the time of maximal nuclear accumulation of the corresponding wild-type STAT. Therefore, the nuclear accumulation dynamics of STAT1wt-GFP, STAT3wt-GFP and STAT5Bwt-GFP were investigated. Quantification of fluorescence signals revealed a maximum for STAT1 nuclear accumulation 90 min after the addition of cytokines which remained stable for another 30 mins (figure 16). Regarding STAT3 nuclear accumulation, a maximum was observed as early as 60 min after cytokine addition, and this remained stable for at least 60 min thereafter (figure 17). Triggering STAT5B nuclear accumulation with GH required co-expression of GHR and Jak2. Using this approach, a

maximal nuclear signal was observed 90 min after treatment with GH and IFN-γ, and this remained stable for 15 min (figure 20). These findings are consistent with previous studies that found STAT1 and STAT3 reaching the maximum cytokine-induced nuclear accumulation between 60-120 min after treatment (Köster and Hauser, 1999; Pranada *et al.*, 2004; Meyer, Begitt and Vinkemeier, 2007).

Lastly, the ΔNLS phenotypes were confirmed for different STATs. The region spanning amino acids 407-413 of STAT1 has been implicated as an NLS that is only active following cytokine stimulation before (Melén, Kinnunen and Julkunen, 2001; McBride *et al.*, 2002; Meyer, 2002). In this work, the STAT1 KK410/413AA mutant first described by Fagerlund *et al*. (2002) was designated as STAT1ΔNLS and its inability to accumulate inside the nucleus following cytokine treatment was confirmed (figure 21). A ΔNLS phenotype as previously reported for STAT3 RR414/417AA (Ma *et al.*, 2003) could be confirmed in this work (figure 22). Thus, this construct was further used as STAT3ΔNLS in Nuclear Translocation Assay experiments.

The stretch of mostly basic amino acids 422-KRIKRADRR-430 has been implicated as a putative cytokine-induced NLS sequence in STAT5A (Iyer and Reich, 2008). Indeed, the homologous residues to STAT1 KK410/413 in STAT5A are KK422/425 and they are also conserved in STAT3 (RR414/417), STAT4 (K408/414K) and STAT5B. Therefore, it was suspected that STAT5B KK422/425AA may be used as STAT5BΔNLS in the Nuclear Translocation Assay. Indeed, the experiments depicted in figure 23 confirmed the ΔNLS phenotype of this construct. Thus, the present study, for the first time, showed that mutating two basic amino acids inside the DBD of STAT5B is sufficient to abrogate cytokine-induced nuclear accumulation of this protein. Moreover, this study confirms that the conserved residues at this position have similar functions in at least two other STAT family members.

Before the work that identified basic residues contained in STAT5A 422-430 as a cytokine-induced NLS, another study showed that cytokine-induced nuclear accumulation of STAT5B was abrogated when amino acids 466-VVVI-469 were mutated to alanine (Herrington *et al.*, 1999). This sequence is quite different from that expected for classical NLS, usually comprising mono- or bipartite stretches of basic amino acids (Lange *et al.*, 2007). Indeed, the authors argue that nuclear accumulation

of STAT5B might be due to high-affinity DNA binding since this stretch of amino acids had also been previously shown to facilitate the binding of STAT5B to the IRF-1 promotor (Luo and Yu-Lee, 1997). Interestingly, STAT5B 466-VVVI-469 is well conserved in STAT1, STAT3 and STAT4. In the crystal structures of phosphorylated STAT1 and STAT3 bound to DNA, the sequences 455-VVVI-458 and 455-VVVI-458, respectively, formed a β-sheet and their side chains protruded towards the hydrophobic cores of the DBDs. Therefore, it is unlikely that these participate in DNA binding directly. However, the following residue S459 was shown to interact with the phosphate backbone of DNA in both structures (Becker, Groner and Müller, 1998; Chen *et al.*, 1998). Residues 466-VVVI-469 of STAT5A were also seen to form a β-sheet and side chains protruded towards the hydrophobic core in the crystal structure of unphosphorylated STAT5A (Neculai *et al.*, 2005). Thus, it is conceivable that these residues do not directly contribute to STAT5 DNA binding either. Importantly, the β-sheet formed by these amino acids can be seen in close three-dimensional proximity, with one side chain even pointing towards the stretch of amino acids comprising the conserved NLS in the crystal structures of STAT1, STAT3 and STAT5A (Becker, Groner and Müller, 1998; Chen *et al.*, 1998; Mao *et al.*, 2005; Neculai *et al.*, 2005; Ren *et al.*, 2008). Hence, an alternative explanation could be that mutating STAT5B residues 466-VVVI-469 to alanines locally destabilised the DBD causing the ΔNLS phenotype and DNA binding defect. However, the overall structural integrity of the protein remained intact, which would explain why the mutant protein could still form active homo and heterodimers and shows both efficient cytokine-induced tyrosine phosphorylation and paracrystal formation (Herrington *et al.*, 1999).

In summary, in this chapter, the reagents used in Nuclear Translocation Assay experiments were characterised. To avoid measurement artefacts entailed due to having a low STATΔNLS-mCherry-to-STAT-GFP ratio present within cells, a system was established by which the relative protein amounts of GFP, and mCherry-tagged STATs could be determined. Furthermore, the length of the cytokine treatment to bring out the maximum effect of nuclear accumulation inhibition was determined, and the ΔNLS phenotype was confirmed for STAT1, STAT3 and STAT5B. The principles behind the Nuclear Translocation Assay required that the ΔNLS used to determine STAT homo- and heterodimers exert a dominant negative effect on cytokine-induced nuclear

accumulation of active STAT dimers. Whether the ΔNLS mutations characterised in this chapter fulfil this criterion will be assessed next (chapter 4.2).

4.1.6 Nuclear Translocation Assay protocol

After characterising Nuclear Translocation Assay reagents in the last chapter, an imaging and image analysis workflow was established and summarised under the Nuclear Translocation Assay protocol. It is important to note that all imaging was done blind to genotype unless otherwise stated. For this, the clear names of STAT combinations were replaced by a number system assigned by Prof. Dr. Uwe Vinkemeier. All experiments were conducted using a HeLa cell line model system. Transfection mixes were then assembled by Dr. Andreas Begitt, so I remained blind to the contents of individual combinations. However, since different imaging settings had to be used for cells containing STAT1, STAT3 or STAT5B, respectively, I was made aware of which settings had to be implemented. Including all control and optimisation experiments, 159 individual transfections were imaged. Samples were only unblinded once all combinations were fully analysed. Typically, four experiments could be imaged per experiment day. The rGFP-mCherry standard was imaged in addition to the four experiments on each day to ensure consistency of experimental conditions. Thus, the 159 individual transfections included 52 rGFP-mCherry standards and 101 samples containing bait and prey STATs, the remaining six transfections were controls such as mock-transfected cells to determine the autofluorescence of cells. Each of the 39 individual STAT combinations was imaged on at least two separate occasions. The Nuclear Translocation protocol was implemented as follows:

Live-cell imaging of HeLa cells transfected with bait- and prey-STATs.

1. The transfected HeLa cells were stained with Hoechst 33342 to highlight the nucleus. Afterwards, the growth medium was replaced with imaging medium and the cells were placed in the microscope's environmental chamber set to 37 °C to equilibrate for at least 10 min.
2. A low-resolution "map" covering a large area of the culture dish was captured to find cells showing GFP and mCherry fluorescence (further details in chapter 3.5.2 and Appendix B). The position of these cells was marked for further imaging. This was repeated until a total of 30 positions were marked.

3. High-resolution images of seven focal planes covering the entire thickness of the cells were captured for the previously marked image frames (see also chapter 3.5.2 and Appendix B). These images were used to assess STAT subcellular distribution before cytokine treatment.
4. Immediately afterwards, cytokines were added without moving the cell culture dishes, and the cells were left undisturbed for 90 min.
5. Another set of high-resolution images was captured of the same cells from earlier (see also chapter 3.5.2 and Appendix B). These images were used to assess STAT subcellular distribution after cytokine treatment.
6. At the end of each experiment set, the rGFP-mCherry standard was imaged with the same settings and workflow used for the samples imaged previously. For obvious reasons, steps 4 and 5 were omitted for cells containing the fluorescent standard. Furthermore, only 10 positions of rGFP-mCherry-containing cells were imaged at high-resolution.

Image analysis step 1 – Determine relative protein ratios.

7. High-resolution images of latent cells were loaded into Fiji/ImageJ and whole-cell ROIs were generated using the Whole Cell Segmentation macro and subsequent manual adjustments (chapter 3.5.4).
8. The GFP and mCherry intensities of the seven focal planes were summed up in a single 2D image and the total fluorescence in each channel was determined for individual cells based on the whole cell ROIs created before.
9. The total mCherry to GFP fluorescence ratio determined for cells containing STATΔNLS/STAT combinations was compared to the one measured for rGFP-mCherry-containing cells to determine relative protein ratios. Only cells harbouring a bait-STAT-mCherry molecular excess of at least 1.5x but no more than 10x progressed towards the second step of the analysis.

Image analysis step 2 – Measuring nuclear fluorescence before and after cytokine treatment.

10. Going back to the original high-resolution images of latent cells, the focal plane crossing the median of the cell nucleus was determined by visual inspection.

Nuclear ROIs were determined within this focal plane based on the Hoechst signal using the semi-automated Nucleus Segmentation macro (chapter 3.5.4).

11. The nuclear fluorescence in latent cells was measured for both the GFP and mCherry channel in the focal plane crossing the median of the cell nucleus using the previously generated nuclear ROIs.
12. High-resolution images of the same cells imaged 90 min after activation were loaded into Fiji/ImageJ.
13. Typically, the cells had moved during the 90 min incubation period. Therefore, new nuclear ROIs had to be generated according to the procedure described in 10.
14. Nuclear fluorescence after activation in the same cells was measured for both channels in the focal plane crossing the median of the cell nucleus using the newly generated nuclear ROIs.
15. To quantify the cytokine-induced change in nuclear fluorescence intensity, the nuclear score was calculated using formula IV.
16. To decide if individual STATs could form active dimers, the nuclear score of the STAT-GFP when co-expressed with a STATΔNLS-mCherry was compared to the one determined for the same STAT-GFP when it was co-expressed with the wild-type variant corresponding to the STATΔNLS. If the STAT-GFP nuclear score was significantly decreased in the former compared to the latter combination, the STATΔNLS-mCherry could cross-inhibit STAT-GFP cytokine induced nuclear accumulation via the formation of active dimers.

4.2 Assay Validation: Testing the Nuclear Translocation Assay using known interactors and known non-interactors.

In the previous chapter, different constructs and reagents required for the Nuclear Translocation Assay were characterised and the Nuclear Translocation Assay protocol was compiled. Importantly, the cross-inhibition of STAT-GFP cytokine-induced nuclear accumulation by direct interaction with a STATΔNLS-mCherry is only possible if the ΔNLS mutation is dominant negative; That is, if deletion of the NLS in one monomer leads to full abrogation of cytokine-induced nuclear accumulation of the whole STAT dimer. The Nuclear Translocation Assay cannot work if this criterion is not fulfilled. Therefore, the tests conducted in this chapter were utilised to confirm the dominant-negative nature of the ΔNLS mutation for STAT1, STAT3 and STAT5B.

At the same time, the accuracy of the Nuclear Translocation Assay was confirmed. For this, STATs that are known to form active dimers (known interactors) and STATs that have been previously shown to be unable to dimerise in response to cytokine (known non-interactors) were subjected to the Nuclear Translocation Assay. The three known interactors assessed here were STAT1 homodimers as a combination of STAT1ΔNLS-mCherry and STAT1wt-GFP, STAT3 homodimers as a combination of STAT3ΔNLS-mCherry and STAT3wt-GFP, and STAT5B homodimers as a combination of STAT5BΔNLS-mCherry and STAT5Bwt-GFP. STAT1 and STAT5B were found to be activated at the same time at different cytokine receptors but could not form active heterodimers (Wehinger *et al.*, 1996; Kirken *et al.*, 1997; Schulze *et al.*, 2000). Therefore, a STAT5BΔNLS-mCherry and STAT1wt-GFP combination represented a known non-interactor pair and was used to assess the accuracy of the Nuclear Translocation Assay.

4.2.1 The STAT1ΔNLS mutation is dominant negative.

As mentioned before, the STAT1ΔNLS was shown to cross-inhibit the cytokine-induced nuclear accumulation of co-expressed STAT2 before, suggesting that the ΔNLS effect is dominant negative (Melén, Kinnunen and Julkunen, 2001; Fagerlund *et al.*, 2002). To test if it has the same influence on STAT1wt-GFP, cells co-expressing this protein alongside STAT1ΔNLS-mCherry were subjected to the Nuclear Translocation Assay as outlined in the protocol (chapter 4.1.6). To minimise STAT1wt-GFP self-assembly, a

molecular excess of the STAT1ΔNLS-mCherry was required within cells. This excess was generated by transfecting a higher amount of plasmid DNA encoding for STAT1ΔNLS-mCherry than for STAT1wt-GFP (0.66 µg/dish vs. 0.33 µg/dish, respectively). To assess the uninhibited cytokine-induced nuclear accumulation of STAT1wt, another cell sample was transfected with STAT1wt-mCherry and STAT1wt-GFP using the same transfection regime.

GFP and mCherry fluorescence resided predominantly in the cytoplasm in either cell sample before cytokine stimulation, but weak fluorescence remained detectable in the nucleus (figure 24A, row 1+3). However, a clear difference between cells expressing the two wild-type STAT1 constructs and cells expressing the STAT1ΔNLS/STAT1wt combination was observed when looking at images taken 90 min after cytokine treatment. In cells expressing the former combination, STAT1wt-GFP and STAT1wt-mCherry localised almost entirely inside the cell nucleus (figure 24A, row 2) as evident by co-localisation with Hoechst stain (figure 24A, row 2). On the contrary, and as observed before in single transfected cells (figure 21), STAT1ΔNLS-mCherry did not show changes in subcellular and almost no co-localisation distribution with the nuclear dye following cytokine treatment. This indicated that the co-expression of a STAT1wt did not rescue the ΔNLS phenotype. Said co-expressed STAT1wt-GFP showed some fluorescence inside the nucleus 90min after cytokine treatment. However, a large amount of fluorescence remained detectable inside the cytoplasm, which suggested that the cytokine-induced nuclear accumulation of STAT1wt-GFP was cross-inhibited by the co-expressed STAT1ΔNLS-mCherry (figure 24A, row 4). The difference between the two conditions was further highlighted when a heatmap pseudo-colour image look-up table was applied to GFP-channel images. In cells containing the two STAT1wt constructs, the STAT1wt-GFP showed high fluorescence intensities within nuclear borders, while pixels located inside cytoplasmic ROIs only showed background-level intensities (figure 24A, panel 8). In comparison, while nuclear fluorescence still showed moderate intensities in the nucleus when STAT1wt-GFP was co-expressed with STAT1ΔNLS-mCherry, pixels in cytoplasmic ROIs were distinguishable from the background due to their elevated intensities (figure 24A, panel 16).

To quantify the degree of cross-inhibition on STAT1wt-GFP nuclear accumulation exerted by the STAT1ΔNLS-mCherry, nuclear scores were calculated for both channels using formula IV and compared to nuclear scores calculated for the STAT1wt-GFP/STAT1wt-mCherry combination. This surmounted to nuclear scores being calculated for 79 and 106 individual cells containing the STAT1wt-GFP/STAT1wt-mCherry and STAT1wt-GFP/STAT1ΔNLS-mCherry combination, respectively (figure 24B; table 23).

Table 21 | Descriptive statistics of nuclear scores determined for cells co-expressing STAT1wt-GFP and STAT1ΔNLS-mCherry or STAT1wt-GFP and STAT1wt-mCherry.

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT1wt-mCherry + STAT1wt-GFP	79	4.24	0.14	4.46	0.14
STAT1ΔNLS-mCherry + STAT1wt-GFP	106	1.12	0.02	2.38	0.05

An average nuclear score of 4.24±0.14 was observed in the mCherry channel and an average nuclear score of 4.46±0.14 was seen in the GFP channel for cells co-expressing the STAT1wt combination. Interestingly, these values were higher than nuclear scores measured for STAT1wt at this time point in time course experiments described in a previous chapter (4.1.3). On average, the mCherry-tagged STAT1ΔNLS showed a nuclear score of 1.12±0.02 which was significantly different to that measured for the STAT1wt-mCherry construct in the other combination. A nuclear score close to 1 indicated that the cytokine treatment could not induce nuclear accumulation of this construct, confirming the ΔNLS phenotype. An average nuclear score of 2.36±0.05 was calculated for STAT1wt-GFP that was co-expressed with STAT1ΔNLS-mCherry. This almost two-fold reduction compared to STAT1wt-GFP when co-expressed with another STAT1wt (figure 24B, table 22) suggested, that the STAT1ΔNLS-mCherry cross-inhibited the cytokine-induced nuclear accumulation of STAT1wt-GFP. This implied that the Nuclear Translocation Assay could reliably detect the formation of STAT1 homodimers.

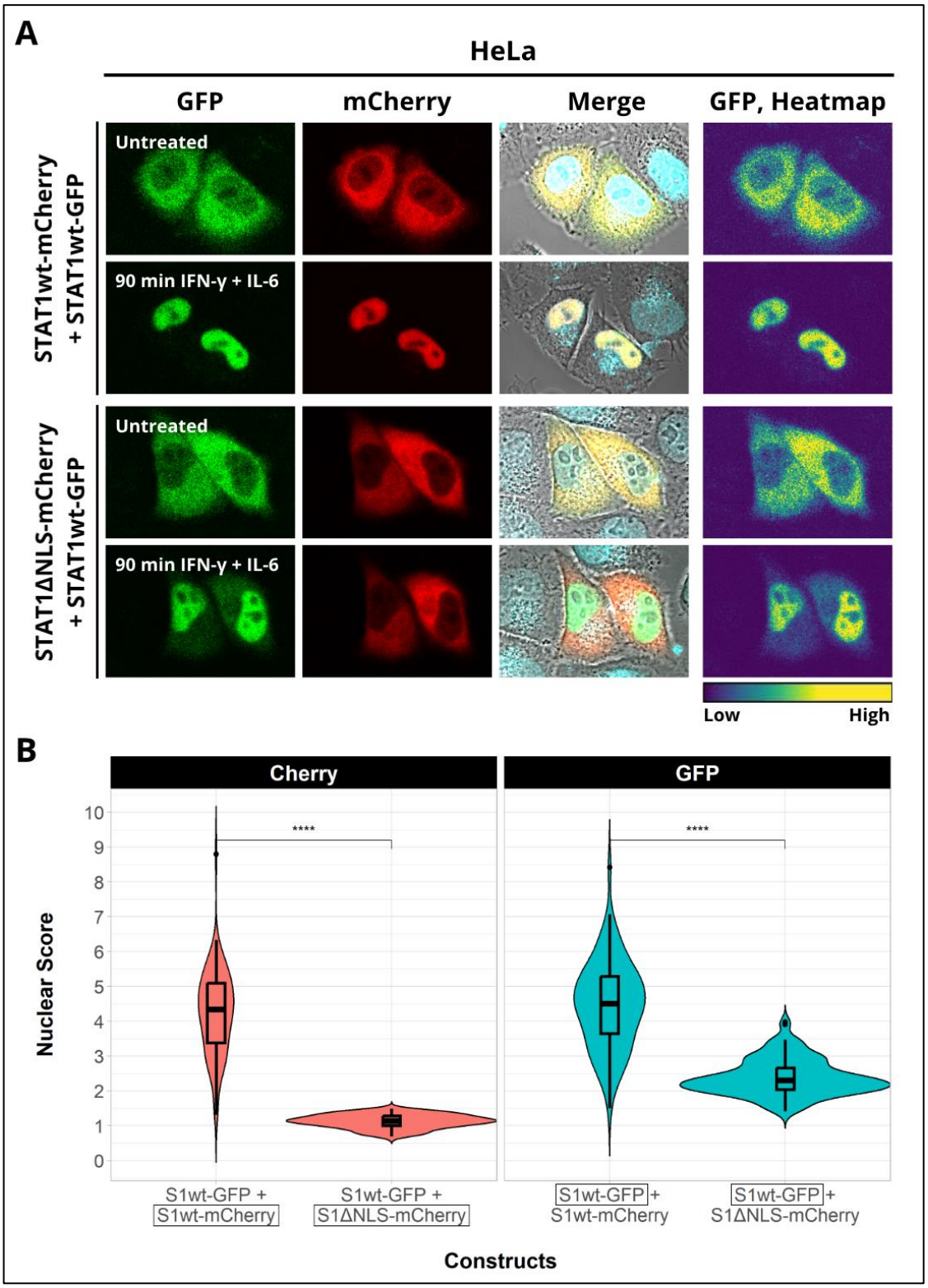


Figure 23 | STAT1ΔNLS cross-inhibits cytokine-induced nuclear accumulation of STAT1wt: ***[A]*** *Fluorescence micrographs of HeLa cells co-expressing either STAT1wt-mCherry and STAT1wt-GFP or STAT1ΔNLS-mCherry and STAT1wt-GFP imaged before and 90 min after cytokine treatment. The last column of panels shows a contrast enhanced depiction of the GFP channel. In this, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).* ***[B]*** *Quantification of mCherry and GFP nuclear scores in 79 or 106 cells co-expressing the two STAT1 combinations described in [A], respectively.*

Efficient phosphorylation of Y701 is a necessity for the induction of STAT1 nuclear accumulation downstream of cytokine receptors (Schindler *et al.*, 1992). To rule out that the inhibition of STAT1wt-GFP nuclear accumulation by co-expressed STAT1ΔNLS-mCherry was caused by diminished phosphorylation rather than the dominant negative accumulation defect, the phosphorylation status of co-expressed proteins was checked in Western blot experiments. For this, an "untagged" STAT1wt construct was co-expressed with either STAT1wt-GFP or STAT1ΔNLS-GFP in U3R cells. The untagged STAT1wt construct was generated by introducing a stop codon after the STAT1 gene to interrupt transcription before the GFP gene. The "untagged" STAT1wt and different GFP-tagged STAT1 constructs were combined so that the two proteins could be told apart by their difference in molecular weight in SDS-PAGE. After transfection, the U3R cells were treated with IFN-γ for 60 min and whole-cell extracts were prepared for Western blot experiments using antibodies targeted against full-length STAT1 or the phosphorylated Y701 phosphopeptide. When detecting with the former antibody, the blots revealed two bands in both extracts containing untagged STAT1wt and STAT1wt-GFP or untagged STAT1wt and STAT1ΔNLS-GFP (figure 25A, top panel, lanes 1+2). The slower of the two migrated slightly faster than the 130 kDa marker band, while the faster of the two was a little faster than the 95 kDa marker band (not shown). The difference in migration speed is consistent with the difference in molecular weights of GFP-tagged and untagged STAT1, respectively. Of note, in both samples the untagged-STAT1wt band was weaker than the one corresponding to GFP-tagged STAT1 constructs (figure 25A, top panel), suggesting that the expression of untagged-STAT1wt was lower than that of the GFP-tagged variants. Detection with the antibody targeting the phosphorylated Y701 of STAT1 revealed that all constructs are phosphorylated after 60 min of IFN-γ stimulation (figure 25A, middle panel). To assess if the co-expression of different GFP-tagged STAT1 constructs affected phosphorylation of the untagged STAT1wt, blots from three independent experiments were quantified as before (table 24).

Phosphorylation of both STAT1wt-GFP and STAT1ΔNLS-GFP co-expressed with untagged STAT1wt decreased by 40% when compared to samples where these constructs were expressed alone (table 24). Still, STAT1ΔNLS-GFP showed elevated levels of phosphorylation compared to STAT1wt-GFP even when co-expressed with

untagged STAT1wt. However, while the difference remained statistically significant, the increase was lower in co-expressing cells compared to single-expressing cells (410% higher phosphorylation in co-expressing vs 755% in single-expressing cells; table 24 vs figure 21C). Interestingly, the phosphorylation of untagged STAT1wt co-expressed with STAT1wt-GFP was similar to single-expressed STAT1wt-GFP (95% of STAT1wt-GFP phosphorylation, p=0.91; table 24 vs. figure 21C). This suggested that, while less untagged STAT1wt is present in these cells, phosphorylation is more efficient compared to co-expressed STAT1wt-GFP. What's more, phosphorylation of untagged STAT1 was significantly increased by 250% when co-expressed with STAT1ΔNLS-GFP than in presence of STAT1wt-GFP (table 24). A graphical representation of this is depicted in figure 25B.

Table 22 | Quantification of STAT1 phosphorylation by Western blotting: *Whole cell extracts from U3R co-expressing untagged STAT1 with STAT1wt-GFP or STAT1ΔNLS were analysed for STAT1 phosphorylation.*

Combination	Construct	Mean ratio $\frac{\text{α-STAT1 pY701}}{\text{α-STAT1 FL}}$
Untagged STAT1wt +	Untagged STAT1wt	0.10 ± 0.01
STAT1wt-GFP	STAT1wt-GFP	0.06 ± 0.02
Untagged STAT1wt +	Untagged STAT1wt	0.25 ± 0.02
STAT1ΔNLS-GFP	STAT1ΔNLS-GFP	0.25 ± 0.04

Taken together, these results suggest that the phosphorylation of STAT1wt downstream of IFN-γ signalling was not adversely affected by co-expression with STAT1ΔNLS. Much rather, it appeared that the STAT1ΔNLS increased phosphorylation of STAT1wt when they were present together. Perhaps this was due to the STAT1wt being trapped inside the cytosol through interaction with STAT1ΔNLS (figure 24) and therefore shielded from dephosphorylation by nuclear phosphatases (ten Hoeve *et al.*, 2002; Wu *et al.*, 2002). A 40% decrease in phosphorylation was detected when STAT1wt-GFP was co-expressed with untagged STAT1wt. However, concerns that this would reduce nuclear accumulation capabilities can be dismissed given the fact that this construct showed high nuclear scores when co-expressed with STAT1wt-mCherry in HeLa cells (figure 24).

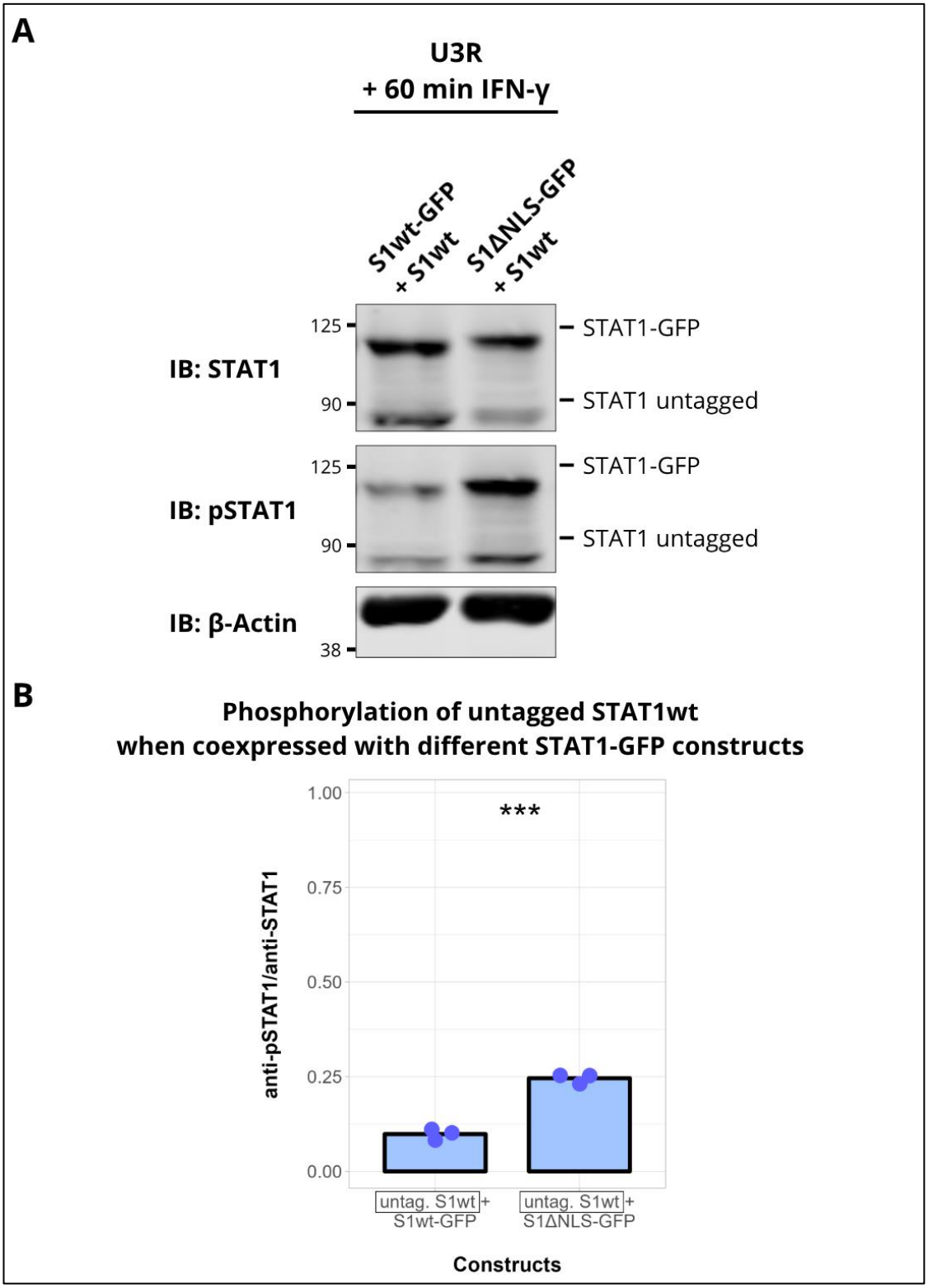


Figure 24 | Co-expression of STAT1ΔNLS does not adversely affect STAT1wt phosphorylation: ***[A]*** *U3R cells coexpressing untagged STAT1wt with either STAT1wt-GFP or STAT1ΔNLS-GFP were treated with IFN-γ for 60 min. Whole cell extracts were prepared and subjected to Western blot analysis using anti-STAT1 and anti-STAT1 pY701 antibodies.* ***[B]*** *Quantification of untagged STAT1wt phosphorylation when co-expressed with either STAT1wt-GFP or STAT1ΔNLS-GFP. Quantifications were performed on data from 3 independent experiments including the one shown in [A]. MW markers in kDa.*

Next, I wondered if the cross-inhibition exerted by STAT1ΔNLS was concentration dependent. To assess this, the imaging data obtained from HeLa cells co-expressing STAT1ΔNLS-mCherry and STAT1wt-GFP for use in the Nuclear Translocation Assay were reviewed and the two variables "Percent of STAT1ΔNLS molecular excess" and "STAT1-GFP nuclear score" probed for correlation. Importantly, progression into the later analysis stages of the Nuclear Translocation Assay was restricted to cells containing a 1.5-fold molar excess of the STAT1ΔNLS construct to maximise the formation of STAT1ΔNLS-mCherry/STAT1wt-GFP dimers (and avoid formation of STAT1wt-GFP/STAT1wt-GFP dimers that would introduce artefacts into nuclear score quantifications). For correlation analysis, this restriction was lifted to be able to also investigate the inhibitory behaviour of STAT1ΔNLS at comparatively low concentrations. Thus, overall, data from 386 cells was used to calculate the Pearson's correlation coefficient between the two variables specified above. As depicted in figure 26, this resulted in a correlation coefficient of R_{S1} = 0.066 which suggested that STAT1ΔNLS-mCherry molecular excess and STAT1wt-GFP nuclear score were not correlated. Indeed, a student t-statistical test confirmed that no significant correlation could be detected between the two variables (p = 0.1959, 95% CI = -0.034-0.165; figure 26). To visualise the relationship between the two variables, a linear regression curve (and its 95% CI) was included in figure 26. All in all, this indicated that the degree of cross-inhibition exerted by the STAT1ΔNLS on STAT1wt-GFP cytokine-induced nuclear accumulation was not dependent on its molar excess over STAT1wt-GFP. Conversely, this would mean that this construct can inhibit the cytokine-induced nuclear accumulation of prey-STAT1s even if it is present in relatively low concentrations.

Indeed, when calculating the mean GFP nuclear score in cells containing less than 1.5x STAT1ΔNLS-mCherry over STAT1wt-GFP it was comparable to the one determined for cells containing a molar excess higher than 1.5 (compare tables 23 and 25). However, the calculation of the STAT1ΔNLS-mCherry nuclear score in these cells revealed a slightly higher value compared to cells that contained more of this protein (less than 1.5-fold excess: 1.31 ± 0.02 vs more than 1.5-fold excess 1.12 ± 0.02; compare table 23 and 25). Nevertheless, the difference between these is too small to suggest that STAT1wt-GFP can rescue the nuclear accumulation defect of STAT1ΔNLS when the former is present at higher concentrations.

Table 23 | Descriptive statistics of nuclear scores determined for cells expressing less than 1.5x the amount of STAT1ΔNLS-mCherry over STAT1wt-GFP.

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT1ΔNLS-mCherry + STAT1wt-GFP	274	1.31	0.02	2.27	0.04

In summary, the data discussed in this chapter confirmed the theory behind the Nuclear Translocation Assay, that a STAT1ΔNLS protein cross-inhibits cytokine-induced nuclear accumulation of a STAT1wt through direct interactions when both are co-expressed in HeLa cells at certain ratios (figure 24). Importantly, Y701 phosphorylation of the STAT1wt protein was not affected by the presence of the STAT1ΔNLS (figure 25), suggesting that the inhibition was indeed exerted by the dominant negative ΔNLS. The fact that the degree of STAT1ΔNLS-mCherry excess over STAT1wt-GFP did not show a significant correlation with the nuclear score of STAT1wt-GFP suggested that the inhibition exerted by STAT1ΔNLS further strengthened this notion.

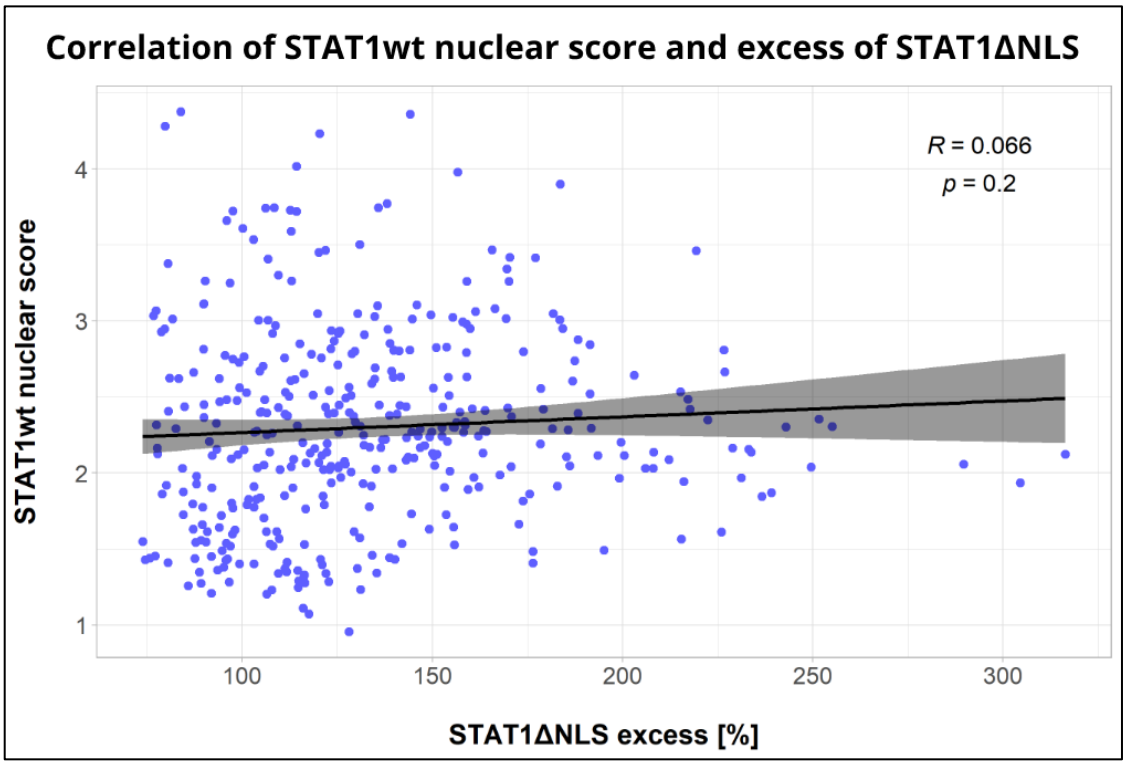


Figure 25 | Correlation analysis of STAT1ΔNLS-mCherry molecular excess and STAT1wt-GFP nuclear score: *Data obtained from HeLa cells expressing both constructs was used to calculate Pearson's correlation coefficient. Hypothesis testing was performed using a student t-test. A linear regression curve is indicated in black with confidence intervals depicted in grey.*

4.2.2 The STAT3ΔNLS mutation is dominant negative.

We have seen in the previous chapter that the STAT1ΔNLS cross-inhibited the cytokine-induced nuclear accumulation of co-expressed STAT1wt. Next, the STAT3ΔNLS was assessed for its dominant negative effect on STAT3wt nuclear accumulation. The assessment of this essentially followed the same protocol as described for STAT1 in the last chapter. First, HeLa cells co-expressing STAT3wt-GFP with either STAT3wt-mCherry or STAT3ΔNLS-mCherry were imaged according to the Nuclear Translocation Assay protocol. In latent cells, all constructs localised pancellularly except for STAT3wt-mCherry co-expressed with STAT3wt-GFP, which showed a more cytoplasmic distribution (figure 27A, row 1+3). Therefore, the distribution of all constructs except the STAT3wt-mCherry matched those of other STAT3 constructs observed before cytokine stimulation in earlier experiments (figure 17+22). This confirmed that the co-expression of STAT3ΔNLS-mCherry did not adversely affect the distribution of STAT3wt-GFP before cytokine stimulation. Treatment with IFN-γ, IL-6 and sIL6Rα triggered the complete nuclear accumulation an co-localisation with Hoechst stain of both constructs in cells containing STAT3wt-GFP and STAT3wt-mCherry. Moreover, paracrystal formation was observed in the nuclei in these cells (figure 27A, row 2). In contrast, cells containing the STAT3wt-GFP/STAT3ΔNLS-mCherry combination had most of their GFP and mCherry fluorescence located in the cytoplasm 90 min after IFN-γ, IL-6 and sIL6Rα addition (figure 27A, row 4). Almost no fluorescence was detected in the nuclei STAT3ΔNLS-mCherry as evident by absent co-localisation with nuclear stain (figure 27A, panel 14 & 15). Instead, the protein localised into paracrystals within the cytoplasm. While most of the co-expressed STAT3wt-GFP colocalised with the STAT3ΔNLS-mCherry paracrystals in the cytoplasm, some nuclear fluorescence remained detectable (figure 27A, panel 13). Still, the subcellular distribution of the STAT3wt-GFP construct appeared considerably different to that observed for the same construct when co-expressed with STAT3wt-mCherry (compare figure 27A, panels 8 and 16). To further confirm these observations, nuclear scores were determined for each sample using formula IV. This was done for both the GFP and mCherry channel individually. A statistical summary of the values calculated is given in table 26 and a graphical representation is depicted in figure 27B. In the case of cells containing the STAT3wt/STAT3wt combination, high nuclear scores were measured in both the GFP

and mCherry channels (2.49±0.07 and 3.37±0.10, respectively; table 26, row 2; figure 27B). The elevated nuclear score of STAT3wt-mCherry is likely explained by the more cytoplasmic distribution of this construct in resting cells observed earlier (figure 27A, panel 10). The considerable difference between GFP and mCherry-tagged wild-type constructs was not observed in similar experiments on STAT1 (figure 24B). In cells containing STAT3wt-GFP alongside STAT3ΔNLS-mCherry, a nuclear score of 1.03±0.03 was calculated for the latter construct (table 26, row 3; figure 27B). This indicated that the subcellular distribution of this protein did not change after treatment with IFN-γ and IL-6. The nuclear score of co-expressed STAT3wt-GFP was determined to be 1.18±0.03 (table 26, row 3), and was thus significantly different to the nuclear score determined for the same construct when co-expressed with STAT3wt-mCherry (figure 27B). This suggested that STAT3ΔNLS cross-inhibited the cytokine-induced nuclear accumulation of co-expressed STAT3wt.

Table 24 | Descriptive statistics of nuclear scores determined for cells co-expressing STAT3wt-GFP and STAT3ΔNLS-mCherry or STAT3wt-GFP and STAT1wt-mCherry.

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT3wt-mCherry + STAT3wt-GFP	140	3.4	0.1	2.5	0.1
STAT3ΔNLS-mCherry + STAT3wt-GFP	51	1.0	0.03	1.2	0.03

Cytokine-induced phosphorylation of residue Y705 is essential for the nuclear accumulation of STAT3 (Yahata *et al.*, 2003). Thus, the phosphorylation status of STAT3wt co-expressed with STAT3ΔNLS was assessed by Western blotting to rule out that diminished phosphorylation caused the diminished cytokine-induced nuclear accumulation observed above. Again, an "untagged" construct was combined with either STAT3wt-mCherry or STAT3ΔNLS-mCherry to distinguish between proteins based on their molecular weight in SDS-PAGE. Accordingly, a STAT3wtΔGFP construct was generated by deleting the amino acids spanning the GFP gene using site-directed mutagenesis. U3R cells expressing the two combinations were treated with IL-6 and sIL6Rα for 60 min and their whole cell extracts were subjected to Western blot analysis using antibodies raised against full-length STAT3 and the pY705 phosphopeptide.

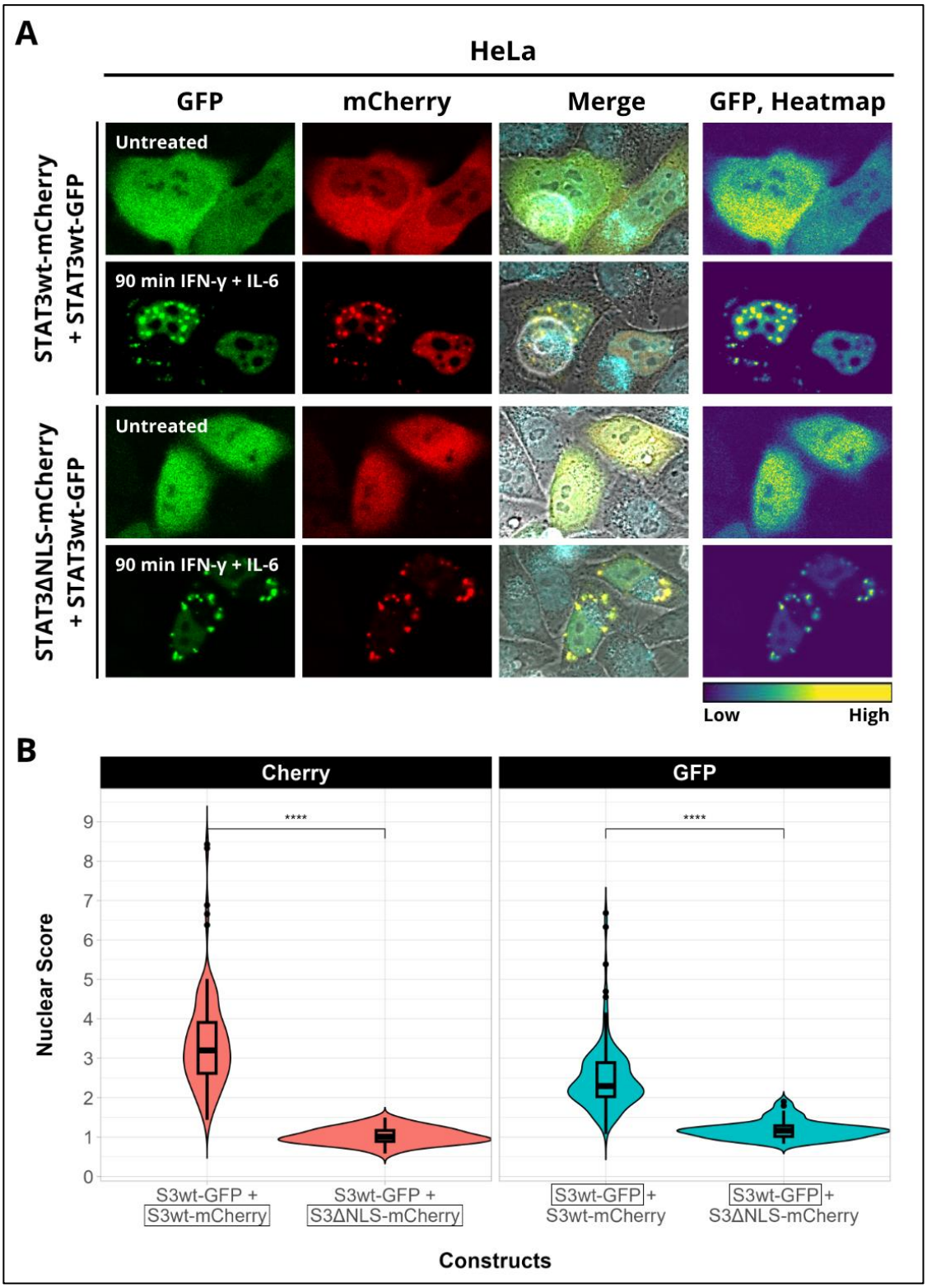


Figure 26 | STAT3ΔNLS cross-inhibits cytokine-induced nuclear accumulation of STAT3wt: ***[A]*** *Fluorescence micrographs of HeLa cells co-expressing either STAT3wt-mCherry and STAT3wt-GFP or STAT3ΔNLS-mCherry and STAT3wt-GFP imaged before and 90 min after cytokine treatment. The last column of panels shows a contrast enhanced depiction of the GFP channel. In this, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).* ***[B]*** *Quantification of mCherry and GFP nuclear scores in 79 or 106 cells co-expressing the two STAT3 combinations described in [A], respectively.*

Two specific bands were detected in extracts containing either untagged-STAT3wt and STAT3wt-mCherry or untagged STAT3wt and STAT3ΔNLS-mCherry (figure 28A, first panel). Based on the expected molecular weight, the slower migrating band was identified as the mCherry-tagged STAT3 variant and the faster band as the untagged STAT3wt. All bands appeared to be of similar strength, indicating that all constructs were equally well expressed. This was further confirmed by quantifying the band intensities for the four bands from three independent experiments and dividing them by their respective band intensities detected with an antibody raised against β-Actin (figure 28A, panel 3). When detecting with an antibody raised against STAT3 pY705, again four bands were visible that co-migrated with the ones detected using the ant-STAT3 antibody. These bands also showed similar strength (figure 28A, panel 2). The proportion of phosphorylated protein was quantified from measurements of three independent Western blot experiments as before. That is, the fluorescence intensities of bands detected with the anti-STAT3 antibody were divided by the intensities measured for the corresponding bands detected with the anti-STAT3 pY705 antibody. Interestingly, all constructs showed considerably lower phosphorylation when co-expressed than when expressed alone (compare table 27 and figure 22). Tough, considering that the same constructs showed high nuclear scores, it can be assumed that the lower phosphorylation did not affect cytokine-induced nuclear accumulation (figure 27, table 27). Both the STAT3wt-mCherry and STAT3ΔNLS-mCherry constructs showed similar proportions of phosphorylated protein ($p=0.71$) that were about 30% lower than for single transfected STAT3wt.

The phosphorylation of the untagged STAT3wt was elevated compared to its co-expressed counterparts, although the difference between co-expressed STAT3ΔNLS-mCherry and untagged STAT3wt was not significant ($p=0.12$). Again, the phosphorylation of untagged STAT3wt was lower than that of single transfected STAT3wt-mCherry from earlier (figure 22). Importantly, even though the mean phosphorylation of untagged STAT3wt was moderately higher when co-expressed with STAT3wt-mCherry than with STAT3ΔNLS-mCherry, this difference was not significant ($p=0.0978$; figure 28B, table 27, rows 1 +3). This indicated that the co-expression with STAT3ΔNLS did not negatively affect the phosphorylation of STAT3wt downstream of cytokine stimulation. This, in turn, suggested that the apparent cross-inhibition of

STAT3wt-GFP cytokine-induced nuclear accumulation exerted by STAT3ΔNLS-mCherry was not a result of decreased phosphorylation.

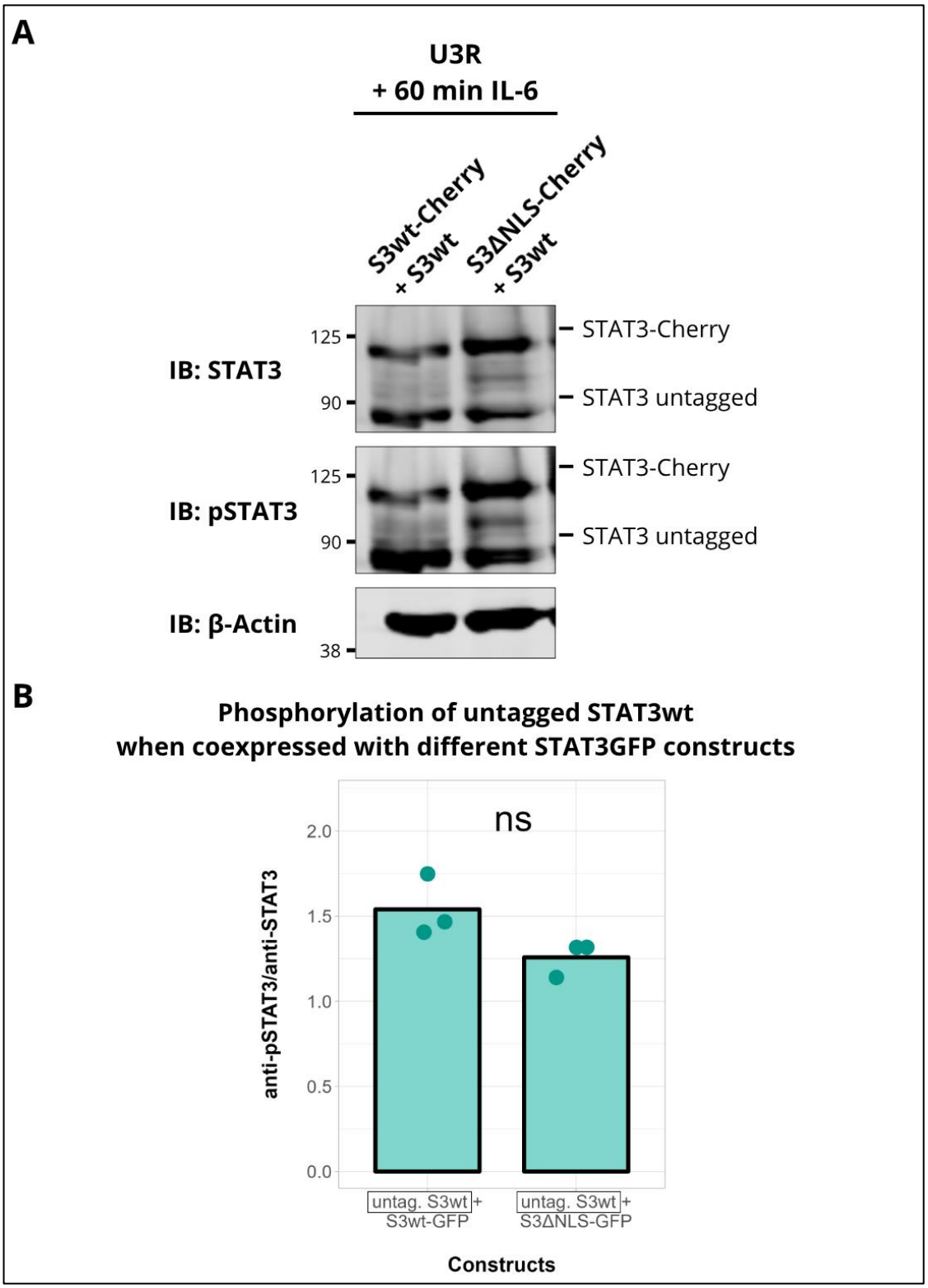


Figure 27 | Co-expression of STAT3ΔNLS does not adversely affect STAT3wt phosphorylation: ***[A]*** *U3R cells coexpressing untagged STAT3wt with either STAT3wt-mCherry or STAT3ΔNLS-mCherry were treated with IL-6 for 60 min. Whole cell extracts were prepared and subjected to Western blot analysis using anti-STAT3 and anti-STAT3 pY705 antibodies.* ***[B]*** *Quantification of untagged STAT3wt phosphorylation when co-expressed with either STAT3wt-mCherry or STAT3ΔNLS-mCherry. Quantifications were performed on data from 3 independent experiments including the one shown in [A]. MW markers in kDa.*

Table 25 | Quantification of STAT3 phosphorylation by Western blotting: *Whole cell extracts from U3R co-expressing untagged STAT3 with STAT3wt-mCherry or STAT3ΔNLS-mCherry were analysed for STAT3 phosphorylation.*

Combination	Construct	Mean ratio $\frac{\text{α-STAT3 pY705}}{\text{α-STAT3 FL}}$
Untagged STAT3wt +	Untagged STAT3wt	1.54 ± 0.18
STAT3wt-mCherry	STAT3wt-mCherry	0.92 ± 0.05
Untagged STAT3wt +	Untagged STAT3wt	1.26 ± 0.1
STAT3ΔNLS-mCherry	STAT3ΔNLS-mCherry	0.97 ± 0.2

In their study describing the STAT3 RR414/417AA ΔNLS mutant for the first time, the authors found that the nuclear accumulation defect of this construct can be rescued by co-expression of STAT3wt (Ma *et al.*, 2003). This is in stark contrast to the results obtained from the Nuclear Translocation Assay presented here that showed that STAT3ΔNLS cross-inhibited the nuclear accumulation of a co-expressed STAT3wt (figure 27). One possible explanation for this discrepancy is that the two proteins were present in different ratios in the two studies, with a relative STAT3ΔNLS excess being present in this work vs relative STAT3wt excess being present in the work by Ma *et al*. Indeed, in their work, they did not have the necessary tools available to quantify relative protein amounts based on fluorescence. As explained earlier, the cells that entered the final analysis stage of the Nuclear Translocation Assay were required to contain at least a 1.5-fold excess of STAT3ΔNLS-mCherry protein to avoid the introduction of measurement artefacts. However, to assess if the ΔNLS phenotype could be rescued when the STAT3wt-GFP was present in higher amounts than the STAT3ΔNLS-mCherry, the correlation between the percentage of STAT3ΔNLS-mCherry excess and the nuclear score of STAT3wt-GFP was calculated. This was explicitly done for all cells that were imaged under the Nuclear Translocation Assay protocol, regardless of the 1.5-fold excess criteria to extend the analysis to cells where high amounts of STAT3wt-GFP were present.

When the Pearson correlation coefficient of the two variables "percentage of STAT3ΔNLS-mCherry excess" and "STAT3wt-GFP nuclear score" was calculated, a value of R_{S3} = -0.21 was found. This indicated a weak negative correlation for the two variables that was statistically significant with p=0.017 (figure 29). To better

understand the relationship between the two variables, linear regression was performed. The result of this showed that the equation $y = -0.00060 * x + 1.419$, with y being the STAT3wt-GFP nuclear score and x being the percentage of STAT3ΔNLS-mCherry molecular excess, best described the relationship of the two variables (figure 29). Together, these results showed that the STAT3ΔNLS exerted cross-inhibitory activity on cytokine-induced nuclear accumulation of a co-expressed STAT3wt construct and that this effect got stronger with increasing concentrations of the STAT3ΔNLS protein. As explained earlier, the discrepancy found between this work and the work by Ma *et al.* could be due to a difference in STAT3ΔNLS-to-STAT3wt ratios in the two studies. Because the data described in this chapter mostly covers cells that contain at least equal amounts of STAT3wt and STAT3ΔNLS, assessment of this theory is limited. However, when the nuclear scores of cells containing less than the 1.5x STAT3ΔNLS excess limit were evaluated, the values for both the GFP and mCherry channels remained low (table 28). This also didn't change when only cells containing more STAT3wt-GFP than STAT3ΔNLS-mCherry were checked for their nuclear scores (1.32 ± 0.13 and 1.75 ± 0.09, respectively). However, it is important to note that only nine cells were included in the latter calculation and that none of them had a greater STAT3wt-GFP excess than 1.29.

Table 26 | Descriptive statistics of nuclear scores determined for cells expressing less than 1.5x the amount of STAT3ΔNLS-mCherry over STAT3wt-GFP.

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT3ΔNLS-mCherry + STAT3wt-GFP	58	1.2	0.04	1.42	0.04

Taken together, the results discussed in this chapter indicate that the STAT3ΔNLS cross-inhibits the cytokine-induced nuclear accumulation of co-expressed STAT3wt by sequestering it in the cytoplasm without affecting its phosphorylation. Importantly, an inhibitory effect was still observed at relatively low concentrations of STAT3ΔNLS which indicated that the effect, like the one exerted by STAT1ΔNLS, was dominant negative. Indeed, based on the data presented here, the notion that co-expressed STAT3wt can rescue the ΔNLS phenotype of STAT3ΔNLS which was published by Ma *et al,* could not be upheld. In any case, the results presented here showed that the

STAT3ΔNLS is a reliable tool to investigate the homo and heterodimer formation of STAT3 via the Nuclear Translocation Assay.

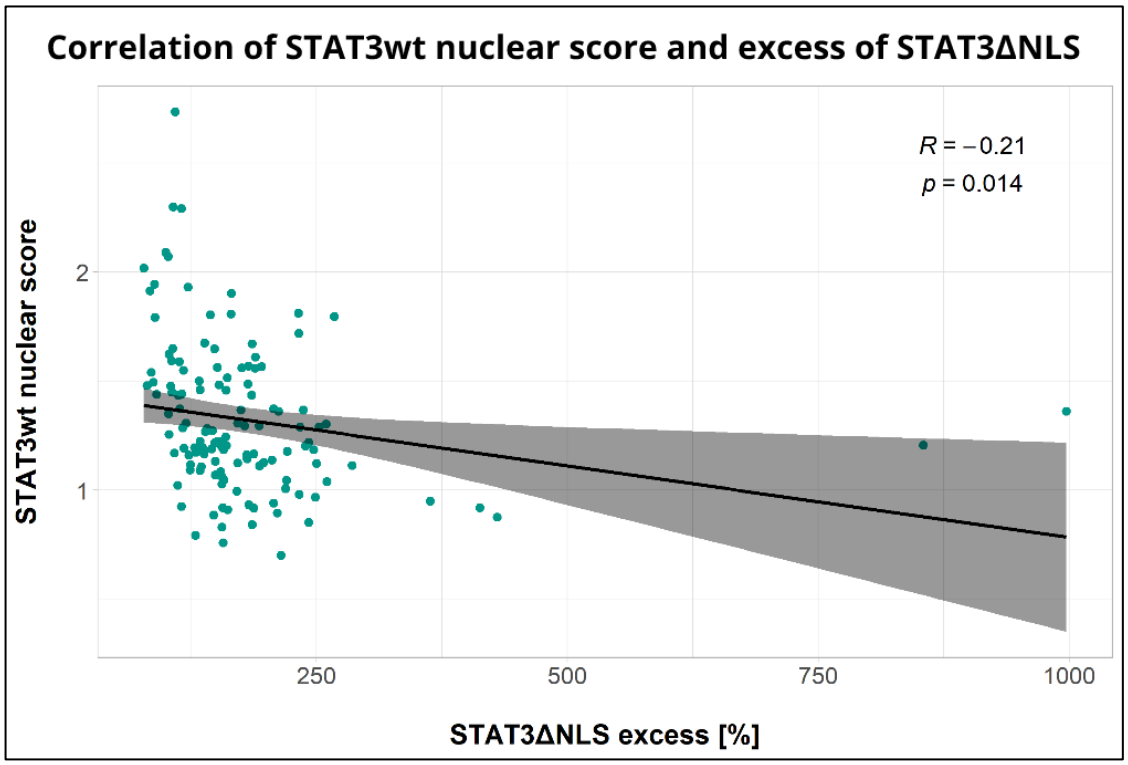


Figure 28 | Correlation analysis of STAT3ΔNLS-mCherry molecular excess and STAT3wt-GFP nuclear score: *Data obtained from HeLa cells expressing both constructs was used to calculate Pearson's correlation coefficient. Hypothesis testing was performed using a student t-test. A linear regression curve is indicated in black with confidence intervals depicted in grey.*

4.2.3 The STAT5BΔNLS mutation is dominant negative.

In this chapter, the ability of STAT5BΔNLS to cross-inhibit the cytokine-induced nuclear accumulation of co-expressed STAT5Bwt will be investigated to confirm that the ΔNLS mutation is dominant negative. First, using the Nuclear Translocation Assay protocol the nuclear score of STAT5Bwt-GFP was determined when co-expressed with either STAT5Bwt-mCherry or STAT5BΔNLS-mCherry. For this, HeLa cells growing on microscopy-grade cell culture dishes were transfected with cDNA coding for said STAT5B combinations, GHR and Jak2. Importantly, these experiments were not conducted blind to genotype. Both cells containing the STAT5Bwt/STAT5Bwt and STAT5Bwt/STAT5BΔNLS combinations showed a pancellular distribution of GFP and mCherry-tagged constructs before cytokine treatment (figure 30A, row 1+3). When the same cells containing the former combination were imaged again 90 min after the

addition of cytokine, a clear shift in subcellular distribution towards the nucleus was detected in both fluorescent channels. Overlaying the Hoechst channel further confirmed this notion. Still, some GFP and mCherry fluorescence remained detectable within the cytoplasm (figure 30A, row 2). The cytokine-induced shift towards nuclear accumulation was not observed in cells co-expressing STAT5Bwt-GFP and STAT5BΔNLS-mCherry. Instead, a slight shift in favour of a cytoplasmic distribution could be seen for STAT5BΔNLS-mCherry in some of these cells (figure 30A, panel 14). This effect was not observed for co-expressed STAT5Bwt-GFP (figure 30A, panel 13). The difference in STAT5Bwt-GFP distribution in cells co-expressing the STAT5Bwt-mCherry or STAT5BΔNLS-mCherry is most obvious when GFP channel images are depicted as contrast-enhanced images (compare figure 30A, panels 8 and 16). In summary, these results indicate that the co-expression of STAT5BΔNLS cross-inhibited the cytokine-induced nuclear accumulation of STAT5Bwt. To further gain evidence for this, the images described just now were quantified, and the GFP and mCherry nuclear scores were calculated for those cells that contained at least 1.5x the amount of mCherry-tagged versus GFP-tagged protein. Based on this, 23 cells containing the STAT5Bwt/STAT5Bwt combination and 27 cells containing the STAT5Bwt/STAT5BΔNLS combination were eligible for this analysis (table 29; figure 30B). For the former cells, average nuclear scores of 1.54±0.05 in the GFP channel and 1.69±0.06 in the mCherry channel were determined (table 29 row 1; figure 30B). Therefore, the nuclear scores observed for STAT5Bwt-GFP here were lower compared to those observed at the 90 min timepoint in STAT5Bwt-GFP time course experiments (figure 20). The average nuclear score for STAT5BΔNLS-mCherry was determined to be 0.99±0.03 (table 29, row 2; figure 30B). This value indicated that the subcellular distribution of this protein did not change despite stimulation with cytokines and thus did not uphold the previously described visual observation that the construct showed partial re-location towards the cytoplasm following cytokine treatment (figure 30A, panel 14). The STAT5Bwt-GFP co-expressed with the STAT5BΔNLS-mCherry showed an average nuclear score of 1.04±0.03. This was significantly different from the STAT5Bwt-GFP determined in STAT5Bwt-mCherry co-expressing cells (table 29, row 2; figure 30B). This indicated that STAT5BΔNLS was indeed able to cross-inhibit the cytokine-induced nuclear accumulation of co-expressed STAT5Bwt-GFP.

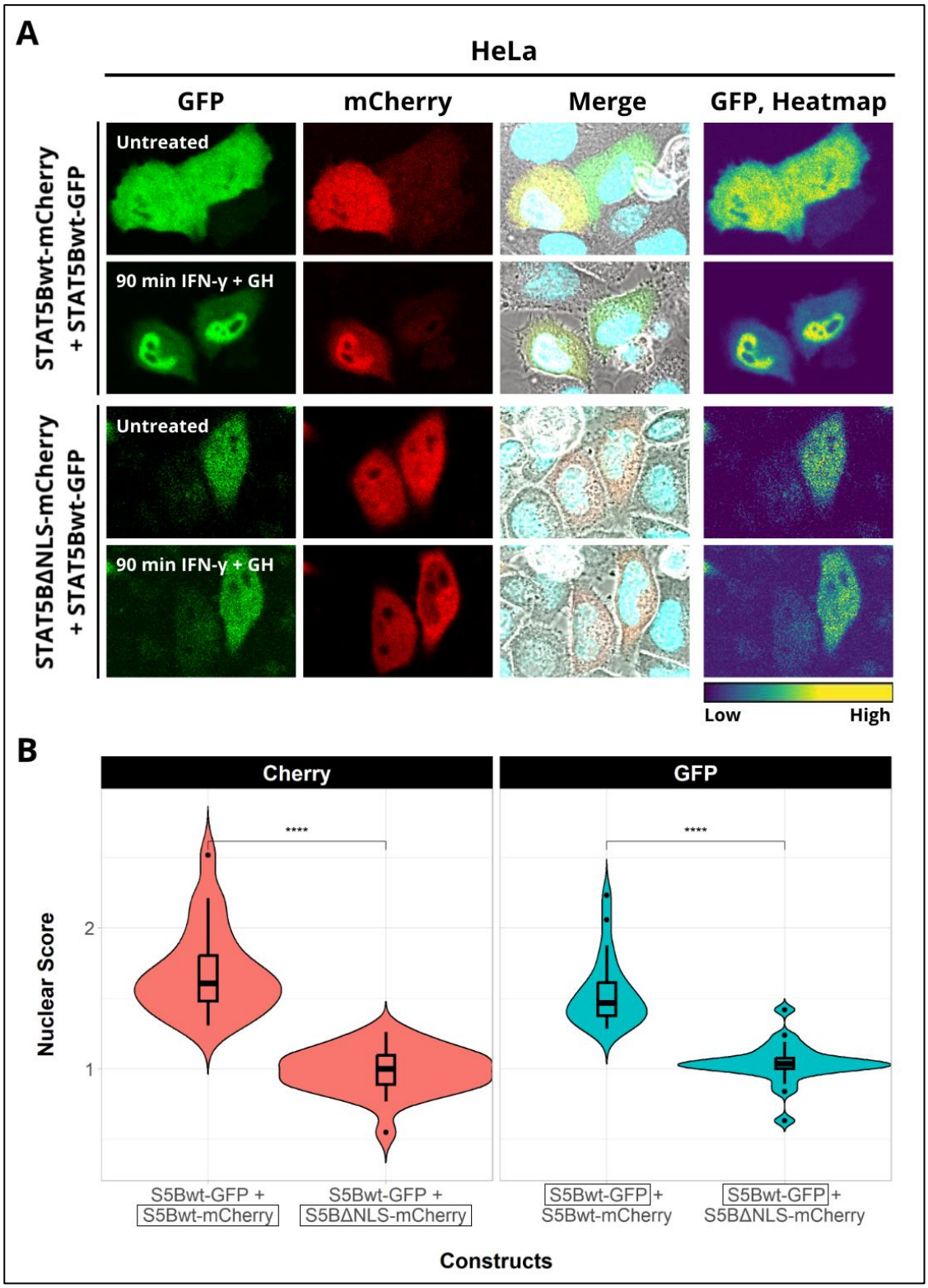


Figure 29 | STAT5BΔNLS cross-inhibits cytokine-induced nuclear accumulation of STAT5Bwt: ***[A]*** *Fluorescence micrographs of HeLa cells co-expressing either STAT5Bwt-mCherry and STAT5Bwt-GFP or STAT5BΔNLS-mCherry and STAT5Bwt-GFP imaged before and 90 min after cytokine treatment. The last column of panels shows a contrast enhanced depiction of the GFP channel. In this, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).* ***[B]*** *Quantification of mCherry and GFP nuclear scores in 23 or 27 cells co-expressing the two STAT5B combinations described in [A], respectively.*

Table 27 | Descriptive statistics of nuclear scores determined for cells co-expressing STAT5Bwt-GFP and STAT5BΔNLS-mCherry or STAT5Bwt-GFP and STAT5Bwt-mCherry.

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT5Bwt-mCherry + STAT5Bwt-GFP	23	1.69	0.06	1.54	0.05
STAT5BΔNLS-mCherry + STAT5Bwt-GFP	27	0.99	0.03	1.04	0.03

Next, the phosphorylation status of each STAT5B construct of the above combinations was assessed to rule out that the observed nuclear accumulation defect was caused by impaired phosphorylation at the cytokine receptor. Importantly, to distinguish the two co-expressed by molecular weight in SDS-PAGE, an untagged STAT5Bwt construct was generated by deleting the oligonucleotide sequence coding for the GFP gene via site-directed mutagenesis. This construct was then co-transfected with either STAT5Bwt-mCherry or STAT5BΔNLS-mCherry alongside GHR and Jak2. Whole-cell extracts from GH-treated cells expressing the two STAT5B combinations were then used in Western blot analysis and the full-length STAT5B and the STAT5B pY694 phosphopeptide were detected simultaneously using specific antibodies. Four bands were detected in each sample when imaging the plot to detect the full-length STAT5. In each sample, two of the bands migrated close together at a molecular weight consistent with that of mCherry-tagged STAT5B and two bands at a molecular weight consistent with that of untagged STAT5B (figure 23). This banding pattern was also observed in earlier bands the two bands per construct likely stemmed from a STAT5B degradation product or a variant of STAT5B that carried an additional post-translational modification. Either way, all bands appeared to have similar strengths, indicating equal expression of tagged and untagged STAT5B in U3R cells (figure 31A, row 1). When the anti-STAT5 pY694 antibody was used only two bands were visible in each sample. The slower band migrated at a molecular weight consistent with phosphorylated STAT5B-mCherry and the faster one at a molecular weight consistent with phosphorylated, untagged STAT5B (figure 31A, row 2). This indicated that the variants that may have arisen from additional post-translational modifications or degradation could not be phosphorylated in response to cytokine treatment.

The presence of such a product that cannot be phosphorylated and as such is expected to remain cytoplasmic after cytokine treatment might explain the poor cytokine-induced nuclear accumulation of STAT5B observed in experiments presented in this and prior chapters. Nevertheless, the phosphorylation of the other bands seemed sufficient. To quantify this, plots from three independent experiments were quantified by measuring band fluorescence intensities and a score calculated as before (table 30).

When comparing these values calculated for the top bands of each sample, a significant increase in STAT5BΔNLS-mCherry over STAT5Bwt-mCherry phosphorylation was noticed ($p = 0.01$; table 30). However, in both cases, the phosphorylation of the proteins was decreased by about 50% in co-expression settings in contrast to earlier experiments where they were expressed alone, albeit this was not significant for the latter construct ($p=0.21$; figure 23). Yet, considering that the STAT5Bwt-mCherry showed a sufficient nuclear score in previous Nuclear Translocation Assay experiments it is unlikely that the reduced levels of phosphorylation may cause severe artefacts in this assay (figure 30, table 29). Still, the nuclear score found for STAT5Bwt-mCherry was not as high as at the 90 min time point in STAT5Bwt-GFP nuclear accumulation time course experiments described in chapter 4.1.3 (figure 20), which might mean that subtle cross-inhibitory effects exerted by the STAT5BΔNLS may be masked when probing certain interaction partners. Another reason why the observed diminished phosphorylation should not be of concern for the application is that the STAT5BΔNLS-mCherry construct exhibited full inhibition on the cytokine-induced nuclear accumulation of co-expressed STAT5Bwt-GFP (figure 30, table 29), despite its decreased phosphorylation.

The untagged STAT5Bwt showed similar phosphorylation as its co-expressed STAT5Bwt-mCherry counterpart ($p=0.31$), and this was also not significantly diminished in cells that co-expressed STAT5BΔNLS-mCherry instead of STAT5Bwt-mCherry ($p=0.14$; figure 31B, table 30). These results suggested that the co-expression of STAT5BΔNLS did not adversely affect STAT5Bwt phosphorylation following cytokine treatment and confirmed this construct as suitable for protein interaction experiments within the Nuclear Translocation Assay.

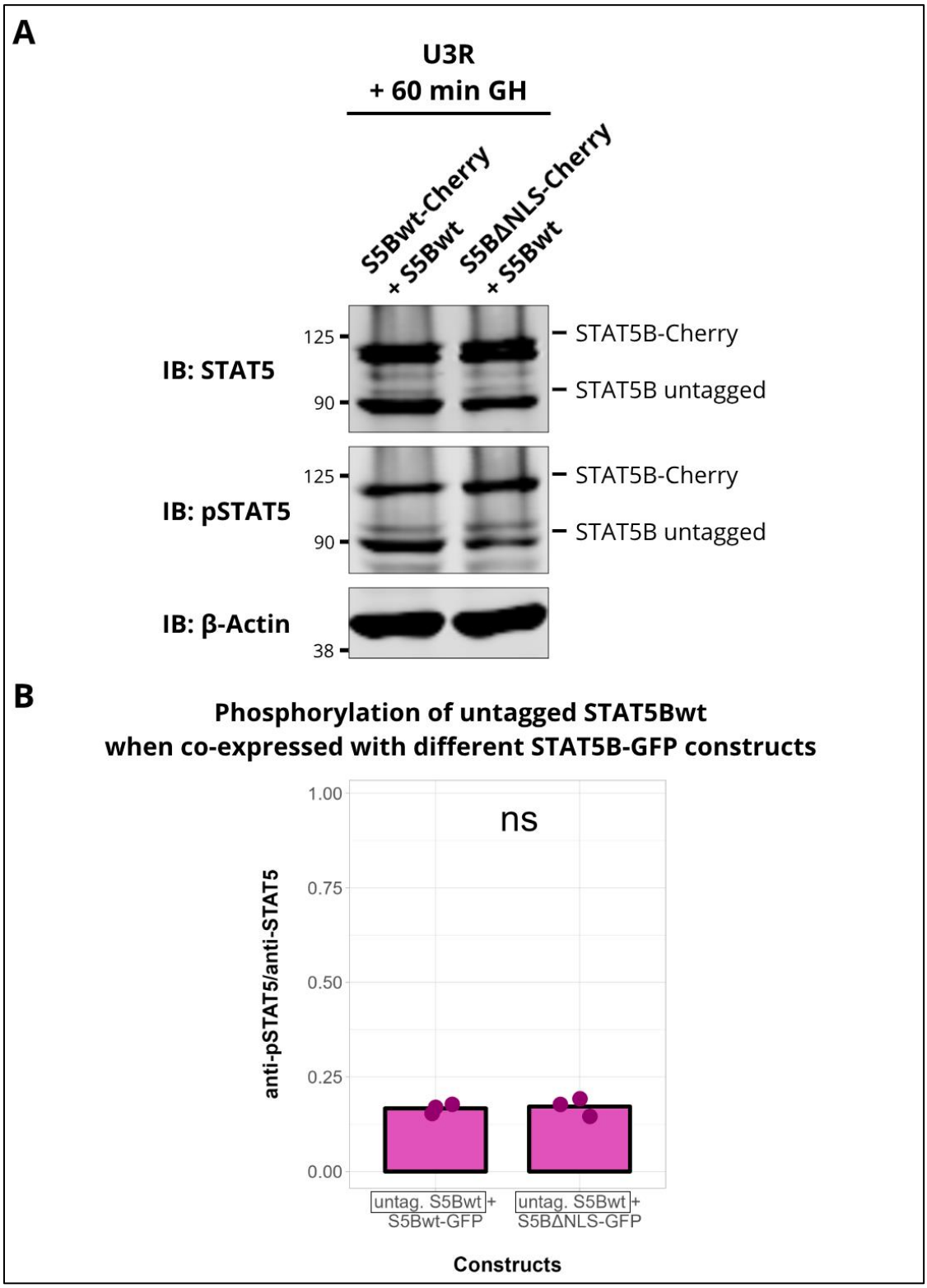


Figure 30 | Co-expression of STAT5BΔNLS does not adversely affect STAT5Bwt phosphorylation: ***[A]*** *U3R cells coexpressing untagged STAT5Bwt with either STAT5Bwt-mCherry or STAT5BΔNLS-mCherry were treated with GH for 60 min. Whole cell extracts were prepared and subjected to Western blot analysis using anti-STAT5B and anti-STAT5B pY694 antibodies.* ***[B]*** *Quantification of untagged STAT5Bwt phosphorylation when co-expressed with either STAT5Bwt-mCherry or STAT5BΔNLS-mCherry. Quantifications were performed on data from 3 independent experiments including the one shown in [A]. MW markers in kDa.*

Table 28 | Quantification of STAT5B phosphorylation by Western blotting: *Whole cell extracts from U3R co-expressing untagged STAT5B with STAT5Bwt-mCherry or STAT5BΔNLS-mCherry were analysed for STAT5B phosphorylation.*

Combination	Construct	Mean ratio $\frac{\text{α-STAT5B pY694}}{\text{α-STAT5B FL}}$
Untagged STAT5Bwt +	Untagged STAT5Bwt	0.17 ± 0.01
STAT5Bwt-mCherry	STAT5Bwt-mCherry	0.16 ± 0.007
Untagged STAT5Bwt +	Untagged STAT5Bwt	0.17 ± 0.02
STAT5BΔNLS-mCherry	STAT5BΔNLS-mCherry	0.2 ± 0.01

As with STAT1 and STAT3, the inhibitory effect exerted on the STAT5B nuclear accumulation by STAT5BΔNLS was further characterised. To this end, the correlation between the STAT5BΔNLS-mCherry molecular excess and the nuclear score of STAT5Bwt-GFP was assessed in co-expressing cells from earlier. Cells containing less than 1.5-fold STAT5BΔNLS-mCherry excess over STAT5Bwt-GFP were also included in this analysis. A Pearson's correlation coefficient of R_{S5B} = -0.15 indicated no significant correlation between the two variables (p = 0.45) which strongly suggested that the effect exerted by the STAT5BΔNLS was dominant-negative (figure 32). However, it is important to note that only one additional cell contained a STAT5BΔNLS below the 1.5x threshold and thus data for cells containing low levels of this protein is limited. Still, the results obtained in the correlation analysis strongly suggest that the inhibitory effect should be similar across STAT5BΔNLS concentrations used typically in Nuclear Translocation Assay experiments.

In summary, the data presented in this chapter showed a dominant-negative effect on STAT5Bwt cytokine-induced nuclear accumulation exerted by co-expressed STAT5BΔNLS. Furthermore, the inhibition was not caused by decreased phosphorylation but instead by the cross-inhibitory effect of the ΔNLS mutation, suggesting that the NLS sequence in both monomers must be intact for efficient cytokine-induced nuclear accumulation of the active dimer. This confirms the STAT5BΔNLS as an ideal candidate for use in protein interaction studies using the Nuclear Translocation Assay. Moreover, these findings validate that active STAT5B homodimers could be detected reliably using the Nuclear Translocation Assay.

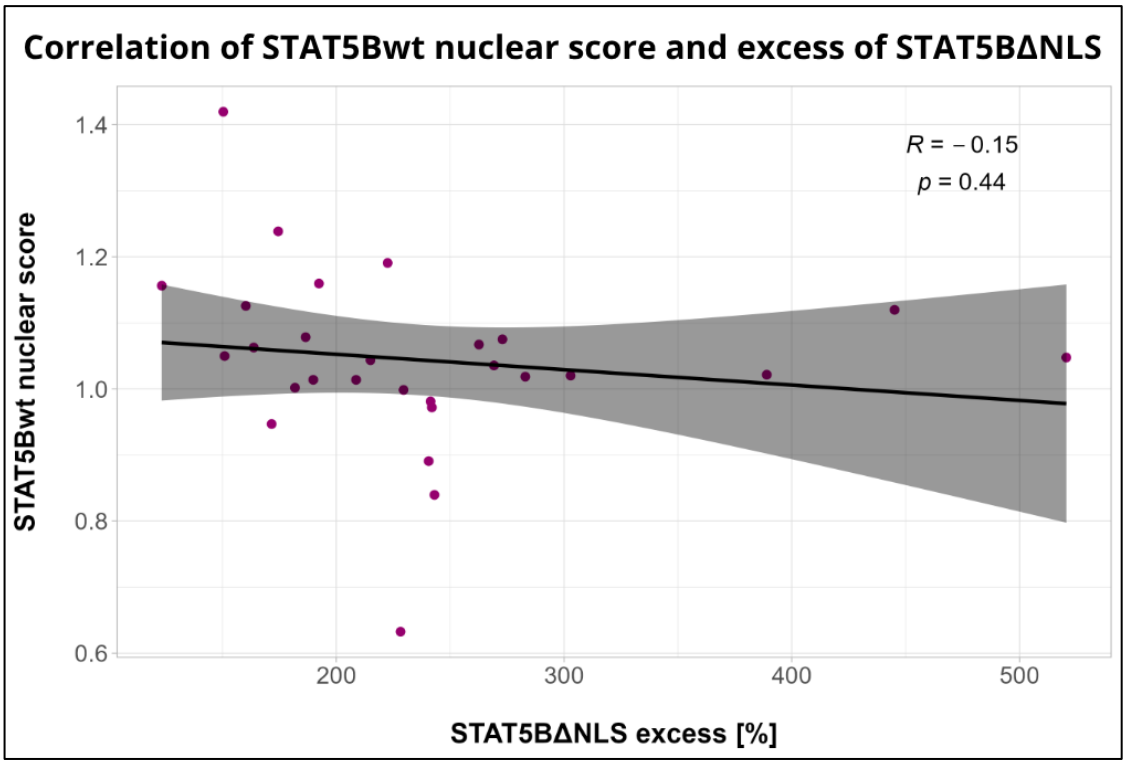


Figure 31 | Correlation analysis of STAT5BΔNLS-mCherry molecular excess and STAT5Bwt-GFP nuclear score: *Data obtained from HeLa cells expressing both constructs was used to calculate Pearson's correlation coefficient. Hypothesis testing was performed using a student t-test. A linear regression curve is indicated in black with confidence intervals depicted in grey.*

4.2.4 STAT5BΔNLS does not inhibit cytokine-induced nuclear import of co-expressed STAT1wt.

Several studies confirm that both STAT1 and STAT5B can form active heterodimers with other STATs. For STAT1, active heterodimer formation with STAT2, STAT3 and STAT4 has been confirmed before. (Delgoffe and Vignali, 2013; Harrison and Moseley, 2020). STAT5B was found to form active heterodimers with STAT3 and STAT5A (Delgoffe and Vignali, 2013; Harrison and Moseley, 2020). However, previous studies have determined several cytokines that activate STAT1 and STAT5B at the same time but could not identify the formation of STAT1/STAT5B heterodimers (Wehinger *et al.*, 1996; Kirken *et al.*, 1997; Schulze *et al.*, 2000). The method of choice for active STAT homo and heterodimer detection in these and other studies was EMSA DNA binding assays. Indeed, DNA binding assays were vital in the discovery of several STAT dimers including IFN-γ-activated factor (GAF), the ISGF3 complex and STAT1/STAT3 heterodimers (Levy *et al.*, 1989; Decker *et al.*, 1991; Zhong, Wen and Darnell, 1994a) and are therefore a well-characterised tool to investigate active STAT dimers.

To replicate these results in this work, EMSA experiments were conducted next. For this, U3R cells were left untransfected, transfected with STAT1wt-GFP alone or co-transfected with STAT5B-mCherry and GHR or STAT1wt-GFP, STAT5B-mCherry and GHR. After a 24 h incubation period, the transfected cells were either left untreated or treated with IFN-γ and GH for 60 min. Afterwards, detergent-free whole-cell extracts were generated. The said extracts were then incubated with an IRDye® labelled DNA probe termed IFN-γ response region (GRR) from the promotor of the high-affinity Fc receptor for IgG (FcγRI) that was previously shown to bind STAT1 and STAT5 (Pearse *et al.*, 1993; Johnston *et al.*, 1995). The complexes were separated on a native polyacrylamide gel electrophoretically, and fluorescent bands were detected using a LICOR Odyssey® imager.

No specific bands were detected in extracts from untransfected U3R cells irrespective of the cytokine treatment (figure 33, lane 1+2). Extracts from cells expressing STAT1wt-GFP alone or STAT5B-mCherry and GHR showed one strong band, respectively. The band detected in the latter sample migrated slower than the one detected in the former sample (figure 33, lane 3+4). When STAT1wt-GFP, STAT5B-mCherry and GHR were expressed together, two bands were observed (figure 33, lane 5). The top band migrated at the same speed as the single band observed in STAT5B-mCherry+GHR containing extracts while the bottom band migrated at the same height as the one observed in extracts only containing STAT1wt-GFP. Indeed, the top band was only supershifted by an antibody specific for STAT5 identifying it as STAT5B homodimer (figure 33, lane 7), while the bottom band was only supershifted by an anti-STAT1 antibody revealing it as STAT1 homodimer (figure 33, lane 6). Importantly, the co-expression of STAT1 and STAT5B did not lead to the formation of an intermediate band. This strongly suggested that no STAT1/STAT5B heterodimers could be formed. Therefore, it was expected that the co-expression of STAT5BΔNLS would not cross-inhibit STAT1wt cytokine-induced nuclear accumulation and thus should not cause a drop in the nuclear score of the latter in the Nuclear Translocation Assay.

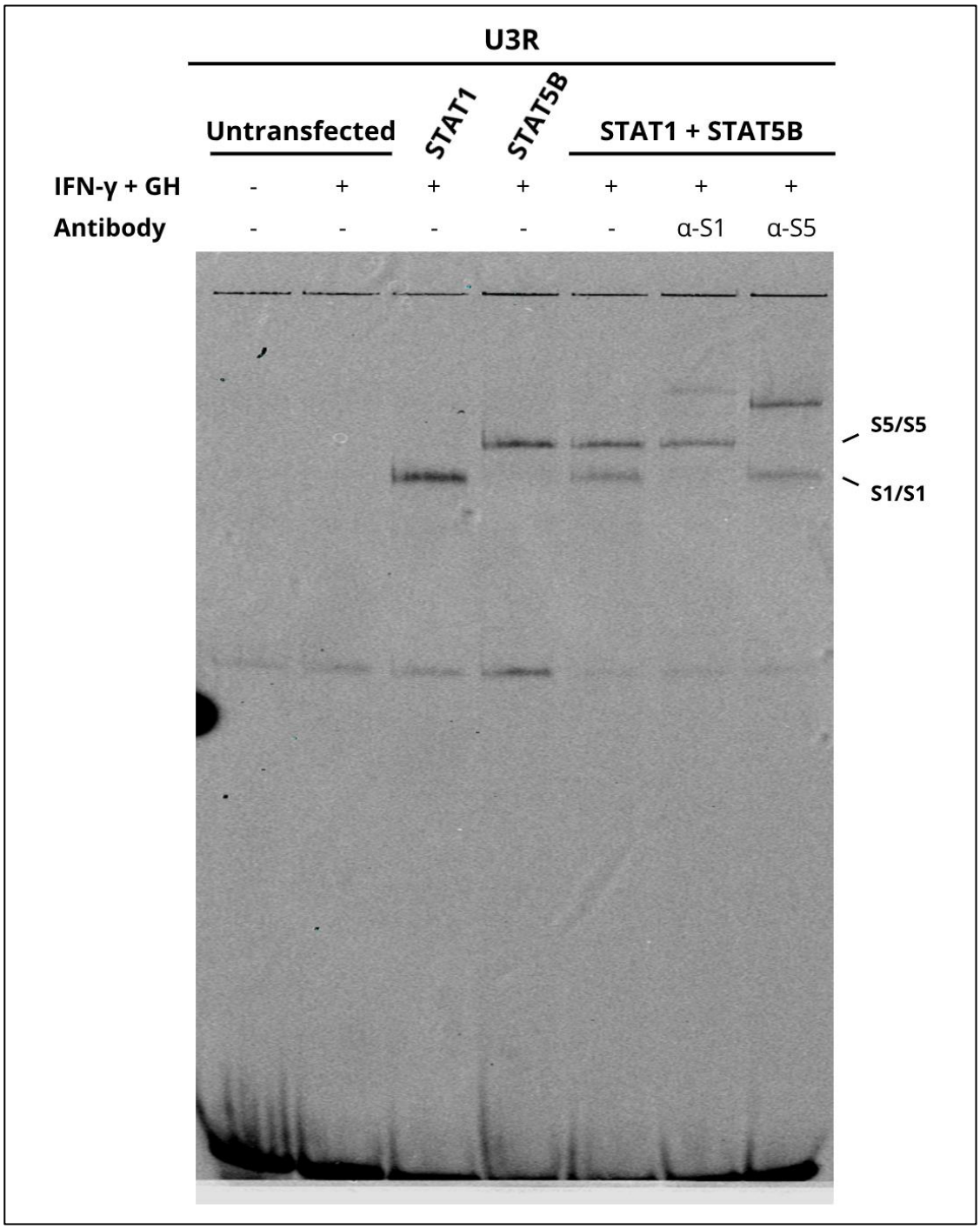


Figure 32 | STAT1 and STAT5B form two sets of active dimers when co-expressed: a STAT1 homodimer and a STAT5B homodimer: *Whole cell extracts of cytokine treated and non-treated U3R cells that were left untransfected, transfected with STAT1wt-GFP alone or co-transfected with STAT5Bwt-mCherry, GHR and Jake or STAT1wt-GFP, STAT5Bwt-mCherry, GHR and Jak2 were subjected to EMSA experiments using the GRR DNA probe. Antibodies raised against native full-length STAT1 and STAT5 were used for supershifts. This figure is representative for two independent experiments.*

To test these assumptions, HeLa cells were co-transfected with cDNA plasmids coding for STAT1wt-GFP and STAT5B-mCherry or STAT1wt-GFP and STAT5BΔNLS-mCherry. After a 24 h long incubation period, the living cells were imaged before and 90 min after treatment with IFN-γ and GH according to the Nuclear Translocation Assay protocol. The images revealed that both STAT5B constructs used in this experiment elicited a pancellular distribution in untreated cells (figure 34A, panel 2 + 10), which was also observed in previous experiments where these constructs were co-expressed with other STAT5B proteins (figure 30). Similarly, the STAT1 distribution observed in this experiment also resembled previous accounts of a more cytoplasmic localisation in latent cells (figure 34A, panel 1 + 9). This confirmed that the co-expression of STAT1 does not alter the subcellular distribution of latent STAT5B proteins or vice-versa. Nuclear accumulation was detected for both STAT1wt-GFP and STAT5B-mCherry when co-expressed together in images taken 90 min after cytokine treatment (figure 34A, row 2 + 4). The accumulation of the STAT5B protein was not complete; a phenotype that was also observed in earlier experiments (figure 30). The cytokine-induced nuclear accumulation of STAT1wt-GFP appeared to be complete even when combined with STAT5BΔNLS-mCherry (figure 34A, panel 13). This was despite the co-expressed STAT5BΔNLS-mCherry not showing any signs of nuclear accumulation after cytokine treatment. Some cells showed a cytokine-induced shift of STAT5BΔNLS towards the cytoplasm (figure 34A, panel 14), a phenomenon that was also observed for this construct in previous experiments, but not for other STATΔNLS proteins (compare figures 24, 27 and 30). These findings suggest that STAT5BΔNLS-mCherry does not cross-inhibit STAT1wt-GFP cytokine-induced nuclear accumulation which in turn would mean that the two proteins do not form active heterodimers.

Table 29 | Descriptive statistics of nuclear scores determined for cells co-expressing STAT1wt-GFP and STAT5BΔNLS-mCherry or STAT1wt-GFP and STAT5Bwt-mCherry.

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT5Bwt-mCherry + STAT1wt-GFP	63	1.31	0.05	3.32	0.14
STAT5BΔNLS-mCherry + STAT1wt-GFP	77	0.88	0.02	2.91	0.08

To confirm these assumptions, microscopy images were quantified according to the Nuclear Translocation Assay protocol. Assessment of relative protein amounts based on GFP and mCherry fluorescence intensities revealed that 63 or 77 cells containing STAT1wt-GFP and STAT5B-mCherry or STAT1wt-GFP and STAT5BΔNLS-mCherry, respectively, fulfilled the criteria dictated by the protocol and could be progressed to quantification of GFP and mCherry nuclear scores. The calculations revealed a nuclear score of 3.32±0.14 for STAT1wt-GFP when combined with STAT5Bwt-mCherry for which a nuclear score of 1.31±0.05 was reported (figure 34B; table 31). While these numbers are indicative of nuclear accumulation, it is worth noting that both constructs were measured with higher nuclear scores previously when co-expressed with other proteins (table 23+29). Image quantifications of cells containing the STAT1wt-GFP/STAT5BΔNLS-mCherry combination revealed a STAT1wt-GFP nuclear score of 2.91 ± 0.08 (figure 34B; table 31). Even though this was significantly different to the score that was determined for the same protein when co-expressed with STAT5Bwt-mCherry, the absolute numbers were still sufficiently similar to rule out a negative cross-inhibition of STAT1wt-GFP cytokine-induced nuclear accumulation by the STAT5BΔNLS-mCherry. Said STAT5BΔNLS-mCherry itself showed a nuclear score of 0.88±0.02 (figure 34B; table 31). A nuclear score lower than $F_n = 1$ indicated a shift towards a more cytoplasmic distribution following cytokine stimulation. While a value of 0.88±0.02 suggested only a slight redistribution, this is notable in fluorescence micrographs as described before (figure 34A, panel 14). Furthermore, a similar phenotype was observed visually in experiments presented in chapter 4.1.4 but could not be verified by quantifications (figure 30; table 29) which could imply that the effect was indeed only subtle. On the other hand, the fact that this STAT5BΔNLS phenotype was observed when co-expressed with either STAT5Bwt-GFP (figure 30) or STAT1wt-GFP (figure 34) clearly shows that the presence of the latter does not impair STAT5BΔNLS function. Consequently, the results presented here indicated that STAT1 and STAT5B did not form active STAT heterodimers and that this was distinguishable in the Nuclear Translocation Assay.

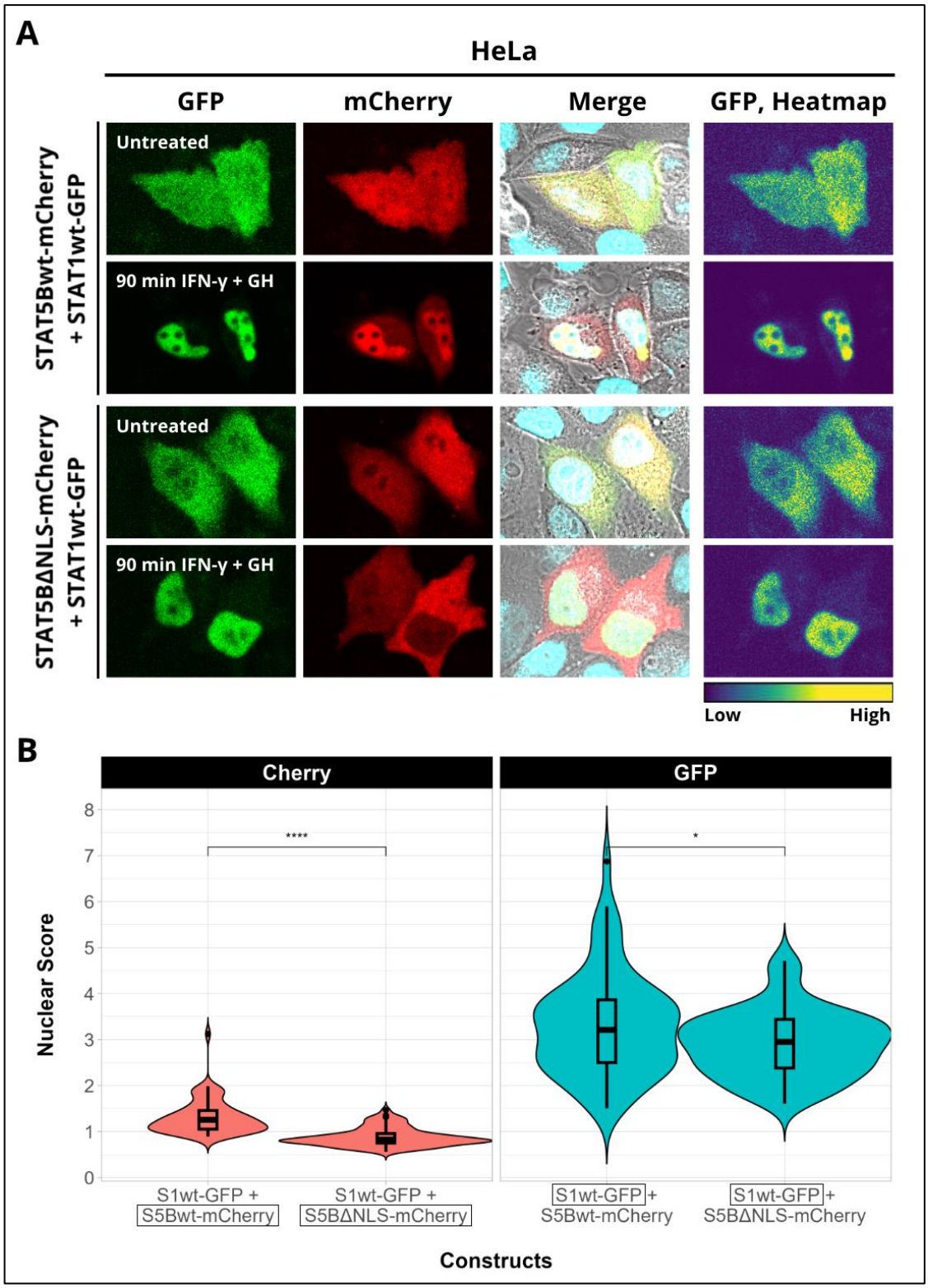


Figure 33 | STAT5BΔNLS does not cross-inhibit cytokine-induced nuclear accumulation of STAT1wt: ***[A]*** *Fluorescence micrographs of HeLa cells co-expressing either STAT5Bwt-mCherry and STAT1wt-GFP or STAT5BΔNLS-mCherry and STAT1wt-GFP imaged before and 90 min after cytokine treatment. The last column of panels shows a contrast enhanced depiction of the GFP channel. In this, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).* ***[B]*** *Quantification of mCherry and GFP nuclear scores in 63 or 77 cells co-expressing the two STAT1/STAT5B combinations described in [A], respectively.*

To rule out that diminished cytokine-induced tyrosine phosphorylation is to blame for the inability of STAT5BΔNLS to cross-inhibit STAT1wt nuclear accumulation, the phosphorylation status of each construct was determined in Western blot experiments. For this, U3R cells were co-transfected with the STAT1/STAT5B combinations from earlier on three independent occasions, treated for 60 min with IFN-γ and GH and lysed to generate whole cell extracts. Each extract was then loaded onto two separate gels for electrophoretic separation followed by blotting onto separate nitrocellulose membranes. After blocking, one of the two membranes was incubated with anti-STAT1 and anti-STAT1 pY701 antibodies while the other was incubated with anti-STAT5 and anti-STAT5B pY694 antibodies.

The Western blot revealed a single band migrating at a molecular weight consistent with STAT1wt-GFP when extracts containing STAT1wt-GFP and STAT5Bwt-mCherry were assessed with the anti-STAT1 antibody (figure 35A, panel 1). When the same extract was probed using the anti-STAT1 pY701 antibody, a single band running at the same height was detected (figure 35A, panel 5). This indicated that STAT1wt-GFP was phosphorylated in response to cytokine treatment. Probing this extract with the anti-STAT5 antibody loaded on the other membrane revealed two strong bands migrating close together at a height that would be expected for STAT5B-mCherry protein (figure 35A, panel 2). As before, only one band was visible when detecting with the anti-STAT5B pY694 antibody (figure 35A, panel 6) and this band migrated at a similar speed to the slower migrating band detected with the antibody raised against full-length STAT5. Thus, the faster migrating STAT5B band might be a product of degradation or a post-translationally modified version of STAT5B that cannot be phosphorylated. However, the slower migrating band seemed to be phosphorylated efficiently. A single band was observed when extracts containing STAT1wt-GFP and STAT5BΔNLS-mCherry were probed with the anti-STAT1 antibody, although this was weaker than the STAT1 band detected in the other extracts (figure 35A, panel 3). Still, a single band was observed with the anti-STAT1 pY701 antibody which was also proportionately weaker (figure 35A, panel 7). Nevertheless, this indicated that the STAT1 contained in these extracts was phosphorylated in response to cytokine treatment. When the same extract was incubated with the anti-STAT5 antibody, again, two bands migrating close together at a size that would be expected for the STAT5BΔNLS-mCherry protein were

found (figure 35A, panel 4). As before, only one band was observed when detecting with the anti-STAT5B pY694 antibody suggesting that only the slower-moving band was phosphorylated (figure 35A, panel 8). Loading was similar in all wells as evidenced by detecting β-Actin with a specific antibody (figure 35A, row 3). All in all, the visual observations described here indicate that both STAT1 and STAT5B constructs get efficiently phosphorylated following treatment with IFN-γ and GH when co-expressed together. This phenotype was not influenced by the STAT5BΔNLS mutation.

Table 30 | Quantification of STAT1 and STAT5B phosphorylation by Western blotting: *Whole cell extracts from U3R co-expressing STAT1wt-GFP with STAT5Bwt-mCherry or STAT5BΔNLS-mCherry were analysed for phosphorylation.*

Combination	Construct	Mean ratio $\frac{\text{α-STAT pY}}{\text{α-STAT FL}}$
STAT1wt-GFP +	STAT1wt-GFP	0.12 ± 0.01
STAT5Bwt-mCherry	STAT5Bwt-mCherry	0.37 ± 0.27
STAT1wt-GFP +	STAT1wt-GFP	0.12 ± 0.01
STAT5BΔNLS-mCherry	STAT5BΔNLS-mCherry	0.19 ± 0.06

To confirm these impressions, Western blots of U3R whole cell extracts that were co-transfected with the two STAT1/STAT5B combinations in three independent experiments were quantified. A value representing the proportion of phosphorylated protein calculated by dividing the fluorescence intensity measured for bands detected with the antibodies raised against the phosphopeptide by the intensity measured for the full-length protein bands for STAT1 and STAT5B. There was no significant difference in STAT1wt-GFP phosphorylation whether the construct was co-expressed with STAT5Bwt-mCherry or STAT5BΔNLS-mCherry (p=0.86; figure 35B; table 32). Furthermore, the phosphorylation was similar to what was found in experiments where U3R cells expressed STAT1wt-GFP alone (p=0.71 and p=0.82; figure 21). When comparing the phosphorylation of STAT5Bwt-mCherry versus STAT5BΔNLS-mCherry no significant difference was found (p=0.37; figure 35B; table 32). However, the phosphorylation of these proteins decreased when co-expressed with STAT1wt-GFP compared to experiments where the proteins were expressed alone (figure 23).

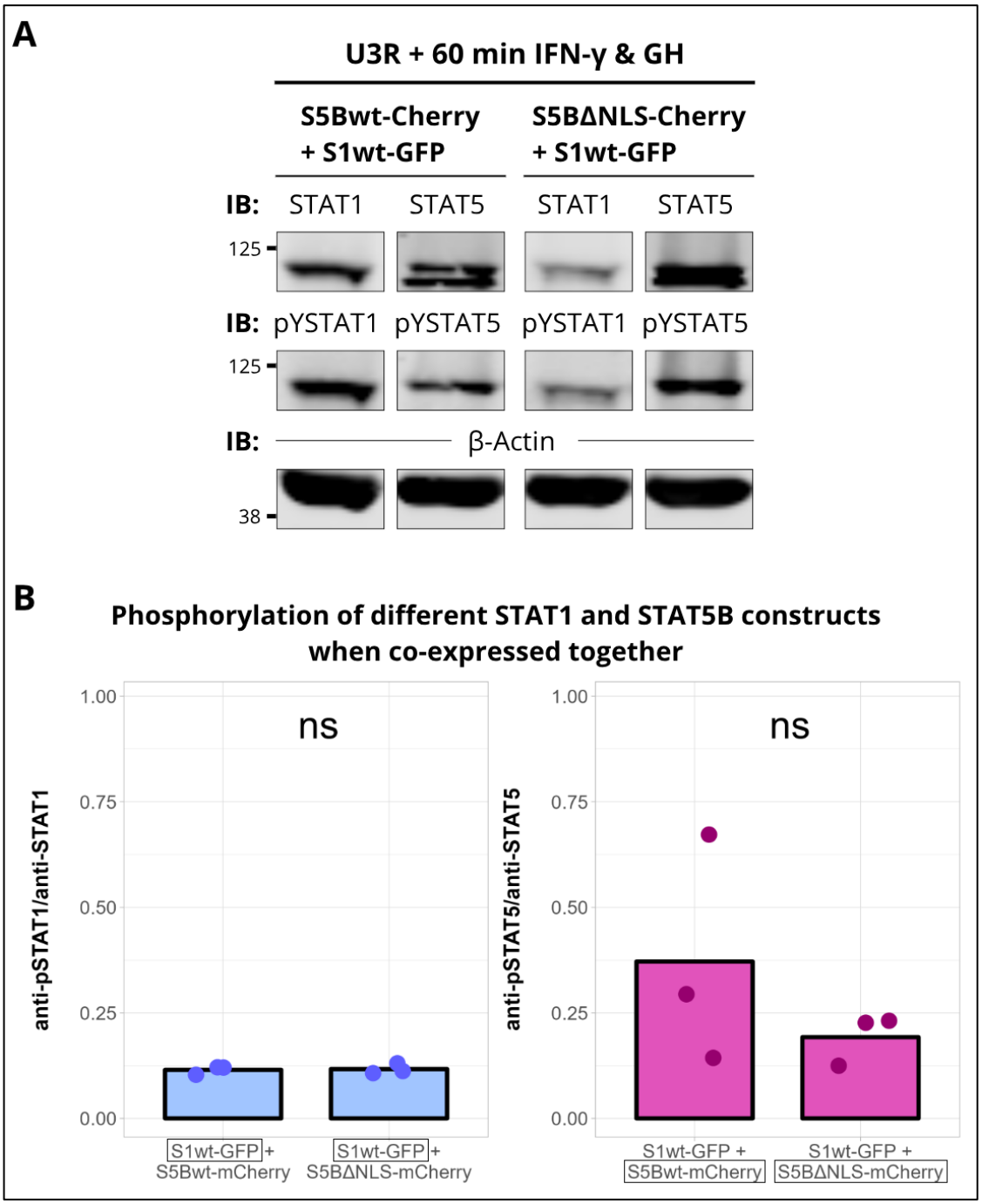


Figure 34 | Co-expression of STAT5BΔNLS does not adversely affect STAT1wt phosphorylation: ***[A]*** *U3R cells coexpressing STAT1wt-GFP with either STAT5Bwt-mCherry or STAT5BΔNLS-mCherry were treated with IFN-γ and GH for 60 min. Whole cell extracts were prepared and subjected to Western blot analysis using anti-STAT1, anti-STAT5B, anti-STAT1 pY701 and anti-STAT5B pY694 antibodies.* ***[B]*** *Quantification of STAT1wt-GFP, STAT5Bwt-mCherry and STAT5BΔNLS-mCherry when co-expressed together. Quantifications were performed on data from three independent experiments including the ones shown in [A]. MW markers in kDa.*

In summary, the data presented in this chapter confirms evidence found in the literature that STAT1 and STAT5B do not form active heterodimers following IFN-γ and GH treatment and that this is recognised by the Nuclear Translocation Assay. Furthermore, Western blot analysis revealed no defect in phosphorylation for all STAT1

and STAT5B constructs investigated which ruled out diminished phosphorylation as the cause for the observations made in the Nuclear Translocation Assay.

4.2.5 Discussion of Chapter 4.2

The main purpose of the experiments described in this chapter was to validate the dominant negative effect of the STAT1ΔNLS, STAT3ΔNLS and STAT5BΔNLS mutations. Furthermore, the accuracy of the Nuclear Translocation Assay was tested by assessing if it could distinguish between known STAT interactors (STAT1, STAT3 and STAT5B homodimers) and known non-interactors (STAT1 and STAT5B heterodimers).

The Nuclear Translocation Assay requires that the STATΔNLS mutation used to assess dimer formation with two STATs is dominant negative, meaning that the cytokine-induced nuclear accumulation of a STATΔNLS/STAT dimer is abrogated even though only one monomer carries the ΔNLS mutation. Indeed, this was shown to be the case for the STAT1ΔNLS, the STAT3ΔNLS and the STAT5BΔNLS. In a previous study, the STAT1ΔNLS was shown to cross-inhibit the type I IFN-induced nuclear accumulation of STAT2 (Melén, Kinnunen and Julkunen, 2001; Fagerlund *et al.*, 2002). This work could now expand these findings to dominant-negative exhibited by STAT1ΔNLS in homodimers (figure 26). A previous study on the effects of the STAT3ΔNLS mutation RR414/417AA indicated that abrogated IL-6-induced nuclear accumulation could be rescued by co-expression of a STAT3wt construct (Ma *et al.*, 2003). This could not be replicated in this work. Instead, the STAT3ΔNLS exerted cross-inhibition on STAT3wt after treatment with IFN-γ and IL-6 (figure 29). In experiments presented in a later chapter (4.4.2), the STAT3ΔNLS was observed to have the same effect on STAT1wt accumulation following cytokine treatment. The residues mediating STAT5B cytokine-induced nuclear accumulation have been narrowed down to RR422/425 for the first time in this work. Additionally, experiments described in this chapter revealed that this mutation is dominant negative since it cross-inhibited nuclear accumulation of co-expressed STAT5Bwt following treatment with IFN-γ and GH over a broad range of concentrations (figure 32). Extensive literature research revealed two other studies describing STAT5B mutants that are unable to accumulate in the nucleus following cytokine treatments and these findings have been discussed in chapter 4.1.5. Upon review of these studies, it was discovered that neither investigated if the ΔNLS

mutation was dominant negative (Herrington *et al.*, 1999; Iyer and Reich, 2008) making this work the first account of such an effect.

The findings described in the previous paragraph also confirmed that STAT1, STAT3 and STAT5B homodimers could be detected reliably (figure 24, 27 and 30). Importantly, when a STAT5BΔNLS-mCherry was co-expressed with STAT1wt-GFP, the latter's cytokine-induced nuclear accumulation was not cross-inhibited (figure 34), confirming previous accounts that these two proteins do not form active heterodimers (Wehinger *et al.*, 1996; Kirken *et al.*, 1997; Schulze *et al.*, 2000). This showed that the Nuclear Translocation Assay could reliably distinguish between non-interactors like STAT1 and STAT5B and known interactors like STAT1, STAT3 and STAT5B homodimers.This novel assay may also be useful for the assessment of interactions between proteins beyond the STATs. The approach employed by the Nuclear Translocation Assay has the advantage that interactions can be assessed in living cells rather than relying on purified proteins or cell extracts. Indeed, using altered subcellular distributions to highlight interactions of other proteins with STATs was reported before, especially in studies investigating the inhibition of STAT functions by viral proteins. For example, using indirect immunofluorescence, it was shown that the Nipah virus V protein interacts with, and retains both STAT1 and STAT2 in the cytoplasm, thereby inhibiting their cytokine-induced nuclear accumulation (Rodriguez, Parisien and Horvath, 2002). Other viruses' V proteins within the same family (*Paramyxoviridae*) likewise interact with STAT1 and STAT2 and this was also uncovered using fluorescence microscopy by showing their type I IFN-induced nuclear accumulation was diminished in virus-infected cells (Palosaari *et al.*, 2003; Hagmaier *et al.*, 2007). Besides these, a similar mechanism has been observed in microscopy experiments for the rabies virus P protein which sequesters STAT1 and STAT3 inside the cytoplasm to inhibit their cytokine-induced nuclear accumulation (Vidy, Chelbi-Alix and Blondel, 2005; Vidy *et al.*, 2007; Lieu *et al.*, 2013), Zika virus NS5 which targets the nuclear accumulation of STAT2 in response to cytokine (Grant *et al.*, 2016), hepatitis C virus core virus that targets STAT1 accumulation (Melén *et al.*, 2004) and Lyssavirus P protein that selectively inhibits STAT1/STAT3 heterodimers by retention in the cytoplasm (Harrison *et al.*, 2020). Designs resembling that of the Nuclear Translocation Assay have also been used outside the STAT field (Knauer *et al.*, 2005; Dixon and Lim, 2010).

4.3 Active STAT1 and STAT3 homodimers form independent of unphosphorylated dimer integrity.

The previous chapters described the development and validation of the Nuclear Translocation Assay. In this chapter, the process of applying the Nuclear Translocation Assay to examine how active homodimers of STAT1 and STAT3 assemble is described. Two alternatives were examined. Either the homodimers pre-form before cytokine stimulation and transition to active signalling complexes via a conformational change without the exchange of monomers or active homodimers form independent of their unphosphorylated counterparts and exchange monomers in the process. To resolve this question, four different scenarios had to be considered which will be outlined in the following paragraphs.

4.3.1 Introduction

Unphosphorylated and active STAT dimers are held together by different interaction surfaces. While the active dimer is mediated by reciprocal SH2/pY interactions involving the C-terminal ends of monomers, it is the N-terminal portion of the proteins that facilitate the formation of unphosphorylated dimers. Consequently, the unphosphorylated dimer can be disrupted by mutation without affecting the formation of the active dimer and vice-versa (Meyer *et al.*, 2004). The experiments presented in the following were designed around this principle to investigate if unphosphorylated dimers influence the formation of active STAT1 and STAT3 homodimers. Briefly, STATs carrying mutations that are known to disrupt latent dimers will be assessed for their effect on active dimer formation using the Nuclear Translocation Assay. Several mutations are known to disrupt the main latent dimer interaction surfaces in the NTD, CCD and DBD, for example, STAT3-L78R, STAT1-F172W and STAT1-Q340W, respectively (Mao *et al.*, 2005; Mertens *et al.*, 2006; Domoszlai *et al.*, 2014). The experiments described here will utilise the former two.

To answer the question at hand, four scenarios have to be considered and investigated using the Nuclear Translocation Assay. If STAT1 and STAT3 homodimers are pre-formed as unphosphorylated dimers and merely change conformation upon activation at the cytokine receptor, then the outcome of these scenarios would be as follows.

In the first scenario, the baseline uninhibited cytokine-induced nuclear accumulation of both interaction partners is assessed. For this, a STATwt-GFP is co-expressed with a STATwt-mCherry. The expected outcome for this would be a high nuclear score for both proteins since none of them carries the inhibitory ΔNLS mutation (figure 36A). This scenario was applied to STAT1, STAT3 and STAT5B homodimers in chapter 4.2.

In the second scenario, the STAT-GFP was combined with the STATΔNLS-mCherry. If the two proteins can interact, a drop in nuclear score would be expected for both proteins (figure 36B). This is because the ΔNLS mutation is dominant negative, meaning that the STATΔNLS-mCherry cross-inhibits the cytokine-induced nuclear accumulation of the STAT-GFP when they are contained in the same active STAT dimer. This outcome was confirmed for STAT1, STAT3 and STAT5B homodimers in chapter 4.2.

In the third scenario, a mutation that is known to split unphosphorylated STAT dimers is introduced into either the STATΔNLS-mCherry or STAT-GFP construct. In the following, the theory behind this scenario will be explained based on an example where the unphosphorylated dimer splitting mutation is introduced into the STAT-GFP construct (termed monomerised or mSTAT-GFP from now on). The theory behind introducing this mutation into the STATΔNLS-mCherry construct is the same. Should the STAT dimer assessed in this study be pre-formed before activation then the STATΔNLS-mCherry would assemble unphosphorylated STATΔNLS-mCherry/STATΔNLS-mCherry dimers (figure 36C). Since the STAT-GFP carries a mutation that disrupts unphosphorylated dimers, STATΔNLS-mCherry/STAT-GFP dimers cannot form in untreated cells. Instead, the STAT-GFP is monomeric before cytokine stimulation (figure 36C). When cytokine is added, the STATΔNLS-mCherry/STATΔNLS-mCherry dimers change their conformation to the active complex. Yet, due to the ΔNLS mutation, these dimers cannot accumulate in the nucleus (figure 36C). Since the mutation introduced into the mSTAT-GFP only affects the formation of unphosphorylated but not active dimers, it can still dimerise upon activation at the cytokine receptor. However, because it is already bound in pre-formed dimers the STATΔNLS-mCherry is not available as an interaction partner. Therefore, the mSTAT-GFP must self-assemble into active mSTAT-GFP/mSTAT-GFP dimers. These dimers would then accumulate inside the nucleus because they do not carry a copy of the

ΔNLS mutation. As a result, a high mSTAT-GFP nuclear score would be expected (figure 36C). The same outcome would be expected when a mSTATΔNLS-mCherry and STAT-GFP would be combined and the STAT dimer in question is pre-formed before cytokine treatment. In this case, it would be the STAT-GFP that pre-forms STAT-GFP/STAT-GFP dimers before activation. These dimers would then change their conformation to their active form and could accumulate inside the nucleus. The mSTATΔNLS-mCherry would only assemble dimers after cytokine stimulation and must interact with another mSTATΔNLS monomer because all STAT-GFP molecules are already bound in dimers.

In the fourth scenario, both mSTAT-GFP and mSTATΔNLS-mCherry are combined. Therefore, all STATs would be monomeric before their activation at the cytokine receptor. Again, because latent dimer disrupting mutations do not affect the interaction surface of active dimers, the STATs would still be expected to assemble active dimers following cytokine stimulation. Furthermore, since none of the dimers would be pre-associated, the assembly of active dimers would be random. Though, because the mSTATΔNLS-mCherry is present in molar excess, the mSTAT-GFP would mostly bind to this protein instead of undergoing self-assembly. Consequently, the nuclear score of both proteins is expected to drop to similar levels as observed in scenario two (figure 36D). Should STAT1 and STAT3 assemble into active homodimers independent of their latent counterparts, then low STAT-GFP nuclear scores would be expected for all scenarios. Specifically considering scenarios two and three, the pre-associated latent dimers would split upon cytokine stimulation, briefly making most of the STATs contained within the treated cells monomeric. This situation would represent that found in scenario four where active dimers can assemble from monomeric STATs at random. Again, in this case, the molar ratios would dictate that the STAT-GFP predominantly assembles dimers with the STATΔNLS-mCherry causing a drop in GFP nuclear score.The implementation of the experimental designs described above is documented in the following. The theories will be tested on both STAT1 and STAT3 homodimers in separate studies. To ensure that the introduction of latent dimer-disrupting mutations does not cause any severe impairment of normal STAT behaviours, tests to characterise the unphosphorylated dimer-disrupting mutations preceded studies based on the four scenarios described here.

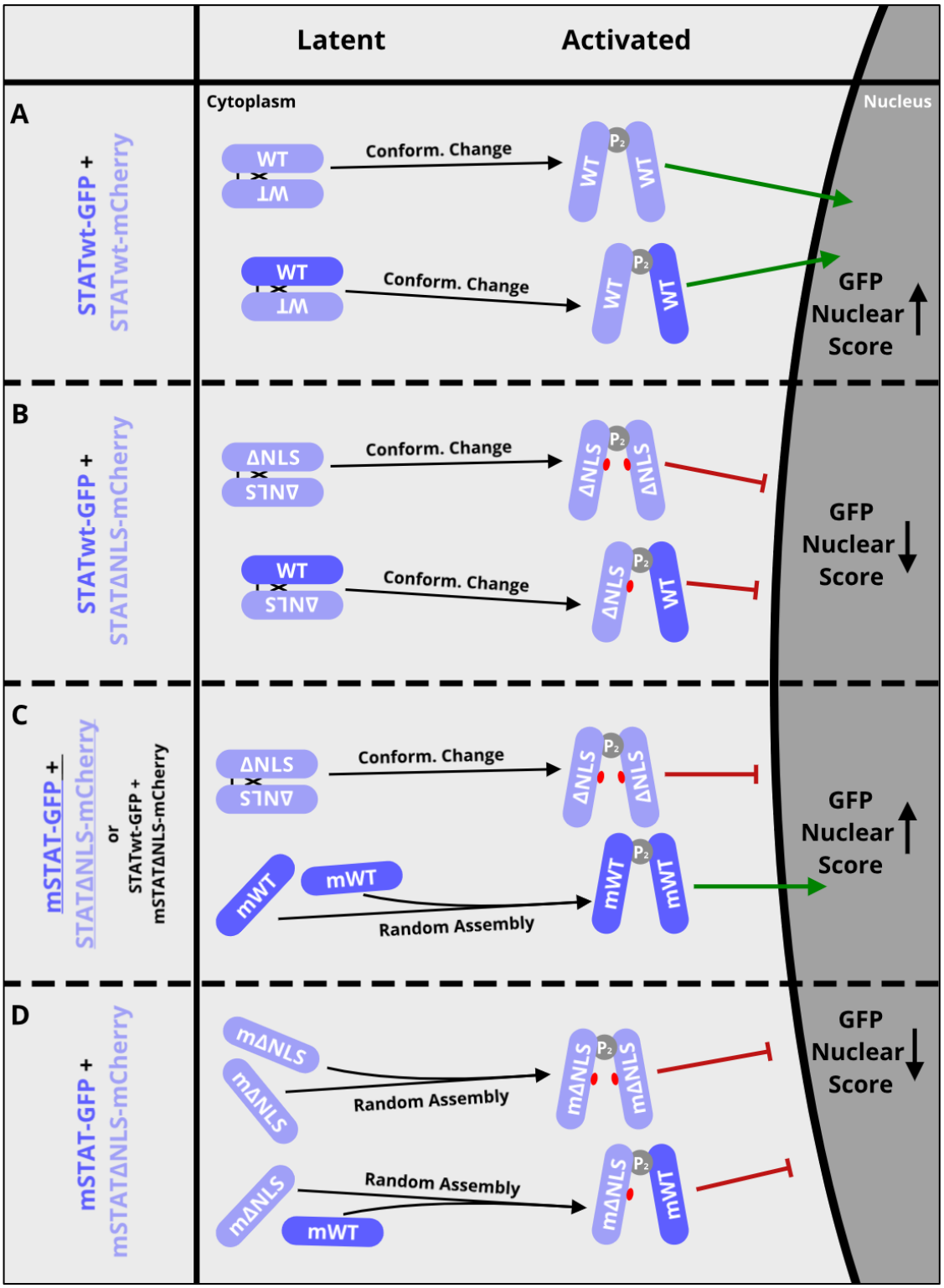


Figure 35 | Four scenarios were considered to investigate the transition of STAT dimers from latent to active: *In the four scenarios different GFP and mCherry-tagged STATs were combined. Some of these carry mutations that disrupt the latent dimer form. These are indicated by an "m" (for monomerised) in front of their names. The cartoons depict the mechanisms expected if active STAT dimer formation depends on the pre-association as unphosphorylated dimers.*

4.3.2 Characterising point mutations of STAT1 residue F172

The importance of the STAT1 residue F172 for the integrity of unphosphorylated STAT1 dimers was recognised for the first time during crystallographic studies. Indeed, the residue situated in the CCD is buried in a hydrophobic binding pocket formed by six residues located in the DBD of its interaction partner (Mao *et al.*, 2005). Furthermore, it was shown that the point mutation F172W disrupts latent dimer formation *in vitro*, an observation that was also confirmed in other studies (Mao *et al.*, 2005; Zhong *et al.*, 2005). Interestingly, in 2013 a clinical study linked the STAT1 F172L mutation to the genetic disorder STAT1 gain-of-function (GOF) disease in a patient carrying a heterozygous copy of this mutation (Sampaio *et al.*, 2013). People suffering from this disease show a broad range of clinical manifestations including chronic mucocutaneous candidiasis (CMC), increased susceptibility to bacterial and viral infection, autoimmune disease, and other comorbidities (Liu *et al.*, 2011; Toubiana *et al.*, 2016; Okada *et al.*, 2020). The molecular mechanism behind this disease is poorly understood, yet it has been recognised that most of the mutations causing STAT1 GOF disease are located at the STAT1 unphosphorylated dimer interaction interface suggesting that this conformation is an important regulator of STAT1 activity (Kagawa *et al.*, 2017a; Okada *et al.*, 2020). Overall, these findings make mutations of STAT1 F172 ideal candidates to characterise the formation of active STAT1 homodimers. In the following, the STAT1 mutations F172L and F172W will be characterised further.

STAT1 F172 mutants show normal nuclear accumulation in response to cytokines.

The ability of STAT1 F172W to accumulate inside the cell nucleus in response to IFN-γ treatment has been confirmed qualitatively before (Staab, Herrmann-Lingen and Meyer, 2013). To accompany this data with quantitative evidence for STAT1 F172W and F172L, HeLa cells were co-transfected with STAT1-F172W-GFP and STAT1-F172W-mCherry or STAT1-F172L-GFP and STAT1-F172L-mCherry. The cells were then imaged live before and 90 min after treatment with IFN-γ, IL-6 and sIL6Rα for analysis in the Nuclear Translocation Assay.

Before cytokine exposure, both the STAT1-F172L-GFP and the STAT1-F172L-mCherry resided predominantly in the cytoplasm. Yet, some fluorescence was still detectable inside the cell nuclei (figure 37A, row 1). The subcellular distribution of both molecules

was vastly different in the images taken of the same cells 90 min after cytokine treatment. Indeed, both proteins showed an almost complete nuclear accumulation (figure 37A, row 2). The subcellular distributions of STAT1-F172W-GFP and STAT1-F172W-mCherry were similar to those observed for the STAT1-F172L constructs (figure 37A, row 3). It is noteworthy that this also resembled the localisation of STAT1wt before activation (figure 24A). Again, both STAT1-F172W constructs showed an almost complete nuclear accumulation in response to the cytokine treatment (figure 37A, row 4). These visual observations suggested that both STAT1-F172L and STAT1-F172W were capable of normal cytokine-induced nuclear accumulation. To further assess this, nuclear scores were calculated for 31 cells containing the STAT1-F172L combination from one experiment and 67 cells containing the STAT1-F172W combination from two individual experiments. For the former, a GFP nuclear score of 3.5±0.2 and a mCherry nuclear score of 4.0±0.2 was found (figure 37B; table 33). For the latter, the nuclear score for the GFP channel was 4.0±0.1 and for the mCherry channel 4.1±0.2 (figure 37B; table 33).

These numbers indicate the uninhibited cytokine-induced nuclear accumulation for all constructs tested. Indeed, there was no significant difference between the nuclear scores of the mCherry-tagged F172L and F172W mutants (figure 37B). On the contrary, a small and significant difference was found between the GFP-tagged STAT1-F172L and STAT1-F172W with the former's nuclear score being lower than the scores of all other proteins investigated in this experiment (figure 37B). Nevertheless, the nuclear score was still high enough to confirm its uninhibited cytokine-induced nuclear accumulation. Indeed, all four constructs showed nuclear scores that were in the same range as those observed for GFP- and mCherry-tagged STAT1wt (figure 24B). Thus, in summary, it can be concluded that STAT1 cytokine-induced nuclear accumulation remains intact when residue F172 is mutated to L or W. Furthermore, in terms of this phenotype, both mutations seem to behave in the same manner as evidenced by very similar nuclear scores (table 33). Following the slightly more consistent behaviour of the STAT1-F172W mutant, the decision was made to progress it into further experiments. On the other hand, a STAT1ΔNLS-F172L variant was used as a latent dimer monomerised version of STAT1ΔNLS, mainly due to availability constraints.

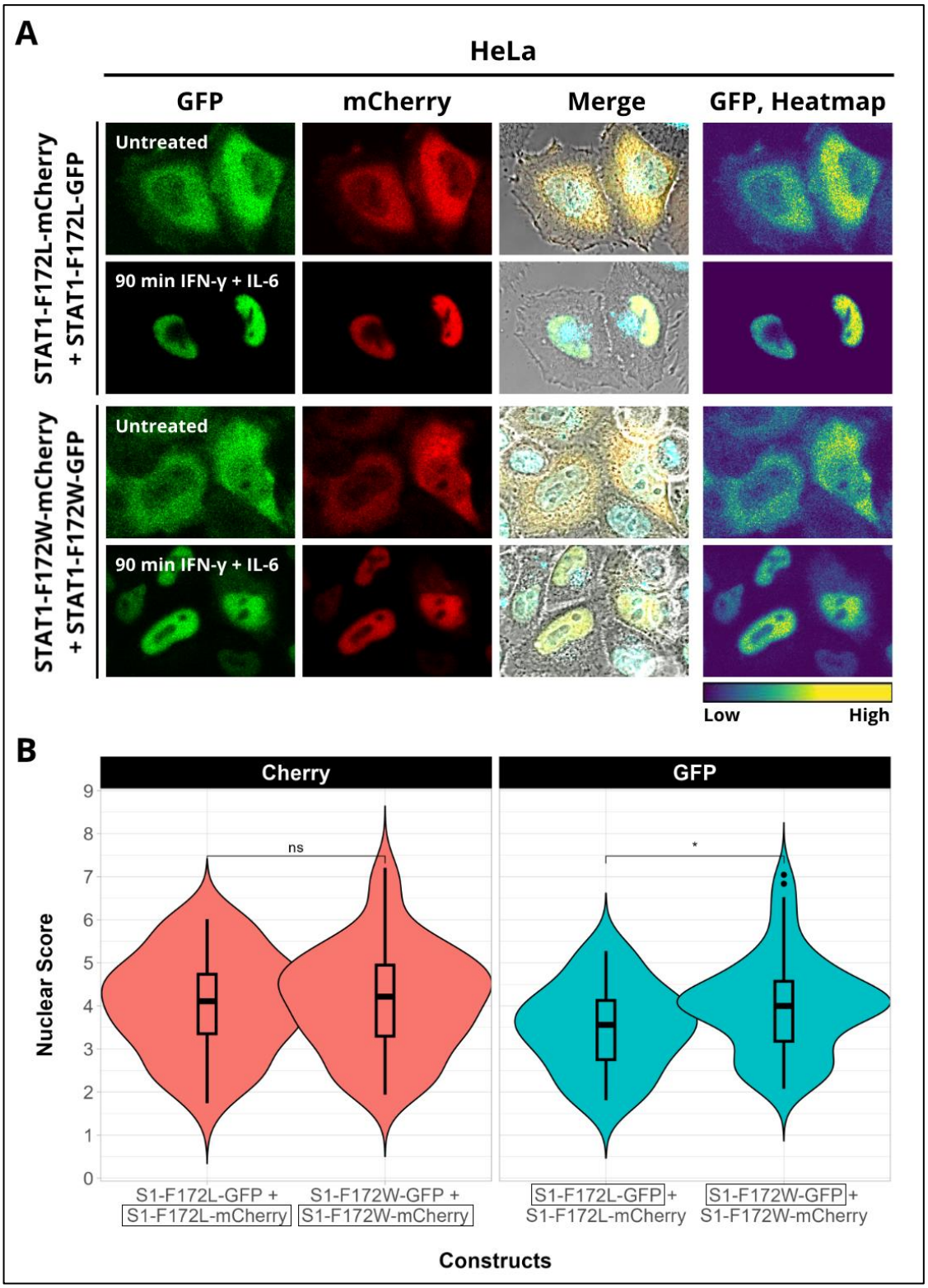


Figure 36 | STAT1-F172L and STAT1-F172W both show strong cytokine-induced nuclear accumulation: ***[A]*** *Fluorescence micrographs of HeLa cells co-expressing either STAT1-F172L-mCherry and STAT1-F172L-GFP or STAT1-F172W-mCherry and STAT1-F172W-GFP imaged before and 90 min after cytokine treatment. The last column of panels shows a contrast enhanced depiction of the GFP channel. In this, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).* ***[B]*** *Quantification of mCherry and GFP nuclear scores in 31 or 67 cells co-expressing the combinations described in [A], respectively.*

Table 31 | ***Descriptive statistics of nuclear scores determined for cells co-expressing STAT1-F172L-GFP and STAT1-F172L-mCherry or STAT1-F172W-GFP and STAT1-F172W-mCherry.***

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT1-F172L-mCherry + STAT1-F172L-GFP	31	3.5	0.2	4.0	0.2
STAT1-F172W-mCherry + STAT1-F172W-GFP	67	4.0	0.1	4.14	0.2

Mutating STAT1 F172 leads to increased tyrosine phosphorylation.

Both STAT1-F172L and STAT1-F172W showed good nuclear accumulation in response to cytokine stimulation. This suggested that these mutations did not decrease cytokine-induced tyrosine phosphorylation. Indeed, accounts in the literature found that alterations of residue 172 cause elevated phosphorylation in response to IFN-γ treatment instead (Zhong *et al.*, 2005; Sampaio *et al.*, 2013). To confirm this and to assess the phosphorylation status of STAT1ΔNLS-F172L, which has not been investigated before, a plasmid construct coding for STAT1-F172W-GFP or STAT1ΔNLS-F172L was expressed in U3R cells, the cells stimulated with IFN-γ for 60 mins and whole cell lysates subjected to Western blot analysis. For this, the same antibodies raised against full-length STAT1 or the STAT1 pY701 peptide were used as before.

When using the anti-STAT1 antibody for detection, a single specific band was observed in both the extract containing STAT1-F172W-GFP and the one containing STAT1ΔNLS-F172L-GFP. The two bands migrated at a speed that was consistent with the expected molecular weight of GFP-tagged STAT1 variants (figure 38A, panel 1). A strong band migrating at the same height was detected in each sample when using the anti-STAT1 pY701 antibody (figure 38A, panel 2). This was indicative of efficient tyrosine phosphorylation in response to IFN-γ treatment. The proportion of phosphorylated protein was calculated for each construct by dividing the fluorescence signal measured for anti-STAT1 pY701 bands by the signals detected for the corresponding anti-STAT1 bands. This was repeated for three independent experiments (figure 38B). As expected from visual observations, the values calculated for STAT1-F172W-GFP and STAT1ΔNLS-F172L-GFP confirmed that both proteins were efficiently phosphorylated in response to IFN-γ treatment. The phosphorylation of STAT1-F172W-GFP was 6.93-fold higher on

average than that of STAT1wt-GFP when expressed alone (compare figures 37B and 21C). Interestingly, the phosphorylation of STAT1ΔNLS-F172L-GFP was even further elevated at 950% of STAT1wt-GFP phosphorylation (compare figures 37B and 21C). We have seen earlier that STAT1ΔNLS-GFP also showed higher phosphorylation than STAT1wt-GFP (755%, figure 21C). Thus, the results presented here suggested that the high-phosphorylation phenotypes of STAT1-F172L and STAT1ΔNLS were additive. Overall, these results confirmed the findings of previous studies that mutations of the STAT1 residue F172 cause elevated phosphorylation and showed that this was also the case when the mutation was introduced into a STAT1ΔNLS construct. Furthermore, these findings indicated that the STAT1-F172W-GFP and STAT1ΔNLS-F172L constructs were suitable for use in the Nuclear Translocation Assay without causing measurement artefacts from impaired phosphorylation.

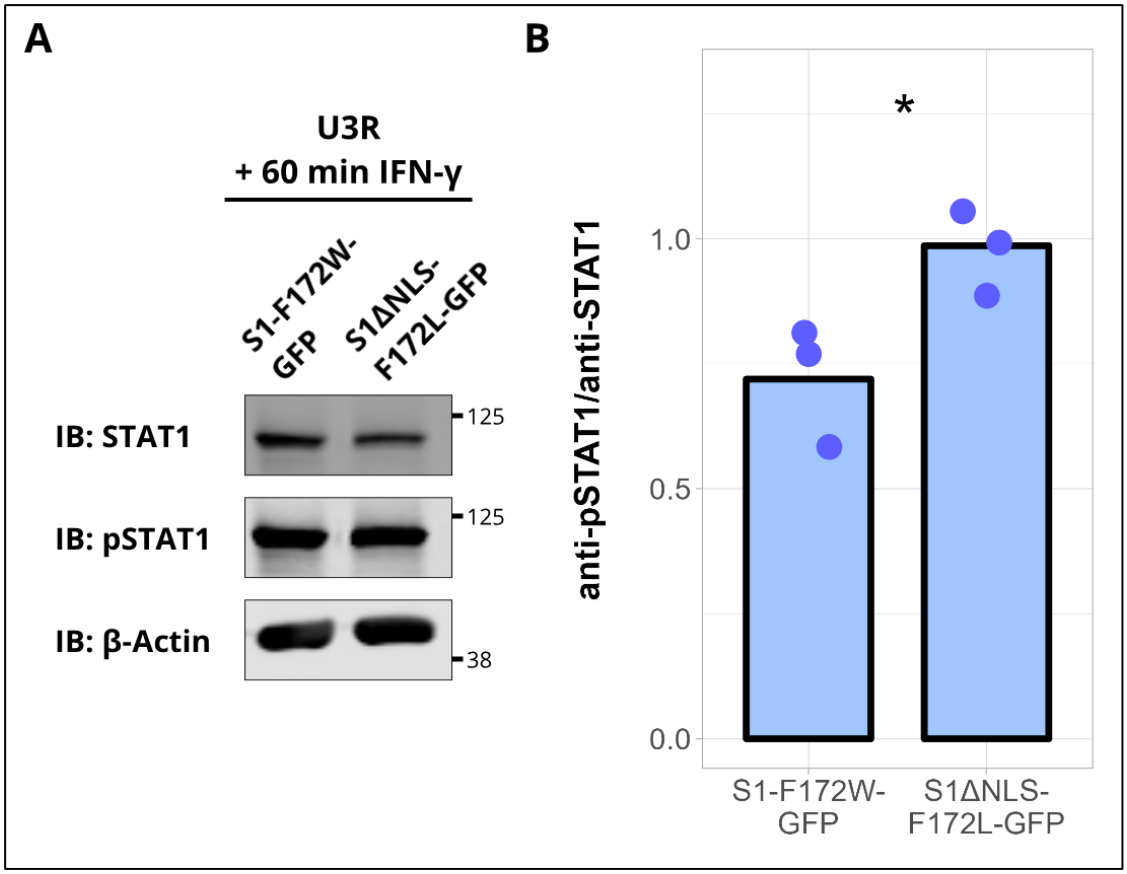


Figure 37 | Mutating residue F172 results in elevated cytokine-induced phosphorylation of STAT1-F172W-GFP and STAT1ΔNLS-F172L-GFP constructs: ***[A]*** *Whole cell extracts from IFN-γ stimulated U3R cells containing recombinant STAT1-F172L-GFP or STAT1ΔNLS-F172L-GFP were subjected to Western analysis using antibodies raised against full-length STAT1, the STAT1 pY701 peptide or ß-Actin.* ***[B]*** *Phosphorylation of the two constructs was quantified by dividing the fluorescence measured for anti-STAT1 pY701 bands by the signal measured for anti-STAT1 bands from three independent experiments. MW markers in kDa.*

4.3.3 The monomerisation of unphosphorylated dimers does not affect the formation of activated STAT1 homodimers.

This part of the study aims to assess whether STAT1 homodimers are pre-formed before activation and merely change their conformation to the active form upon cytokine stimulation without the exchange of monomers. Chapter 4.3.1 outlines the rational and experimental design necessary to answer this research question. The following paragraphs describe the application of these principles to STAT1 homodimers using the constructs characterised in the previous chapter.

Investigating STAT1 latent dimer splitting mutations in the Nuclear Translocation Assay

Nuclear scores were determined using the Nuclear Translocation Assay for the following combinations: STAT1-F172W-GFP + STAT1ΔNLS-mCherry, STAT1wt-GFP + STAT1ΔNLS-F172L-mCherry and STAT1-F172W-GFP + STAT1ΔNLS-F172L-mCherry. For this, HeLa cells co-expressing the three combinations were imaged before and 90 min after the addition of IFN-γ, IL-6 and sIL6Rα according to the Nuclear Translocation Assay protocol (chapter 4.1.6). All the constructs showed similar subcellular distributions in images taken before cytokine stimulation (figure 39, rows 1, 3 and 5). The predominantly cytoplasmic localisation with weak fluorescence signals detected in the nucleus also resembled that observed for the STAT1wt-GFP + STAT1wt-mCherry combination before activation (figure 24A). This suggested that the mechanism responsible for this distribution was intact in mutated STAT1 constructs. In images taken 90 min after the addition of cytokine, the mCherry-tagged STAT1ΔNLS showed no signs of nuclear accumulation or any other form of redistribution (figure 39, panel 6) even when it carried the F172L mutation (figure 39, panel 10 + 14). While the co-expressed wild-type and mutant STAT1-GFP proteins showed some degree of re-distribution to the nucleus in images taken after cytokine treatment, nuclear accumulation seemed incomplete when compared to STAT1wt-GFP/STAT1wt-mCherry expressing cells (see figures 24A and 38). Furthermore, cross-inhibition of cytokine-induced nuclear accumulation of STAT1-F172W-GFP and STAT1wt-GFP was not affected when an F172L mutation was introduced into the STAT1ΔNLS-mCherry construct (compare figure 39, panels 12 and 16). Overall, these findings suggested that STAT1 homodimers formed independent of unphosphorylated dimer integrity.

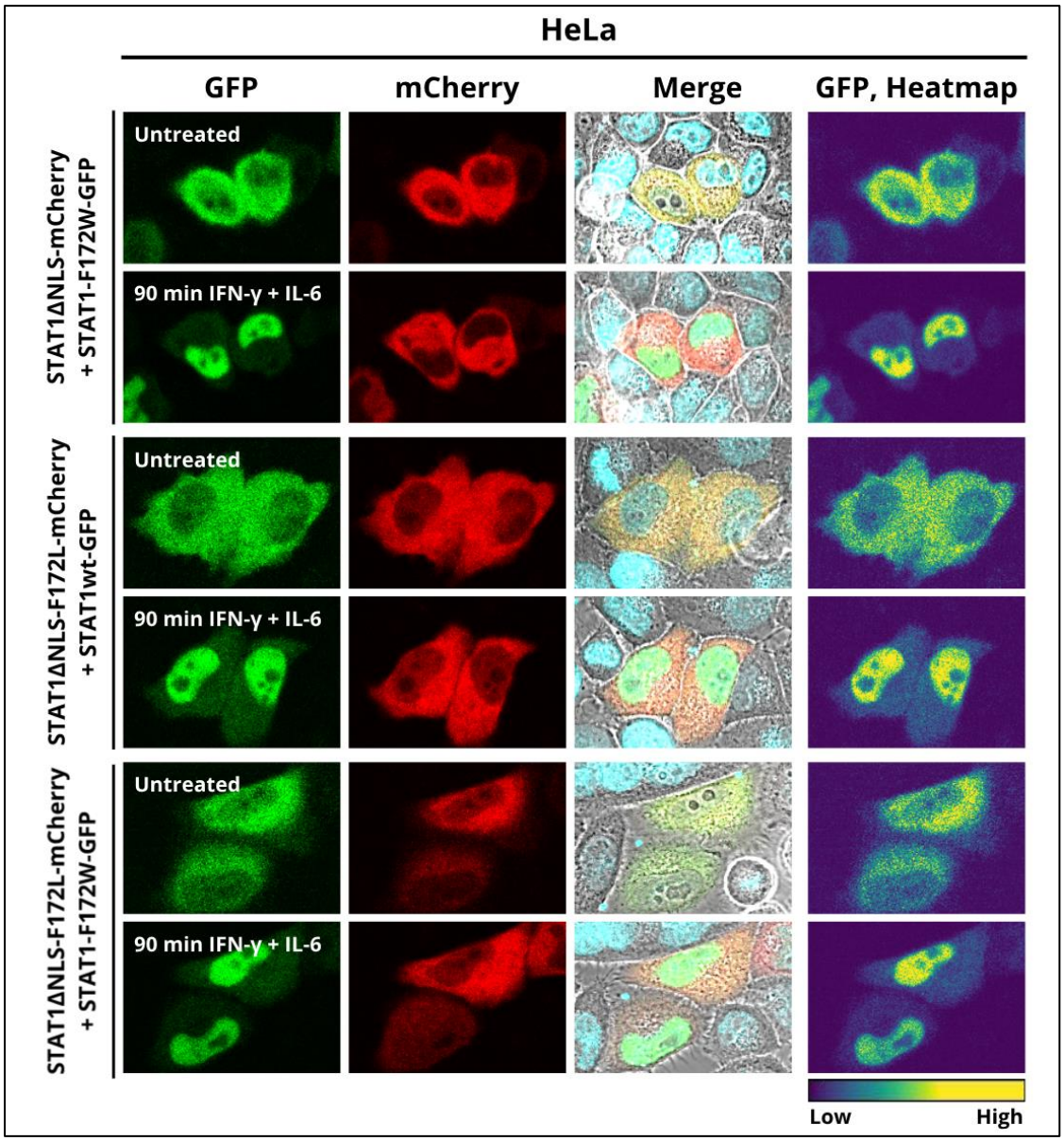


Figure 38 | STAT1ΔNLS-mCherry cross-inhibits STAT1-GFP cytokine-induced nuclear accumulation irrespective of unphosphorylated dimer integrity: *Fluorescence micrographs of HeLa cells co-expressing either STAT1ΔNLS-mCherry and STAT1-F172W-GFP, STAT1ΔNLS-F172L-mCherry and STAT1wt-GFP or STAT1ΔNLS-F172L-mCherry and STAT1-F172W-GFP imaged before and 90 min after cytokine treatment. The right column shows a contrast enhanced depiction of the GFP channel. Here, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).*

To quantify these observations, images taken from eligible cells in different experiments were subjected to nuclear score calculations according to formula IV. Eligibility was assessed by determining the relative protein amounts in individual cells and this amassed to 43 cells from 2 individual experiments for the STAT1-F172W-GFP + STAT1ΔNLS-mCherry combination, 66 cells from 5 individual experiments for the STAT1wt-GFP + STAT1ΔNLS-F172L-mCherry combination and 28 cells from 2 individual experiments for the STAT1-F172W-GFP + STAT1ΔNLS-F172L-mCherry combination

being progressed to nuclear score calculations. As expected from the visual inspection, the respective wild-type and mutant STAT1ΔNLS-mCherry constructs exerted considerable inhibition on the cytokine-induced nuclear accumulation of STAT1wt-GFP and STAT1-F172W-GFP when compared to cells expressing STAT1wt-GFP or STAT1-F172W-GFP in the absence of a STATΔNLS (GFP nuclear scores of 4.5±0.1 and 4.1±0.2, respectively; table 34; figure 40). Inhibition was the lowest for the STAT1-F172W-GFP + STAT1ΔNLS-mCherry combination with a GFP nuclear score of 2.7±0.1 and the highest for the STAT1wt-GFP + STAT1ΔNLS-F172L-mCherry combination with a GFP nuclear score of 1.8±0.1 (table 34; figure 40). The inhibition exerted on the GFP-tagged STAT1-F172W by STAT1ΔNLS-F172L-mCherry ranged somewhere in the middle compared to the above combinations (GFP nuclear score: 2.3±0.1; table 34; figure 40). No significant difference between mCherry nuclear scores of mutated and non-mutated STAT1ΔNLS-mCherry constructs was found (figure 40). Furthermore, the low absolute values for their nuclear scores were just slightly higher than 1, indicating that the cytokine treatment did not introduce any redistribution of these proteins (figure 40, table 34).

Table 32 | Descriptive statistics of nuclear scores determined for cells co-expressing STAT1ΔNLS-mCherry and STAT1-F172W-GFP, STAT1ΔNLS-F172L-mCherry and STAT1wt-GFP or STAT1ΔNLS-F172L-mCherry and STAT1-F172L-GFP. *Descriptive statistics of cells co-expressing STAT1wt-mCherry and STAT1wt-GFP, STAT1ΔNLS-mCherry and STAT1wt-GFP or STAT1-F172W-mCherry and STAT1-F172W-GFP are also shown in table 23 and 33 but were added again here for comparison.*

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT1wt-mCherry + STAT1wt-GFP	79	4.2	0.1	4.5	0.1
STAT1-F172W-mCherry + STAT1-F172W-GFP	67	4.0	0.1	4.1	0.2
STAT1ΔNLS -mCherry + STAT1wt-GFP	106	1.1	0.02	2.4	0.1
STAT1ΔNLS- mCherry + STAT1-F172W-GFP	43	1.2	0.03	2.7	0.1
STAT1ΔNLS-F172L-mCherry + STAT1wt-GFP	66	1.2	0.02	1.8	0.1
STAT1ΔNLS-F172L-mCherry + STAT1-F172W-GFP	28	1.2	0.02	2.3	0.1

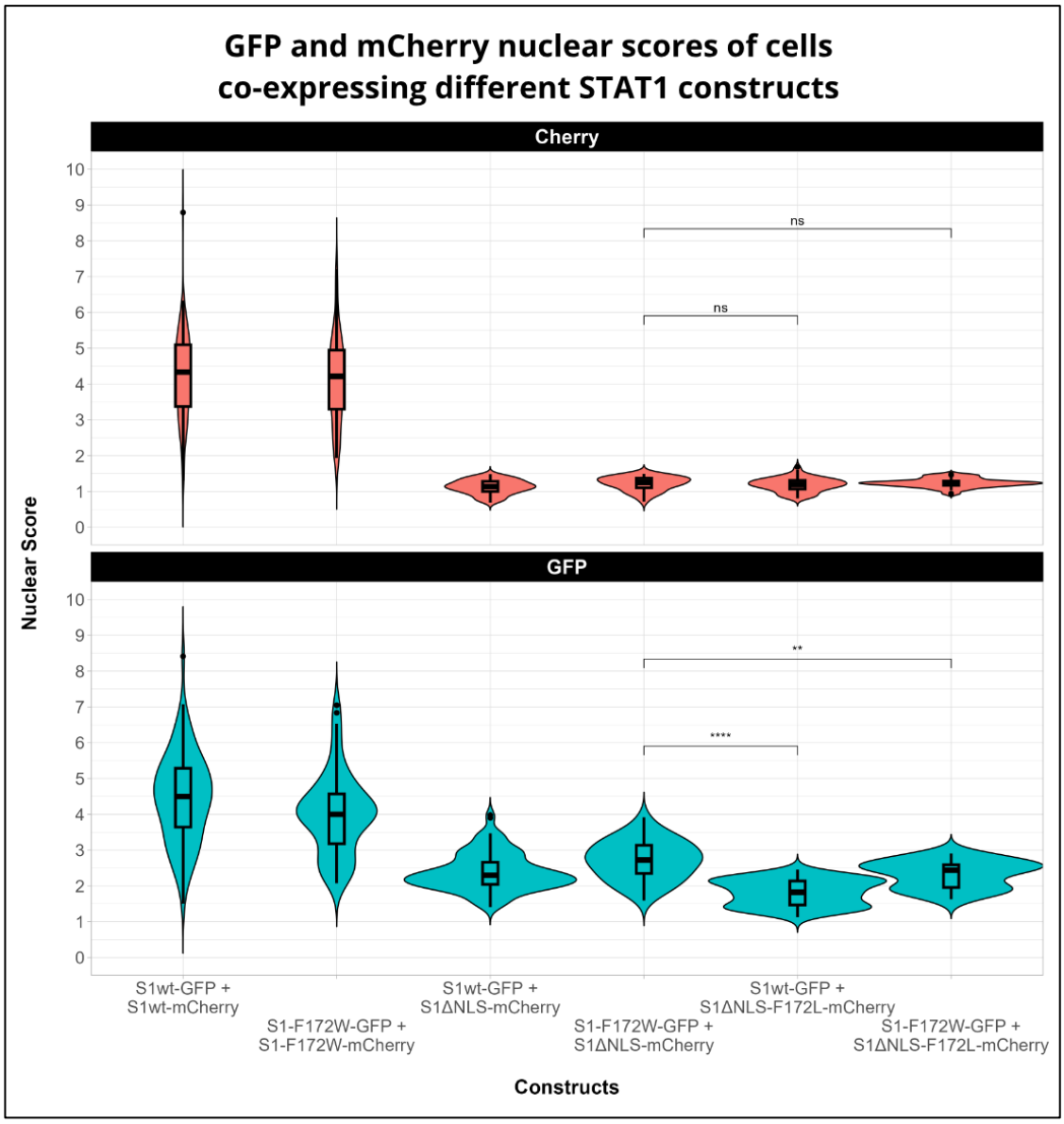


Figure 39 | Quantification of GFP and mCherry nuclear scores of cells expressing STAT1ΔNLS-mCherry and STAT1-F172W-GFP, STAT1ΔNLS-F172L-mCherry and STAT1wt-GFP or STAT1ΔNLS-F172L-mCherry and STAT1-F172W-GFP: *Nuclear scores were quantified in 43 cells expressing the former combination, 66 cells expressing the STAT1ΔNLS-F172L-mCherry and STAT1wt-GFP and 28 cells expressing the latter combination including the cells shown in figure 38. The first three columns represent data that was already presented and discussed in earlier chapters (see figures 24 and 36) but was included here for comparison.*

As explained in chapter 4.3.1, if STAT1 homodimers were pre-formed before cytokine exposure and did not exchange subunits upon activation, a specific inhibition pattern should have emerged in the above experiments. That is, the inhibition of STAT1-GFP nuclear accumulation would have been alleviated when one of the two interaction partners carries a mutation in residue F172 (figure 36). However, the results presented in the previous paragraph showed inhibition of STAT1-GFP nuclear accumulation in all three combinations irrespective if one or both interaction partners carried the

unphosphorylated dimer splitting mutation (figure 40). This suggested that active dimers could form between different STAT1ΔNLS-mCherry and STAT1-GFP in every case, thus pointing towards a mechanism where STAT1 homodimer composition is not pre-determined before their activation at the cytokine receptor.

Decreased tyrosine phosphorylation is not the cause for inhibitory effects observed in STAT1 Nuclear Translocation Assay experiments.

The inhibition of cytokine-induced nuclear accumulation of STAT1wt-GFP and STAT1-F172W-GFP by STAT1ΔNLS-mCherry and STAT1ΔNLS-F172L-mCherry could be caused by decreased phosphorylation at the cytokine receptor. To rule this out, the phosphorylation status of co-expressed STAT1-F172W and STAT1ΔNLS-F172L was assessed via Western blot analysis. As before, an untagged STAT1-F172W construct (pcDNA 3.1 zeo (-) + STAT1 F172W) was co-expressed with the STAT1ΔNLS-F172L-mCherry in U3R cells to be able to distinguish between the two proteins by molecular weight difference in SDS-PAGE. The cells were stimulated with IFN-γ for 60 min, harvested and lysed to generate whole cell extracts. Western blotting was performed on extracts from three independent experiments by incubation with anti-STAT1 and anti-STAT1 pY701 antibodies used in other experiments presented in this work before.

Two specific bands were observed on membranes loaded with extracts from cells co-expressing untagged STAT1-F172W and STAT1ΔNLS-F172L-GFP when detecting with the anti-STAT1 antibody. The faster migrating band was identified as the former and the slower migrating one as the latter based on expected molecular weights for untagged and GFP-tagged STAT1 variants (figure 41A, panel 1). Two bands were also observed when detecting with an anti-STAT1 pY701 antibody. These bands migrated at the same speed as the bands detected by the anti-STAT1 antibody, suggesting that both constructs were phosphorylated efficiently (figure 41A, panel 2).

The proportion of phosphorylated protein was quantified from the three experiments by dividing the fluorescence measured for anti-STAT1 pY701 bands by the fluorescence measured for anti-STAT1 bands. Both constructs showed high phosphorylation levels on average (figure 41B; table 35). More specifically, when compared to cells co-expressing untagged STAT1wt and STAT1ΔNLS-GFP (figure 25), the phosphorylation of untagged STAT1-F172W was higher than that of the former which already showed

elevated phosphorylation than when it was co-expressed with STAT1wt-GFP. The phosphorylation of STAT1ΔNLS-GFP and STAT1ΔNLS-F172L-GFP was similar (p=0.84). However, the phosphorylation of untagged STAT1-F172W and STAT1ΔNLS-F172L-GFP was lower when they were co-expressed than when they were expressed alone (figure 38). Still, the data presented here indicated that both constructs were sufficiently phosphorylated when co-expressed, ruling out diminished phosphorylation as the cause for the reduced nuclear score of STAT1-F172W-GFP and STAT1wt-GFP observed in Nuclear Translocation Assay experiments described above (figure 40).

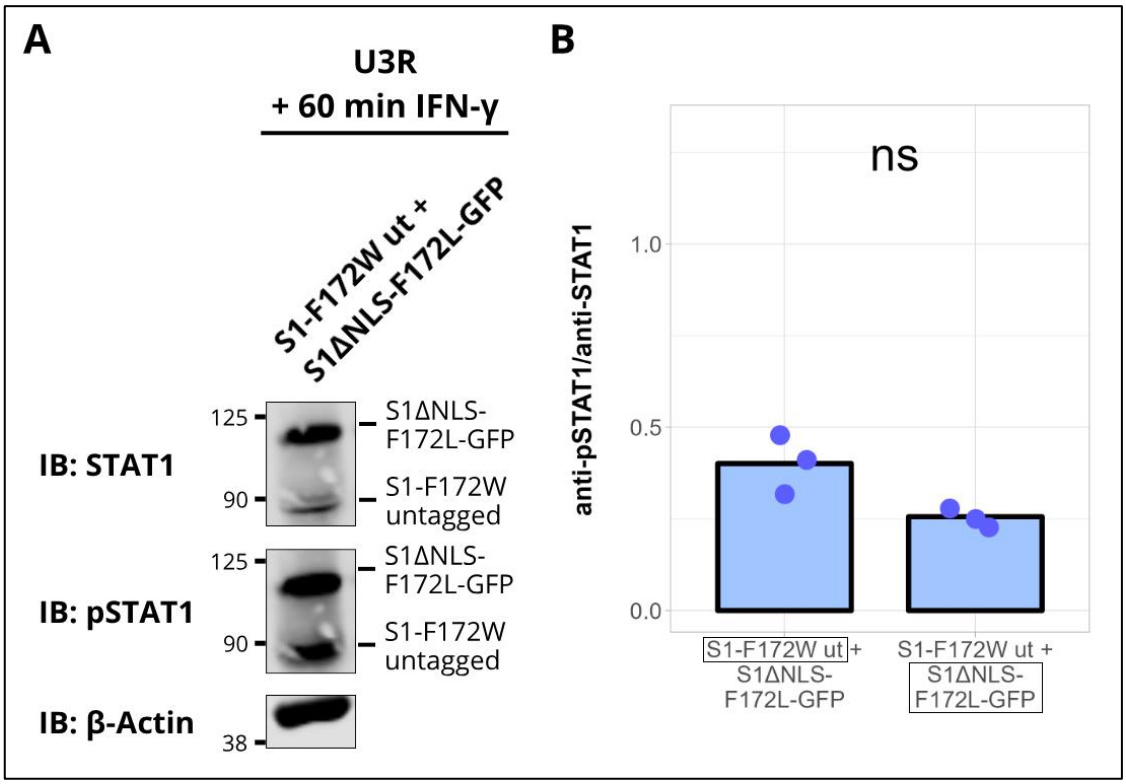


Figure 40 | Co-expression of STAT1-F172W and STAT1ΔNLS-F172L does not decrease cytokine-induced tyrosine phosphorylation: ***[A]*** *Whole cell extracts from IFN-γ stimulated U3R cells co-expressing recombinant untagged STAT1-F172W and STAT1ΔNLS-F172L-GFP were subjected to Western analysis using antibodies raised against full-length STAT1, the STAT1 pY701 peptide or β-Actin.* ***[B]*** *Phosphorylation of the two constructs was quantified by dividing the fluorescence measured for anti-STAT1 pY701 bands by the signal measured for anti-STAT1 bands from three independent experiments. MW markers in kDa.*

Table 33 | ***Quantification of STAT1 phosphorylation by Western blotting:*** *Whole cell extracts from U3R co-expressing untagged STAT1-F172W with STAT1ΔNLS-F172L-GFP were analysed for STAT1 phosphorylation.*

Combination	Construct	Mean ratio $\frac{\text{α-STAT1 pY701}}{\text{α-STAT1 FL}}$
Untagged STAT1-F172W +	Untagged STAT1-F172W	0.40 ± 0.08
STAT1ΔNLS-F172L-GFP	STAT1ΔNLS-F172L-GFP	0.26 ± 0.03

Taken together, the results presented in this chapter showed that the assembly of active homodimers is independent of unphosphorylated dimer integrity. The Nuclear Translocation Assay experiments showed cross-inhibition exerted by different STAT1ΔNLS construct on STAT1-GFP cytokine-induced nuclear accumulation, irrespective if one or both of the interaction partners were mutated in residue F172. Furthermore, decreased phosphorylation was ruled out as the cause for the observed inhibition, confirming that this effect stemmed from the dominant negative effect of the ΔNLS mutation. Based on the results described in this chapter a mechanism could be proposed where unphosphorylated STAT1 dimers must split during or after activation at the cytokine receptor before reforming as the active signalling complex.

4.3.4 Characterising STAT3-L78R

The binding surface that facilitates reciprocal hydrophobic interactions between the CCD and DBD of latent STAT1 dimers plays a lesser role in the formation of latent STAT3 dimers (Ren *et al.*, 2008). In contrast, the NTD interaction surface involved in the formation of these dimers is well conserved amongst STAT1 and STAT3 (Vinkemeier *et al.*, 1998). Indeed, a homologous mutation to STAT1 L78A, that was shown to disrupt STAT1 latent dimers before (Chen, 2003; Mao *et al.*, 2005), had the same effect on latent STAT3 dimers (Domoszlai *et al.*, 2014). Furthermore, this mutation (STAT3-L78R) was found to be clinically relevant since it was associated with inflammatory hepatocellular adenomas (Pilati *et al.*, 2011). Together, these properties make STAT3-L78R an ideal candidate to assess the mechanism by which STAT3 forms active homodimers.

STAT3L78R cytokine-induced nuclear accumulation resembles that of STAT3wt.

Previous studies have already shown that STAT3-L78R accumulates inside the nucleus after stimulation with recombinant IL-6 and sIL6Rα (Domoszlai *et al.*, 2014). To quantify the degree of this, HeLa cells co-expressing STAT3-L78R-GFP and STAT3-L78R-mCherry were subjected to live-cell imaging under the Nuclear Translocation Assay protocol (chapter 4.1.6). Images taken before the addition of cytokine revealed a pancellular distribution of fluorescent protein in both the GFP and mCherry channels (figure 42A, row 1). Interestingly, the distribution observed for the two STAT3-L78R constructs here differed from that of STAT3wt constructs investigated in earlier

experiments (figure 27). The fluorescence in cells containing the L78R mutant appeared equally strong in both the nucleus and the cytoplasm, while the fluorescence in STAT3wt-containing cells was more pronounced in the cytoplasm (figure 27A). However, in pictures taken 90 min after the addition of IFN-γ, IL-6 and sIL6Rα, the STAT3-L78R-GFP and STAT3-L78R-mCherry fusion proteins almost completely re-located to cell nuclei, suggesting that these constructs behave similarly to STAT3wt in terms of cytokine-induced nuclear accumulation. Importantly, paracrystal formation was almost completely absent in stimulated cells expressing the STAT3-L78R mutants (figure 42A, row 2). Previous studies have shown that NTD-NTD interactions are essential for the paracrystal formation of SUMO-free STAT1 (Droescher *et al.*, 2011). A similar mechanism is likely at play for STAT3 which would explain the absence of paracrystals in cytokine-stimulated cells expressing the NTD-NTD dimer disrupting STAT3-L78R mutants.

To quantify the degree of STAT3-L78R cytokine-induced nuclear accumulation, nuclear scores of 63 cells imaged from three independent experiments were calculated as outlined by the Nuclear Translocation Assay protocol (chapter 4.1.6). While STAT3-L78R mCherry and GFP nuclear scores were both reasonably high (figure 42B; table 35), they were considerably lower than the ones calculated from cells expressing a STAT3wt-GFP/STAT3wt-mCherry combination (figure 27; table 25). This is likely explained by the observation made earlier that STAT3-L78R constructs show a more pancellular distribution than STAT3wt before its activation at the cytokine receptor. A more pancellular distribution before cytokine exposure will inherently limit the fold increase of nuclear fluorescence upon cytokine-induced nuclear accumulation.

Table 34 | Descriptive statistics of nuclear scores determined for cells co-expressing STAT3-L78R-mCherry and STAT3-L78R-GFP.

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT3-L78R-mCherry + STAT3-L78R-GFP	63	2.48	0.09	2.08	0.06

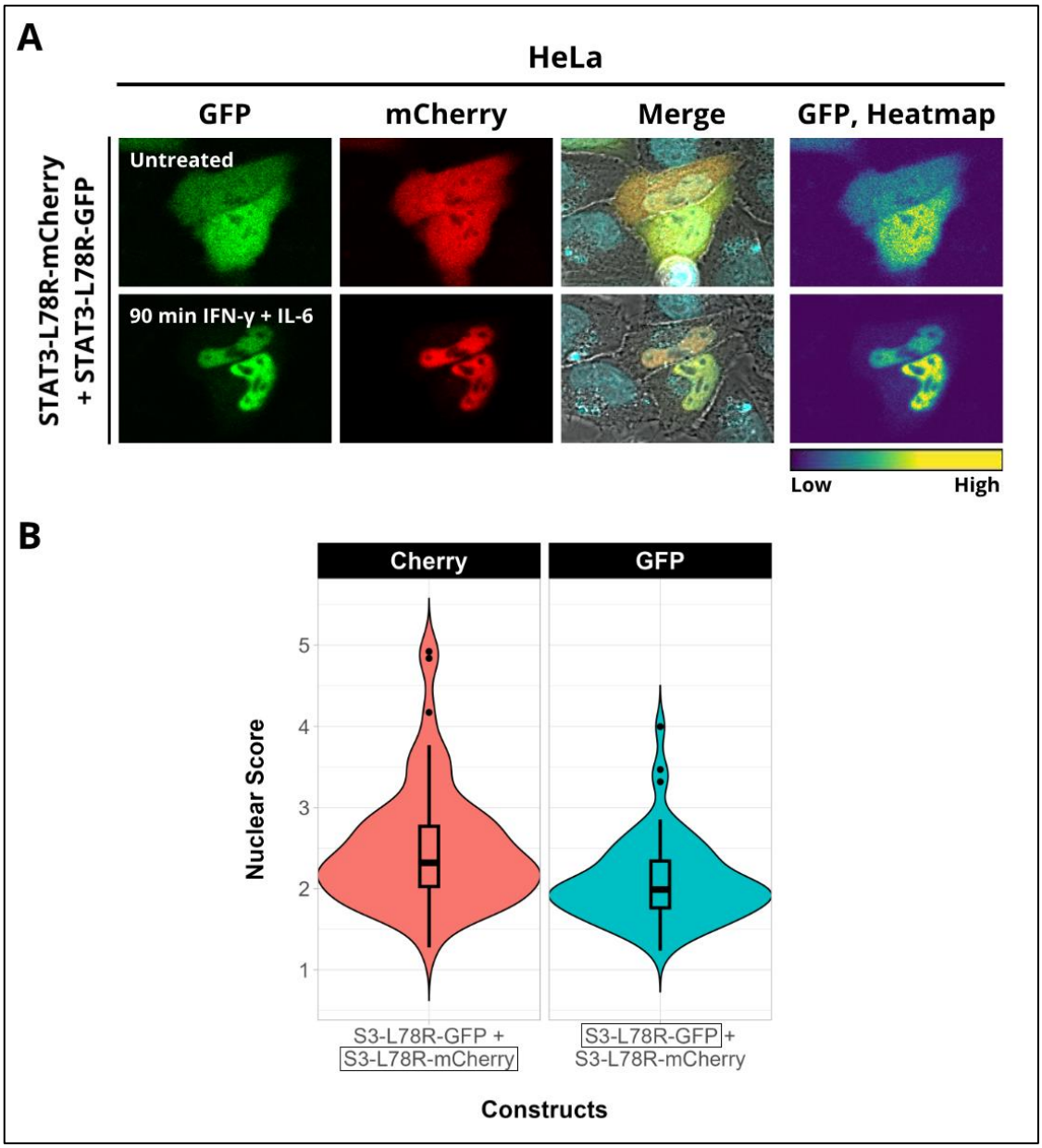


Figure 41 | STAT3-L78R shows strong cytokine-induced nuclear accumulation: ***[A]*** *Fluorescence micrographs of HeLa cells co-expressing STAT3-L78R-mCherry and STAT3-L78R-GFP imaged before and 90 min after cytokine treatment. The last column of panels shows a contrast enhanced depiction of the GFP channel. In this, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).* ***[B]*** *Quantification of mCherry and GFP nuclear scores in 63 cells co-expressing the combinations described in [A], respectively.*

STAT3L78R shows elevated phosphorylation in response to cytokine.

Another explanation for the reduced nuclear scores measured for STAT3-L78R constructs above could be decreased tyrosine phosphorylation at the cytokine receptor. To investigate this, whole cell extracts from IL-6 stimulated U3R cells expressing STAT3-L78R-mCherry or STAT3ΔNLS-L78R-mCherry were subjected to Western blot analysis, detecting both full-length STAT3 and the STAT3 pY705 phosphopeptide with specific antibodies. A single band was observed in both extracts when detecting with the former antibody (figure 43A, row 1). Detection with the latter antibody also revealed a single band in each extract migrating at the same height as the bands detected with the anti-STAT3 antibody (figure 43A, row 2). Since the speed at which these bands migrated was consistent with the theoretical molecular weights of STAT3-mCherry, it was concluded that both STAT3-L78R-mCherry and STAT3ΔNLS-L78R-mCherry were efficiently phosphorylated after 60 min of IL-6 treatment. To confirm this suspicion, anti-STAT3 and anti-STAT3 pY705 bands measured in three independent U3R cell extracts containing either of the two STAT3 constructs were quantified, and the proportion of phosphorylated protein scored as before. A small but non-significant difference was found between phosphorylation scores of STAT3-L78R-mCherry and STAT3ΔNLS-L78R-mCherry ($p=0.098$; figure 43B). Furthermore, a small but non-significant difference was found when these phosphorylation scores were compared to those calculated for cells expressing STAT3wt-mCherry (figure 22C). This suggested that the reduced nuclear accumulation of STAT3-L78R observed earlier was not caused by a decrease in phosphorylation at the cytokine receptor. Furthermore, the results also confirmed that the introduction of ΔNLS and L78R mutations in conjunction did not affect tyrosine phosphorylation in response to IL-6 treatment.

Taken together, the STAT3-L78R-GFP and STAT3-L78R-mCherry mutants suffered from lower nuclear scores compared to STAT3wt-GFP and STAT3-mCherry. However, it was unlikely that this was caused by a defect in cytokine-induced nuclear accumulation but instead was caused by a measurement artefact. Furthermore, reduced phosphorylation at the cytokine receptor could not explain the nuclear accumulation phenotype. Since nuclear scores of the mutant STAT3 construct were still sufficiently high to detect a cytokine-induced response, the decision was made to use this construct in future Nuclear Translocation Assay experiments regardless.

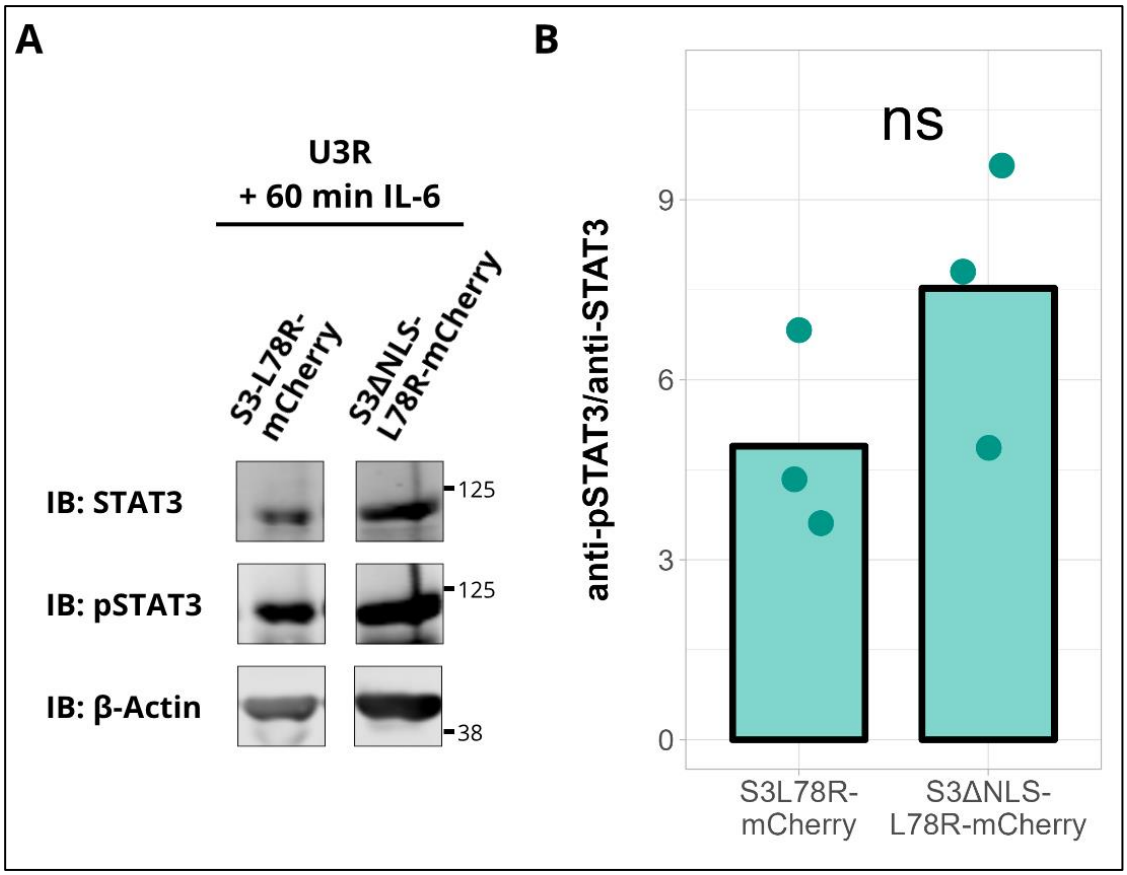


Figure 42 | STAT3-L78R-mCherry and STAT3ΔNLS-L78R-mCherry show elevated cytokine-induced phosphorylation: ***[A]*** *Whole cell extracts from IFN-γ stimulated U3R cells containing recombinant STAT3-L78R-mCherry or STAT3ΔNLS-L78R-mCherry were subjected to Western analysis using antibodies raised against full-length STAT3, the STAT3 pY705 peptide or β-Actin.* ***[B]*** *Phosphorylation of the two constructs was quantified by dividing the fluorescence measured for anti-STAT3 pY705 bands by the signal measured for anti-STAT3 bands from three independent experiments. MW markers in kDa.*

4.3.5 The monomerisation of unphosphorylated dimers does not affect the dimerisation of activated STAT3 constructs.

In chapter 4.3.3, Nuclear Translocation Assay data was presented which suggested that active STAT1 homodimers form independent of unphosphorylated dimer integrity. The following experiments aimed to assess if STAT3 homodimers behave similarly. The experiments utilised to answer this question were designed in principle as outlined in chapter **Error! Reference source not found.**. The STAT3-L78R mutation was used to monomerise unphosphorylated dimers. In the previous two chapters, the cytokine-induced nuclear accumulation and tyrosine phosphorylation of this mutant was characterised. Thereby, it was found that STAT3-L78R showed decreased nuclear scores compared to STAT3wt. Because tyrosine phosphorylation at the cytokine receptor was not negatively affected by this mutation, it was concluded that this

decrease is likely caused by an altered subcellular distribution in latent cells compared to STAT3wt. Therefore, the results presented in the following had to be interpreted with caution.

Investigating STAT3 latent dimer splitting mutations in the Nuclear Translocation Assay

To determine if active STAT3 homodimer formation depends on unphosphorylated dimer integrity, cells co-expressing the following STAT3 combinations were subjected to two-step live-cell imaging and analysis according to the Nuclear Translocation Assay protocol (chapter 4.1.6): STAT3-L78R-GFP + STAT3ΔNLS-mCherry, STAT3wt-GFP + STAT3ΔNLS-L78R-mCherry and STAT3-L78R-GFP + STAT3ΔNLS-L78R-mCherry.

In latent cells expressing the former combination, a pancellular distribution was observed for STAT3-L78R-GFP while the STAT3ΔNLS-mCherry was localised more cytoplasm (figure 44, row 1). In contrast, cells expressing the other two combinations showed pancellular distributions for both the GFP and mCherry-tagged constructs (figure 44, row 3 + 5). When the same cells co-expressing STAT3-L78R-GFP and STAT3ΔNLS-mCherry were re-imaged 90 min after the addition of cytokine, cytoplasmic paracrystals were observed in both fluorescent channels for most cells. Almost all of STAT3ΔNLS-mCherry fluorescence localised inside these paracrystals. STAT3-L78R-GFP fluorescence co-localised in the same paracrystals for the most part but some fluorescence was also seen to accumulate in the nucleus without paracrystal formation (figure 44, row 2). Still, co-localisation of the GFP and mCherry-tagged construct in the cytoplasmic paracrystals indicated an interaction between the two proteins in this compartment. Ninety minutes after cytokine addition to cells co-expressing STAT3wt-GFP + STAT3ΔNLS-L78R-mCherry, the former showed moderate accumulation inside the cell nucleus with some fluorescence remaining detectable in the cytoplasm. The STAT3ΔNLS-L78R-mCherry showed no change in subcellular distribution (figure 44, row 4). Substantial co-localisation of both proteins in the cytoplasm suggested that the mutated STAT3ΔNLS construct can cross-inhibit the cytokine-induced nuclear accumulation of STAT3wt-GFP. In cytokine-treated cells harbouring the STAT3-L78R-GFP + STAT3ΔNLS-L78R-mCherry combination, while some STAT3-L78R-GFP protein appeared to accumulate in the nucleus, a considerable amount of GFP fluorescence was retained in the cytoplasm (figure 44, row 6). The

retained amount thereby was estimated higher visually than that seen in cells containing the STAT3wt-GFP + STAT3ΔNLS-L78R-mCherry combination (figure 44, row 4). STAT3ΔNLS-L78R-mCherry did not redistribute following cytokine stimulation (figure 44, row 6).

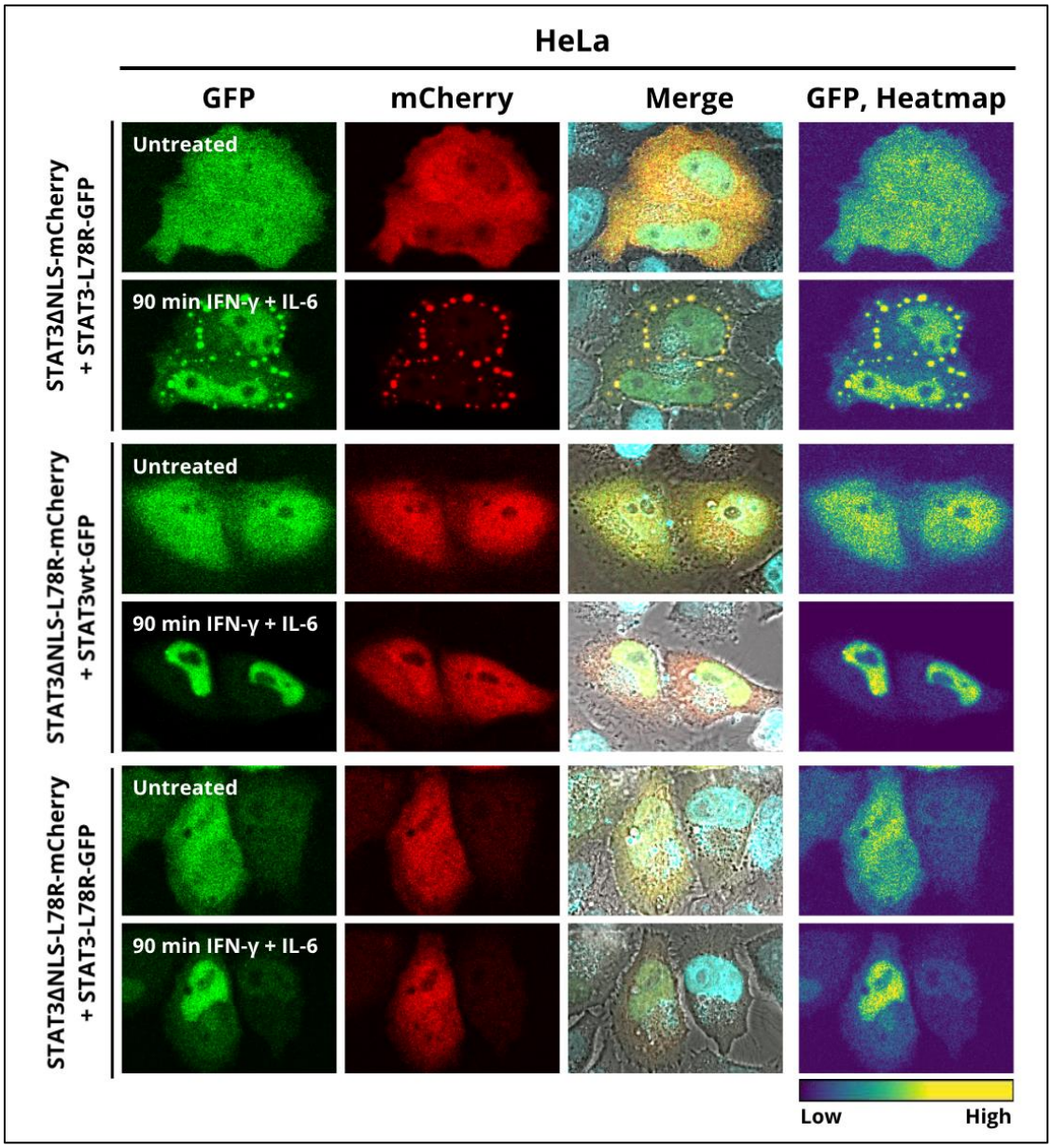


Figure 43 | STAT3ΔNLS-mCherry cross-inhibits STAT3-GFP cytokine-induced nuclear accumulation irrespective of unphosphorylated dimer integrity: *Fluorescence micrographs of HeLa cells co-expressing either STAT3ΔNLS-mCherry and STAT3-L78R-GFP, STAT3ΔNLS-L78R-mCherry and STAT3wt-GFP or STAT3ΔNLS-L78R-mCherry and STAT3-L78R-GFP imaged before and 90 min after cytokine treatment. The right column shows a contrast enhanced depiction of the GFP channel. Here, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).*

Nuclear scores were calculated for multiple cells expressing one of the three STAT3 combinations (41, 60 and 47 cells for the STAT3-L78R-GFP + STAT3ΔNLS-mCherry, STAT3wt-GFP + STAT3ΔNLS-L78R-mCherry and STAT3-L78R-GFP + STAT3ΔNLS-L78R-

mCherry combinations, respectively) and compared to the nuclear scores of other STAT3 combinations discussed in previous chapters. The nuclear scores determined for mutated and non-mutated bait-STAT3ΔNLS-mCherry were below 1.3 in all combinations (figure 45; table 36), suggesting that the ΔNLS phenotype was not influenced by the L78R mutation. STAT3-L78R-GFP showed marked decreases in nuclear scores when co-expressed with either STAT3ΔNLS-mCherry or STAT3ΔNLS-L78R-mCherry (1.39±0.03 or 1.28±0.03, respectively; figure 45; table 37). This indicated that both constructs interacted to form active STAT3 dimers in both combinations. The GFP nuclear score determined from cells containing STAT3wt-GFP and STAT3ΔNLS-L78R-mCherry was 2.07±0.03 and therefore significantly higher than was observed for the other two combinations. Still, compared to STAT3wt-GFP/STAT3wt-mCherry co-expressing cells (STAT3wt-GFP nuclear score of 2.49±0.07; figure 45; table 36), the nuclear score was significantly lower. This could indicate a small inhibitory effect exerted on STAT3wt-GFP by STAT3ΔNLS-L78R-mCherry. Interestingly, the GFP nuclear score of this combination was not significantly different from the one determined for cells co-expressing STAT3-L78R-GFP and STAT3-L78R-mCherry (figure 45). Moreover, the GFP-tagged proteins in both of these combinations showed, a similar, almost perfectly pancellular distribution in latent cells which was suggested as the cause for a reduced nuclear score of STAT3-L78R-GFP and STAT3-L78R-mCherry when co-expressed in the previous chapter. Therefore, the reduced nuclear score determined for STAT3wt-GFP co-expressed with STAT3ΔNLS-L78R-mCherry could be a detection artefact due to an altered subcellular distribution in favour of the cytoplasm before activation.

All in all, the data gained from the Nuclear Translocation Assay experiments alone neither confirmed nor disproved that STAT3 homodimers pre-form in latent cells and remain intact while transitioning to their active form. According to theoretical considerations from earlier (chapter 4.3.1), if this was the case, then rescue of the dominant negative effect exerted by wild-type and mutant STAT3ΔNLS constructs on the cytokine-induced nuclear accumulation of mutant STAT3-GFP constructs would be expected when only one of the two interaction partners carries L78R (figure 36C). Instead, the combination where the STAT3ΔNLS-mCherry carried this mutation behaved oppositely to the combination where the STAT3-GFP was mutated in L78.

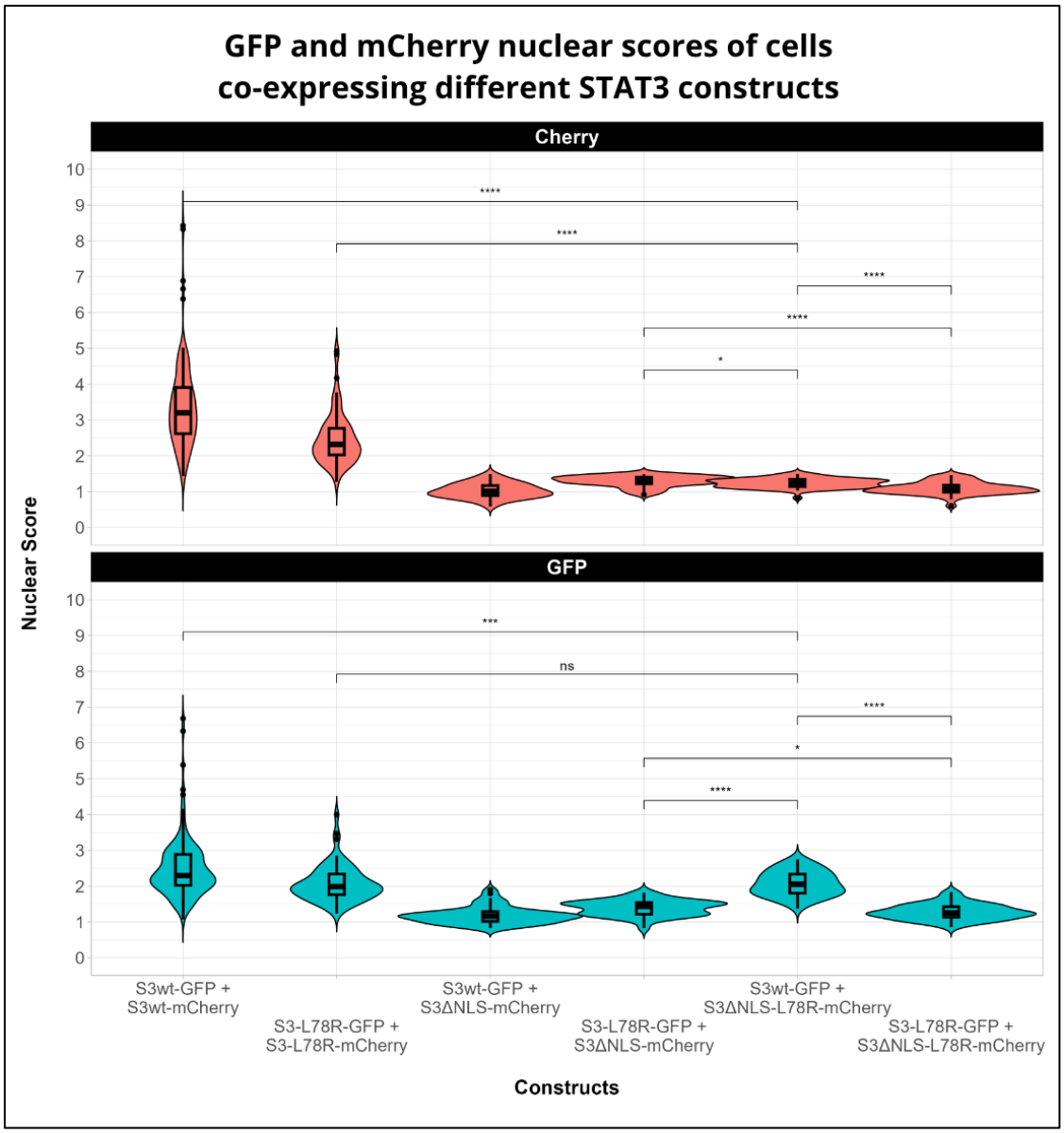


Figure 44 | Quantification of GFP and mCherry nuclear scores of cells expressing STAT3ΔNLS-mCherry and STAT3-L78R-GFP, STAT3ΔNLS-L78R-mCherry and STAT3wt-GFP or STAT3ΔNLS-L78R-mCherry and STAT3-L78R-GFP: *Nuclear scores were quantified in 41 cells expressing the former combination, 60 cells expressing the STAT3ΔNLS-L78R-mCherry and STAT3wt-GFP and 47 cells expressing the latter combination including the cells shown in figure 44. The first three columns represent data that was already presented and discussed in earlier chapters (see figures 27 and 41) but was included here for comparison.*

Table 35 | *Descriptive statistics of nuclear scores determined for cells co-expressing STAT3ΔNLS-mCherry and STAT3-L78R-GFP, STAT3ΔNLS-L78R-mCherry and STAT3wt-GFP or STAT3ΔNLS-L78R-mCherry and STAT3-L78R-GFP.* *Descriptive statistics of cells co-expressing STAT3wt-mCherry and STAT3wt-GFP, STAT3ΔNLS-mCherry and STAT3wt-GFP or STAT3-L78R-mCherry and STAT3-L78R-GFP were also shown in table 26 and 36 but were added here for comparison.*

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT3wt-mCherry + STAT3wt-GFP	140	3.4	0.1	2.5	0.1
STAT3-L78R-mCherry + STAT3-L78R-GFP	63	2.5	0.1	2.1	0.06
STAT3ΔNLS-mCherry + STAT3wt-GFP	51	1.0	0.03	1.2	0.03
STAT3ΔNLS-mCherry + STAT3-L78R-GFP	41	1.3	0.02	1.4	0.03
STAT3ΔNLS-L78R-mCherry + STAT3wt-GFP	60	1.2	0.02	2.1	0.04
STAT3ΔNLS-L78R-mCherry + STAT3-L78R-GFP	47	1.1	0.03	1.3	0.03

Decreased tyrosine phosphorylation is not the cause for inhibitory effects observed in STAT3 Nuclear Translocation Assay experiments.

A possible explanation for the discrepancy between theoretical considerations and actual experimental outcomes described for the STAT3 study above could be a defect in cytokine-induced tyrosine phosphorylation when STAT3-L78R and STAT3ΔNLS-L78R mutants are co-expressed. To investigate this, whole cell extracts from IL-6-stimulated U3R cells co-expressing an untagged STAT3-L78R variant (pmGFP-N1 + STAT3wt ΔGFP) and STAT3ΔNLS-L78R-mCherry were subjected to Western blot analysis. The phosphorylation status of individual constructs was assessed by specifically detecting proteins using the anti-STAT3 and anti-STAT3 pY705 antibodies from before. When utilising the anti-STAT3 antibody, two bands were observed (figure 46A, panel 1). The slower one migrated at a speed that was consistent with the molecular weight of STAT3ΔNLS-L78R-mCherry while the faster one was identified as untagged STAT3-L78R based on its molecular mass. Two bands moving at the same speeds were detected when using the anti-STAT3 pY705 antibody (figure 46A, panel 2), suggesting that STAT3 constructs were phosphorylated upon cytokine treatment.

The proportion of phosphorylated protein was scored by dividing the average band fluorescence measured over three independent experiments with the latter antibody by the average fluorescence measured with the former antibody for the corresponding band migrating at the same speed. No significant difference in specific phosphorylation was found between the untagged STAT3-L78R and STAT3ΔNLS-L78R-mCherry (p=0.095; figure 46B). Furthermore, the phosphorylation of both constructs was not significantly different to the phosphorylation of co-expressed untagged STAT3wt and STAT3wt-mCherry or co-expressed STAT3wt and STAT3ΔNLS-mCherry (compare to figure 28). This suggested that phosphorylation of STAT3-L78R and STAT3ΔNLS-L78R-mCherry was not adversely affected when these proteins were co-expressed. Thus, decreased phosphorylation could be ruled out as the cause for diminished nuclear scores in Nuclear Translocation Assay experiments.

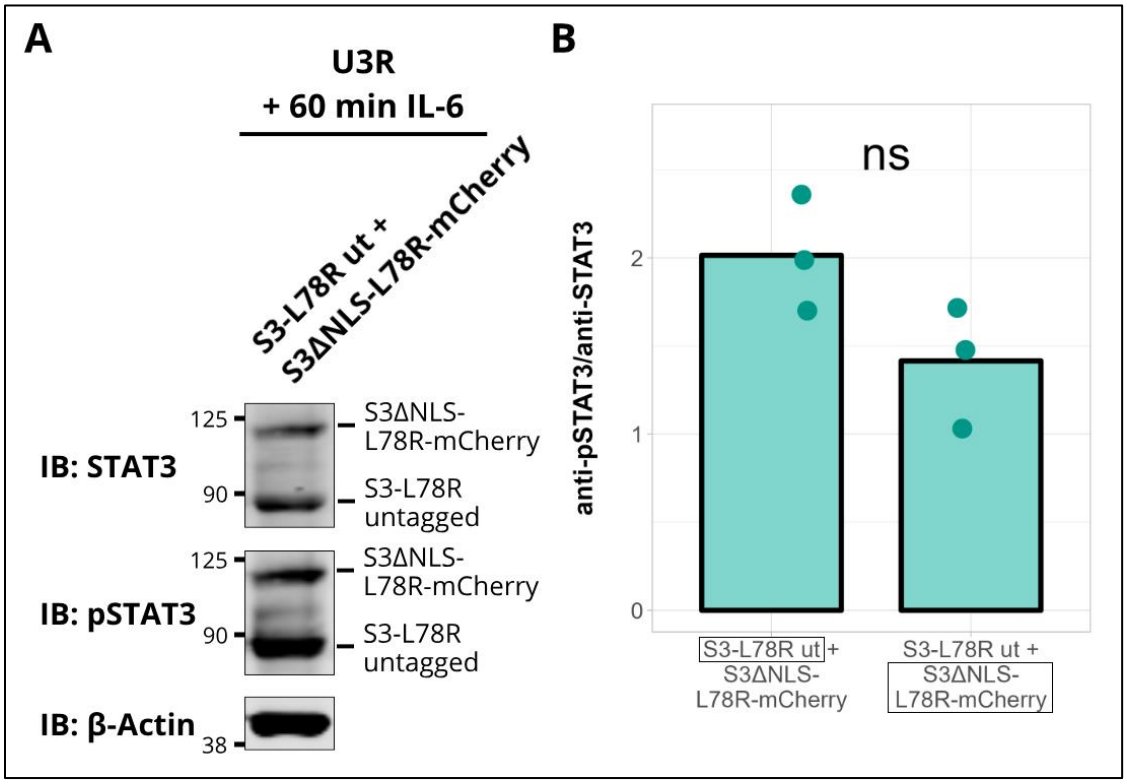


Figure 45 | Co-expression of STAT3-L78R and STAT3ΔNLS-L78R does not decrease cytokine-induced tyrosine phosphorylation: ***[A]*** *Whole cell extracts from IL-6 stimulated U3R cells co-expressing recombinant untagged STAT3-L78R and STAT3ΔNLS-L78R-GFP were subjected to Western analysis using antibodies raised against full-length STAT3, the STAT3 pY705 peptide or β-Actin.* ***[B]*** *Phosphorylation of the two constructs was quantified by dividing the fluorescence measured for anti-STAT3 pY705 bands by the signal measured for anti-STAT3 bands from three independent experiments. MW markers in kDa.*

4.3.6 Discussion of Chapter 4.3

The purpose of experiments presented in this chapter was to discern whether STAT1 and STAT3 homodimer formation requires the integrity of their unphosphorylated counterparts. For this, four different STAT-GFP/STAT-mCherry combinations had to be considered in which either none, one or both of the interaction partners carried an unphosphorylated dimer splitting mutation. Cells containing these combinations were then treated with cytokines and the nuclear scores were assessed as outlined in the Nuclear Translocation Assay protocol.

For STAT1, all STAT1ΔNLS-mCherry constructs cross-inhibited the cytokine-induced nuclear accumulation of STAT1-GFP regardless of whether one or both interaction partners carried a mutation in residue F172. These findings suggested that active STAT1 homodimers formed independent of latent dimer integrity because activation likely involved dissociation and re-formation of dimeric molecules.

When assessing the consequences of latent dimer splitting on the formation of active STAT3 homodimers, a less clear picture emerged. If STAT3 active homodimer formation was dependent on unphosphorylated dimer integrity, then no cross-inhibition of STAT3-GFP cytokine-induced nuclear accumulation by STAT3ΔNLS-mCherry would have been expected in experiments where one but not the other interaction partner carried the L78R mutation. Accordingly, if the opposite was true and active STAT3 homodimers form independent of their latent counterparts then cross-inhibition of STAT3-GFP accumulation exerted by STAT3ΔNLS-mCherry should have been high when one but not the other interaction partner carried the L78R mutation. However, while STAT3-GFP cytokine-induced nuclear accumulation was markedly reduced when itself but not the STAT3ΔNLS-mCherry carried the L78R mutation, the cross-inhibitory effect exerted on STAT3wt-GFP by STAT3ΔNLS-L78R-mCherry was low. What's more, the weak effect could not for definite be attributed to inhibition exerted by the bait-STAT3ΔNLS-L78R but instead could be an artefact due to altered subcellular distributions before cytokine exposure. The absence of inhibitory activity of STAT3ΔNLS-L78R-mCherry could be explained by its inability to form active dimers. Though, since the same construct cross-inhibited the cytokine-induced nuclear accumulation of STAT3-L78R-GFP, this notion could be discarded. Further experiments

are required to clarify whether active STAT3 homodimer formation requires pre-formation as unphosphorylated dimers. Nevertheless, because inhibition was detected when STAT3-L78R-GFP and STAT3ΔNLS-mCherry were combined, the findings hint towards a mechanism where active STAT3 dimers can form independent of their unphosphorylated assemblies.

The findings described for STAT1 presented in this chapter contradicted some assumptions of the current model for STAT1 transitioning from unphosphorylated to active dimers proposed following the solution of the crystal structure of an unphosphorylated STAT1 dimer (Mao *et al.*, 2005). While the unphosphorylated STAT1 dimer predominantly adopts an antiparallel conformation (Wenta *et al.*, 2008), the authors hypothesised that the dimer is likely recruited to the cytokine receptor as a parallel dimer held together by NTD-NTD interactions (Mao *et al.*, 2005). The NTD-NTD interaction would then only be split once the pY-SH2 interaction is established following STAT activation, thus making the exchange of monomers during activation at the cytokine receptor impossible. Importantly, since the flexible tether linking NTD and core fragment (CF) could not be resolved in this structure, the authors were unable to determine which NTD was linked to which CF. Thus, it is unclear whether the parallel unphosphorylated STAT1 dimer held together by NTD-NTD interactions is merely an artefact of crystallisation. Furthermore, they were unable to provide direct evidence that proves that the split of the NTD-NTD interaction only occurs once the pY-SH2 interaction is established. Still, the study implies that both the NTD-NTD and reciprocal CCD-DBD interaction must be broken at some point to allow for the formation of the active dimer. Considering the evidence presented in this work, a scenario where the unphosphorylated dimer is split before or shortly after receptor docking allowing for the exchange of monomers before, during or after phosphorylation would be favoured. This is also supported by findings showing that the phosphorylation of STAT1wt, STAT1-F172W and STAT1-F77A (a mutation that disrupts the NTD-NTD dimer interaction) was equally rapid after IFN-γ treatment, suggesting that the integrity of latent dimers is not required for this process (Zhong *et al.*, 2005). The receptor docking model suggested for unphosphorylated STAT5A dimers based on its crystal structure and accompanying FRET experiments also supported the splitting of unphosphorylated dimers before active dimer formation (Neculai *et al.*, 2005). A detailed model of STAT3

receptor docking and activation by phosphorylation has not been reported. The first crystal structure solved for an unphosphorylated STAT3 dimer indicated a lesser role for the CCD-DBD interface in latent dimer interactions (Ren *et al.*, 2008). Instead, the NTD-NTD was found to be essential for the integrity of the unphosphorylated STAT3 dimer (Vogt *et al.*, 2011). Another crystal structure of unphosphorylated STAT3 bound to DNA indicated that this dimer adopts a parallel conformation more readily than the unphosphorylated STAT1 homodimer (Nkansah *et al.*, 2013), a notion that was also confirmed by *in situ* FRET experiments (Domoszlai *et al.*, 2014). Indeed, the ability of STAT3 to effortlessly form parallel unphosphorylated dimers may be a direct consequence of the weak reciprocal CCD-DBD interactions. Moreover, it was shown that the deletion of the STAT3 NTD does not affect tyrosine phosphorylation in response to IL-6 treatment (Zhang *et al.*, 2006) and that the bond between isolated STAT3 NTDs is weaker than observed for STAT1 NTDs (Wenta *et al.*, 2008). This, in line with the results shown here, may mean that the latent STAT3 dimer is even more easily disrupted at the cytokine receptor to allow for subunit exchange before re-forming as an active STAT3 homodimer.

4.4 Unphosphorylated dimers do not pre-determine the formation of active STAT1/STAT3 heterodimers.

In the previous chapter, the Nuclear Translocation Assay was applied to investigate whether the formation of active STAT1 and STAT3 homodimers is regulated by their unphosphorylated counterparts. In experiments presented in this chapter, the same principles were applied to examine active STAT1/STAT3 heterodimer interactions and to find out if these entities require pre-formation before cytokine exposure. Importantly, to detect active STAT1/STAT3 heterodimers in the Nuclear Translocation Assay, the STAT3ΔNLS must be able to exert a dominant negative effect on STAT1 constructs. While the dominant negative effect of STAT3ΔNLS has been confirmed regarding STAT3 homodimers in chapter 4.2.2 it is yet unclear if its effect is also dominant negative regarding STAT1/STAT3 heterodimers. Therefore, this was confirmed first before proceeding with experiments to assess the mechanism of active STAT1/STAT3 heterodimer formation. Furthermore, the Nuclear Translocation Assay was used to investigate STAT1 and STAT5B for their ability to form active heterodimers when their unphosphorylated dimers are disrupted. The microscopy assays were accompanied by *in vitro* EMSA experiments.

4.4.1 Introduction

Chapter 4.3.1, outlined a four-step experimental design that made use of the fact that latent STAT homodimers can be disrupted without affecting the formation of their active forms, and vice-versa, to investigate if unphosphorylated STAT dimers regulate the formation of their active STAT signalling complexes. The following paragraphs describe how these principles were applied to STAT1/STAT3 heterodimers.

Figure 47 outlines the experimental design and theory behind the four-step study. Importantly, the STAT3-mCherry constructs will carry the ΔNLS mutation required for the Nuclear Translocation Assay. Again, the first outlined the baseline for uninhibited cytokine-induced STAT1wt/STAT3wt nuclear accumulation (figure 47A) The second scenario outlined the baseline cross-inhibitory effect exerted on STAT1wt-GFP accumulation by STAT3ΔNLS-mCherry, should this effect be dominant negative and the assembly of active STAT1/STAT3 heterodimers in living cells be confirmed (figure 47B). In the third scenario, either the STAT3ΔNLS-mCherry or the STAT1-GFP carried an

unphosphorylated dimer-splitting mutation. If STAT1/STAT3 heterodimers are pre-formed before cytokine exposure and stay intact during activation at the cytokine receptor, then an alleviation of STAT3ΔNLS-mCherry inhibition would be expected for this scenario. This is due to the pre-association of the non-mutated interaction partner into unphosphorylated dimers, meaning that it was no longer available to form active dimers with the mutated protein upon cytokine stimulation. Thus, the mutated protein would be forced to self-assemble. Irrespective of which interaction partner carried the unphosphorylated dimer splitting mutation, this would then result in active STAT1-GFP/STAT1-GFP and STAT3ΔNLS-mCherry/STAT3ΔNLS-mCherry but not STAT1-GFP/STAT3ΔNLS dimer formation. Of course, since the former does not carry the ΔNLS mutation, it could accumulate inside the nucleus and show a high nuclear score (figure 47C). In the final scenario, both interaction partners carry the unphosphorylated dimer splitting mutation and are therefore exclusively monomeric in latent cells. Therefore, upon cytokine treatment, all proteins would assemble into active dimers at random. Because cells would be transfected so that they contain a relative excess of STAT3ΔNLS-L78R-mCherry over STAT1-F172W-GFP the assembly of active STAT1-GFP/STAT3ΔNLS would be maximised (figure 47D). Therefore, the GFP nuclear score measured in this scenario should be similar to that measured for scenario two.

Alternatively, if active STAT1/STAT3 heterodimers form independent of unphosphorylated dimer integrity, then low GFP nuclear scores would be expected in all scenarios. Specifically considering scenarios two and three, regardless of their composition, the unphosphorylated dimers would split upon activation at the cytokine receptor, briefly making all STAT subunits monomeric. These monomers could then randomly assemble into active dimers. Because the mutant or wild-type STAT3ΔNLS-mCherry constructs would be present in molar excess, the probability that the co-expressed mutant or wild-type STAT1-GFP constructs would bind to a STAT3ΔNLS-mCherry would be maximised. Thus, should it be true that active STAT1/STAT3 heterodimers form independent of unphosphorylated dimer integrity, then the mechanism for active dimer assembly observed in scenarios two and three would resemble that of scenario four (figure 47D).

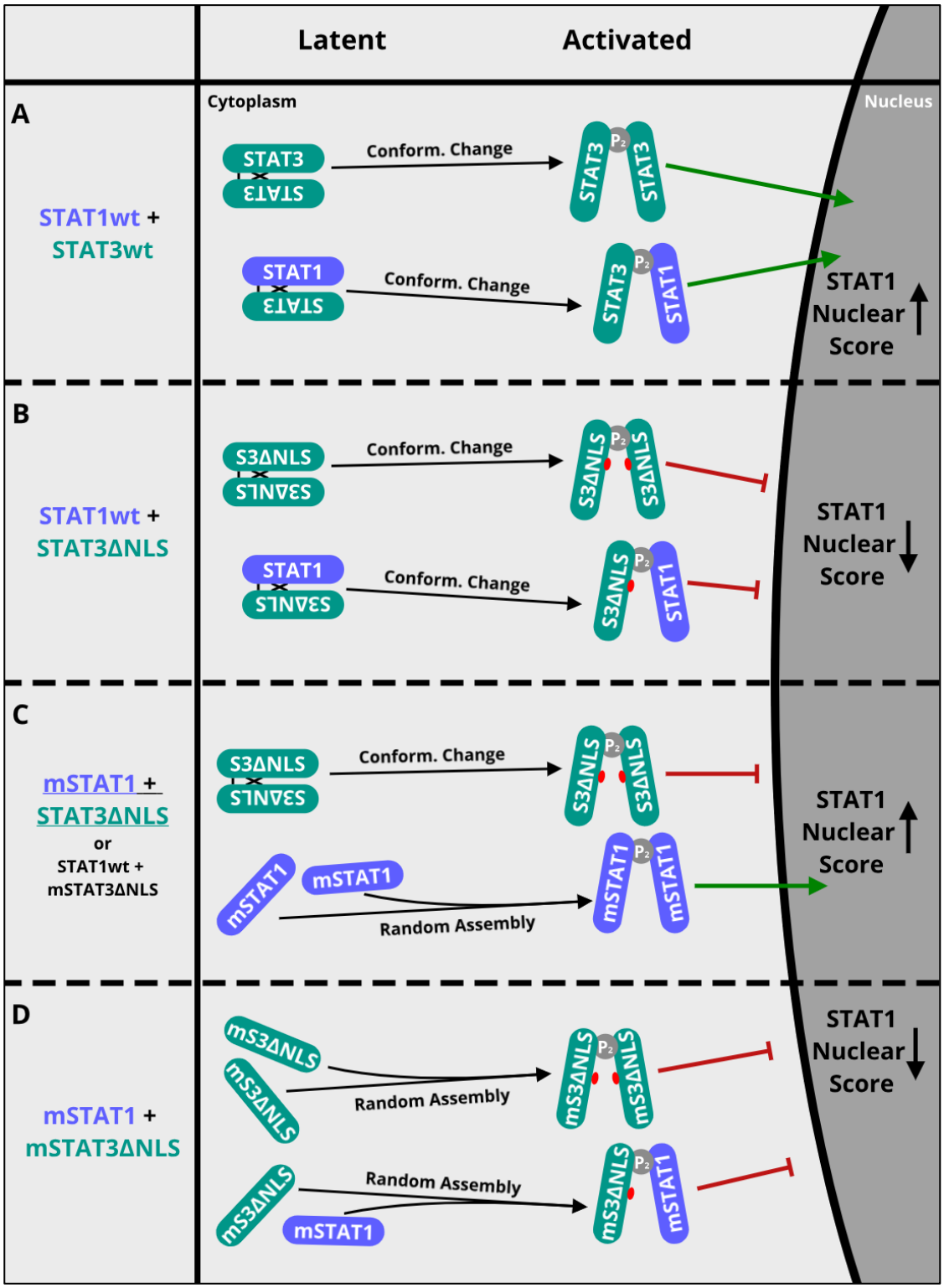


Figure 46 | Four scenarios were considered to investigate the transition of STAT dimers from latent to active: *In the four scenarios different STAT1-GFP and STAT3-mCherry constructs were combined. Some of these carry mutations that disrupt the latent dimer form. These are indicated by an "m" (for monomerised) in front of their names. The cartoons depict the mechanisms expected if active STAT dimer formation depends on the pre-association as unphosphorylated dimers.*

We know from several previous accounts that STAT1 and STAT5B are activated by multiple cytokines simultaneously but do not form active heterodimers (Wehinger *et al.*, 1996; Kirken *et al.*, 1997; Schulze *et al.*, 2000; Delgoffe and Vignali, 2013). Multiple factors may be at play to facilitate this effect and the composition of unphosphorylated STAT dimers could be one of them. Indeed, should active STAT dimers pre-form before cytokine exposure and merely undergo a monomer exchange-less conformational change upon activation, then unphosphorylated dimers would determine the composition of the cellular pool of STAT homo and heterodimers. Therefore, it is conceivable that the reason for the absence of active STAT1/STAT5B heterodimers is that they cannot form the corresponding unphosphorylated dimer. Instead, it is possible that in latent cells most STAT1 molecules are bound in STAT1 homodimers and most STAT5B molecules are bound in STAT5B homodimers (or in heterodimers with other STATs). Consequently, it might be possible to create artificial active STAT1/STAT5B heterodimers by disrupting these reservoirs of pre-associated dimers through unphosphorylated dimer splitting mutations. Essentially, a scenario similar to the one described in figure 47D for STAT1/STAT3 heterodimers may be created where mutated STAT1 and STAT5B which were previously monomeric before cytokine exposure, randomly assemble into active STAT1/STAT5B heterodimers upon their activation at the cytokine receptor. This scenario will be tested below both via Nuclear Translocation Assay assessments as well as *in vitro* EMSA experiments.

4.4.2 Active STAT1 and STAT3 readily interact despite monomerization of their latent dimers.

The formation of active STAT1/STAT3 heterodimers in response to several cytokines has been well documented. Despite this, the mechanisms that determine the formation of this dimer over others are poorly understood. As stated above, it is conceivable that STAT1/STAT3 heterodimers pre-assemble before activation and merely change their conformation to form the active signalling complex upon cytokine stimulation without the exchange of subunits. This notion was tested using the Nuclear Translocation Assay via the scheme outlined in the previous chapter in the following.

Active STAT1/STAT3 heterodimers are detectable *in vitro.*

The interaction of active STAT1 with active STAT3 in response to IL-6 and EGF treatment has been documented early on *in vitro* via EMSA experiments (Zhong, Wen and Darnell, 1994a, 1994b).

To replicate these EMSA results, U3R cells were left untransfected, transfected with cDNA coding for STAT1wt-GFP or STAT3wt-mCherry alone or co-transfected with both STAT1wt-GFP and STAT3wt-mCherry, treated with IFN-γ, IL-6 and sIL6Rα for 60 min or left untreated and lysed to generate detergent free whole cell extracts. The extracts were then incubated with an IRDye®- labelled M67 probe, a mutated oligonucleotide of the *sis*-inducible element found in the promotor of the c-*fos* gene that shows a high affinity to STAT1 and STAT3 and can also bind the STAT1/STAT3 heterodimer (Wagner *et al.*, 1990; Horvath, Wen and Darnell, 1995). Afterwards, the mixture was loaded onto a native polyacrylamide gel for electrophoretic separation. Following this, the gel was imaged using a LICOR Odyssey® fluorescence imager.

No specific bands were detected in the lanes containing extracts from untransfected cells irrespective of cytokine treatment (figure 48, lane 1+2). A single band was detected in both extracts from cytokine-treated cells containing the single transfected STAT1wt-GFP or single transfected STAT3wt-mCherry (figure 48, lane 3+4), whereby the STAT3wt-mCherry band migrated slower. Extracts from cytokine-treated cells co-expressing the two proteins showed three specific bands in the native gel. The top and bottom bands migrated at the same heights as STAT3wt-mCherry and STAT1wt-GFP, respectively, while the middle band migrated in between (figure 48, lane 6). This banding pattern was similar to the one observed in the study that first reported STAT1/STAT3 heterodimer binding activity (Zhong, Wen and Darnell, 1994a). Indeed, the middle band was supershifted by both anti-STAT1 and anti-STAT3 antibodies while the top band was only supershifted by the latter and the bottom band was only supershifted by the former (figure 48, lane 7+8). This clearly shows that STAT1 and STAT3 form active homo and heterodimers in response to cytokine treatment, replicating the findings found in the literature.

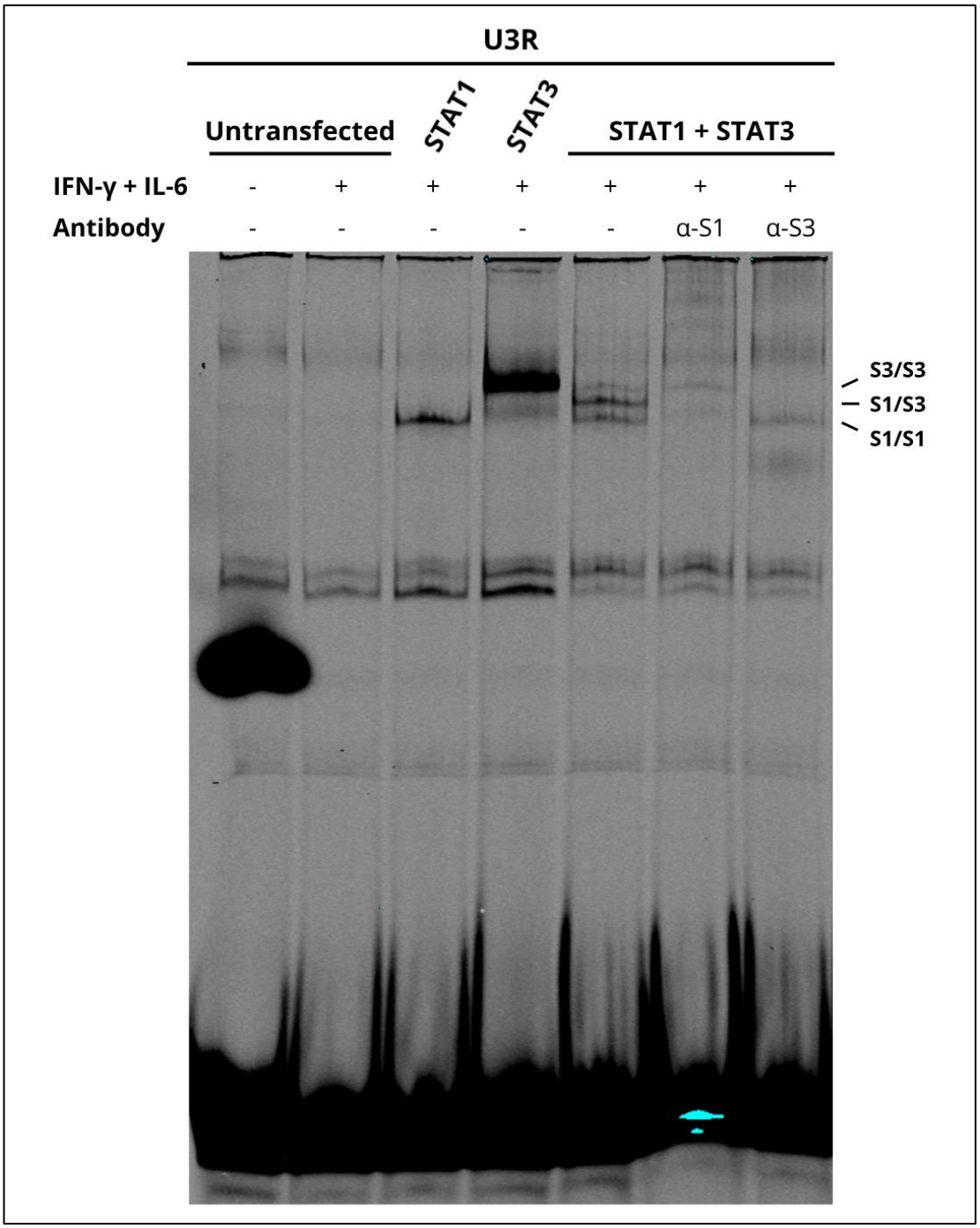


Figure 47 | STAT1 and STAT3 form three sets of dimers when co-expressed: a STAT1 homodimer, a STAT3 homodimer and a STAT1/STAT3 heterodimer: *Whole cell extracts of cytokine treated and non-treated U3R cells that were left untransfected, transfected with STAT1-GFP or STAT3-mCherry alone or co-transfected with STAT1-GFP and STAT3-mCherry were subjected to EMSA experiments using the M67 DNA probe. Antibodies raised against native full-length STAT1 and STAT3 were used for supershifts. This figure is representative for three individual EMSA experiments.*

Active STAT1/STAT3 heterodimers are detectable in living cells.

In the Nuclear Translocation Assay, the formation of active STAT dimers between a STATΔNLS-mCherry and a STAT-GFP construct was declared when the STATΔNLS-mCherry cross-inhibited the cytokine induced nuclear accumulation of the STAT-GFP. Thereby, cross-inhibition would only be possible if the ΔNLS mutation is dominant negative, meaning that introducing this mutation into one monomer of a STAT dimer is sufficient to render the entire dimer incapable of cytokine-induced nuclear accumulation. In chapter 4.2.2 it was shown that the STAT3ΔNLS mutation showed a dominant negative effect in STAT3 homodimers. In the following, the process of testing if the STAT3ΔNLS also shows a dominant negative effect in active STAT1/STAT3 heterodimers is described.

Just like the EMSA experiment shown in the previous chapter, most accounts of STAT1/STAT3 heterodimers are based on *in vitro* experiments which can suffer from experimental artefacts. For example, the use of different detergents in the whole cell extract preparation can drastically alter the signals of different STAT dimer bands in EMSA experiments (Haan and Haan, 2013). This is also the reason why whole cell extracts for EMSA analyses in this work were prepared without the use of any detergent. Therefore, besides the reasoning for the following experiments described in the previous paragraph, they were also used to confirm the existence of active STAT1/STAT3 heterodimers in living cells. For this, HeLa cells co-expressing STAT1wt-GFP and STAT3wt-mCherry or STAT1wt-GFP and STAT3ΔNLS-mCherry were subjected to Nuclear Translocation Assay analysis.

Before cytokine exposure, STAT1wt-GFP and STAT3wt-mCherry showed a similar distribution accentuating the cytoplasm (figure 49A, row 1). The STAT1wt-GFP and STAT3ΔNLS-mCherry combination behaved similarly (figure 49A, row 3). In images of the same cells taken 90 min after the addition of IFN-γ and IL-6, both STAT1wt-GFP and STAT3wt-mCherry showed complete nuclear accumulation (figure 49A, row 2). On the contrary, in cytokine-treated cells containing the other combination, STAT3ΔNLS-mCherry remained predominantly cytoplasmic where it formed paracrystals (figure 49A, panel 14). The STAT1wt-GFP showed some re-distribution towards the nucleus, but a considerable amount of fluorescence remained detectable in the cytoplasm and

showed incorporation into STAT3ΔNLS paracrystals (figure 49A, panel 13). This suggested that the latter could inhibit STAT1wt-GFP cytokine-induced nuclear accumulation and, therefore, that STAT1 and STAT3 formed heterodimers. In both combinations, STAT3 paracrystal formation was markedly reduced (STAT1wt-GFP+STAT3ΔNLS-mCherry; figure 49A, panel 14) or completely absent (STAT1wt-GFP+STAT3wt-mCherry, figure 49A, panel 6) compared to cells expressing only STAT3 constructs (figure 27+44). This could be explained by the incorporation of SUMOylated STAT1wt into STAT3 paracrystals. The solubilising effect of SUMO on STAT paracrystals has been implicated in STAT1 and STAT3 before (Droescher *et al.*, 2011).

To quantify these observations, images of 61 cells containing STAT1wt-GFP and STAT3wt-mCherry and 65 cells containing STAT1wt-GFP and STAT3ΔNLS-mCherry were analysed under the Nuclear Translocation Assay protocol. In the former combination, both STAT1wt-GFP and STAT3wt-mCherry showed a high nuclear score (4.0±0.1 and 3.7±0.1, respectively; figure 49B; table 38). In the cells containing the latter combination, the nuclear score of STAT3ΔNLS-mCherry (1.14±0.04) suggested that no cytokine-induced redistribution had taken place, confirming the ΔNLS phenotype (figure 49B; table 38). In these cells, the STAT1wt-GFP nuclear score (2.3±0.1) was markedly reduced compared to that measured for this construct in the STAT1wt-GFP/STAT3wt-mCherry combination. Thus, the STAT3ΔNLS-mCherry was able to cross-inhibit the cytokine-induced nuclear accumulation of STAT1wt-GFP. This confirms that active STAT1/STAT3 assemble active heterodimers in living cells and that the STAT3ΔNLS mutation exerts a dominant negative effect on these dimers. Furthermore, the STAT3ΔNLS-mCherry was able to inhibit STAT1wt-GFP nuclear accumulation to a similar degree as the STAT1ΔNLS from earlier (figure 24; table 23) which may suggest that STAT1 homodimers and STAT1/STAT3 heterodimers can form at equal rates.

Table 36 | Descriptive statistics of nuclear scores determined for cells co-expressing STAT1wt-GFP and STAT1ΔNLS-mCherry or STAT1wt-GFP and STAT1wt-mCherry.

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT3wt-mCherry + STAT1wt-GFP	61	3.7	0.1	4.0	0.1
STAT3ΔNLS-mCherry + STAT1wt-GFP	65	1.14	0.04	2.3	0.1

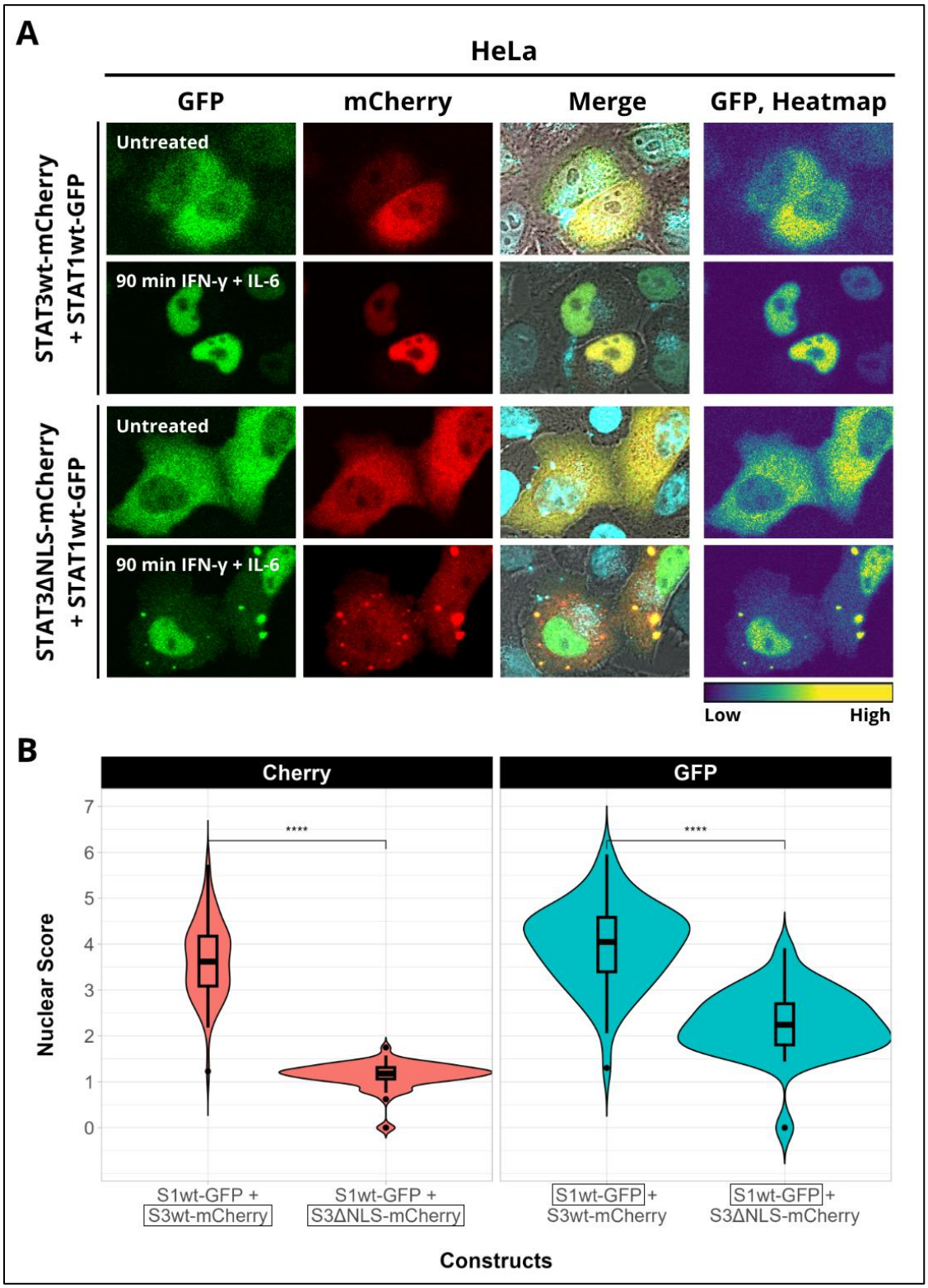


Figure 48 | STAT3ΔNLS inhibits cytokine-induced nuclear accumulation of STAT1wt: ***[A]*** *Fluorescence micrographs of HeLa cells co-expressing either STAT3wt-mCherry and STAT1wt-GFP or STAT3ΔNLS-mCherry and STAT1wt-GFP imaged before and 90 min after cytokine treatment. The last column of panels shows a contrast enhanced depiction of the GFP channel. In this, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).* ***[B]*** *Quantification of mCherry and GFP nuclear scores in 61 or 65 cells co-expressing the two STAT1/STAT3 combinations described in [A], respectively.*

Active STAT1/STAT3 heterodimers form independent of unphosphorylated dimers.

After confirming the existence of active STAT1/STAT3 dimers in the previous paragraphs, whether these dimers form independent of unphosphorylated dimers was investigated next. This was done using the Nuclear Translocation Assay following the experimental design consisting of four individual steps as outlined in chapter4.4.1. The first two scenarios representing baseline uninhibited cytokine-induced nuclear accumulation of STAT1wt-GFP, and the baseline cross-inhibition exerted when it was co-expressed with STAT3ΔNLS-mCherry were presented in the previous chapter. Here, cross-inhibition exerted by STAT3ΔNLS-mCherry on STAT1-GFP cytokine-induced nuclear accumulation was assessed when one or both of the interaction partners carried a unphosphorylated dimer-splitting mutation. For this, cells co-expressing the following STAT1/STAT3 combinations were subjected to the Nuclear Translocation Assay: STAT1-F172W-GFP + STAT3ΔNLS-mCherry, STAT1wt-GFP + STAT3ΔNLS-L78R-mCherry and STAT1-F172W-GFP + STAT3ΔNLS-L78R-mCherry.

The two proteins contained in the first combination distributed cytoplasmic before cytokine exposure with some nuclear fluorescence remaining detectable (figure 50, row 1). This is consistent with previous experiments involving STAT1-F172W-GFP and STAT3ΔNLS-mCherry seen elsewhere in this work. STAT1wt-GFP and STAT3ΔNLS-L78R-mCherry of the second combination showed the same distribution before cytokine exposure as described for the first combination, except, the distribution of the latter looked more pancellular compared to that observed for STAT3ΔNLS-mCherry before (figure 50, row 3). This indicated that introducing the L78R mutation changed the latent distribution of the STAT3ΔNLS construct, a phenomenon that was also observed in chapter 4.3.5. When STAT1-F171W-GFP and STAT3ΔNLS-L78R-mCherry were combined, they distributed similarly within latent cells as when they were co-expressed with other constructs (figure 50, row 5).

Ninety minutes after treatment with IFN-γ, IL-6 and sIL6Rα, STAT3ΔNLS-mCherry of the first combination did not accumulate inside the nucleus but formed cytoplasmic paracrystals. The co-expressed STAT1-F172W-GFP showed incomplete nuclear accumulation and some cytoplasmic GFP fluorescence co-localised with the cytoplasmic STAT3ΔNLS-mCherry paracrystals (figure 50, row 2). In the images taken

90 min after cytokine treatment of cells containing the second combination, the STAT3ΔNLS-L78R-mCherry did not show altered subcellular distribution compared to images taken before stimulation. The accompanying STAT1wt-GFP, however, exhibited limited redistribution to the nucleus with considerable fluorescence being detectable inside the cytoplasm (figure 50, row 4). The same outcome was observed for cells containing STAT1-F172W-GFP and STAT3ΔNLS-L78R-mCherry (figure 50, row 6).

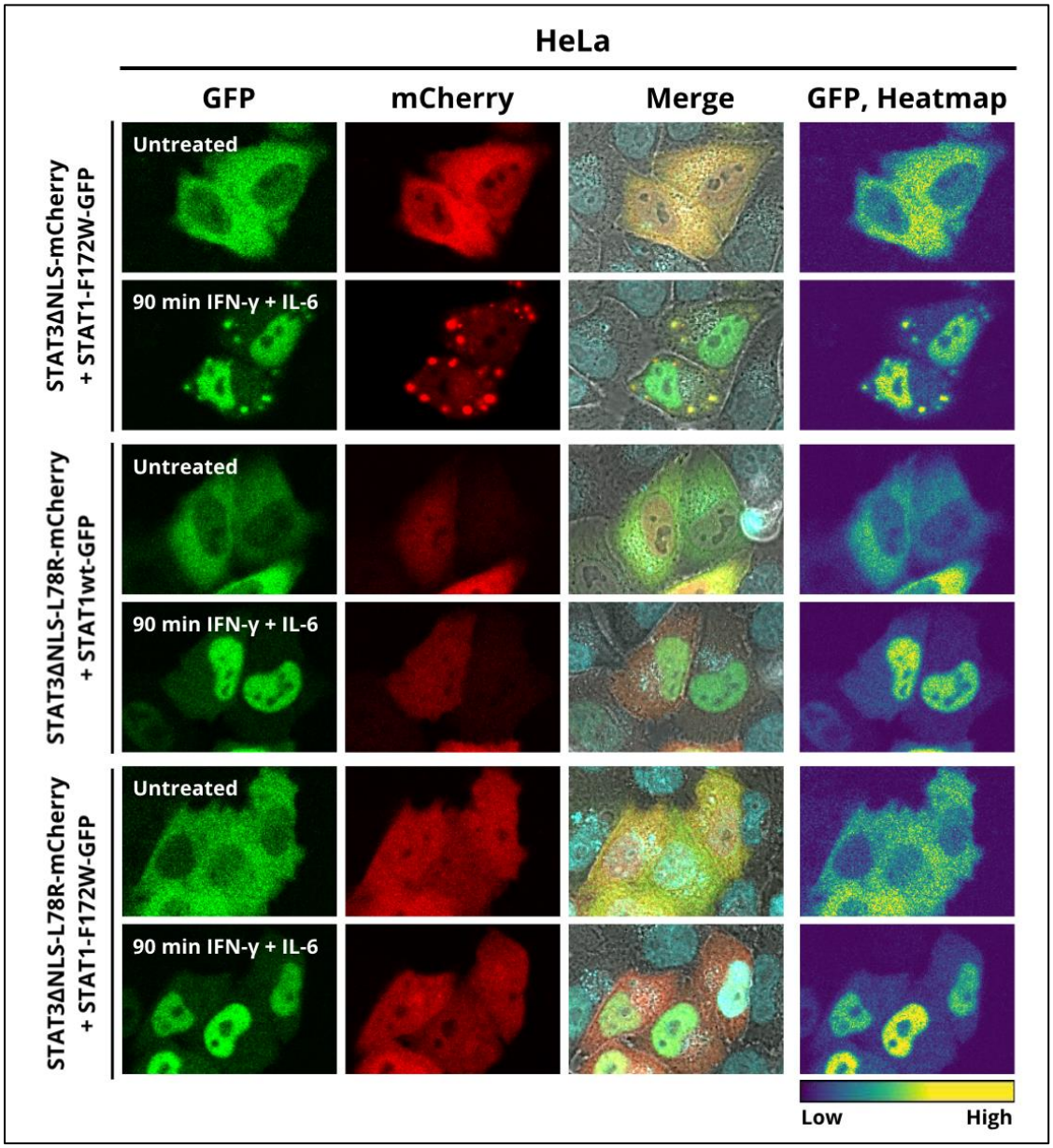


Figure 49 | STAT3ΔNLS-mCherry inhibits STAT1-GFP cytokine-induced nuclear accumulation irrespective of unphosphorylated dimer integrity: *Fluorescence micrographs of HeLa cells co-expressing either STAT3ΔNLS-mCherry and STAT1-F172W-GFP, STAT3ΔNLS-L78R-mCherry and STAT1wt-GFP or STAT3ΔNLS-L78R-mCherry and STAT3-F172W-GFP imaged before and 90 min after cytokine treatment. The right column shows a contrast enhanced depiction of the GFP channel. Here, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).*

The observations described above indicate that the different STAT3ΔNLS constructs were able to interrupt cytokine-induced nuclear accumulation of different STAT1 constructs in all three combinations. To quantify these observations, nuclear scores were determined for 105, 47 or 50 cells expressing the three combinations, respectively. Indeed, all STAT1-GFP nuclear scores were significantly reduced compared to cells co-expressing STAT1wt (nuclear score 3.99±0.11) and STAT3wt (nuclear score 3.66±0.11). The nuclear score of STAT1-F172W-GFP was lowest (1.73±0.09) when combined with STAT3ΔNLS-mCherry. STAT1wt-GFP showed the highest nuclear score (2.52±0.07) when combined with STAT3ΔNLS-L78R-mCherry. The nuclear score of STAT1-F172W-GFP, when combined with STAT3ΔNLS-L78R-mCherry, ranged in-between (2.08±0.06). The nuclear scores of the three bait-STAT3ΔNLS constructs ranged from 1.15±0.02 to 1.23±0.07, confirming their ΔNLS phenotype (figure 51; table 39).

Table 37 | Descriptive statistics of nuclear scores determined for cells co-expressing STAT3ΔNLS-mCherry and STAT1-F172W-GFP, STAT3ΔNLS-L78R-mCherry and STAT1wt-GFP or STAT3ΔNLS-L78R-mCherry and STAT1-F172L-GFP. *Descriptive statistics of cells co-expressing STAT3wt-mCherry and STAT1wt-GFP or STAT3ΔNLS-mCherry and STAT1wt-GFP are also shown in table 38 but were added again here for comparison.*

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT3wt-mCherry + STAT1wt-GFP	61	3.7	0.1	4.0	0.1
STAT3ΔNLS-mCherry + STAT1wt-GFP	65	1.14	0.04	2.3	0.1
STAT3ΔNLS-mCherry + STAT1-F172W-GFP	105	1.2	0.02	1.7	0.1
STAT3ΔNLS-L78R-mCherry + STAT1wt-GFP	47	1.2	0.02	2.5	0.1
STAT3ΔNLS-L78R-mCherry + STAT1-F172W-GFP	50	1.23	0.02	2.1	0.1

Overall, the quantifications confirmed the visual observations that different STAT3ΔNLS constructs could inhibit the cytokine-induced nuclear accumulation of different STAT1 constructs whether none, one or both interaction partners carried a unphosphorylated dimer disrupting mutation. Following the logic described in chapter 4.4.1, this implied that active STAT1/STAT3 heterodimers formed independent of

unphosphorylated dimer integrity. Consequently, the findings indicate that unphosphorylated dimers were disrupted upon cytokine stimulation and underwent monomer exchange before reforming as the active signalling complex.

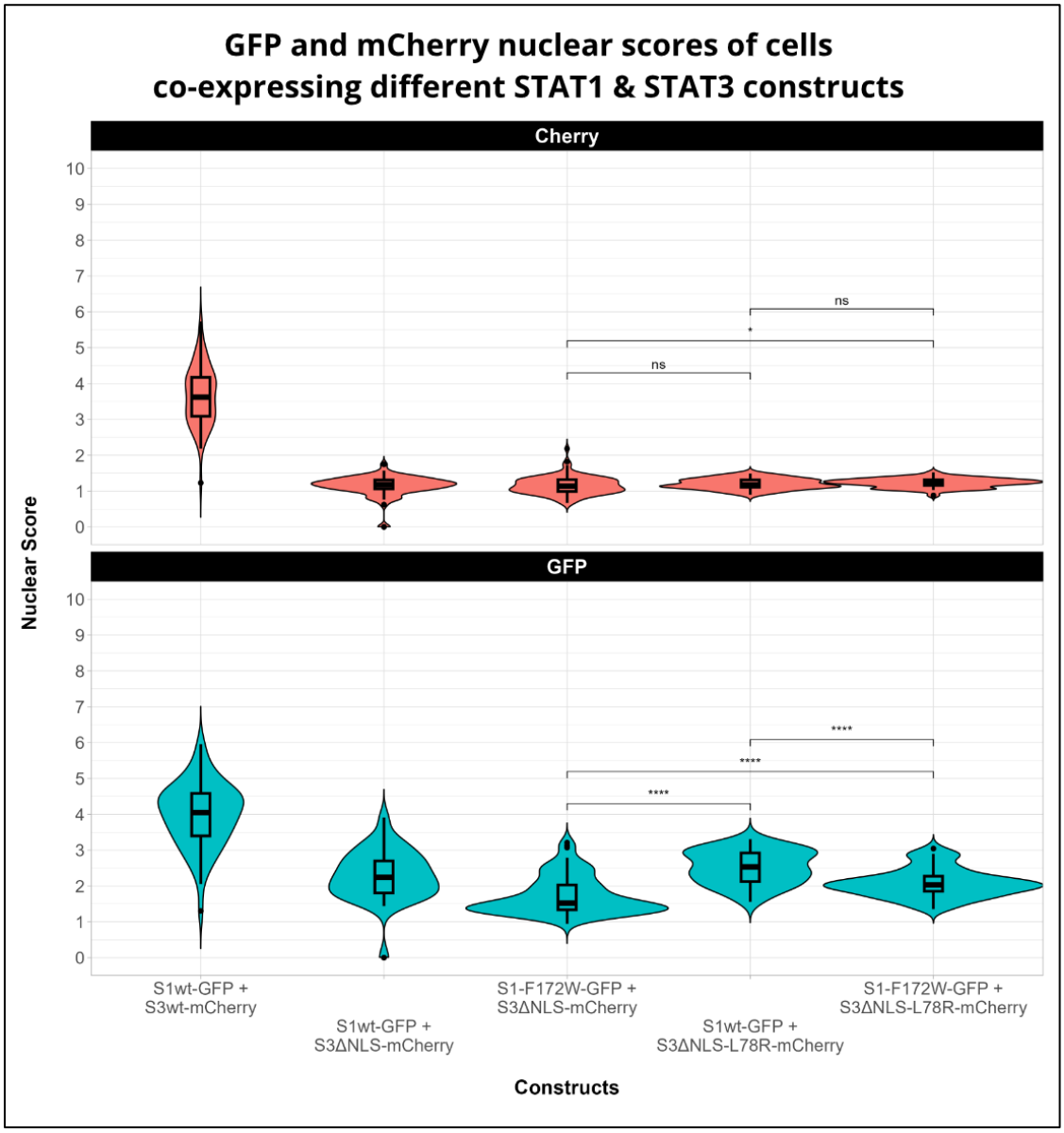


Figure 50 | Quantification of GFP and mCherry nuclear scores of cells expressing STAT3ΔNLS-mCherry and STAT1-F172W-GFP, STAT3ΔNLS-L78R-mCherry and STAT1wt-GFP or STAT3ΔNLS-L78R-mCherry and STAT1-F172W-GFP: *Nuclear scores were quantified in 105 cells expressing the former combination, 47 cells expressing the STAT1ΔNLS-F172L-mCherry and STAT1wt-GFP and 50 cells expressing the latter combination including the cells shown in figure 50. The first two columns represent data that was already presented and discussed in earlier chapters (see figures 49) but was included here for comparison.*

Co-expression of STAT1 and STAT3 does not abrogate cytokine-induced tyrosine phosphorylation of either protein.

The diminished cytokine-induced nuclear accumulation of different STAT1-GFP constructs observed in co-expressing cells observed in the previous chapter, could be explained by reduced tyrosine phosphorylation at the cytokine receptor. Therefore, the phosphorylation status of the following STAT1/STAT3 combinations was assessed: STAT1wt-GFP + STAT3wt-mCherry, STAT1wt-GFP + STAT3-L78R-mCherry, STAT1-F172W-GFP + STAT3-L78R-mCherry, STAT1wt-GFP + STAT3ΔNLS-mCherry and STAT1-F172W-GFP + STAT3ΔNLS-L78R-mCherry. In three independent experiments, said combinations were expressed in U3R cells that were then treated with IFN-γ, IL-6 and sIL6Rα for 60 min. The cells were lysed and whole extracts were loaded onto two separate SDS-PAGE gels per independent repeat. After electrophoretic partitioning, the contents of the two gels were blotted onto separate nitrocellulose membranes. One membrane was incubated with antibodies raised against full-length STAT1 and the STAT1 pY701 phosphopeptide while the other was incubated with anti-STAT3 and anti-STAT3 pY705 phosphopeptide antibodies. Afterwards, the blots were incubated with IRDye®-conjugated secondary antibodies and blots were developed using a LICOR Odyssey® fluorescence imager.

Figure 52 shows Western blot images from one of the three independent experiments. The images showing signals from anti-STAT1 antibodies revealed a single specific band in each sample migrating at a height consistent with the theoretical molecular weight of GFP-tagged STAT1 (figure 52, panel 1). A single strong band migrating at the same height was detected in each sample when viewing the images representing anti-STAT1 pY701 signals (figure 52, panel 2). This suggested that all STAT1 constructs were phosphorylated at the cytokine receptor when co-expressed with different STAT3 constructs. When viewing the blot that was incubated with the anti-STAT3, a single specific band migrating at the theoretical molecular weight of STAT3-mCherry was observed in each sample (figure 52, panel 3). In the corresponding anti-STAT3 pY705 blot, a single band migrating at the same height was detected confirming phosphorylation of the different STAT3 constructs in response to cytokine treatment (figure 52, panel 4). Detection of β-Actin with a specific antibody was used to control the uniform loading of each sample among the two blots. Indeed, the β-Actin signals

detected for the blot that was incubated with the two antibodies raised against STAT1 (figure 52, panel 5) were similar to those detected in the blot that was incubated with STAT3 antibodies (figure 52, panel 6).

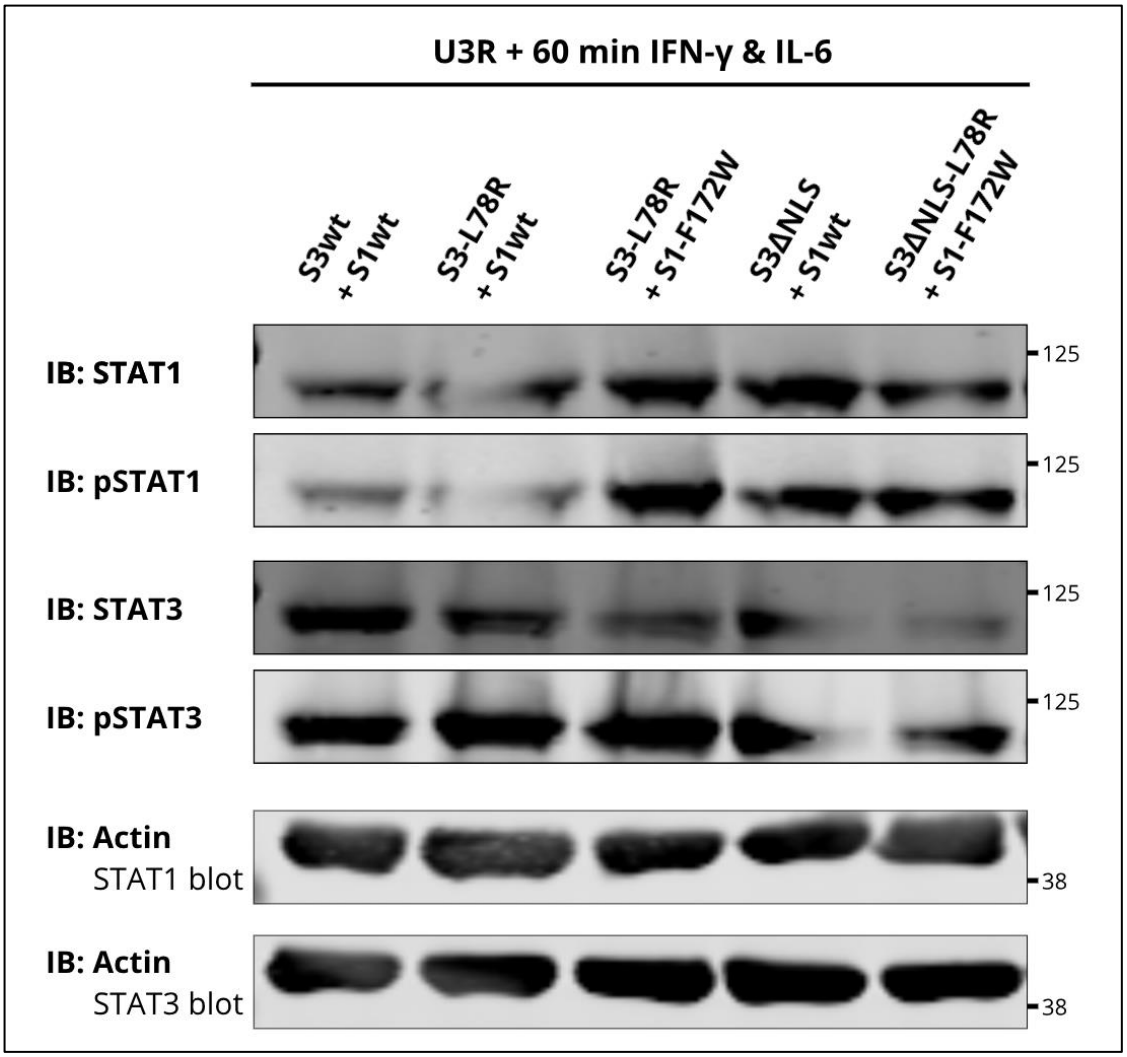


Figure 51 | Phosphorylation of different STAT1 and STAT3 constructs is not affected when they are co-expressed in different combinations: *U3R cells were co-transfected with different STAT1-GFP/STAT3-mCherry combinations and treated with cytokines. Whole cell extracts were loaded onto two different gels and subjected to Western blot analysis using anti-STAT1 and anti-STAT1 pY701 antibodies for one and anti-STAT3 and anti-STAT3 pY705 antibodies for the other. This figure is representative for three independent Western blot experiments. MW markers in kDa.*

To assess any differences in proportional phosphorylation among different STAT1/STAT3 combinations, Western blots from all three independent experiments were quantified and the ratio of phosphorylated STAT was calculated by dividing the signals detected using the antibodies raised against the STAT phosphopeptides by the signals detected using the corresponding antibodies raised against the full-length version of the proteins.

The phosphorylation of STAT1wt-GFP was not significantly different whether it was co-expressed with STAT3wt-mCherry, STAT3-L78R-mCherry (p=0.61) or STAT3ΔNLS-mCherry (p=0.19). STAT1-F172W-GFP overall showed significantly higher phosphorylation than STAT1wt-GFP, yet there was no significant difference whether it was co-expressed with STAT3-L78R-mCherry or STAT3ΔNLS-L78R-mCherry (p=0.49; figure 53, top graph). Overall, the STAT1 constructs were phosphorylated at comparative levels as when they were expressed alone (figures 21+38).

When focusing on the ratio of phosphorylated STAT3 co-expressed with above mentioned STAT1 constructs, the following observations were made. Compared to STAT3wt-mCherry co-expressed with STAT1wt-GFP, both STAT3-L78R-mCherry and STAT3ΔNLS-mCherry co-expressed with the same STAT1 construct showed significantly elevated phosphorylation. Moreover, introducing the L78R mutation in the latter construct resulted in a further increase in phosphorylation when co-expressed with STAT1-F172W-GFP, though this was not statistically significant (p=0.052; figure 53, bottom graph). Indeed, the absolute numbers calculated for the different STAT3 constructs co-expressed with STAT1 proteins were not much lower than those calculated for the same proteins when expressed alone (figure 22+43). However, it is worth pointing out that STAT3-L78R-mCherry showed significantly higher phosphorylation when co-expressed with STAT1-F172W-GFP than with STAT1wt-GFP (p=0.044; figure 53, bottom graph). Overall, the results presented here indicated that the co-expression of different STAT1 constructs did not affect the cytokine-induced tyrosine phosphorylation of specific STAT3 constructs.

In conclusion, given that neither different STAT1 nor STAT3 constructs showed diminished cytokine-induced phosphorylation in the experiments presented here, it is unlikely that this was the cause for reduced nuclear scores observed when different STAT1-GFP proteins were combined with different STAT3ΔNLS-mCherry constructs in the Nuclear Translocation Assay. In turn, this strengthens the evidence that said observations were indeed caused by the dominant negative cross-inhibitory effect exerted on the cytokine-induced nuclear accumulation of different STAT1-GFP constructs by different STAT3ΔNLS-mCherry constructs.

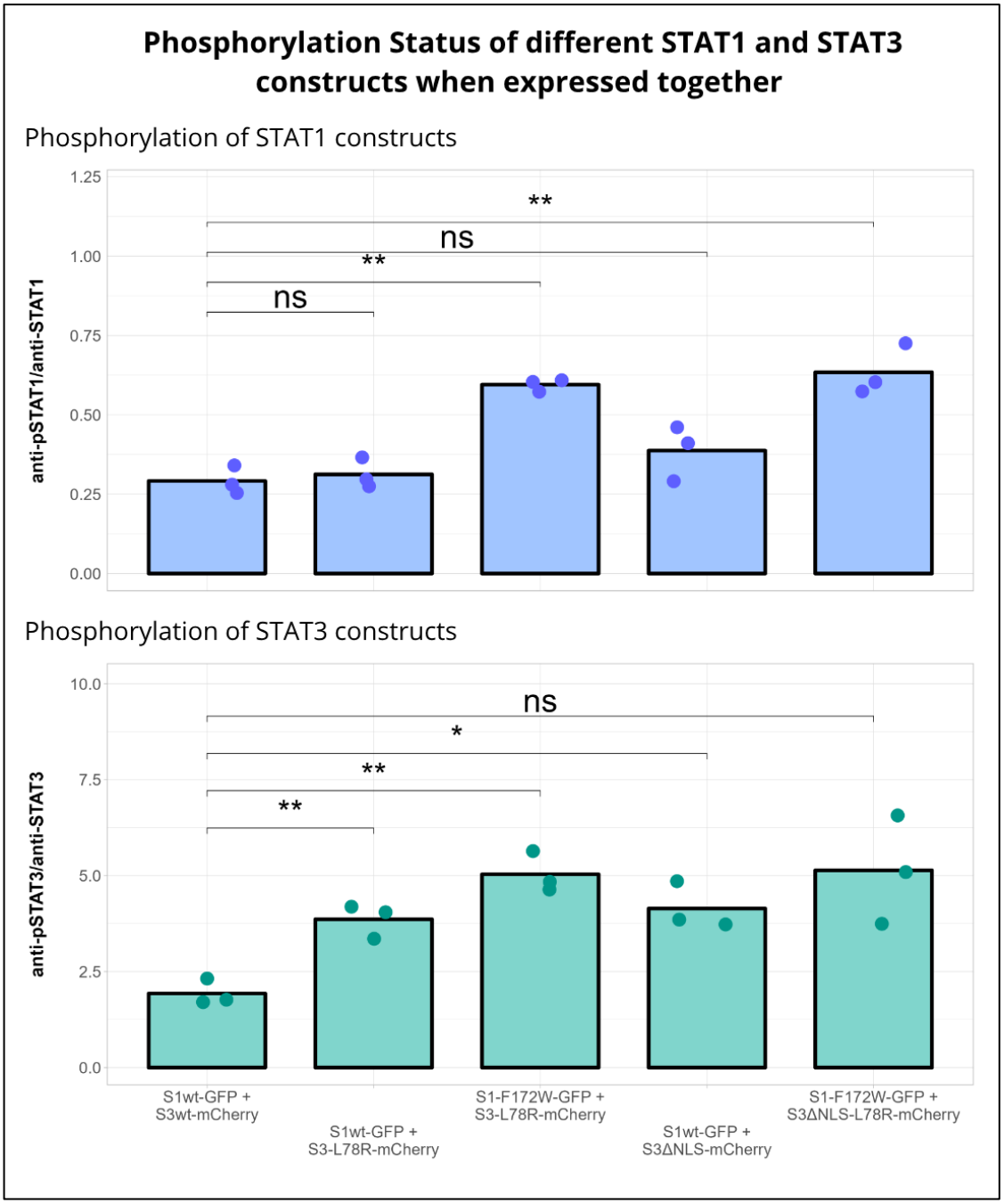


Figure 52 | Quantification of proportional phosphorylation of different STAT1 and STAT3 constructs when co-expressed together: *Western blots from three independent experiments were quantified and the ratio of phosphorylated protein calculated as indicated in the text. Different STAT1-GFP/STAT3-mCherry combinations were co-expressed in U3R cells which were subsequently treated with cytokine to induce STAT phosphorylation.*

Monomerising unphosphorylated STAT1 and STAT3 dimers does not affect the three-band pattern observed in EMSA experiments.

We have seen earlier and in previous studies that EMSA gels of extracts from cytokine-treated cells containing STAT1-GFP and STAT3-GFP show a three-band pattern. The three bands represented the STAT3-GFP homodimer (top band), STAT1-GFP homodimer (bottom band) and STAT1-GFP/STAT3-GFP heterodimer (intermediate band; figure 48). We have also seen from Nuclear Translocation Assay experiments that STAT1 and STAT3 remained capable of forming active heterodimers even when they carried an unphosphorylated dimer-splitting mutation (figure 50+51). However, the latter experiments inherently could not resolve if unphosphorylated dimer splitting changed the relative abundance of STAT1 and STAT3 homodimers in favour of STAT1/STAT3 heterodimers. To investigate this, U3R cells that, after co-transfection of either STAT1wt-GFP and STAT3wt-mCherry or STAT1-F172W-GFP and STAT3-L78R-mCherry, were treated with IFN-γ, IL-6 and sIL6Rα were subjected to EMSA experiments. Again, an IRDye®-conjugated M67 probe was used to assess DNA binding.

In the extract containing the wild-type STATs, a three-band pattern was detected that migrated at the same relative heights as in a previous experiment, indicating the presence of STAT1 and STAT3 homodimers and STAT1/STAT3 heterodimers (figure 54, lane 2). The lane loaded with extracts from cells containing STAT1-F172W-GFP and STAT3-L78R-mCherry likewise showed three bands that migrated at the same height as the previous sample (figure 54, lane 3). Indeed, incubating the extract with an antibody specific to native STAT1 before loading resulted in supershifts of the bottom two bands (figure 54, lane 4). Correspondingly, incubating with an antibody specific to native STAT3 supershifted the top two bands (figure 54, lane 5). This confirmed the top band as active STAT3 homodimers, the bottom band as active STAT1 homodimers and the intermediate band as active STAT1/STAT3 heterodimers. Interestingly, the bands observed in extracts containing the mutated constructs were much stronger than those observed for wild-type dimers. This is likely attributed to increased phosphorylation of the mutant proteins (figure 53). When comparing individual bands in the extract containing mutated STATs, the intermediate band appeared strongest with the top and bottom bands showing roughly equal fluorescence (figure 54, lane 3). The same pattern emerged in the extract containing the wild-type STAT (figure 54, lane

3 + figure 48 lane 5). This indicated that the introduction of latent dimer splitting mutations did not change the ratio between active STAT1 and STAT3 homodimers and STAT1/STAT3 heterodimers.

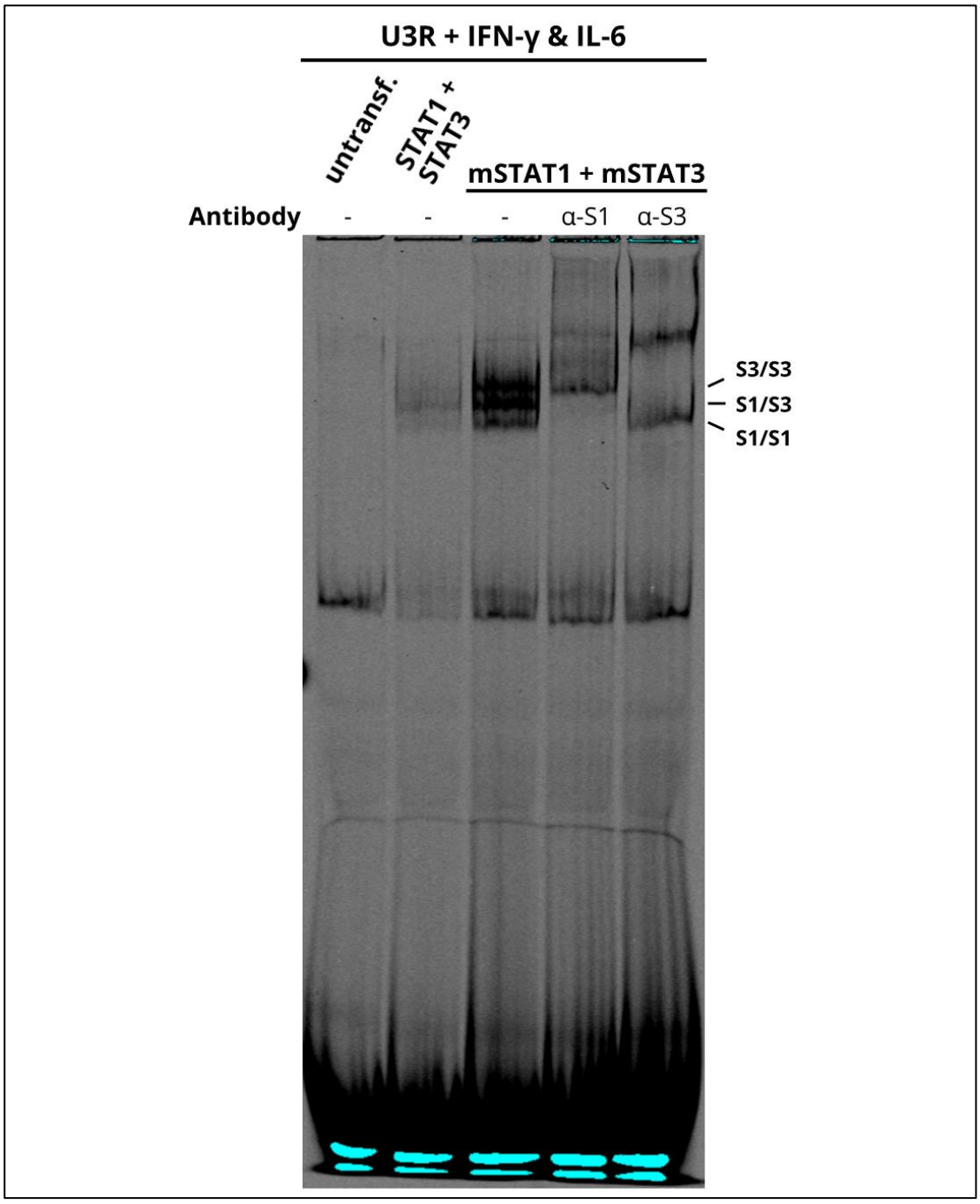


Figure 53 | The three band pattern seen in STAT1/STAT3 co-transfected cell extracts is not changed by unphosphorylated dimer splitting mutations: *Whole cell extracts of cytokine treated U3R cells that were left untransfected, co-transfected with STAT1-GFP and STAT3-mCherry or STAT1-F172W-GFP and STAT3-L78R-mCherry were subjected to EMSA experiments using the M67 DNA probe. Antibodies raised against native full-length STAT1 and STAT3 were used for supershifts.This figure is representative for three independent experiments.*

Active STAT1 and STAT3 homodimers do not readily exchange monomers once formed

Following the results from above which indicated that unphosphorylated dimers do not determine the relative abundance of active STAT1 and STAT3 homodimers over heterodimers or vice-versa, the focus was shifted towards investigating the stability of these dimers. To assess this, extracts from cytokine-treated U3R cells containing STAT1wt-GFP or STAT3wt-mCherry individually were mixed and then incubated with an M67 probe for 0, 30 or 90 min alongside extracts from cytokine-treated cells co-expressing the two proteins that were incubated with the probe for the same times. After individual incubation periods, the proteins were loaded onto a native polyacrylamide gel and separated electrophoretically to resolve their banding patterns. The reasoning behind this is that should active STAT1 and STAT3 homodimers and STAT1/STAT3 heterodimers readily exchange their subunits, the three-band pattern seen in extracts from co-transfected cells should also become detectable over time in the samples containing mixes of extracts containing STAT1-GFP or STAT3-mCherry individually. However, this was not the case. Instead, only two bands were observed in samples where STAT1 extracts were mixed with STAT3 extracts and the DNA probe for 0, 30 or 90 min (figure 55, lane 5, 7 and 9, respectively).On the contrary, the three-band pattern was detectable throughout the entire time course when extracts containing co-expressed STAT1-GFP and STAT3-mCherry were incubated with M67 for the same times (figure 55, lane 4, 6 and 8, respectively). In all cases, the two bands seen in the former samples co-migrated with the top and bottom bands of the latter, suggesting that these correspond to active STAT3 and STAT1 homodimers, respectively. The extracts containing only STAT1wt-GFP or STAT3wt-mCherry were also incubated with M67 individually and loaded onto the gel at time point 0. This confirms that both extracts indeed only contained STAT1wt-GFP or STAT3wt-mCherry (figure 55, lane 2+3, respectively).

Overall, these results suggested that active STAT1 and STAT3 homodimers are stable once formed at the cytokine receptor and do not readily exchange subunits with STAT1/STAT3 heterodimers.

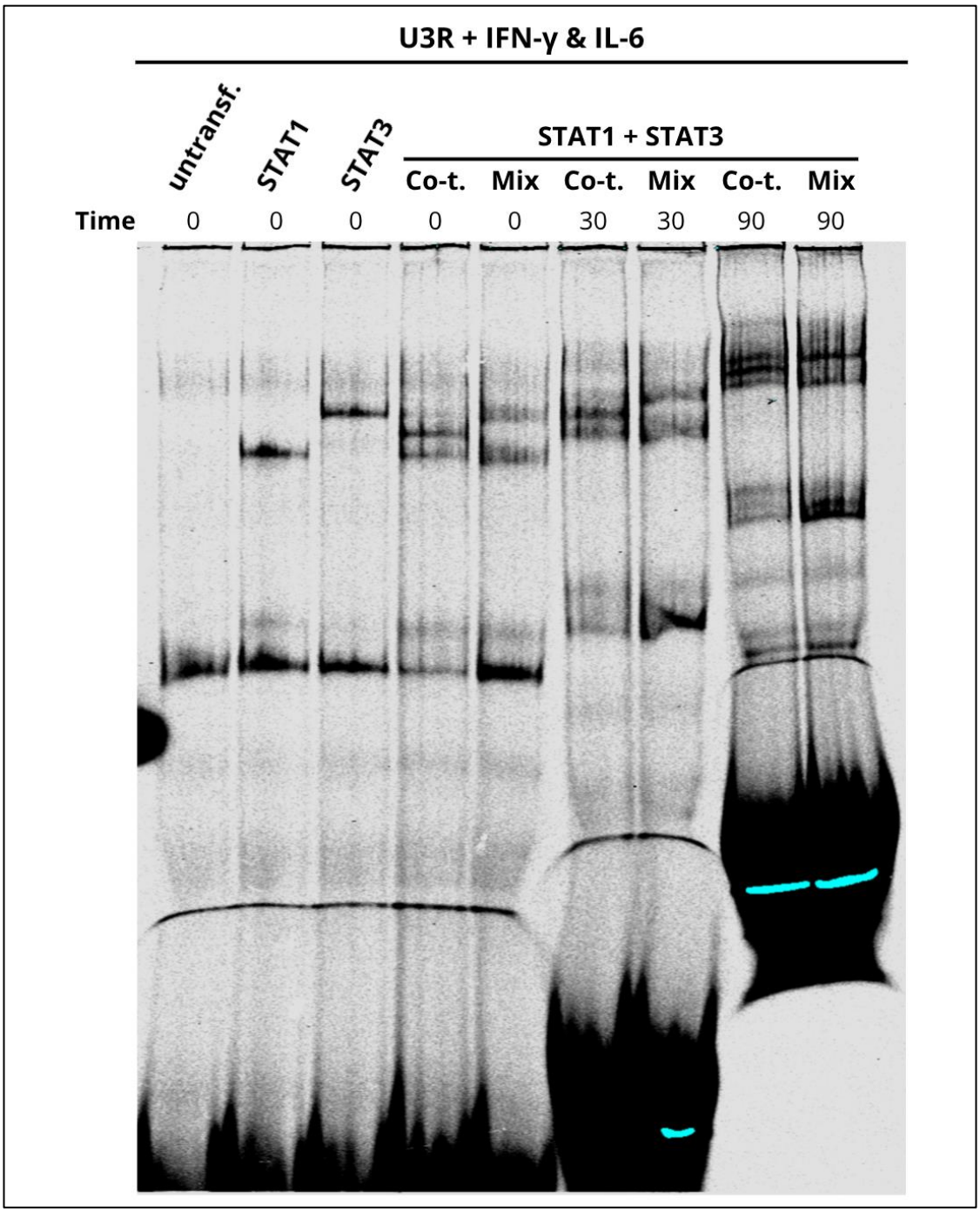


Figure 54 | Active STAT1 and STAT3 homodimers do not exchange subunits: *U3R cells were left untransfected, transfected with STAT1-GFP or STAT3-mCherry alone or co-transfected with both and subsequently treated with cytokines. Extracts from cells only containing one STAT were mixed (Mix) and incubated with M67 for the indicated times before loading on EMSA gels alongside extracts from co-transfected cells (Co-t.) that were incubated for M67 for the same time. The extracts only containing one STAT were also mixed with M67 individually and loaded on the same gel at time point 0. Electrophoretic separation was started at time point 0 and stopped at time point 30 min and 90 min for loading of samples, explaining the altered path of the last four lanes.*

4.4.3 Monomerization of latent dimers does not lead to the creation of artificial active STAT1/STAT5B heterodimers.

Evidence presented in chapter 4.2.4 confirmed previous literature accounts that STAT1 and STAT5B do not form active heterodimers. Considering the hypothesis that STAT dimers require the pre-association in unphosphorylated cells and merely undergo a conformation change to transition to the active form, the composition of the unphosphorylated dimers may pre-empt what active dimers can exist. Therefore, it may be possible to create artificial active heterodimers comprised of two STATs that normally do not form active dimers by the splitting of their unphosphorylated STATs. This is because the unphosphorylated dimer-splitting mutation should not affect the interaction surface of active STATs. Therefore, STATs that are monomeric before cytokine exposure should still be able to form active dimers. In the case when STAT1 and STAT5B, both carrying an unphosphorylated dimer splitting mutation, were combined, the two may randomly assemble into STAT1/STAT5B heterodimers. This investigated next and results are presented in the following.

Characterising the STAT5B mutation L82A

A set of conserved non-polar amino acids in the NTD of STATs (figure 56, red box) have been implicated in NTD-NTD interactions of unphosphorylated STAT1, STAT2, STAT3 and STAT4 (Vinkemeier *et al.*, 1998; Chen, 2003; Ota *et al.*, 2004; Mao *et al.*, 2005; Domoszlai *et al.*, 2014; Ho *et al.*, 2016). The same interface spanning STAT1 F77 and L78 has also been shown to be essential for the tetramerisation of different STATs, including STAT5B, on tandem DNA recognition sequences (Meyer *et al.*, 2004; Ota *et al.*, 2004; Lin *et al.*, 2012). While there is no specific evidence available showing that the corresponding residues STAT5B F81 and L82 are required for STAT5B NTD-NTD unphosphorylated dimer integrity, the fact that this interface is equally important for unphosphorylated dimer formation and tetramerization on tandem DNA sequences in other STAT paralogues facilitates the assumption that this likewise could be the case for STAT5B. Therefore, since STAT5B L82 showed the highest conservation amongst STATs (excluding STAT2 and STAT6; figure 56, violet) the decision was made to generate a STAT5B-L82A mutant that was assumed (and later confirmed; see chapter 4.5.4) to split STAT5B unphosphorylated homodimers. The following paragraphs

describe properties of this mutation regarding its cytokine-induced nuclear accumulation and tyrosine phosphorylation in detail.

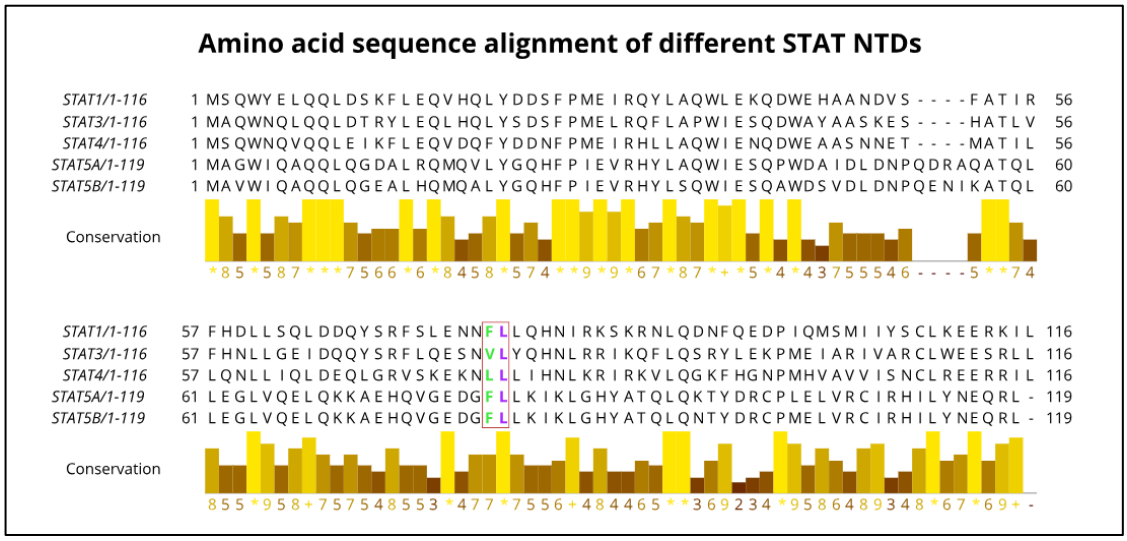


Figure 55 | Residues essential for STAT NTD-NTD interactions are conserved in most STATs: *The NTD of STAT1, STAT3, STAT4, STAT5A and STAT5B were aligned using the Clustal Omega alignment tool on the ebi.ac.uk and viewed in JalView software. Conserved residues corresponding to STAT1 F77 (green) and L78 (violet) which have been implicated in NTD-NTD interactions are indicated by a red box.*

To quantify STAT5B-L82A cytokine-induced nuclear accumulation in the presence of STAT1-F172W, the two proteins were co-expressed with GHR and Jak2 in HeLa cells and subjected to the Nuclear Translocation Assay. Before cytokine exposure, the STAT1-F172W-GFP showed a subcellular distribution accentuating the cytoplasm, which was previously observed for this protein in other contexts within this work (figure 57A, panel 1). The STAT5B-L82A-mCherry protein distributed pancellularly in latent cells which resembled what was observed for other STAT5B constructs before (figure 57A, panel 2). When the same cells were re-imaged 90 min after the addition of IFN-γ and GH, the STAT1-F172W-GFP was fully accumulated in the nucleus, indicating intact cytokine-induced nuclear accumulation (figure 57A, panel 5). In the same cells, the STAT5B-L82A-mCherry also showed a shift towards a more nuclear distribution. However, the accumulation was far less complete than seen for the STAT1 construct (figure 57A, panel 6). Still, this behaviour was not uncommon among other STAT5B constructs (figures 30+34), indicating that the L82A mutation did not negatively influence STAT5B in this regard. To confirm these qualitative observations, nuclear scores were determined for 44 cells co-expressing the two constructs at the right molar ratios. As expected, STAT5B-L82A-mCherry showed a nuclear score (1.3±0.05; figure 57B; Table 40) that was not significantly different from the one

observed for STAT5Bwt-mCherry (1.31±0.05) when co-expressed with STAT1wt-GFP (figure 34). As before, while its nuclear score was always lower than those of other STATs it should be noted that STAT5Bwt-mCherry showed a slightly higher nuclear score when it was co-expressed with other STAT5B constructs or expressed alone (figures 20+30). Still, these numbers and visual observations were clear indicators of uninhibited cytokine-induced nuclear accumulation.

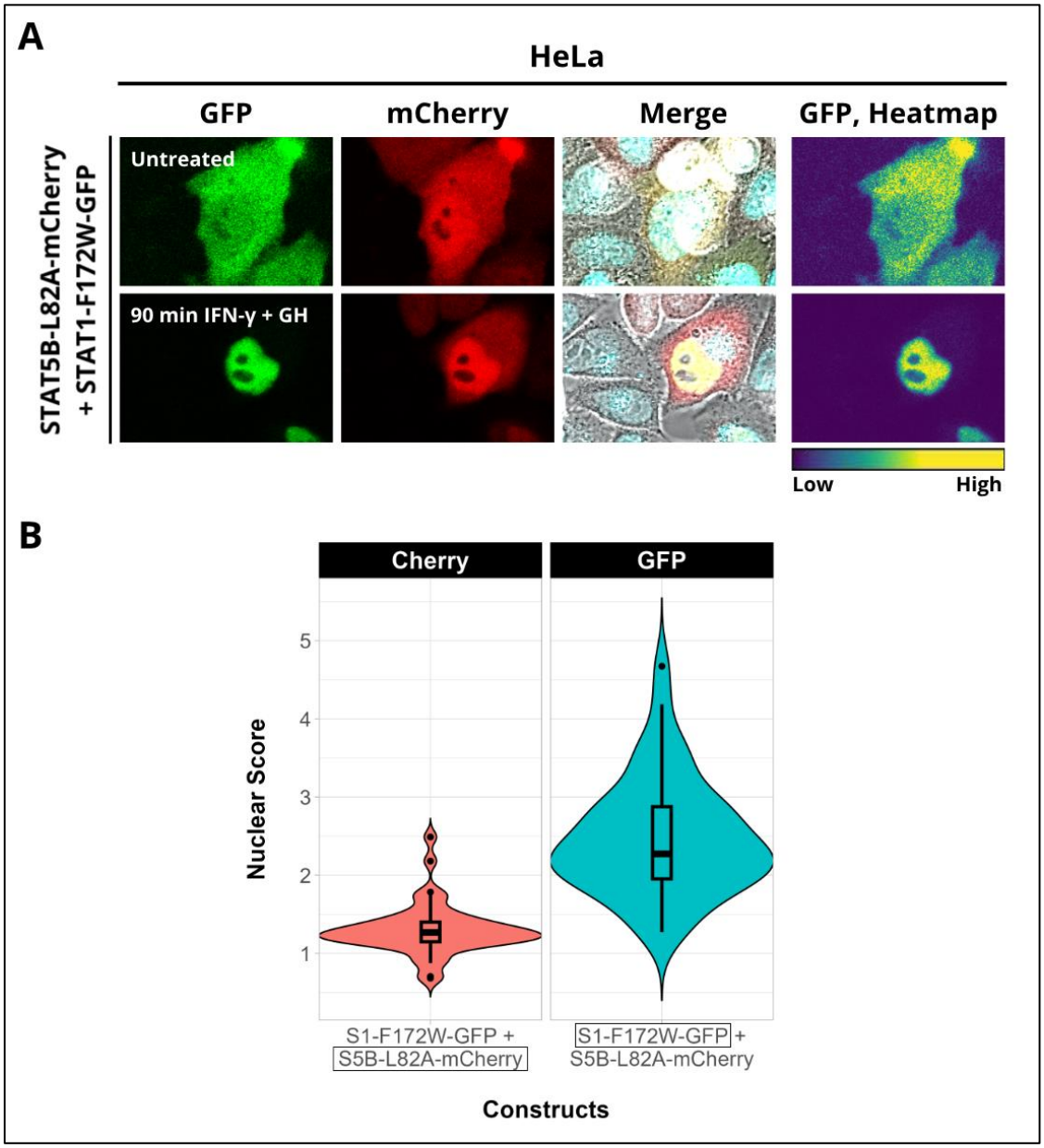


Figure 56 | STAT5B-L82A and STAT1-F172W accumulate inside the nucleus following cytokine treatment when expressed together. ***[A]*** *Fluorescence micrographs of HeLa cells co-expressing STAT5B-L82A-mCherry and STAT1-F72W-GFP imaged before and 90 min after cytokine treatment. The last column of panels shows a contrast enhanced depiction of the GFP channel. In this, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).* ***[B]*** *Quantification of mCherry and GFP nuclear scores in 44 cells co-expressing the combinations described in [A], respectively.*

The nuclear score of co-expressed STAT1-F172W-GFP (2.47±0.11) was significantly lower than that of STAT1wt-GFP (3.32±0.14) when co-expressed with STAT5Bwt-mCherry ($p<0.001$; figure 57B; table 40). However, visual observations again clearly showed full nuclear accumulation in response to cytokines. Overall, the results presented here showed that STAT5B-L82A-mCherry co-expressed with the STAT1 mutant accumulated inside the nucleus following cytokine treatment to a similar degree as STAT5Bwt-mCherry when it was co-expressed with STAT1wt-GFP. On the other hand, STAT1-F172W-GFP behaved slightly worse than its wild-type counterpart.

Table 38 | Descriptive statistics of nuclear scores determined for cells co-expressing STAT5B-L82A-mCherry and STAT1-F172W-GFP.

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT5B-L82A-mCherry + STAT1-F172W-GFP	44	1.3	0.05	2.5	0.1

STAT5B-L82A-mCherry shows sufficient cytokine-induced tyrosine phosphorylation when expressed alone or co-expressed with STAT1-F172W-GFP.

Next, to rule out any defects in cytokine-induced tyrosine phosphorylation, the phosphorylation status of STAT5B-L82A-mCherry was assessed when U3R cells expressed this construct alone or in conjunction with STAT1-F172W-GFP.

When extracts from cytokine-treated cells expressing the STAT5B-L82A-mCherry construct alone were subjected to Western blotting and incubated with antibodies raised against full-length STAT5 and the STAT5 pY694 phosphopeptide, the former revealed two strong bands running closely together at the expected height (figure 58A, panel 1). However, incubation with the anti-STAT5 pY694 antibody revealed only one band that co-migrated with the top band detected by the other antibody (figure 58A, panel 2). This pattern was also observed for other STAT5B constructs in single transfected cells (figure 31+35). To evaluate the ratio of phosphorylated STAT5B-L82A protein, Western blots from different experiments were quantified and the signal detected with the anti-STAT5 pY694 antibody was divided by the signal detected with the anti-STAT5 antibody (figure 58B). This revealed a non-significant elevation of

STAT5B-L82A-mCherry phosphorylation compared to single transfected STAT5wt-mCherry from earlier (p = 0.26; figure 23).

In extracts from cytokine-treated cells co-expressing STAT5B-L82A-mCherry and STAT1-F172W-GFP, detection with an anti-STAT1 antibody revealed one strong band migrating at the molecular weight expected for this construct (figure 58C, panel 1). Likewise, detection with an anti-STAT1 pY701 antibody also revealed a single band that co-migrated with the band detected with the other antibody (figure 58C, panel 2). When an anti-STAT5 antibody was used for detection on the same extracts, the two-band pattern also observed in extracts containing STAT5B-L78R-mCherry alone was observed (figure 58C, panel 4). Again, the incubation with an anti-STAT5 pY694 antibody showed that only the top band was phosphorylated (figure 58C, panel 5).

Western analysis for U3R cells co-expressing STAT5-L82A-mCherry and STAT3-F172W-GFP using STAT5 and STAT1 antibodies was done separately by loading the same extracts onto two different gels. Detection of β-Actin using a specific antibody confirmed that similar amounts of extract were loaded onto each gel (figure 58C, panel 3+6). As before, the ratio of phosphorylated protein was quantified for both co-expressed constructs. The co-expressed STAT5B-L82A-mCherry (figure 58D) showed a significant decrease in phosphorylation compared to cells expressing this construct alone (figure 58B). However, compared to STAT5Bwt-mCherry co-expressed with STAT1wt-GFP no significant difference was detected (p = 0.47; figure 35). This indicated that the mutation of residue L82 itself did not affect cytokine-induced tyrosine phosphorylation negatively but that the simultaneous expression of a STAT1 construct might have done to some degree. The STAT1-F172W-GFP that was co-expressed with the STAT5B-L82A-mCherry construct, showed a significant, approximately 2-fold, increase in phosphorylation compared to a STAT1wt-GFP that was co-expressed with STAT5Bwt-mCherry (figure 58D+35). Though compared to cells only expressing the STAT1-F172W-GFP the phosphorylation was roughly 2-fold decreased (p = 0.028). These results show that the co-expression of both STAT5B-L82A-mCherry and STAT1-F172W-GFP decreased the phosphorylation of both constructs. This may explain the reduced nuclear scores observed for both constructs in Nuclear Translocation Assay

experiments. However, since cytokine-induced nuclear accumulation was still sufficiently high, the STAT5B-L82A was progressed into further experiments.

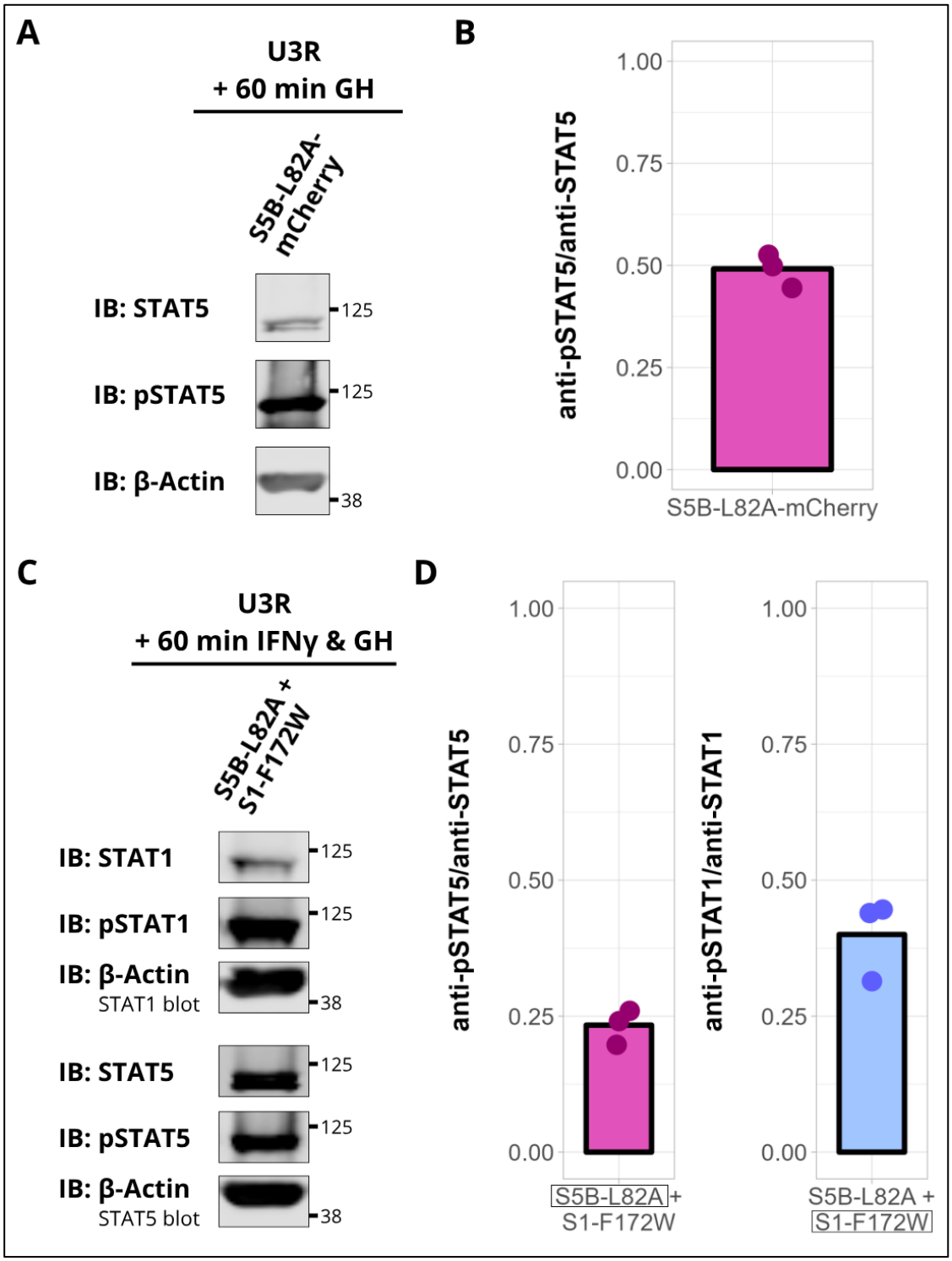


Figure 57 | STAT5B-L82A does not show deficient cytokine-induced tyrosine phosphorylation: ***[A]*** *Extracts from cytokine-treated U3R cells containing STAT5B-L82A-mCherry were subjected to Western blot analysis using the indicated antibodies.* ***[B]*** *Quantification of single-transfected STAT5B-L82A phosphorylation from three independent experiments.* ***[C]*** *STAT5B-L82A-mCherry and STAT1-F172W-GFP were co-expressed and extracts from cytokine-treated cells subjected to STAT1 and STAT5 specific Western blot analysis on two separate gels, respectively.* ***[D]*** *Quantification of co-expressed STAT5B-L82A-mCherry and STAT1-F172W-GFP phosphorylation from three independent experiments. MW markers in kDa.*

Nuclear Translocation Assay: Can active STAT1/STAT5B heterodimers be created artificially?

To test if artificial active STAT1/STAT5B heterodimers can be created by disrupting the ability of STAT1 and STAT5B to pre-form as unphosphorylated homodimers, HeLa cells co-expressing STAT5BΔNLS-L82A-mCherry and STAT1-F172W-GFP (and GHR and Jak2) were subjected to the Nuclear Translocation Assay. Images of these cells taken before the addition of cytokine showed the typical subcellular distribution expected for STAT1-F172W-GFP (pancellular with stronger fluorescence in the cytoplasm than in the nucleus) and STAT5BΔNLS-L82A-mCherry (pancellular with equally strong fluorescence throughout the cell; figure 59A, row 1). When the same cells were imaged again, 90 min after the addition of IFN-γ and GH, the STAT1-F172W-GFP protein was seen fully accumulated inside the nucleus while STAT5BΔNLS-L82A-mCherry did not change subcellular distribution compared to images taken before cytokine exposure (figure 59A, row 2). This suggested that the latter construct could not cross-inhibit the cytokine-induced nuclear accumulation of STAT1-F172W because active heterodimers could not form between the two proteins.

To further investigate this notion, images of 38 cells co-expressing the two proteins at the right molar ratios were quantified and nuclear scores were calculated according to the Nuclear Translocation Assay protocol. Figure 59B shows a graph reporting said nuclear scores for this combination alongside other STAT1/STAT5 combinations that have already been discussed in other chapters for comparison. In cells expressing STAT5BΔNLS-L82A-mCherry and STAT1-F72W-GFP, the latter showed a high nuclear score (3.15±0.14) indicative of unimpaired cytokine-induced nuclear accumulation (figure 59B, fourth column; table 41). Indeed, the STAT1-F172W-GFP nuclear score in this combination was not significantly different from nuclear scores determined for STAT1wt-GFP when co-expressed with STAT5Bwt-mCherry (GFP nuclear score 3.32±0.14) or STAT5BΔNLS-mCherry (GFP nuclear score 2.91±0.08; figure 59B; table 41). Interestingly, the STAT1-F172W-GFP nuclear score was significantly higher when it was combined with STAT5BΔNLS-L82A-mCherry than when it was combined with STAT5B-L82A-mCherry (GFP nuclear score 2.47±0.11; figure 59B; table 41). While this is a counterintuitive result, the difference was low and the STAT1 variant still showed

a sufficiently high nuclear score to assume good cytokine-induced nuclear accumulation, especially when considering visual observations of individual images.

STAT5BΔNLS-L82A-mCherry showed a nuclear score of 1.18±0.02 when combine with STAT1-F172W-GFP (figure 59B; table 41). This was significantly higher than that observed for STAT5BΔNLS-mCherry when combined with STAT1wt-GFP, which showed a nuclear score of 0.88±0.02, indicative of cytokine-induced redistribution towards the cytoplasm (figure 59B; table 41). Indeed, based on its nuclear score and visual observations, the STAT5BΔNLS-L82A-mCherry did not show this phenotype (figure 59). The difference between the nuclear scores of STAT5BΔNLS-L82A-mCherry, STAT5Bwt-mCherry and STAT5B-L82A-mCherry (all in combination with different STAT1 constructs) were either non-significant or of low significance (figure 59; table 41). This further highlights the issue with full cytokine-induced nuclear accumulation of different uninhibited STAT5B constructs in this model. However, a visual inspection of the images clearly showed some degree of cytokine-induced nuclear accumulation of the latter two constructs and the lack thereof in the former's case (figure 34, 57 and 59).

Table 39 | Descriptive statistics of nuclear scores determined for cells co-expressing STAT5BΔNLS-L82A-mCherry and STAT1-F172W-GFP. *Descriptive statistics of cells co-expressing other STAT1/STAT5B combinations were discussed before (table 31+40) but are shown here again for reasons of comparison.*

Constructs	Number of cells	Nuclear Score mCherry		Nuclear Score GFP	
		Mean	Std Error	Mean	Std Error
STAT5Bwt-mCherry + STAT1wt-GFP	63	1.3	0.05	3.3	0.1
STAT5B-L82A-mCherry + STAT1-F172W-GFP	44	1.3	0.05	2.5	0.1
STAT5BΔNLS-mCherry + STAT1wt-GFP	77	0.9	0.02	2.9	0.1
STAT5BΔNLS-L82A-mCherry + STAT1-F172W-GFP	38	1.2	0.02	3.2	0.1

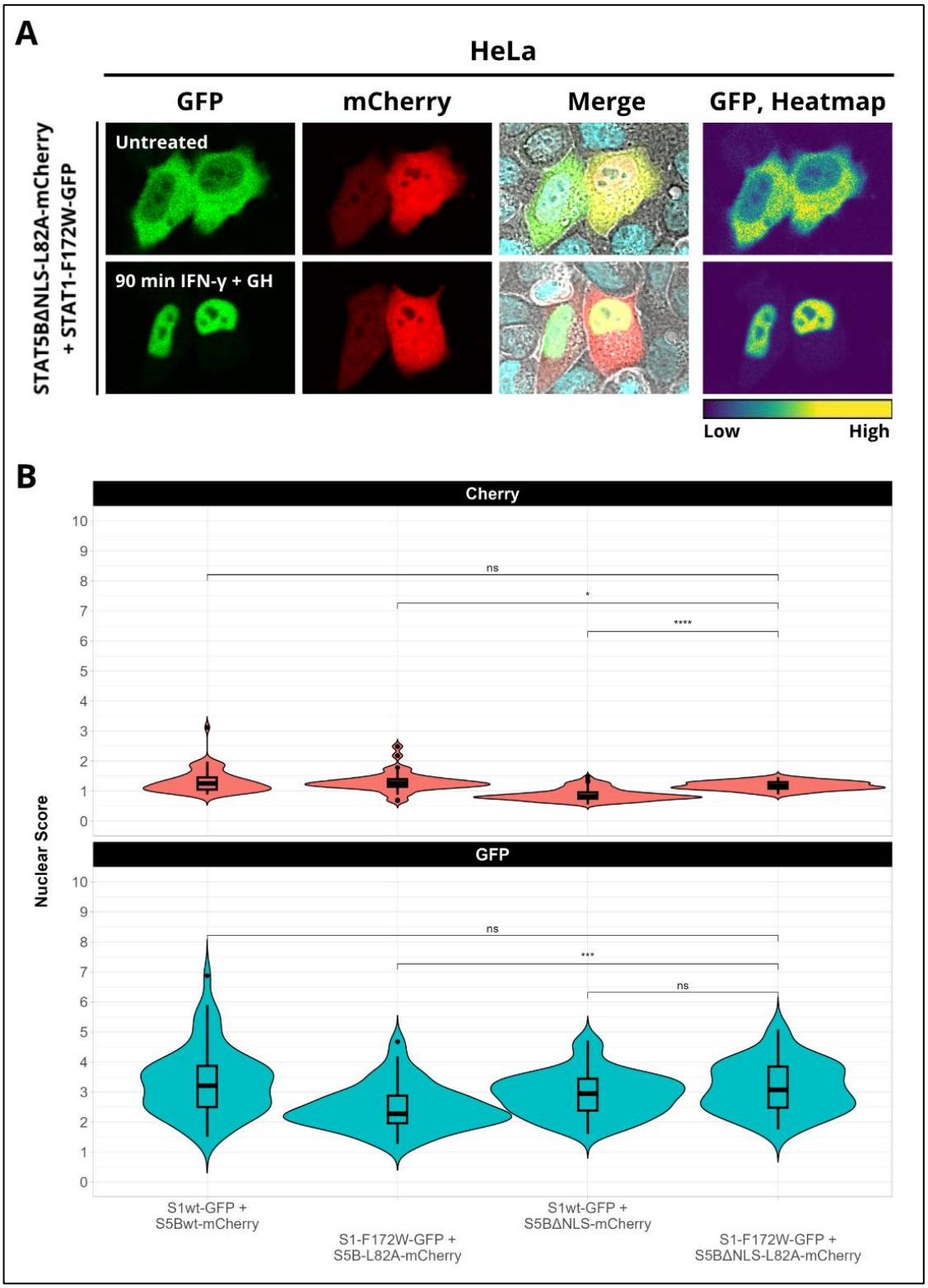


Figure 58 | Active STAT5BΔNLS-L82A/STAT1-F172W could not be detected in the Nuclear Translocation Assay: ***[A]*** *Fluorescence micrographs of HeLa cells co-expressing STAT5B-L82A-mCherry and STAT1-F172W-GFP imaged before and 90 min after cytokine treatment. The last column of panels shows a contrast enhanced depiction of the GFP channel. In this, the continuous colour gradient highlights the fluorescence intensity detected for individual pixels (low intensity: blue; high intensity: yellow).* ***[B]*** *Quantification of mCherry and GFP nuclear scores in 44 cells co-expressing the combinations described in [A]. The first three columns represent different STAT1/STAT5 combination that have been shown before (figure 34+57) and were included for comparisons only.*

STAT5BΔNLS-L82A-mCherry and STAT1-F172W-GFP show sufficient cytokine-induced tyrosine phosphorylation when co-expressed.

To rule out that diminished cytokine-induced tyrosine phosphorylation was the cause for the inability of STAT5BΔNLS-L82A-mCherry to cross-inhibit STAT1-F172W-GFP nuclear accumulation via direct interaction within active heterodimers, the phosphorylation of the two proteins was checked when co-expressed or when the former protein was expressed alone (with GHR and Jak2) in U3R cells. For this, said cells were activated by stimulation with GH only or a combination of IFN-γ and GH for 60 min and lysed thereafter to generate whole cell extracts.

Detecting full-length STAT5B with a specific antibody in cell extracts containing only STAT5BΔNLS-L82A-mCherry revealed two specific bands running very close together at a speed consistent with its molecular weight (figure 60A, panel 1). When detecting the STAT5B pY694 phosphopeptide, only one band was visible that co-migrated with the top band seen with the other antibody suggesting that only this STAT5B band was phosphorylated (figure 60A, panel 2). The phosphorylation of this protein was quantified from three independent experiments and revealed a non-significant increase compared to STAT5Bwt-mCherry being expressed alone before ($p=0.23$; figure 60B). The same extracts from activated cells co-expressing STAT5BΔNLS-L82A-mCherry and STAT1-F172W-GFP were loaded onto two different SDS-PAGE gels for separate detection with anti-STAT5 and anti-STAT5 pY694 or anti-STAT1 or anti-STAT1 pY701 antibodies. The additional detection of β-Actin confirmed equal loading of the same extracts (figure 60C, panel 3+6). In these extracts, the two STAT1 antibodies each revealed a single strong band migrating at the speed expected for the STAT1-F172W-GFP construct (figure 60C, panel 1+2). The banding pattern detected with the STAT5 antibodies resembled that detected for STAT5BΔNLS-L82A-mCherry when expressed alone (figure 60C, panel 4+5). This suggested that both the STAT5B and STAT1 constructs were tyrosine phosphorylated efficiently upon cytokine exposure when expressed together. To confirm these observations, phosphorylation was quantified as before from three independent experiments. Thereby, it was revealed that the STAT5BΔNLS-L82A-mCherry showed significantly lower phosphorylation when co-expressed with STAT1-F172W-GFP than when expressed alone ($p=0.023$; figure 60D). However, the level of phosphorylation of this construct in the former combination was

not significantly different from other STAT5B constructs co-expressed with various STAT1 proteins (versus STAT5Bwt-mCherry+STAT1wt-GFP: p=0.39, versus STAT5BΔNLS-mCherry+STAT1wt-GFP: p=0.86, versus STAT5B-L82A-mCherry+STAT1-F172W-GFP: p=0.3; see also figures 35+58). Furthermore, it is worth pointing out that a STAT5BΔNLS-mCherry co-expressed with untagged STAT5Bwt showed similar phosphorylation levels (p=0.94; figure 31), which is also lower than what STAT5Bwt-mCherry when expressed alone. Yet, this protein was still able to interact with and inhibit the cytokine-induced nuclear accumulation of STAT5Bwt-GFP. Thus, while the proportion of phosphorylated STAT5BΔNLS-L82A-mCherry co-expressed with STAT1-F172W-GFP could be higher, it was deemed sufficient to rule out that reduced phosphorylation was the reason for the inability to inhibit the cytokine-induced nuclear accumulation of the STAT1 construct. The STAT1-F172W-GFP itself showed good phosphorylation when combined with STAT5BΔNLS-L82A-mCherry (figure 60D). Indeed, its phosphorylation was not significantly different from when it was combined with STAT5B-L82A-mCherry (p=0.83; figure 58) or when untagged STAT1-F172W was combined with STAT1ΔNLS-F172L (p=0.84; figure 21). However, it is worth noting that STAT1-F172W-GFP was significantly better phosphorylated when it was expressed alone (p=0.036; figure 38). Still, considering the similar values found for various STAT1-F172W co-expressions and the good nuclear accumulation, this was deemed not detrimental.

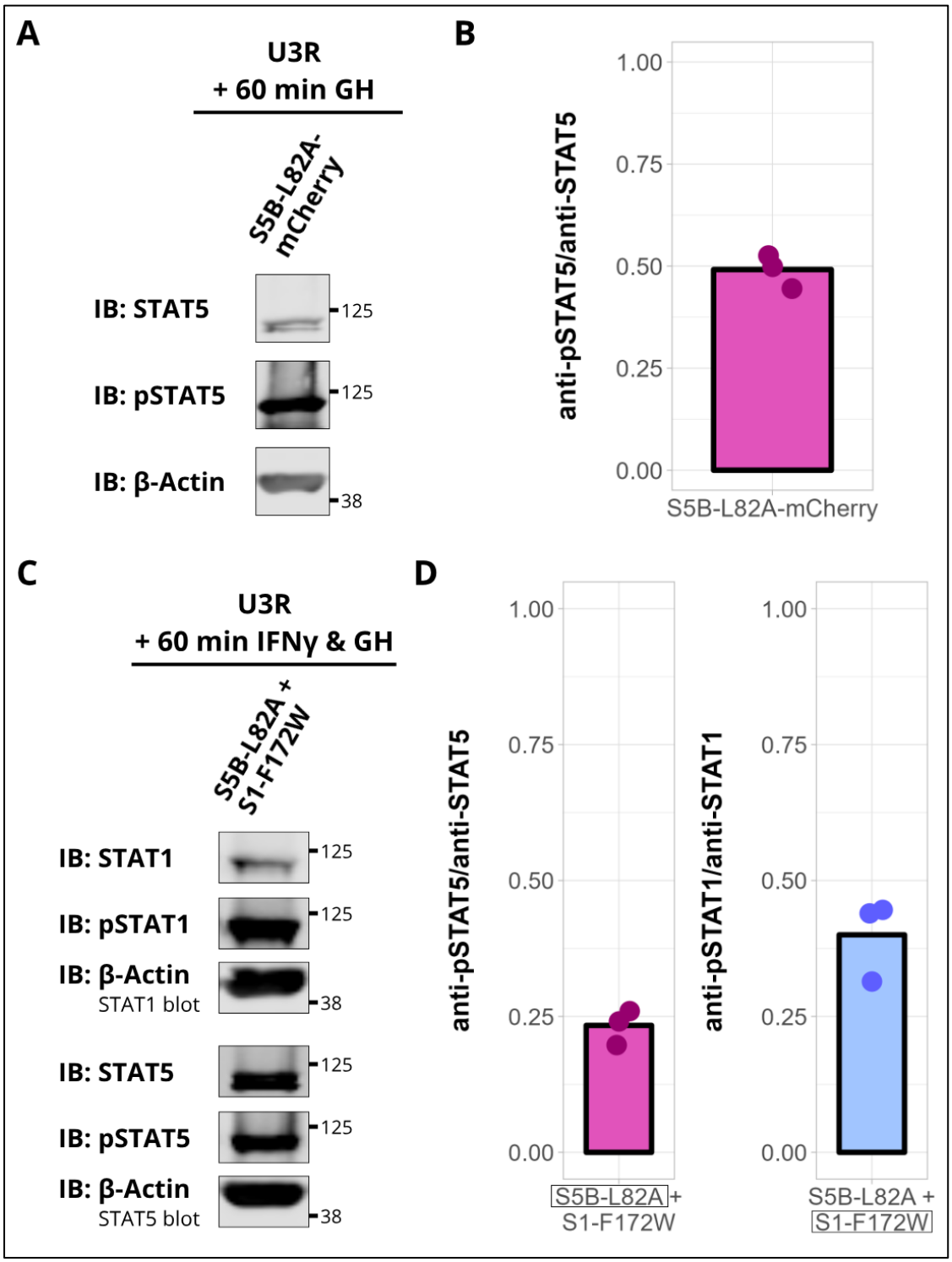


Figure 59 | STAT5BΔNLS-L82A does not show deficient cytokine-induced tyrosine phosphorylation: ***[A]*** *Extracts from cytokine-treated U3R cells containing STAT5BΔNLS-L82A-mCherry were subjected to Western blot analysis using the indicated antibodies.* ***[B]*** *Quantification of single-transfected STAT5BΔNLS-L82A phosphorylation from three independent experiments.* ***[C]*** *STAT5BΔNLS-L82A-mCherry and STAT1-F172W-GFP were co-expressed and extracts from cytokine-treated cells subjected to STAT1 and STAT5 specific Western blot analysis on two separate gels, respectively.* ***[D]*** *Quantification of co-expressed STAT5BΔNLS-L82A-mCherry and STAT1-F172W-GFP phosphorylation from three independent experiments. MW markers in kDa.*

Disruption of STAT1 and STAT5 unphosphorylated dimers does not lead to the appearance of an active STAT1/STAT5 heterodimer band in EMSA experiments.

The evidence presented in the previous few paragraphs indicated that STAT1 and STAT5B did not interact as active heterodimers even when their unphosphorylated dimers were monomerised by mutation. To confirm these results *in vitro,* active heterodimer formation was assessed in EMSA experiments. For this, U3R cells were either left untransfected or co-transfected with STAT1wt-GFP and STAT5B-mCherry or STAT1-F172W-GFP and STAT5B-L82A-mCherry (both also with plasmid DNA coding for GHR and Jak2). After 24 h of incubation, the cells were treated with IFN-γ and GH for 60 min before lysis to generate detergent-free whole-cell extracts. Said extracts were then incubated with an IRDye®680-conjugated GRR DNA probe, antibodies added for supershift experiments where appropriate, and samples ran on a native gel for electrophoretic separation. When the gel was developed, a double band was observed in extracts containing STAT1wt-GFP and STAT5B-mCherry (figure 61, lane 2). The two bands migrated at a similar pace as observed in previous experiments where extracts containing the same combination of proteins were investigated (figure 33). Therefore, the top band was identified as STAT5Bwt-mCherry and the bottom band as STAT1wt-GFP homodimer. A double-band pattern was also observed in cell extracts containing STAT1-F172W-GFP and STAT5B-L82A-mCherry and both bands co-migrated with the ones observed in the previous extracts (figure 61, lane 3). Indeed, the bottom band was only supershifted using an anti-STAT1 antibody (figure 61, lane 4) and the top band was exclusively supershifted with an anti-STAT5 antibody (figure 61, lane 5), identifying them as the mutant active STAT1 and STAT5B homodimers, respectively. The two bands visible in extracts containing STAT mutants were notably stronger than those observed in wild-type STAT1 and STAT5B-containing extracts. This is likely due to the increased phosphorylation that STAT5B-L82A and STAT1-F172W showed in comparison to the wild-type constructs. More importantly, no intermediate band was detected in either extract, suggesting that no STAT1/STAT5B heterodimers could be formed. This confirmed earlier observations made in Nuclear Translocation Assay experiments that disruption of unphosphorylated STAT1 and STAT5B dimers did not lead to the formation of artificial active STAT1/STAT5B heterodimers.

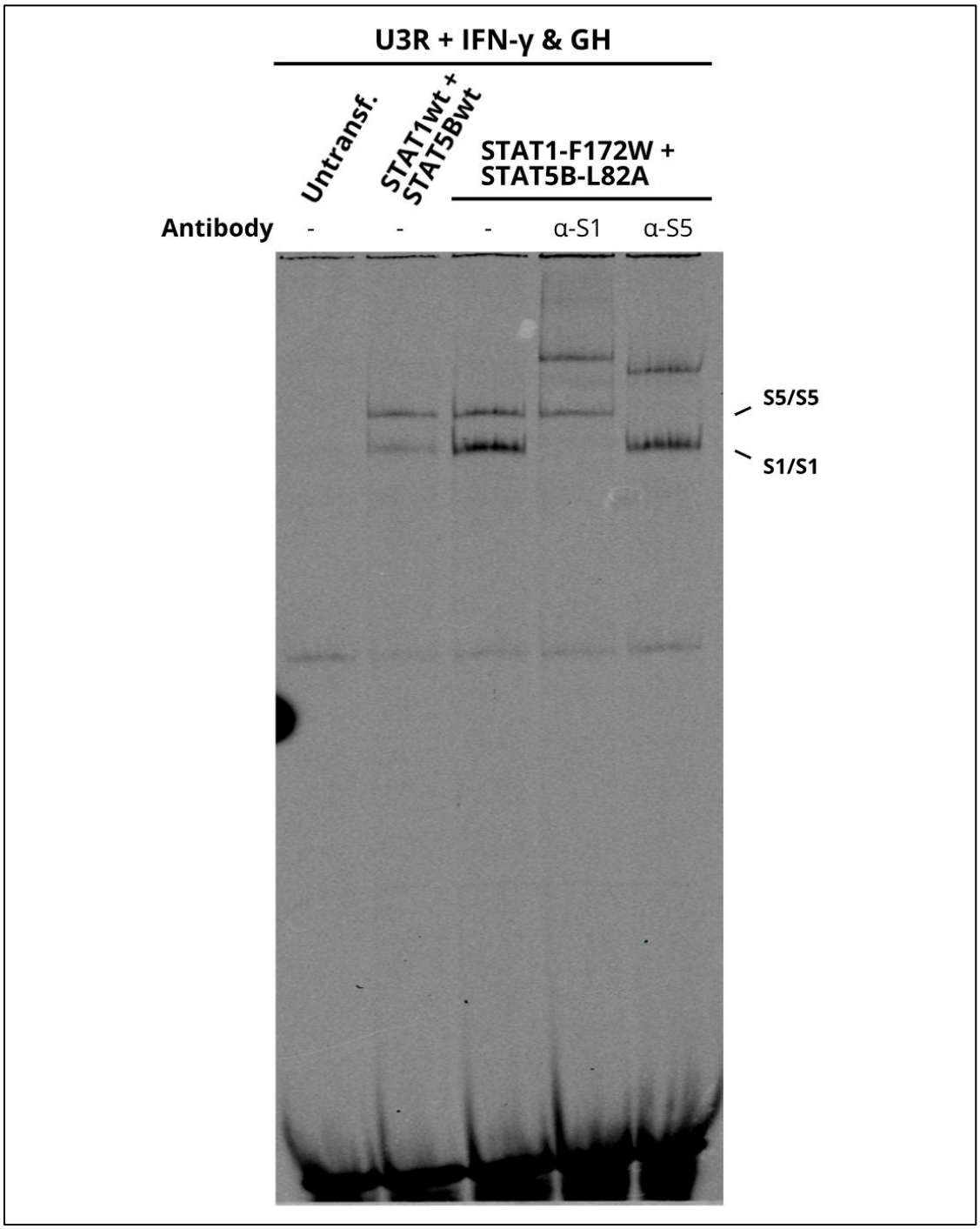


Figure 60 | Co-expression of STAT1-F172W and STAT5B-L82A leads to the formation of active homodimers but not heterodimers: *Whole cell extracts of cytokine treated U3R cells that were left untransfected or co-transfected with STAT1wt-GFP, STAT5Bwt-mCherry, GHR and Jak2 or STAT1-F172W-GFP, STAT5B-L82A-mCherry, GHR and Jak 2were subjected to EMSA experiments using the GRR DNA probe. Antibodies raised against native full-length STAT1 and STAT5 were used for supershifts. This figure is representative of three independent experiments.*

4.4.4 Discussion of Chapter 4.4

By applying both *in vitro* investigations and the Nuclear Translocation Assay to STAT heterodimers, this chapter sought to assess whether the pre-formation of unphosphorylated STAT dimers determines the contents of active STAT signalling complexes. However, the experiments discussed here indicated that that was not the case. Rather, it was found that active STAT1/STAT3 heterodimers formed readily independent of unphosphorylated dimer splitting in STAT1 F172W and/or STAT3 L78R mutants. What's more, is that artificial active STAT1/STAT5B heterodimers could not be created by disrupting their respective unphosphorylated homodimers.

STAT1/STAT3 heterodimers transition by disruption of latent dimers and reformation as their active form

Active STAT1/STAT3 heterodimer formation was first observed in EMSA experiments as early as when STAT3 itself was discovered (Zhong, Wen and Darnell, 1994a, 1994b). These findings were replicated in several reports in the following using different methods (Delgoffe and Vignali, 2013) and could also be reproduced in this work using both EMSA and the Nuclear Translocation Assay. The latter experiments showed that the STAT3ΔNLS-mCherry could cross-inhibit the STAT1wt-GFP cytokine-induced nuclear accumulation to a similar degree as STAT1ΔNLS-mCherry (figure 24+49). This may indicate that the probability for STAT1 and STAT3 to be integrated into active STAT1/STAT3 heterodimers vs active homodimers is the same. Random assembly at equal rates was previously also found for STAT5A homodimers, STAT5B homodimers and STAT5A/STAT5B heterodimers (Boehm *et al.*, 2014). The finding that the STAT3ΔNLS-mCherry could cross-inhibit the cytokine-induced nuclear accumulation of STAT1wt-GFP further confirmed that the STAT3ΔNLS mutation was dominant negative in active STAT1/STAT3 heterodimers.

To test the notion that STAT1/STAT3 heterodimers transition from unphosphorylated to active dimers via conformational change without the exchange of monomers, four different scenarios were considered as outlined in the theoretical considerations in chapter 4.4.1. Essentially, if this notion was true then the nuclear accumulation of STAT1-GFP in response to cytokines would be rescued when either the STAT3ΔNLS-mCherry or STAT1-GFP carries a unphosphorylated dimer splitting mutation. However,

this was contradicted by experiments presented in this chapter. Instead, the nuclear accumulation of different STAT1-GFP was cross-inhibited by different STAT3ΔNLS-mCherry regardless of whether the former, the latter or both carried the F172W or L78R mutation (figure 50+51). Hence, a scenario where unphosphorylated dimers split and reform as active dimers is favoured. Consequently, active STAT1/STAT3 heterodimers formed independent of unphosphorylated dimer integrity. Importantly, the results obtained here also suggested that STAT1/STAT3 heterodimers can form from monomers that were originally bound in latent STAT1 or STAT3 homodimers. Yet, the experiments could not resolve at which point in the STAT activation cycle unphosphorylated dimers were split and whether STATs recruited to different cytokine receptors could mix.

In their review, Delgoffe and Vignali (2013) propose that the structure of cytokine receptors or the proximity of different receptor subunits or even whole receptors may contribute to the formation of STAT heterodimers. Either scenario would indeed constitute a valid explanation for the findings of this work concerning STAT1/STAT3 heterodimers. Both the IL-6 receptor complex and IFNGR have been shown to activate STAT1 and STAT3 simultaneously (Lütticken *et al.*, 1994; Heinrich *et al.*, 1998; Qing and Stark, 2004; Kishimoto and Kang, 2022). In any case, the results generated in this work indicate that mixing of monomers must occur during the activation cycle and not thereafter since active STAT1 and STAT3 homodimers were stable once formed (figure 55). This is contrary to STAT1 unphosphorylated dimers which were shown to readily change their conformation from antiparallel to parallel via dissociation and re-association, thus potentially enabling the exchange of monomers in before cytokine exposure (Wenta *et al.*, 2008).

Delgoffe and Vignali (2013) further propose post-translational modifications as a third mechanism that may determine the formation of heterodimers over homodimers. Indeed, a recent study found that phosphorylation of STAT2 T404 destabilises the otherwise very stable (Wenta *et al.*, 2008; Ho *et al.*, 2016) latent STAT1/STAT2 heterodimer and leads to a boost in type I IFN responses (Wang *et al.*, 2021). It is conceivable that this mechanism may also alter the ratio of active STAT1/STAT2 heterodimers to STAT1 homodimers.

More experiments are required to assess the exact timing and mechanism for monomer exchange at the cytokine receptor for STAT1 and STAT3.

Latent dimer splitting did not lead to the creation of artificial active STAT1/STAT5B heterodimers.

It is well established that active STAT1/STAT5B heterodimers do not form in response to cytokines (Delgoffe and Vignali, 2013). As stated in the hypothesis section of this work, the notion that STAT dimers may not exchange monomers during the transition from latent to active dimers means that the existence of individual active dimer species may be limited by the cellular pool of STAT dimers formed before cytokine exposure. Therefore, it was reasoned that artificial heterodimers may be created by disrupting the unphosphorylated homodimers formed by individual interaction partners. Yet, this theory was discarded, at least for STAT1 and STAT5B, based on results obtained from EMSA and Nuclear Translocation Assay experiments involving STAT1-F172W and STAT5B-L82A, which failed to detect *de novo* active STAT1/STAT5B heterodimers (figure 59+61). Thus, regulatory elements beyond unphosphorylated dimers that determine the contents of active heterodimers are likely to exist. Indeed, the compatibility of different STAT phosphopeptides with SH2 domains probably plays a major role in coordinating the formation of active heterodimers. A STAT5 phosphopeptide (AVDGY^{P}VKPQIK) showed a limited ability to interact with STAT1 compared to phosphopeptides derived from other family members in a surface plasmon resonance experiment (Greenlund *et al.*, 1995). Moreover, multiple computational models have identified several residues in the SH2-domain of STAT5 that provide selectivity for phosphopeptide binding compared to other STAT family members including STAT1 and STAT3 (Gianti and Zauhar, 2015; Langenfeld *et al.*, 2015; de Araujo *et al.*, 2019). However, more research is required to confirm this hypothesis.

4.5 Detecting unphosphorylated STAT dimers using the Co-localisation Assay

Parallel to experiments focusing on active STAT dimer formation, we were also interested in finding out more about unphosphorylated STAT dimers. In general, this form is notably less well-researched compared to its active counterpart. While there is considerable evidence that shows that unphosphorylated dimers play important biological roles, the existence or absence of an unphosphorylated dimeric form has only been confirmed for a handful of possible homo and heterodimeric combinations (Sehgal, 2008; Yang and Stark, 2008). As explained earlier, EMSA gel shift experiments have been utilised a lot for investigations into active STAT dimers due to their accessibility and low cost. However, since unphosphorylated STAT dimers do not or only weakly bind to specific DNA sequences (Wenta *et al.*, 2008), this tool cannot be used for functional assessment of these entities. Therefore, we set out to close this knowledge gap by creating a novel cell-based assay termed the Co-localisation Assay This Assay was used to systematically assess all proteins within the STAT family for their ability to form unphosphorylated homo- and heterodimers. It is worth noting that all experiments shown in the following were conducted by Dr. Andreas Begitt. My contributions to this project focused on conceptualising the microscopy and image analysis workflow based on experience gained from the Nuclear Translocation Assay experiments. Furthermore, I led the statistical analysis and data processing of experimental outputs. Thus, my contributions and the relevance of the results to other aspects of this work presented so far warranted the inclusion of selected aspects of this study. The work presented below, and more is published in Begitt et al., 2023.

4.5.1 Introduction

The concept behind the Nuclear Translocation Assay as applied to active STAT dimerization was based on cross-inhibition of cytokine-induced nuclear accumulation of a STAT-GFP exerted by the dominant negative effect of a STATΔNLS-mCherry mutant if the two proteins could assemble active dimers. However, when investigating latent STATs, a readout that is based on cytokine-induced phenomena cannot be used. Instead, we turned towards the intrinsic ability of unphosphorylated STATs to constitutively shuttle between the nucleus and cytoplasm via different mechanisms

(Vinkemeier, 2004; Reich, 2013) resulting in a mostly pancellular distribution as apparent in various images presented in this work. Work by the Köster group shows that the natural distribution of STATs can be altered in favour of the cytoplasm or nucleus by fusion with either an external NES or NLS (Frahm, Hauser and Köster, 2006). Indeed, we were able to replicate these results by fusing different STATs to either a strong NES (aa sequence: NSNELALKLAGLDINK) derived from the protein kinase A inhibitor protein (PKI) (Fu *et al.*, 2018) to generate a cytoplasmic, or to a strong NLS (aa sequence: PKKKRKV) derived from simian virus 40 large T-antigen (sv40) (Hodel *et al.*, 2006) to generate a nuclear distribution (exemplified by STAT1 and STAT3 in figure 62A). Thus, we hypothesised that we could use the STAT-NES and STAT-NLS constructs as bait-STAT proteins in a bait-and-test interaction assay and assess unphosphorylated STAT dimer formation by observing if the bait-STAT causes relocation of a test-STAT towards the same compartment (figure 62B). Therefore, the Co-localisation assay used the extent of colocalization of bait- and test-STAT within the entire cell as readout to make a statement about unphosphorylated dimer formation (figure 62B). For the Co-localisation Assay to work accurately, knowledge about relative protein amounts within individual cells was necessary to guarantee that the bait-STAT is present in sufficient molecular excess and measurement artefacts could be avoided. Like in the Nuclear Translocation Assay, this was achieved by comparing the GFP/mCherry fluorescence ratio to the one measured for the rGFP-mCherry fluorescence standard using the same imaging setting. Lastly, it is worth noting that in the Co-localisation Assay, the bait-STAT protein that exerts an effect on a putative interaction partner was fused to a GFP fluorophore as opposed to the STATΔNLS constructs that exerted a dominant negative cross-inhibitory effect on the cytokine-induced nuclear accumulation of a putative interaction partner used in the Nuclear Translocation Assay that were fused to an mCherry moiety.

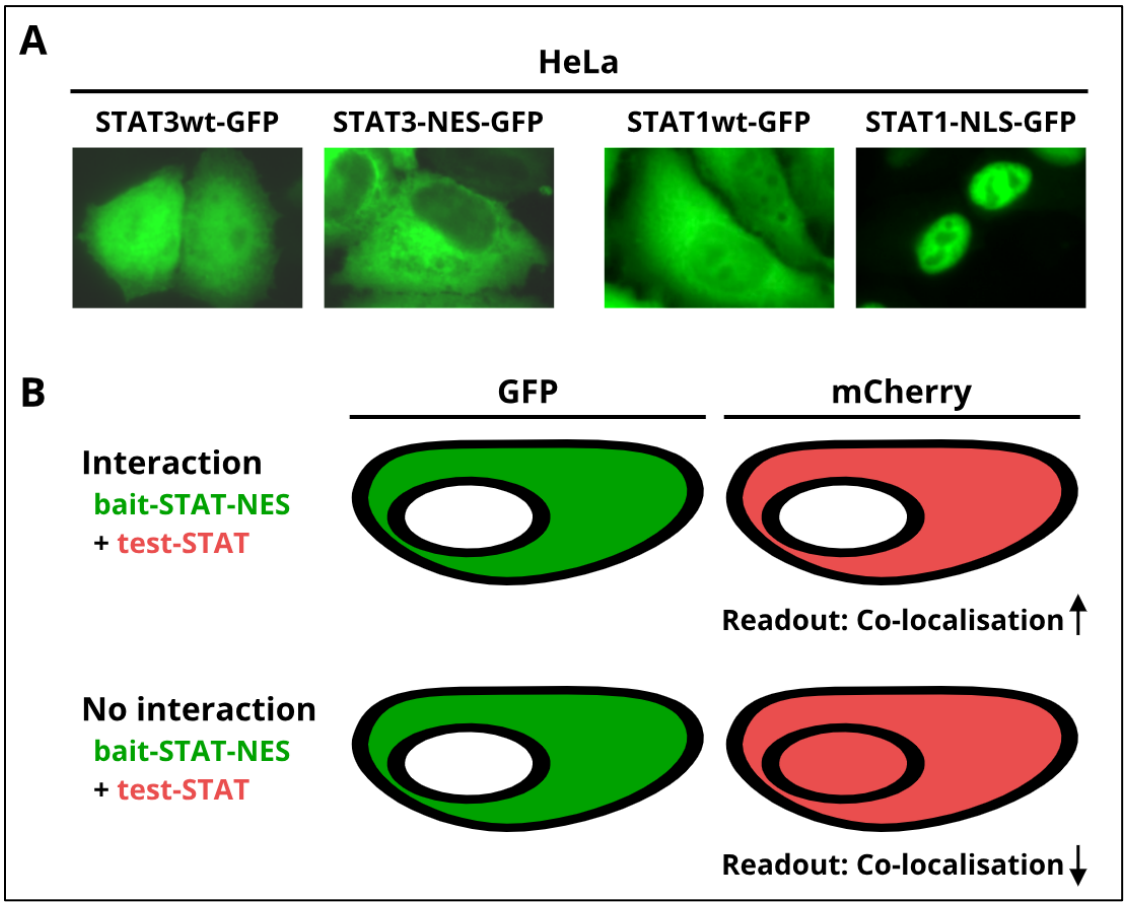


Figure 61 | Co-localisation of a bait-STAT-NES and test-STAT was used as readout to determine latent dimer formation: ***[A]*** *STAT-NES and STAT-NLS fusion proteins were created by introducing an NES sequence derived from PKI or an NLS derived from sv40 to the C-terminus of STATs (upstream of GFP). Expression in HeLa cells showed that fusion with an NES or NLS redistributes STATs to the cytoplasm or nucleus, respectively.* ***[B]*** *Theoretical concept behind the assay modification. If bait- and test-STATs interact, GFP and mCherry would show a high degree of co-localisation. If the two proteins don't interact, the co-localisation score would be low.*

Accordingly, the following Co-localisation Assay protocol was implemented:

Fixed HeLa cells co-expressing the bait and test combinations and cells expressing the rGFP-mCherry fluorescence standard were imaged using a widefield fluorescence microscope. Exposure times were set such that the GFP and mCherry intensities were roughly equal when imaging rGFP-mCherry. The same settings were applied whenever cells containing bait and test combinations were imaged. The microscopy images were subjected to a two-step image analysis workflow. For this, images of samples and the rGFP-mCherry standard were segmented by drawing ROIs around individual cells. Then, total GFP and mCherry fluorescence was measured for all cell ROIs. To determine which cells were eligible for further analysis, the GFP/mCherry fluorescence ratios of sample cells were compared to ones that contained the fluorescence standard. If this indicated that the bait-STAT-GFP was present in equimolar or at most 4-fold excess,

then the cells were progressed to the next image analysis step. There, the level of colocalization of bait- and test-STATs was assessed. This was done via the PSC Colocalization plugin for ImageJ (French *et al.*, 2008), which calculates the Pearson's correlation coefficient for co-localisation. Values can range from -1 which indicates a perfect negative correlation, i.e., distribution of the two proteins to different compartments, through 0 which indicates that distributions of the two proteins are completely independent and up to 1 which indicates perfect co-localisation of the two constructs. The co-localisation coefficient was calculated for at least ten cells co-expressing the same bait and test combination. We declared the formation of unphosphorylated dimers between bait- and test-STATs if the correlation coefficient on average was larger than 0.9.

4.5.2 Assay Validation

As described earlier, the unphosphorylated dimer splitting activity of the STAT3-L78R mutation is well-characterised and disease-relevant (Pilati *et al.*, 2011; Domoszlai *et al.*, 2014). To validate that the Co-localisation assay can reliably distinguish STATs that form unphosphorylated dimers from STATs that don't, we co-expressed a bait-STAT3-NES-GFP with either STAT3wt-mCherry or STAT3-L78R-mCherry in HeLa cells. Transfected cells were fixed after a 24 h incubation period, Hoechst stained and mounted onto microscopy slides. After imaging using calibrated settings, relative GFP to mCherry protein amounts were determined by comparing the total fluorescence ratio of sample images to that of images taken of the rGFP-mCherry standard. This resulted in 11 and 23 cells containing the STAT3-NES-GFP/STAT3wt-mCherry and STAT3-NES-GFP/STAT3-L78R-mCherry combination, respectively, being progressed to the second stage of image analysis. The cells containing the former combination showed a cytoplasmic distribution in both the GFP and mCherry channels. The overlay of the two channels was indicative of a high degree of co-localisation (figure 63A, row 1). Indeed, quantification of co-localisation revealed a Pearson's correlation coefficient of 0.94±0.01, clearly indicating the formation of unphosphorylated STAT3 homodimers (figure 63B). In contrast, cells containing STAT3-NES-GFP and STAT3-L78R-mCherry showed a cytoplasmic distribution in the GFP channel, but the mCherry-tagged protein localised pancellularly. When the two channels were overlayed, co-localisation was

only observed in the cytoplasm and STAT3-L78R-mCherry showed additional fluorescence in the nucleus (figure 63A, row 2). As expected, quantification of this revealed a Pearson's correlation coefficient for co-localisation of 0.64±0.03, which was notably below the threshold required to consider unphosphorylated dimer formation between these two proteins.

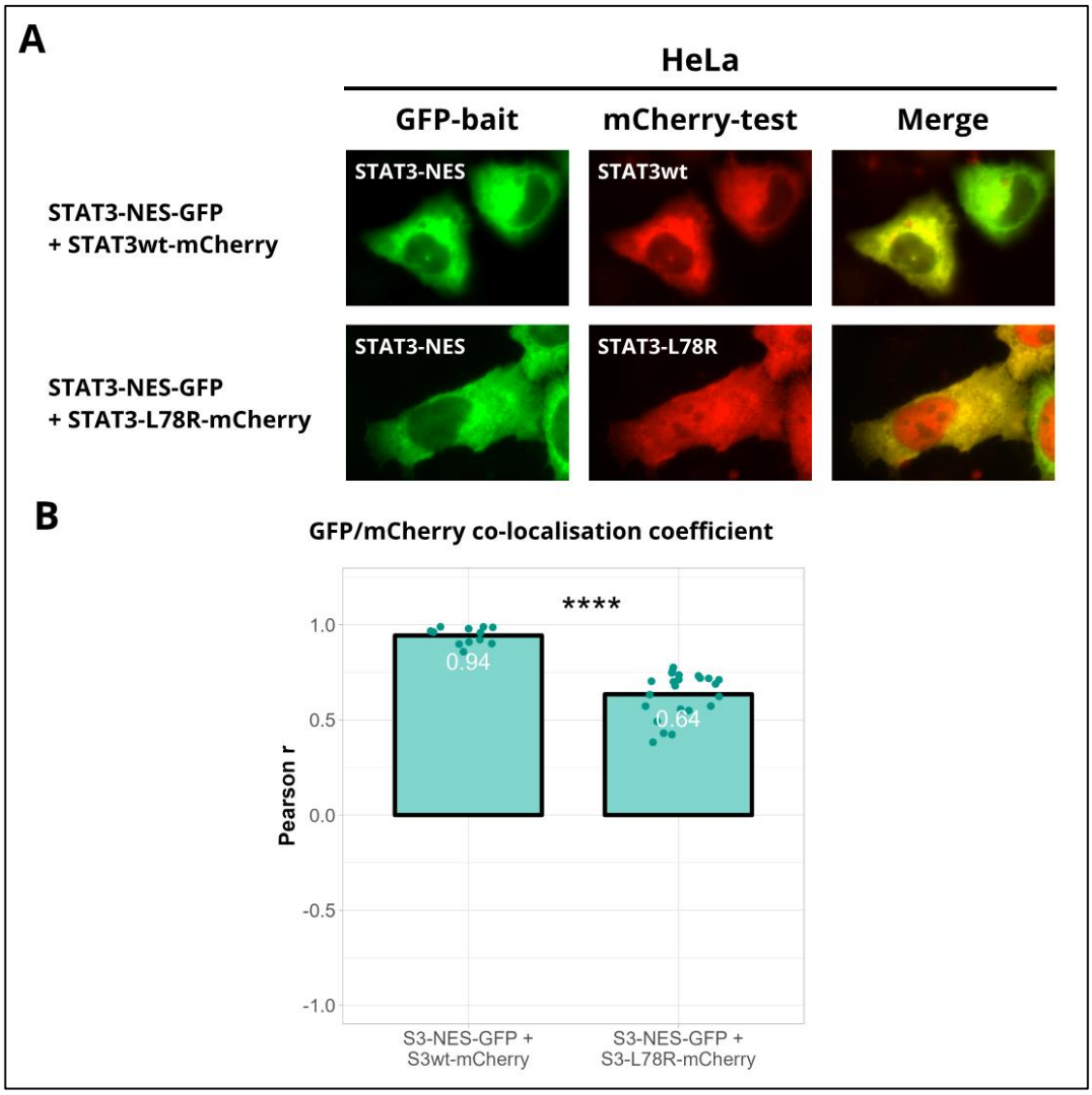


Figure 62 | The modified Nuclear Translocation Assay can reliably distinguish latent STAT dimers: ***[A]*** *HeLa cells were co-transfected with STAT3-NES-GFP and either STAT3wt-mCherry or STAT3-L78R-mCherry and imaged using a widefield microscope according to the modified Nuclear Translocation Assay protocol.* ***[B]*** *GFP/mCherry co-localisation coefficients were calculated for 12 or 23 cells expressing the bait- and test-STAT combinations described in [A]. A Kruskal-Wallis test was used for hypothesis testing.*

In summary, we found a significant drop in co-localisation scores when the STAT3-NES was combined with a STAT3 construct carrying the L78R mutation that is known to disrupt latent dimers compared to the STAT3-NES/STAT3wt combination (figure 63B).

This validated that the Co-localisation Assay could reliably distinguish between STAT constructs that form unphosphorylated dimers versus STATs that don't.

4.5.3 STATs form six latent homodimers and two latent heterodimers.

Following the validation of the modified Nuclear Translocation Assay, we assessed unphosphorylated STAT homo- and heterodimerisation among all 28 possible interaction partners. A summary of Pearson's correlation coefficients calculated for individual bait-STAT/test-STAT combinations is given in figure 64A. Generally, test-STATs (indicated in the first row of figure 64A) were combined with bait-STAT-NES proteins (indicated in the first column of figure 64A), except for STAT2 where the wild-type was used as bait because it naturally contains a strong NES (Frahm, Hauser and Köster, 2006). Furthermore, STAT2 homodimerisation in latent cells was investigated by combining a C-terminally truncated test-STAT2 (STAT2-ΔC, containing amino acids 1-703) with either bait-STAT2wt (figure 64A, †) or bait-STAT2-ΔC-NLS (figure 64A, ††). Overall, we found that STAT6 is the only member of the family that did not form unphosphorylated dimers. For all other STATs, except STAT2, the results revealed homodimerisation (figure 64A, light green cells). Furthermore, unphosphorylated STAT1/STAT2 and STAT5A/STAT5B heterodimers were detected (figure 64A, dark green cells). The findings most relevant to other aspects described in earlier chapters of this work are the results indicating that STAT1, STAT3 and STAT5B form unphosphorylated homodimers. We have already seen in chapter 4.2 and know from prior studies that said STATs can also form active homodimers (Becker, Groner and Müller, 1998; Chen *et al.*, 1998; Grimley, Dong and Rui, 1999). Of further interest is that we were unable to detect unphosphorylated STAT1/STAT3 and STAT1/STAT5B heterodimers. In both cases, a cytoplasmic distribution of the STAT1-NES-GFP was observed but co-expressed STAT3wt-mCherry and STAT5Bwt-mCherry showed a pancelullar distribution resulting in poor overall co-localisation (figure 64B).

While also no active heterodimer formation was detected for the STAT1/STAT5B combination using the Nuclear Translocation Assay (chapter 4.2.4), it is well known that STAT1/STAT3 forms active heterodimers in response to a variety of cytokines (Delgoffe and Vignali, 2013) and this was also picked up in earlier experiments in this work (chapter 4.4.2). Importantly, the non-existence of unphosphorylated

STAT1/STAT3 heterodimers specifically contradicted the notion that this complex is pre-formed before exposure to cytokine and merely changes its conformation to the active signalling complex.

A

Pearson's correlation coefficients calculated for the co-localisation of different bait-STAT/test-STAT combinations

	STAT1	STAT2	STAT3	STAT4	STAT5A	STAT5B	STAT6
STAT1	**0.98 ± 0.007 (12)**	**0.98±0.002 (13)**	-0.01 ± 0.101 (10)	0.42 ± 0.029 (12)	0.29 ± 0.054 (12)	0.24 ± 0.059 (11)	0.31 ± 0.064 (12)
STAT2		0.71 ± 0.03† (19) -0.45 ± 0.1†† (13)	0.58 ± 0.035 (15)	0.26 ± 0.063 (19)	0.05 ± 0.109 (14)	0.54 ± 0.035 (12)	0.40 ± 0.043 (20)
STAT3			**0.94 ± 0.013 (12)**	0.45 ± 0.032 (20)	0.30 ± 0.060 (21)	0.34 ± 0.086 (15)	0.43 ± 0.037 (23)
STAT4				**0.95 ± 0.012 (14)**	0.37 ± 0.063 (17)	0.54 ± 0.026 (14)	0.30 ± 0.040 (15)
STAT5A					**0.97 ± 0.006 (13)**	**0.97 ± 0.007 (17)**	0.40 ± 0.054 (12)
STAT5B						**0.97 ± 0.004 (10)**	0.37 ± 0.066 (13)
STAT6							0.41 ± 0.042 (18)

B

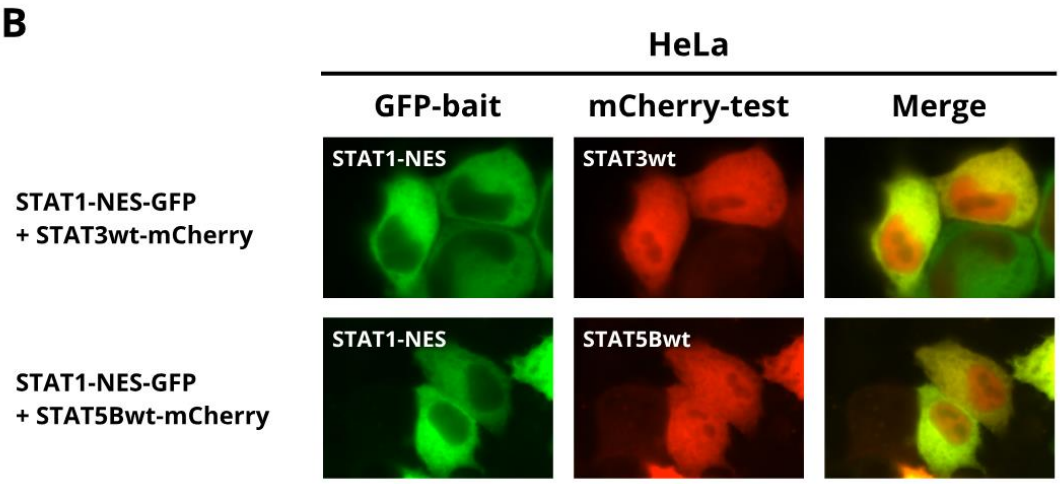


Figure 63 | Six STAT homodimers and two heterodimers were discovered in latent cells: ***[A]*** *Bait-STAT-NES proteins (columns) were combined with different test-STATs (rows). For combinations with STAT2, the STAT2wt was used as bait. To assess STAT2 latent homodimerisation, test-STAT2-ΔC was combined with either bait-STAT2wt (†) or bait-STAT2-ΔC-NLS (††).* ***[B]*** *Fluorescence micrographs of HeLa cells co-expressing STAT1-NES-GFP with either STAT3wt-mCherry or STAT5Bwt-mCherry.*

4.5.4 STAT1, STAT3 and STAT5B latent homodimers are split by mutating a conserved hydrophobic residue in the NTD.

In addition to the systematic analysis of unphosphorylated STAT dimer formation, we were interested in characterising the structural entity that determines unphosphorylated dimer formation. As explained in chapter 4.4.3, the importance of

STAT1 residue F77 for unphosphorylated STAT1 homodimer and unphosphorylated STAT1/STAT2 heterodimer formation is well established (Chen, 2003; Meyer *et al.*, 2004; Ho *et al.*, 2016). To test if we can replicate these results and confirm that this mechanism is relevant in living cells, we co-expressed STAT1-NLS-GFP with either STAT1wt-mCherry or STAT1-F77A-mCherry in HeLa cells. We used the bait-STAT1-NLS here because STAT1wt naturally shows a more cytoplasmic distribution before cytokine exposure. On the contrary, the STAT1-NLS-GFP localised exclusively to the nucleus (figure 65A, panel 1+4). Most of the co-expressed STAT1wt-mCherry was also seen co-localising in the nucleus but some fluorescence was still detectable in the cytoplasm (figure 65A, panel 2+3). The STAT1-F77A-mCherry co-expressed with STAT1-NLS-GFP in a different sample showed a pancellular distribution meaning that co-localisation with the bait-STAT1-NLS was limited (figure 65A, panel 5+6). To confirm, these observations the Pearson's correlation coefficients for GFP/mCherry co-localisation were calculated for 12 cells containing the wild-type combination and 20 cells containing the mutant combination. We found that in the former samples, GFP and mCherry distribution showed a very strong positive correlation (0.98±0.003; figure 65B). In contrast, the fluorescent-tagged STATs in the latter combination showed a significantly lower correlation in their subcellular distributions (0.18±0.04; figure 65B). These results confirm previous observations that mutating residue F77 splits unphosphorylated STAT1 homodimers. Furthermore, the co-localisation of STAT1-NLS-GFP and STAT1wt-mCherry in the nucleus showed that unphosphorylated dimer formation is not restricted to the cytoplasmic compartment.

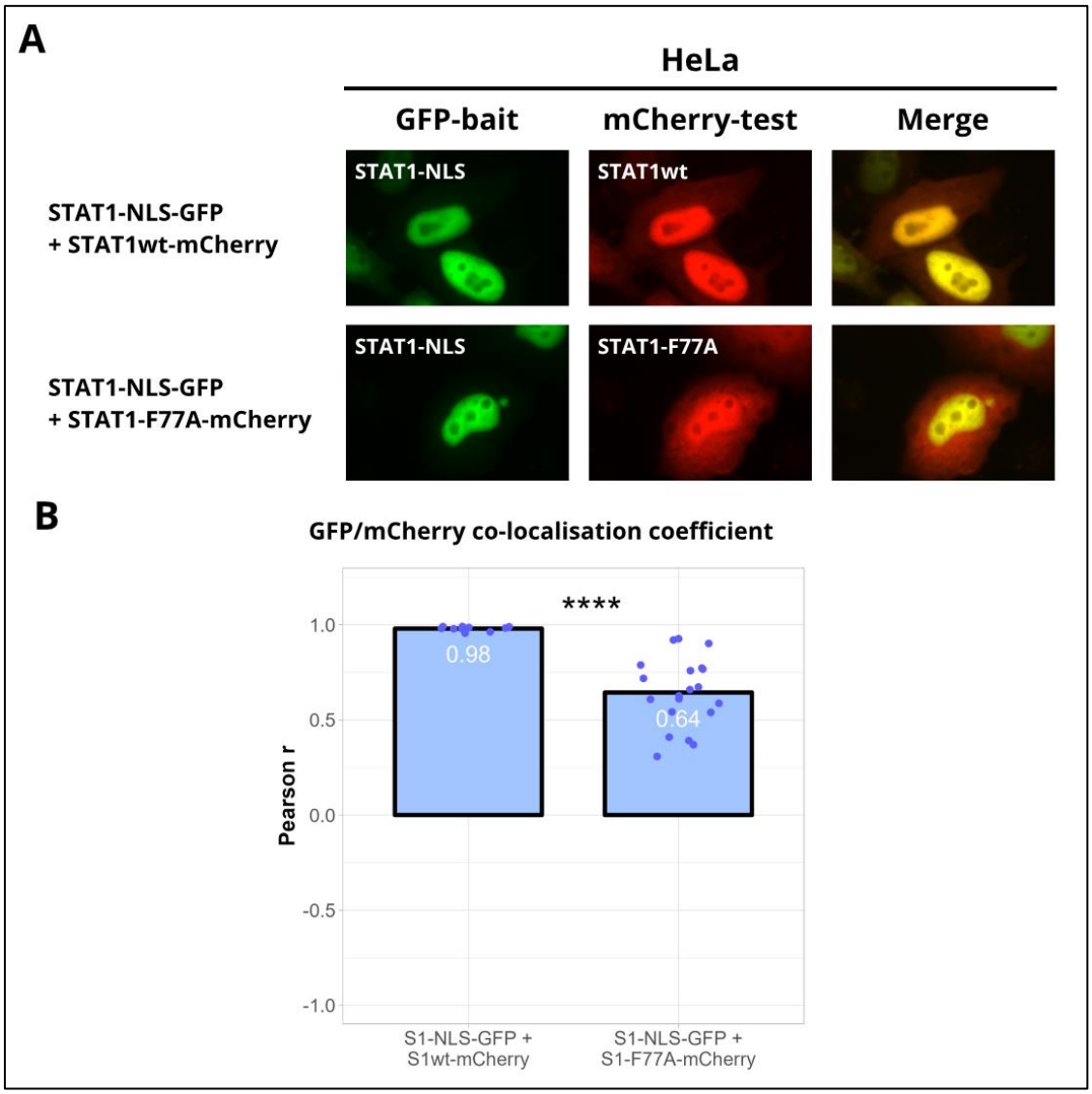


Figure 64 | The mutation F77A disrupts STAT1 latent homodimers: ***[A]*** *HeLa cells were co-transfected with STAT1-NLS-GFP and either STAT1wt-mCherry or STAT1-F77A-mCherry and imaged using a widefield microscope according to the modified Nuclear Translocation Assay protocol.* ***[B]*** *GFP/mCherry co-localisation coefficients were calculated for 12 or 20 cells expressing the bait- and test-STAT combinations described in [A]. A Kruskal-Wallis test was used for hypothesis testing.*

We have seen in chapter 4.5.2 that mutations of an amino acid in STAT3, residue L78R, neighbouring the homologous residue of STAT1 F77, has a similar unphosphorylated dimer splitting effect and that this is reliably detected using the Co-localisation Assay (figure 63). The paralogous residue L78 has also been shown to disrupt STAT1 and STAT4 unphosphorylated homodimers (Vinkemeier *et al.*, 1998; Chen, 2003; Ota *et al.*, 2004). Furthermore, we have seen in the sequence alignment earlier that this residue is conserved in STAT5A and STAT5B (figure 56). Thus, the STAT5B-L82A-mCherry mutant construct from earlier was subjected to the Co-localisation Assay to determine if it disrupts latent STAT5B homodimers. In HeLa cells expressing STAT5B-NES-GFP and STAT5Bwt-mCherry, both proteins co-localised in the cytoplasm (figure 66A, row 1).

On the contrary, STAT5B-NES-GFP and STAT5B-L82A-mCherry showed reduced co-localisation. While the former still showed a cytoplasm distribution, the latter was distributed equally among all cell compartments (figure 66A, row 2). When co-localisation was quantified in 9 cells containing the STAT5Bwt combination, a high positive correlation was determined (0.97±0.004; figure 66B). However, the correlation between GFP and mCherry distributions was significantly reduced in the STAT5B-L82A combination (0.63±0.027; figure 66B). This indicated that the mutation STAT5B-L82A indeed disrupted unphosphorylated STAT5B homodimers.

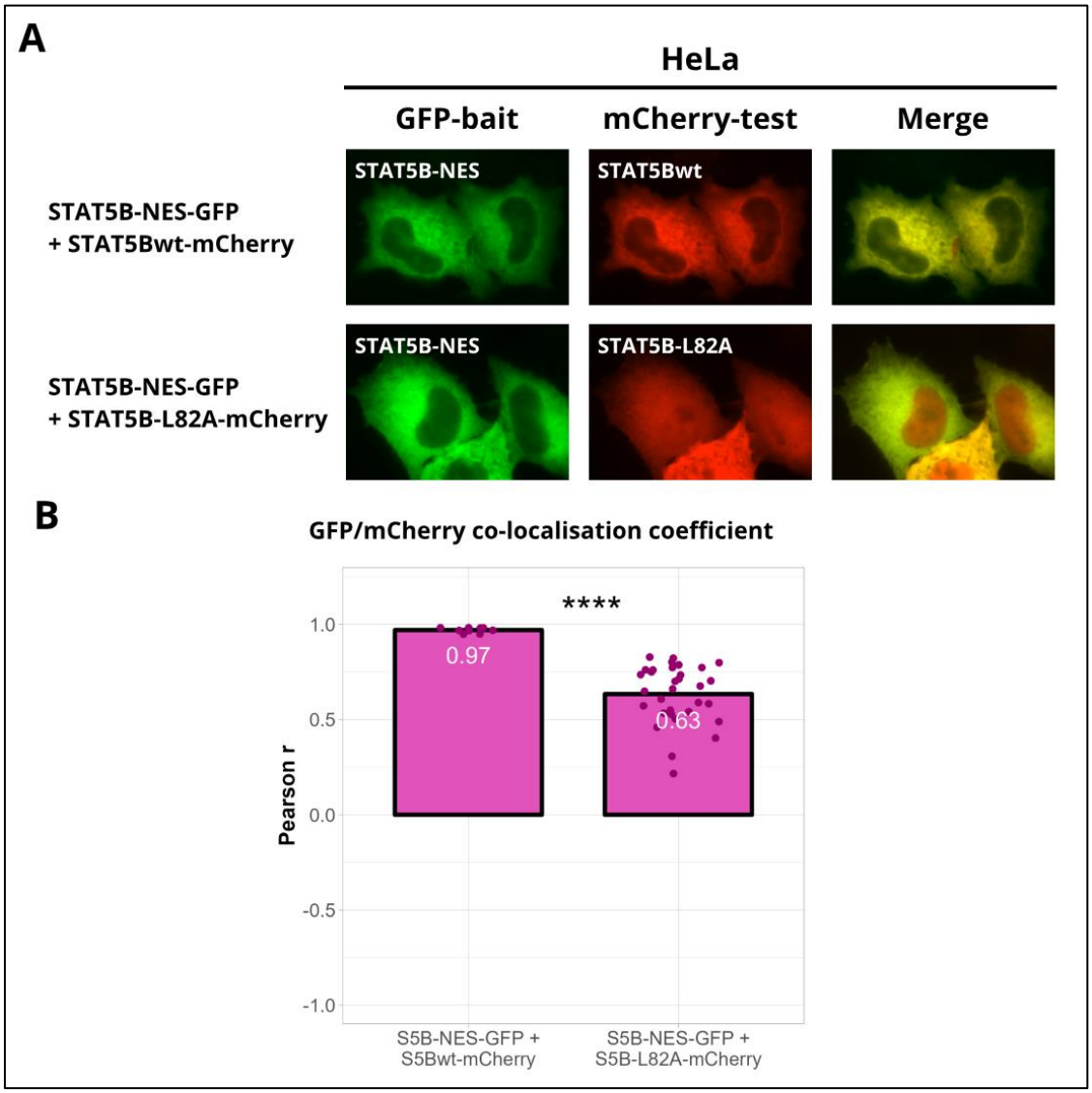


Figure 65 | The mutation L82A disrupts STAT5B latent homodimers: ***[A]*** *HeLa cells were co-transfected with STAT5B-NES-GFP and either STAT5Bwt-mCherry or STAT5B-L82A-mCherry and imaged using a widefield microscope according to the modified Nuclear Translocation Assay protocol.* ***[B]*** *GFP/mCherry co-localisation coefficients were calculated for 9 or 32 cells expressing the bait- and test-STAT combinations described in [A]. A Kruskal-Wallis test was used for hypothesis testing.*

Using the Co-localisation Assay, we also tested other mutations of homologous residues to STAT3-L78 on their effects on unphosphorylated STAT homo and heterodimer formation. Besides its unphosphorylated STAT5B homodimer splitting effect, the STAT5B-L82A mutation also disrupted unphosphorylated STAT5A/STAT5B heterodimers. The homologous mutation STAT5A-L82A further disrupted latent STAT5A homodimers. In addition, we were able to confirm the unphosphorylated STAT4 homodimer splitting activity of STAT4-L78R and showed that STAT2-L82A dissolves unphosphorylated STAT1/STAT2 heterodimers (Begitt *et al.*, 2023).

4.5.5 Discussion of Chapter 4.5

While it was initially believed that all STATs were monomeric before activation by phosphorylation, unphosphorylated dimers have now been reported for several STATs. Still, the experiments discussed in this chapter represent the first effort to assess unphosphorylated dimer formation for all 28 possible STAT combinations. Furthermore, most of the previous studies investigating unphosphorylated dimers relied on invasive techniques such as pull-down assays. However, the co-localisation Assay allowed us to assess unphosphorylated dimer formation under physiological conditions. With this, we found six unphosphorylated STAT homodimers and two unphosphorylated STAT heterodimers (figure 64A). STAT6 was the only family member that did not dimerise before cytokine exposure, which is consistent with previous size exclusion chromatography experiments from another group (Lackmann *et al.*, 1998). However, the isolated STAT6 NTD was found to interact in a yeast-two-hybrid interaction assay (Ota *et al.*, 2004) which is surprising considering that NTD interactions play a major role in stabilising the unphosphorylated dimers of other STATs (Meyer *et al.*, 2004; Mao *et al.*, 2005). STAT2 did not form unphosphorylated homodimers but was associated with STAT1 before cytokine exposure, also confirming previous findings (Stancato *et al.*, 1996; Ho *et al.*, 2016; Wang *et al.*, 2021). Unphosphorylated STAT1, STAT3, STAT4 and STAT5 homodimers found here have also been reported before using various techniques (Braunstein *et al.*, 2003; Kretzschmar *et al.*, 2004; Ota *et al.*, 2004; Schroder *et al.*, 2004; Mao *et al.*, 2005; Neculai *et al.*, 2005; Ren *et al.*, 2008; Bernado *et al.*, 2009; Vogt *et al.*, 2011). The second unphosphorylated heterodimer that was found using our cell-based assay was comprised of STAT5A and STAT5B. This

finding opposes a previous report that failed to detect unphosphorylated STAT5A/STAT5B heterodimers using immunoprecipitations followed by LC-ESI-MS analysis (Boehm *et al.*, 2014).

The following active heterodimers have been confirmed in previous studies STAT1/STAT2, STAT1/STAT3, STAT1/STAT4, STAT2/STAT6, STAT3/STAT4, STAT3/STAT5A, STAT3/STAT5B and STAT5A/STAT5B (Delgoffe and Vignali, 2013; Harrison and Moseley, 2020). Apart from STAT1/STAT2 and STAT5A/STAT5B, we could not detect their unphosphorylated counterparts. The unphosphorylated STAT1/STAT3 heterodimer is especially relevant to this work so will be discussed in more detail below.

Several studies have previously recorded the occurrence of STAT1/STAT3 pre-association before cytokine exposure. However, this unphosphorylated heterodimer was only detectable in cytoplasmic extracts (Haan *et al.*, 2000), at low abundance (Blumert *et al.*, 2013) or only under low-salt and detergent-free extraction conditions (Stancato *et al.*, 1996; Nivarthi *et al.*, 2016). Using a single-cell pull-down assay, another group found weak interactions between unphosphorylated STAT1 and STAT3 in living cells (Wedeking *et al.*, 2015) which may explain the findings of the above reports and our failure to detect these.

The absence or low abundance of latent STAT1/STAT3 heterodimers is in strict violation of the hypothesis that STAT heterodimers pre-form before cytokine exposure and transition to their active form via conformational change and without the exchange of monomers. However, the findings for STAT1 and STAT3 discussed in this chapter agree with the results obtained using the Nuclear Translocation Assay on active STAT1/STAT3 heterodimers (figure 50+51). The fact that a STAT3ΔNLS-mCherry could cross-inhibit STAT1-GFP cytokine-induced nuclear accumulation to a similar degree as the STAT1ΔNLS-mCherry (figure 24) suggested that the probability of active STAT1 being incorporated into homodimers versus STAT1/STAT3 heterodimers was the same. Therefore, it is likely that STAT1 and STAT3 homodimers and STAT1/STAT3 heterodimers form at equal rates. Moreover, the signal strength of individual bands in EMSA experiments on cytokine-treated cell extracts containing co-expressed STAT1 and STAT3 also hinted at this. This would not be achievable should active STAT1/STAT3

heterodimers exclusively stem from weak pre-formed STAT1/STAT3 unphosphorylated dimers of low abundance. Instead, it is more likely that the STAT1 and STAT3 monomers originally formed homodimers before their activation at the cytokine receptor and that these unphosphorylated dimers split upon cytokine stimulation. This may then be followed by randomly assemble into active homodimers and STAT1/STAT3 heterodimers at equal rates.

Disruption of unphosphorylated dimers leads to prolonged phosphorylation of STAT1 with serious biological consequences as evident from STAT1 gain-of-function disorders (Toubiana *et al.*, 2016; Kagawa *et al.*, 2017b; Okada *et al.*, 2020). Work by the Darnell group showed that this is because active STAT1 dimers must adopt the antiparallel conformation also seen in various unphosphorylated STATs to facilitate access of phosphatases to the STAT1 pY701 phosphopeptide (Zhong *et al.*, 2005). Prolonged phosphorylation of unphosphorylated dimer splitting STAT1 mutants is therefore due to a dephosphorylation defect. In a follow-up study, they further showed that STAT1 homodimers do not exchange monomers during the dephosphorylation process (Mertens *et al.*, 2006).

STAT3 dephosphorylation seems to work differently from what has been explained for STAT1 above because unphosphorylated STAT3 dimer splitting did not result in a dephosphorylation defect as evidenced by several studies. In this work, STAT3-L78R only showed minor non-significant elevations of phosphorylation compared to STAT3wt and this was also found in previous reports (Domoszlai *et al.*, 2014). A further study found that deleting the first 133 amino acids of STAT3 did not affect dephosphorylation patterns (Zhang *et al.*, 2006). However, this was contradicted in a different paper reporting prolonged phosphorylation of a STAT3-ΔNTD and prolonged binding of active STAT3-ΔNTD/STAT1 heterodimers to specific DNA sequences. Still, the authors attributed these effects to abrogated nuclear accumulation rather than a dephosphorylation defect (Martincuks *et al.*, 2016).

The phosphatase TC45 has been shown to dephosphorylate both STAT1 and STAT3 in the nucleus (ten Hoeve *et al.*, 2002). Therefore, it is conceivable that this phosphatase also deactivates STAT1/STAT3 heterodimers. However, it is not obvious how this would work in detail. Since STAT1 and STAT3 did not readily form unphosphorylated

heterodimers as evidenced by results in this work and elsewhere (figure 64) (Wedeking *et al.*, 2015), a dephosphorylation mechanism similar to that observed for STAT1, where the formation of an antiparallel dimer mediated by NTD-NTD and CCD-DBD is required for TC45 access (Mertens *et al.*, 2006), is unlikely. For the same reason, it is also unlikely that dephosphorylation does not lead to an exchange of monomers as observed for STAT1 (Mertens *et al.*, 2006). Instead, active heterodimers must be split upon deactivation and monomers may then form unphosphorylated STAT1 and STAT3 homodimers. Two alternative mechanisms for the dephosphorylation of STAT1/STAT3 heterodimers are plausible: (i) the active STAT1/STAT3 heterodimer adopts a conformation that naturally enables phosphopeptide access for a phosphatase such as TC45 (this mechanism may be similar for STAT3 homodimer dephosphorylation (Martincuks *et al.*, 2016)) or (ii) two dimers are participating in the dephosphorylation at the same time whereby association of the two STAT1 NTDs disrupt the STAT1/STAT3 heterodimers in a way that lets TC45 or a different phosphatase approach the phosphopeptides. Either way, more experiments are needed to discern the exact mechanism for STAT1/STAT3 heterodimer deactivation.

5. Concluding remarks

This overarching goal of this work was to determine the mechanism that regulate the formation of active STAT homo- over heterodimers or vice-versa. This aspect of STAT signalling is important because cytokines usually act in concert, meaning that multiple STATs get activated at the same time. A drastic example of this is cytokine storm. It is unclear if the STATs can uphold signals of individual cytokines via the formation of specific active signalling dimers in these contexts. Otherwise, if the formation of active STAT signalling complexes is largely unregulated, then the signals of individual cytokines may blend into one because the activated STATs assemble into different homo and heterodimers at random.

Experiments designed here were successfully applied to resolve this research question for the STAT family members STAT1, STAT3 and STAT5B. However, while the STAT amino acid structure is generally well conserved among the different family members (Schindler, 1995) the findings presented here may not be applicable for other STAT paralogues. Instead, more research is required to confirm that active dimer formation of other STAT family members is regulated in the same way as uncovered for STAT1, STAT3 and STAT5B in this work.

Two alternative hypotheses were stipulated at the beginning of this text. One alternative stated that active STAT dimers pre-form before the exposure of cells to cytokines and convert to the active signalling complexes via a conformational change that does not involve the exchange of monomers. Consequently, active STAT dimers would arise from a pre-defined pool of STAT homo- and heterodimer present in cells before cytokine treatment. The other alternative stated that active STAT dimers assemble independent of their unphosphorylated counterparts. Accordingly, the transition from unphosphorylated to active STAT dimers would involve the splitting of unphosphorylated upon their activation at the cytokine receptor which allows for the free recombination of unphosphorylated STAT homodimers into active heterodimers and vice-versa.

Findings presented in chapter 4.2 and chapter 4.5 suggested that both STAT1 and STAT3 could assemble unphosphorylated and active homodimers. However, evidence

presented in chapter 4.3 showed that the assembly of active STAT1 and STAT3 homodimers was not regulated by their unphosphorylated counterparts. Instead, the transition of unphosphorylated STAT1 and STAT3 homodimers to their active form involved splitting of the former dimers and random assembly into the latter complex.

It was found in chapter 4.4 that active STAT1/STAT3 heterodimer formation is also not regulated by unphosphorylated STAT dimers. Again, the active heterodimeric signalling complex assembled from monomeric subunits which implied that unphosphorylated STAT1 and STAT3 dimers had split during activation at the cytokine receptor. Moreover, evidence presented in chapter 4.5 suggested that unphosphorylated STAT1/STAT3 heterodimers cannot form or show a low affinity. Consequently, the majority, if not all, of the active STAT1/STAT3 heterodimers detected in the previous chapter must have arisen from monomers bound in unphosphorylated STAT1 and STAT3 homodimers before the exposure to cytokine. Further findings suggested that the active STAT1 and STAT3 homodimers do not readily exchange subunits once formed at the cytokine receptor. Thus, the only time in the STAT activation/inactivation cycle where active STAT1/STAT3 heterodimers can form must be during activation at the cytokine receptor. Based on experiments presented in chapters 4.2, 4.3, and 4.4 it was concluded that active STAT1 homodimers and active STAT1/STAT3 heterodimers may form at an equal rate. This means that the probability of a STAT1 monomer being incorporated into an active STAT1 homodimer may be the same as the probability of it assembling active STAT1/STAT3 heterodimers. A few reports, showing that active STAT1/STAT3 heterodimers have important biological roles which are not redundant to the roles of active STAT1 and STAT3 homodimers, exist (Ho and Ivashkiv, 2006; Dimberg *et al.*, 2012; Harrison *et al.*, 2020; Kim, Kim and Lee, 2021). Still, considering the findings regarding the rate of active STAT1/STAT3 heterodimer assembly, the importance of active STAT1/STAT3 heterodimers on biological processes may be greater than previously assumed. While the evidence presented here showed that active STAT1/STAT3 heterodimer formation was not dependent on unphosphorylated dimer integrity, these biological activities may be restricted to specific cell types by regulating active STAT1/STAT3 heterodimer formation in other ways. Such mechanisms may include cell-type specific expression levels of individual cytokine receptors or expression levels of STAT1 and STAT3, which may alter the proportion of

activated STAT1 and STAT3 within cells. In experiments conducted on active STAT1/STAT3 heterodimers in this work, cells were generally stimulated with high concentrations of IFN-γ and IL-6 to ensure efficient activation of STAT1 and STAT3. However, it is likely that cells are exposed to different concentrations of STAT1 and STAT3 activating cytokines *in vivo* depending on the biological context. Thus, differential activation by certain cytokines in concert or alone may also regulate the assembly of active STAT1/STAT3 heterodimers.

It was shown before that STAT1 and STAT5B get activated simultaneously in response to several cytokines but cannot assemble active heterodimers (Wehinger *et al.*, 1996; Kirken *et al.*, 1997; Schulze *et al.*, 2000; Delgoffe and Vignali, 2013). Experiments conducted for STAT1 and STAT5B in chapter 4.2 could confirm this notion *in situ*. Based on the hypothesis that the formation of active heterodimers is pre-determined by unphosphorylated dimers it was tested if splitting of unphosphorylated STAT1 and STAT5B would lead to the assembly of active heterodimers after activation at the cytokine receptor (chapter 4.4). However, this was not the case suggesting that the inability of these two paralogues to form active heterodimers is caused by a different regulatory element. Indeed, previous research indicated that the STAT5B phosphopeptide is not compatible with the STAT1 SH2 domain (Greenlund *et al.*, 1995). We sought to investigate this further by generating a chimeric STAT5B protein that has its SH2-domain and phosphotyrosine tail segment exchanged to that of STAT1 since it was reasoned that this should enable active dimer formation with STAT1wt. A similar approach has been applied to STAT2 which was made responsive to IFN-γ by exchanging its SH2 domain with that of STAT1 (Heim *et al.*, 1995). Initial experiments indicate that the STAT5B chimaera is well activated by IFN-γ and that this is followed by noticeable but imperfect nuclear accumulation. The ability to form heterodimers with STAT1wt shall be assessed in EMSA and Nuclear Translocation Assay experiments in the future.

Most of the findings regarding active STAT dimer formation discussed so far rooted from experiments using the novel cell-based interaction assay termed the Nuclear Translocation Assay. The development of this assay itself uncovered several, previously unknown characteristics of STATs. The functionality of the assay was based on the

cross-inhibition of the STAT-GFP cytokine-induced nuclear accumulation exerted by a STATΔNLS-mCherry (chapter 4.1). Consequently, the effect exerted by the ΔNLS mutation had to be dominant negative, meaning that the presence of this mutation in one monomer is sufficient to render the whole dimer unable to accumulate inside the nucleus in response to cytokine. So far, such a dominant negative effect has only been reported for STAT1ΔNLS exerted on active STAT1/STAT2 heterodimers (Melén, Kinnunen and Julkunen, 2001; Fagerlund *et al.*, 2002). Experiments in this work showed that the STAT1ΔNLS phenotype is also dominant negative in active STAT1 homodimers (chapter 4.2.1). Furthermore, the paralogues STAT3ΔNLS was found to be dominant negative regarding STAT3 homodimers (chapter 4.2.2) and STAT1/STAT3 heterodimers (chapter 4.3.3). The paralogues STAT5BΔNLS was dominant negative in STAT5B homodimers (chapter 4.2.3). Moreover, the discovery that residues RR422/425 mediate the cytokine-induced nuclear accumulation of STAT5B was novel (chapter 4.1.3). Importantly, the Nuclear Translocation Assay could reliably distinguish "known interactors" from "known non-interactors". This was validated when the Nuclear Translocation Assay was able to detect the formation of active STAT1, STAT3 and STAT5B homodimers, and correctly showed that STAT1 and STAT5B could not assemble active heterodimers (chapter 4.2).

Using another novel cell-based STAT interaction termed the Co-localisation Assay, we were able to report the first comprehensive analysis of unphosphorylated STAT dimer formation (chapter 4.5). In this, we found that STAT6 was the only family member that could not assemble unphosphorylated dimers. STAT2 was the only paralogue that could not form unphosphorylated homodimers, but it could form an unphosphorylated heterodimer with STAT1. All other STATs formed unphosphorylated homodimers and unphosphorylated STAT5A/STAT5B heterodimerised additionally. This further showed that cellular pools of unphosphorylated and active STAT dimers overlap only partially. Unphosphorylated dimer counterparts for active STAT1/STAT3, STAT1/STAT4, STAT2/STAT6, STAT3/STAT4, STAT3/STAT5A and STAT3/STAT5B could not be detected. This may suggest that the formation of these active STAT heterodimers is not regulated by unphosphorylated and that they can only form because unphosphorylated dimers of these paralogues split upon cytokine treatment in a similar fashion as observed for active STAT1/STAT3 heterodimers.

References

Algar, W.R. *et al.* (2019) 'FRET as a biomolecular research tool — understanding its potential while avoiding pitfalls', *Nature Methods*, 16(9), pp. 815–829. Available at: https://doi.org/10.1038/s41592-019-0530-8.

Arai, R. *et al.* (2001) 'Design of the linkers which effectively separate domains of a bifunctional fusion protein', *Protein Engineering*, 14(8), pp. 529–532. Available at: https://doi.org/10.1093/protein/14.8.529.

Arai, R. *et al.* (2004) 'Conformations of variably linked chimeric proteins evaluated by synchrotron X-ray small-angle scattering', *Proteins: Structure, Function, and Bioinformatics*, 57(4), pp. 829–838. Available at: https://doi.org/10.1002/PROT.20244.

de Araujo, E.D. *et al.* (2019) 'Structural Implications of STAT3 and STAT5 SH2 Domain Mutations', *Cancers 2019, Vol. 11, Page 1757*, 11(11), p. 1757. Available at: https://doi.org/10.3390/CANCERS11111757.

Argetsinger, L.S. *et al.* (1993) 'Identification of JAK2 as a growth hormone receptor-associated tyrosine kinase', *Cell*, 74(2), pp. 237–244. Available at: https://doi.org/10.1016/0092-8674(93)90415-M.

Bach, E.A., Aguet, M. and Schreiber, R.D. (1997) 'The IFNγ receptor: A paradigm for cytokine receptor signaling', *Annual Review of Immunology*, 15(1), pp. 563–591. Available at: https://doi.org/10.1146/annurev.immunol.15.1.563.

Becker, S., Groner, B. and Müller, C.W. (1998) 'Three-dimensional structure of the Stat3β homodimer bound to DNA', *Nature*. Nature Publishing Group, pp. 145–151. Available at: https://doi.org/10.1038/28101.

Begitt, A. *et al.* (2000) 'Nucleocytoplasmic translocation of Stat1 is regulated by a leucine-rich export signal in the coiled-coil domain', *Proceedings of the National Academy of Sciences of the United States of America*, 97(19), pp. 10418–10423. Available at: https://doi.org/10.1073/pnas.190318397.

Begitt, A. *et al.* (2023) 'A family-wide assessment of latent STAT transcription factor interactions reveals divergent dimer repertoires', *Journal of Biological Chemistry*, 299(5). Available at: https://doi.org/10.1016/j.jbc.2023.104703.

Bernado, P. *et al.* (2009) 'Structural characterization of unphosphorylated STAT5a oligomerization equilibrium in solution by small-angle X-ray scattering', *Protein Science*, 18(4), pp. 716–726. Available at: https://doi.org/10.1002/PRO.83.

Di Blasi, R. *et al.* (2021) 'A call for caution in analysing mammalian co-transfection experiments and implications of resource competition in data misinterpretation', *Nature Communications*, 12(1), pp. 1–6. Available at: https://doi.org/10.1038/s41467-021-22795-9.

Blumert, C. *et al.* (2013) 'Analysis of the STAT3 interactome using in-situ biotinylation and SILAC', *Journal of Proteomics*, 94, pp. 370–386. Available at: https://doi.org/10.1016/J.JPROT.2013.08.021.

Boehm, M.E. *et al.* (2014) 'Identification of isoform-specific dynamics in phosphorylation-dependent STAT5 dimerization by quantitative mass spectrometry and mathematical modeling', *Journal of Proteome Research*, 13(12), pp. 5685–5694. Available at: https://doi.org/10.1021/pr5006923.

Braunstein, J. *et al.* (2003) 'STATs Dimerize in the Absence of Phosphorylation', *Journal of Biological Chemistry*, 278(36), pp. 34133–34140. Available at: https://doi.org/10.1074/jbc.M304531200.

Chen, C. *et al.* (2012) 'Modulation of IFN-γ receptor 1 expression by AP-2α influences IFN-γ sensitivity of cancer cells', *American Journal of Pathology*, 180(2), pp. 661–671. Available at: https://doi.org/10.1016/j.ajpath.2011.10.040.

Chen, X. *et al.* (1998) 'Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA', *Cell*, 93(5), pp. 827–839. Available at: https://doi.org/10.1016/S0092-8674(00)81443-9.

Chen, X. (2003) 'A reinterpretation of the dimerization interface of the N-terminal Domains of STATs', *Protein Science*, 12(2), pp. 361–365. Available at: https://doi.org/10.1110/ps.0218903.

Cherkas, V. *et al.* (2018) 'Measurement of intracellular concentration of fluorescently-labeled targets in living cells', *PLoS ONE*, 13(4), p. e0194031. Available at: https://doi.org/10.1371/journal.pone.0194031.

Chiu, C.S. *et al.* (2001) 'Single-molecule measurements calibrate green fluorescent protein surface densities on transparent beads for use with "knock-in" animals and other expression systems', *Journal of Neuroscience Methods*, 105(1), pp. 55–63. Available at: https://doi.org/10.1016/S0165-0270(00)00354-X.

Chiu, C.S. *et al.* (2002) 'Number, density, and surface/cytoplasmic distribution of GABA transporters at presynaptic structures of knock-in mice carrying GABA transporter subtype 1-green fluorescent protein fusions', *Journal of Neuroscience*, 22(23), pp. 10251–10266. Available at: https://doi.org/10.1523/jneurosci.22-23-10251.2002.

Cimica, V. *et al.* (2011) 'Dynamics of the STAT3 transcription factor: Nuclear import dependent on ran and importin-β1', *PLoS ONE*, 6(5). Available at: https://doi.org/10.1371/journal.pone.0020188.

Collison, L.W. *et al.* (2012) 'The composition and signaling of the IL-35 receptor are unconventional', *Nature Immunology 2012 13:3*, 13(3), pp. 290–299. Available at: https://doi.org/10.1038/ni.2227.

Costa-Pereira, A.P. *et al.* (2002) 'Mutational switch of an IL-6 response to an interferon-γ-like response', *Proceedings of the National Academy of Sciences of the*

United States of America, 99(12), pp. 8043–8047. Available at: https://doi.org/10.1073/pnas.122236099.

Darnell, J.E. (1997) 'STATs and gene regulation', *Science*, 277(5332), pp. 1630–1635. Available at: https://doi.org/10.1126/science.277.5332.1630.

Decker, T. *et al.* (1991) 'Cytoplasmic activation of GAF, an IFN-γ-regulated DNA-binding factor', *EMBO Journal*, 10(4), pp. 927–932. Available at: https://doi.org/10.1002/j.1460-2075.1991.tb08026.x.

Decker, T. and Kovarik, P. (2000) 'Serine phosphorylation of STATs', *Oncogene*, pp. 2628–2637. Available at: https://doi.org/10.1038/sj.onc.1203481.

Decker, T., Kovarik, P. and Meinke, A. (1997) 'GAS elements: A few nucleotides with a major impact on cytokine-induced gene expression', *Journal of Interferon and Cytokine Research*, 17(3), pp. 121–134. Available at: https://doi.org/10.1089/jir.1997.17.121.

Decker, T., Müller, M. and Stockinger, S. (2005) 'The Yin and Yang of type I interferon activity in bacterial infection', *Nature Reviews Immunology*. Nature Publishing Group, pp. 675–687. Available at: https://doi.org/10.1038/nri1684.

Delgoffe, G.M. and Vignali, D.A.A. (2013) 'STAT heterodimers in immunity', *Jak-Stat*, 2(1), p. e23060. Available at: https://doi.org/10.4161/jkst.23060.

Dimberg, L.Y. *et al.* (2012) 'Stat1 activation attenuates IL-6 induced Stat3 activity but does not alter apoptosis sensitivity in multiple myeloma', *BMC Cancer*, 12(1), p. 318. Available at: https://doi.org/10.1186/1471-2407-12-318.

Dixon, A.S. and Lim, C.S. (2010) 'The nuclear translocation assay for intracellular protein-protein interactions and its application to the Bcr coiled-coil domain', *BioTechniques*, 49(1), pp. 519–524. Available at: https://doi.org/10.2144/000113452.

Domoszlai, T. *et al.* (2014) 'Consequences of the disease-related l78R mutation for dimerization and activity of STAT3', *Journal of Cell Science*, 127(9), pp. 1899–1910. Available at: https://doi.org/10.1242/jcs.137422.

Droescher, M. *et al.* (2011) 'Cytokine-induced paracrystals prolong the activity of Signal Transducers and Activators of Transcription (STAT) and provide a model for the regulation of protein solubility by Small Ubiquitin-like Modifier (SUMO)', *Journal of Biological Chemistry*, 286(21), pp. 18731–18746. Available at: https://doi.org/10.1074/jbc.M111.235978.

Dundr, M. *et al.* (2002) 'Quantitation of GFP-fusion proteins in single living cells', *Journal of Structural Biology*, 140(1–3), pp. 92–99. Available at: https://doi.org/10.1016/S1047-8477(02)00521-X.

Eriksen, K.W. *et al.* (2004) 'Bi-phasic effect of interferon (IFN)-α. IFN-α up- and down-regulates interleukin-4 signaling in human T cells', *Journal of Biological Chemistry*, 279(1), pp. 169–176. Available at: https://doi.org/10.1074/jbc.M310472200.

Fagerlund, R. *et al.* (2002) 'Arginine/lysine-rich nuclear localization signals mediate interactions between dimeric STATs and importin alpha 5.', *The Journal of biological chemistry*, 277(33), pp. 30072–30078. Available at: https://doi.org/10.1074/jbc.M202943200.

Fahrenkrog, B., Köser, J. and Aebi, U. (2004) 'The nuclear pore complex: A jack of all trades?', *Trends in Biochemical Sciences*, 29(4), pp. 175–182. Available at: https://doi.org/10.1016/j.tibs.2004.02.006.

Feng, J. *et al.* (1997) 'Activation of Jak2 catalytic activity requires phosphorylation of Y1007 in the kinase activation loop', *Molecular and Cellular Biology*, 17(5), pp. 2497–2501. Available at: https://doi.org/10.1128/mcb.17.5.2497.

Frahm, T., Hauser, H. and Köster, M. (2006) 'IFN-type-I-mediated signaling is regulated by modulation of STAT2 nuclear export', *Journal of Cell Science*, 119(6), pp. 1092–1104. Available at: https://doi.org/10.1242/jcs.02822.

Frederick Wheelock, E. (1965) 'Interferon-like virus-inhibitor induced in human leukocytes by phytohemagglutinin', *Science*, 149(3681), pp. 310–311. Available at: https://doi.org/10.1126/science.149.3681.310.

French, A.P. *et al.* (2008) 'Colocalization of fluorescent markers in confocal microscope images of plant cells', *Nature Protocols*, 3(4), pp. 619–628. Available at: https://doi.org/10.1038/nprot.2008.31.

Fu, S.C. *et al.* (2018) 'Correlation of CRM1-NES affinity with nuclear export activity', *Molecular Biology of the Cell*, 29(17), pp. 2037–2044. Available at: https://doi.org/10.1091/mbc.E18-02-0096.

Fu, X.Y. *et al.* (1990) 'ISGF3, the transcriptional activator induced by interferon α, consists of multiple interacting polypeptide chains', *Proceedings of the National Academy of Sciences of the United States of America*, 87(21), pp. 8555–8559. Available at: https://doi.org/10.1073/pnas.87.21.8555.

Furtado, A. and Henry, R. (2002) 'Measurement of green fluorescent protein concentration in single cells by image analysis', *Analytical Biochemistry*, 310(1), pp. 84–92. Available at: https://doi.org/10.1016/S0003-2697(02)00281-6.

Garbers, C. *et al.* (2018) 'Interleukin-6: Designing specific therapeutics for a complex cytokine', *Nature Reviews Drug Discovery*, 17(6), pp. 395–412. Available at: https://doi.org/10.1038/nrd.2018.45.

Gianti, E. and Zauhar, R.J. (2015) 'An SH2 domain model of STAT5 in complex with phospho-peptides define "sTAT5 Binding Signatures"', *Journal of Computer-Aided Molecular Design*, 29(5), pp. 451–470. Available at: https://doi.org/10.1007/s10822-015-9835-6.

Grant, A. *et al.* (2016) 'Zika Virus Targets Human STAT2 to Inhibit Type i Interferon Signaling', *Cell Host and Microbe*, 19(6), pp. 882–890. Available at: https://doi.org/10.1016/j.chom.2016.05.009.

Green, D.S., Young, H.A. and Valencia, J.C. (2017) 'Current prospects of type II interferon γ signaling & autoimmunity', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, pp. 13925–13933. Available at: https://doi.org/10.1074/jbc.R116.774745.

Greenlund, A.C. *et al.* (1995) 'Stat recruitment by tyrosine-phosphorylated cytokine receptors: An ordered reversible affinity-driven process', *Immunity*, 2(6), pp. 677–687. Available at: https://doi.org/10.1016/1074-7613(95)90012-8.

Grimley, P.M., Dong, F. and Rui, H. (1999) 'Stat5a and Stat5b: Fraternal twins of signal transduction and transcriptional activation', *Cytokine and Growth Factor Reviews*, 10(2), pp. 131–157. Available at: https://doi.org/10.1016/S1359-6101(99)00011-8.

Guesdon, F. *et al.* (2012) 'Expression of a glycosylphosphatidylinositol-anchored ligand, growth hormone, blocks receptor signalling', *Bioscience Reports*, 32(6), pp. 653–660. Available at: https://doi.org/10.1042/BSR20120088.

Gupta, S., Jiang, M. and Pernis, A.B. (1999) 'IFN-α Activates Stat6 and Leads to the Formation of Stat2:Stat6 Complexes in B Cells', *The Journal of Immunology*, 163(7), pp. 3834–3841. Available at: https://doi.org/10.4049/jimmunol.163.7.3834.

Haan, S. *et al.* (2000) 'Cytoplasmic STAT proteins associate prior to activation', *Biochemical Journal*, 345(3), pp. 417–421. Available at: https://doi.org/10.1042/bj3450417.

Haan, S. *et al.* (2005) 'Multiple reasons for an inefficient STAT1 response upon IL-6-type cytokine stimulation', *Cellular Signalling*, 17(12), pp. 1542–1550. Available at: https://doi.org/10.1016/j.cellsig.2005.03.010.

Haan, S. and Haan, C. (2013) 'Detection of activated stat species using electrophoretic mobility shift assay (EMSA) and potential pitfalls arising from the use of detergents', in *Methods in Molecular Biology*. Humana Press Inc., pp. 147–159. Available at: https://doi.org/10.1007/978-1-62703-242-1_10.

Hagmaier, K. *et al.* (2007) 'Mapuera virus, a rubulavirus that inhibits interferon signalling in a wide variety of mammalian cells without degrading STATs', *Journal of General Virology*, 88(3), pp. 956–966. Available at: https://doi.org/10.1099/vir.0.82579-0.

Han, Y. *et al.* (1996) 'Participation of JAK and STAT proteins in growth hormone-induced signaling', *Journal of Biological Chemistry*, 271(10), pp. 5947–5952. Available at: https://doi.org/10.1074/jbc.271.10.5947.

Harrison, A.R. *et al.* (2020) 'Lyssavirus P-protein selectively targets STAT3-STAT1 heterodimers to modulate cytokine signalling', *PLoS Pathogens*, 16(9), p. e1008767. Available at: https://doi.org/10.1371/JOURNAL.PPAT.1008767.

Harrison, A.R. and Moseley, G.W. (2020) 'The Dynamic Interface of Viruses with STATs', *Journal of Virology*, 94(22). Available at: https://doi.org/10.1128/jvi.00856-20.

Heim, M.H. *et al.* (1995) 'Contribution of STAT SH2 groups to specific interferon signaling by the Jak-STAT pathway', *Science*, 267(5202), pp. 1347–1349. Available at: https://doi.org/10.1126/science.7871432.

Heink, S. *et al.* (2017) 'Trans-presentation of IL-6 by dendritic cells is required for the priming of pathogenic T H 17 cells', *Nature Immunology*, 18(1), pp. 74–85. Available at: https://doi.org/10.1038/ni.3632.

Heinrich, P.C. *et al.* (1998) *Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway*, *Biochemical Journal*. Portland Press Limited. Available at: https://doi.org/10.1042/bj3340297.

Hellgren, G. *et al.* (1999) 'The growth hormone receptor associates with Jak1, Jak2 and Tyk2 in human liver', *Growth Hormone and IGF Research*, 9(3), pp. 212–218. Available at: https://doi.org/10.1054/ghir.1999.0111.

Herrington, J. *et al.* (1999) 'A functional DNA binding domain is required for growth hormone-induced nuclear accumulation of Stat5B', *Journal of Biological Chemistry*, 274(8), pp. 5138–5145. Available at: https://doi.org/10.1074/jbc.274.8.5138.

Herrmann, A. *et al.* (2004) 'STAT3 is enriched in nuclear bodies', *Journal of Cell Science*, 117(2), pp. 339–349. Available at: https://doi.org/10.1242/jcs.00833.

Hirano, T. (2021) 'IL-6 in inflammation, autoimmunity and cancer', *International immunology*, 33(3), pp. 127–148. Available at: https://doi.org/10.1093/intimm/dxaa078.

Hirschberg, K. *et al.* (1998) 'Kinetic analysis of secretory protein traffic and characterization of Golgi to plasma membrane transport intermediates in living cells', *Journal of Cell Biology*, 143(6), pp. 1485–1503. Available at: https://doi.org/10.1083/jcb.143.6.1485.

Ho, H.H. and Ivashkiv, L.B. (2006) 'Role of STAT3 in type I interferon responses: Negative regulation of STAT1-dependent inflammatory gene activation', *Journal of Biological Chemistry*, 281(20), pp. 14111–14118. Available at: https://doi.org/10.1074/jbc.M511797200.

Ho, J. *et al.* (2016) 'STAT2 Is a Pervasive Cytokine Regulator due to Its Inhibition of STAT1 in Multiple Signaling Pathways', *PLoS Biology*. Edited by H. Ploegh, 14(10), p. e2000117. Available at: https://doi.org/10.1371/journal.pbio.2000117.

Hodel, A.E. *et al.* (2006) 'Nuclear localization signal receptor affinity correlates with in vivo localization in Saccharomyces cerevisiae', *Journal of Biological Chemistry*, 281(33), pp. 23545–23556. Available at: https://doi.org/10.1074/jbc.M601718200.

ten Hoeve, J. *et al.* (2002) 'Identification of a Nuclear Stat1 Protein Tyrosine Phosphatase', *Molecular and Cellular Biology*, 22(16), pp. 5662–5668. Available at: https://doi.org/10.1128/mcb.22.16.5662-5668.2002.

Horvath, C.M. (2000) 'STAT proteins and transcriptional responses to extracellular signals', *Trends in Biochemical Sciences*. Elsevier Current Trends, pp. 496–502. Available at: https://doi.org/10.1016/S0968-0004(00)01624-8.

Horvath, C.M., Wen, Z. and Darnell, J.E. (1995) 'A STAT protein domain that determines DNA sequence recognition suggests a novel DNA-binding domain', *Genes and Development*, 9(8), pp. 984–994. Available at: https://doi.org/10.1101/gad.9.8.984.

Huliák, I. *et al.* (2012) 'The objectivity of reporters: Interference between physically unlinked promoters affects reporter gene expression in transient transfection experiments', *DNA and Cell Biology*, 31(11), pp. 1580–1584. Available at: https://doi.org/10.1089/dna.2012.1711.

Hüntelmann, B. *et al.* (2014) 'A conserved motif in the linker domain of STAT1 transcription factor is required for both recognition and release from high-affinity DNA-binding sites', *PLoS ONE*, 9(5), p. e97633. Available at: https://doi.org/10.1371/journal.pone.0097633.

Im, H. (2011) 'The Inoue Method for Preparation and Transformation of Competent E. coli: "Ultra Competent" Cells', *Bio-Protocol*, 1(20). Available at: https://doi.org/10.21769/bioprotoc.143.

Iyer, J. and Reich, N.C. (2008) 'Constitutive nuclear import of latent and activated STAT5a by its coiled coil domain', *The FASEB Journal*, 22(2), pp. 391–400. Available at: https://doi.org/10.1096/fj.07-8965com.

Johnston, J.A. *et al.* (1995) 'Tyrosine phosphorylation and activation of STAT5, STAT3, and Janus kinases by interleukins 2 and 15', *Proceedings of the National Academy of Sciences of the United States of America*, 92(19), pp. 8705–8709. Available at: https://doi.org/10.1073/pnas.92.19.8705.

Kagawa, R. *et al.* (2017a) 'Alanine-scanning mutagenesis of human signal transducer and activator of transcription 1 to estimate loss- or gain-of-function variants', *Journal of Allergy and Clinical Immunology*, 140(1), pp. 232–241. Available at: https://doi.org/10.1016/J.JACI.2016.09.035.

Kagawa, R. *et al.* (2017b) 'Alanine-scanning mutagenesis of human signal transducer and activator of transcription 1 to estimate loss- or gain-of-function variants', *Journal of Allergy and Clinical Immunology*, 140(1), pp. 232–241. Available at: https://doi.org/10.1016/j.jaci.2016.09.035.

Kak, G., Raza, M. and Tiwari, B.K. (2018) 'Interferon-gamma (IFN-γ): Exploring its implications in infectious diseases', *Biomolecular Concepts*, 9(1), pp. 64–79. Available at: https://doi.org/10.1515/bmc-2018-0007.

Kaminski, C.F., Rees, E.J. and Schierle, G.S.K. (2014) 'A quantitative protocol for intensity-based live cell FRET imaging', *Methods in Molecular Biology*, 1076, pp. 445–454. Available at: https://doi.org/10.1007/978-1-62703-649-8_19.

Kanai, T., Jenks, J. and Nadeau, K.C. (2012) 'The STAT5b pathway defect and autoimmunity', *Frontiers in Immunology*, 3(AUG), p. 234. Available at: https://doi.org/10.3389/fimmu.2012.00234.

Kany, S., Vollrath, J.T. and Relja, B. (2019) 'Cytokines in inflammatory disease', *International Journal of Molecular Sciences*, 20(23), p. 6008. Available at: https://doi.org/10.3390/ijms20236008.

Kaur, S. *et al.* (2020) 'A panoramic review of IL-6: Structure, pathophysiological roles and inhibitors', *Bioorganic and Medicinal Chemistry*. Pergamon, p. 115327. Available at: https://doi.org/10.1016/j.bmc.2020.115327.

Kelley, J.B. and Paschal, B.M. (2019) 'Fluorescence-based quantification of nucleocytoplasmic transport', *Methods*, 157, pp. 106–114. Available at: https://doi.org/10.1016/j.ymeth.2018.11.002.

Khakh, B.S. *et al.* (2001) 'Activation-dependent changes in receptor distribution and dendritic morphology in hippocampal neurons expressing P2X2-green fluorescent protein receptors', *Proceedings of the National Academy of Sciences of the United States of America*, 98(9), pp. 5288–5293. Available at: https://doi.org/10.1073/pnas.081089198.

Kim, D.H., Kim, H.Y. and Lee, W.W. (2021) 'Induction of Unique STAT Heterodimers by IL-21 Provokes IL-1RI Expression on CD8+ T Cells, Resulting in Enhanced IL-1β Dependent Effector Function', *Immune Network*, 21(5). Available at: https://doi.org/10.4110/IN.2021.21.E33.

Kimura, A. and Kishimoto, T. (2010) 'IL-6: Regulator of Treg/Th17 balance', *European Journal of Immunology*, 40(7), pp. 1830–1835. Available at: https://doi.org/10.1002/eji.201040391.

Kirken, R.A. *et al.* (1997) 'Prolactin stimulates serine/tyrosine phosphorylation and formation of heterocomplexes of multiple Stat5 isoforms in Nb2 lymphocytes', *Journal of Biological Chemistry*, 272(22), pp. 14098–14103. Available at: https://doi.org/10.1074/jbc.272.22.14098.

Kishimoto, T. and Kang, S. (2022) 'IL-6 Revisited: From Rheumatoid Arthritis to CAR T Cell Therapy and COVID-19', *Annual Review of Immunology*, 40, pp. 323–348. Available at: https://doi.org/10.1146/annurev-immunol-101220-023458.

Knauer, S.K. *et al.* (2005) 'Translocation biosensors to study signal-specific nucleo-cytoplasmic transport, protease activity and protein-protein interactions', *Traffic*, 6(7), pp. 594–606. Available at: https://doi.org/10.1111/J.1600-0854.2005.00298.X.

Köster, M. and Hauser, H. (1999) 'Dynamic redistribution of STAT1 protein in IFN signaling visualized by GFP fusion proteins', *European Journal of Biochemistry*, 260(1), pp. 137–144. Available at: https://doi.org/10.1046/j.1432-1327.1999.00149.x.

Kovarik, P. *et al.* (1999) 'Stress-induced phosphorylation of STAT1 at Ser727 requires p38 mitogen-activated protein kinase whereas IFN-γ uses a different signaling

pathway', *Proceedings of the National Academy of Sciences of the United States of America*, 96(24), pp. 13956–13961. Available at: https://doi.org/10.1073/pnas.96.24.13956.

Kretzschmar, A.K. *et al.* (2004) 'Analysis of Stat3 (signal transducer and activator of transcription 3) dimerization by fluorescence resonance energy transfer in living cells', *Biochemical Journal*, 377(2), pp. 289–297. Available at: https://doi.org/10.1042/BJ20030708.

Lackmann, M. *et al.* (1998) 'Biomolecular interaction analysis of IFNγ-induced signaling events in whole-cell lysates: Prevalence of latent STAT1 in high-molecular weight complexes', *Growth Factors*, 16(1), pp. 39–51. Available at: https://doi.org/10.3109/08977199809017490.

Lange, A. *et al.* (2007) 'Classical nuclear localization signals: Definition, function, and interaction with importin α', *Journal of Biological Chemistry*, 282(8), pp. 5101–5105. Available at: https://doi.org/10.1074/jbc.R600026200.

Langenfeld, F. *et al.* (2015) 'How Intrinsic Molecular Dynamics Control Intramolecular Communication in Signal Transducers and Activators of Transcription Factor STAT5', *PLOS ONE*, 10(12), p. e0145142. Available at: https://doi.org/10.1371/JOURNAL.PONE.0145142.

Levy, D.E. *et al.* (1989) 'Cytoplasmic activation of ISGF3, the positive regulator of interferon-alpha-stimulated transcription, reconstituted in vitro.', *Genes & development*, 3(9), pp. 1362–1371. Available at: https://doi.org/10.1101/gad.3.9.1362.

Levy, D.E. and Darnell, J.E. (2002) 'STATs: Transcriptional control and biological impact', *Nature Reviews Molecular Cell Biology*, 3(9), pp. 651–662. Available at: https://doi.org/10.1038/nrm909.

Li, J. *et al.* (2016) 'Structural basis for DNA recognition by STAT6', *Proceedings of the National Academy of Sciences of the United States of America*, 113(46), pp. 13015–13020. Available at: https://doi.org/10.1073/pnas.1611228113.

Lieu, K.G. *et al.* (2013) 'The Rabies Virus Interferon Antagonist P Protein Interacts with Activated STAT3 and Inhibits Gp130 Receptor Signaling', *Journal of Virology*, 87(14), pp. 8261–8265. Available at: https://doi.org/10.1128/jvi.00989-13.

Lim, C.P. and Cao, X. (2006) 'Structure, function, and regulation of STAT proteins', *Molecular BioSystems*, pp. 536–550. Available at: https://doi.org/10.1039/b606246f.

Lin, J.X. *et al.* (2012) 'Critical Role of STAT5 Transcription Factor Tetramerization for Cytokine Responses and Normal Immune Function', *Immunity*, 36(4), pp. 586–599. Available at: https://doi.org/10.1016/j.immuni.2012.02.017.

Lippitz, B.E. (2013) 'Cytokine patterns in patients with cancer: A systematic review', *The Lancet Oncology*. Elsevier, pp. e218–e228. Available at: https://doi.org/10.1016/S1470-2045(12)70582-X.

Liu, L. *et al.* (2011) 'Gain-of-function human STAT1 mutations impair IL-17 immunity and underlie chronic mucocutaneous candidiasis', *Journal of Experimental Medicine*, 208(18), pp. 1635–1648. Available at: https://doi.org/10.1084/jem.20110958.

Luo, G. and Yu-Lee, L. (1997) 'Transcriptional Inhibition by Stat5', *Journal of Biological Chemistry*, 272(43), pp. 26841–26849. Available at: https://doi.org/10.1074/jbc.272.43.26841.

Lütticken, C. *et al.* (1994) 'Association of transcription factor APRF and protein kinase Jak1 with the interleukin-6 signal transducer gp130', *Science*, 263(5143), pp. 89–92. Available at: https://doi.org/10.1126/science.8272872.

Ma, J. *et al.* (2003) 'A novel sequence in the coiled-coil domain of Stat3 essential for its nuclear translocation', *Journal of Biological Chemistry*, 278(31), pp. 29252–29260. Available at: https://doi.org/10.1074/jbc.M304196200.

Macara, I.G. (2001) 'Transport into and out of the Nucleus', *Microbiology and Molecular Biology Reviews*, 65(4), pp. 570–594. Available at: https://doi.org/10.1128/mmbr.65.4.570-594.2001.

Mahmudpour, M. *et al.* (2020) 'COVID-19 cytokine storm: The anger of inflammation', *Cytokine*, 133, p. 155151. Available at: https://doi.org/10.1016/j.cyto.2020.155151.

Mao, X. *et al.* (2005) 'Structural bases of unphosphorylated STAT1 association and receptor binding', *Molecular Cell*, 17(6), pp. 761–771. Available at: https://doi.org/10.1016/j.molcel.2005.02.021.

Marg, A. *et al.* (2004) 'Nucleocytoplasmic shuttling by nucleoporins Nup153 and Nup214 and CRM1-dependent nuclear export control the subcellular distribution of latent Stat1', *Journal of Cell Biology*, 165(6), pp. 823–833. Available at: https://doi.org/10.1083/jcb.200403057.

Marqusee, S. and Baldwin, R.L. (1987) 'Helix stabilization by Glu-...Lys+ salt bridges in short peptides of de novo design.', *Proceedings of the National Academy of Sciences of the United States of America*, 84(24), pp. 8898–8902. Available at: https://doi.org/10.1073/pnas.84.24.8898.

Martincuks, A. *et al.* (2016) 'Dissecting functions of the N-terminal domain and GAS-site recognition in STAT3 nuclear trafficking', *Cellular Signalling*, 28(8), pp. 810–825. Available at: https://doi.org/10.1016/j.cellsig.2016.03.011.

McBride, K.M. *et al.* (2002) 'Regulated nuclear import of the STAT1 transcription factor by direct binding of importin-α', *EMBO Journal*, 21(7), pp. 1754–1763. Available at: https://doi.org/10.1093/emboj/21.7.1754.

McNab, F. *et al.* (2015) 'Type I interferons in infectious disease', *Nature Reviews Immunology*, 15(2), pp. 87–103. Available at: https://doi.org/10.1038/nri3787.

Meinke, A. *et al.* (1996) 'Activation of different Stat5 isoforms contributes to cell-type-restricted signaling in response to interferons', *Molecular and Cellular Biology*, 16(12), pp. 6937–6944. Available at: https://doi.org/10.1128/mcb.16.12.6937.

Melén, K. *et al.* (2004) 'Expression of hepatitis C virus core protein inhibits interferon-induced nuclear import of STATs', *Journal of Medical Virology*, 73(4), pp. 536–547. Available at: https://doi.org/10.1002/jmv.20123.

Melén, K., Kinnunen, L. and Julkunen, I. (2001) 'Arginine/Lysine-rich Structural Element is Involved in Interferon-induced Nuclear Import of STATs', *Journal of Biological Chemistry*, 276(19), pp. 16447–16455. Available at: https://doi.org/10.1074/jbc.M008821200.

Merkx, M. (2010) 'Rational Design of FRET-Based Sensor Proteins', in *Reviews in Fluorescence 2008*. Springer, New York, NY, pp. 69–87. Available at: https://doi.org/10.1007/978-1-4419-1260-2_3.

Mertens, C. *et al.* (2006) 'Dephosphorylation of phosphotyrosine on STAT1 dimers requires extensive spatial reorientation of the monomers facilitated by the N-terminal domain', *Genes and Development*, 20(24), pp. 3372–3381. Available at: https://doi.org/10.1101/gad.1485406.

Mertens, C. *et al.* (2015) 'Mutations in the linker domain affect phospho-STAT3 function and suggest targets for interrupting STAT3 activity', *Proceedings of the National Academy of Sciences of the United States of America*, 112(48), pp. 14811–14816. Available at: https://doi.org/10.1073/pnas.1515876112.

Meyer, T. (2002) 'Constitutive and IFN-γ-induced nuclear import of STAT1 proceed through independent pathways', *EMBO Journal*, 21(3), pp. 344–354. Available at: https://doi.org/10.1093/emboj/21.3.344.

Meyer, T. *et al.* (2003) 'DNA binding controls inactivation and nuclear accumulation of the transcription factor Stat1', *Genes and Development*, 17(16), pp. 1992–2005. Available at: https://doi.org/10.1101/gad.268003.

Meyer, T. *et al.* (2004) 'A Single Residue Modulates Tyrosine Dephosphorylation, Oligomerization, and Nuclear Accumulation of Stat Transcription Factors', *Journal of Biological Chemistry*, 279(18), pp. 18998–19007. Available at: https://doi.org/10.1074/jbc.M400766200.

Meyer, T., Begitt, A. and Vinkemeier, U. (2007) 'Green fluorescent protein-tagging reduces the nucleocytoplasmic shuttling specifically of unphosphorylated STAT1', *FEBS Journal*, 274(3), pp. 815–826. Available at: https://doi.org/10.1111/j.1742-4658.2006.05626.x.

Meyer, T. and Vinkemeier, U. (2004) 'Nucleocytoplasmic shuttling of STAT transcription factors', *European Journal of Biochemistry*, 271(23–24), pp. 4606–4612. Available at: https://doi.org/10.1111/j.1432-1033.2004.04423.x.

Milward, A. *et al.* (2004) 'Growth hormone (GH) insensitivity syndrome due to a GH receptor truncated after Box1, resulting in isolated failure of STAT 5 signal transduction', *Journal of Clinical Endocrinology and Metabolism*, 89(3), pp. 1259–1266. Available at: https://doi.org/10.1210/jc.2003-031418.

Mohr, A. *et al.* (2012) 'Dynamics and non-canonical aspects of JAK/STAT signalling', *European Journal of Cell Biology*, pp. 524–532. Available at: https://doi.org/10.1016/j.ejcb.2011.09.005.

Morris, R., Kershaw, N.J. and Babon, J.J. (2018) 'The molecular details of cytokine signaling via the JAK/STAT pathway', *Protein Science*, pp. 1984–2009. Available at: https://doi.org/10.1002/pro.3519.

Muller, M. *et al.* (1993) 'Complementation of a mutant cell line: Central role of the 91 kDa polypeptide of ISGF3 in the interferon-α and -(γ) signal transduction pathways', *EMBO Journal*, 12(11), pp. 4221–4228. Available at: https://doi.org/10.1002/j.1460-2075.1993.tb06106.x.

Neculai, D. *et al.* (2005) 'Structure of the unphosphorylated STAT5a dimer', *Journal of Biological Chemistry*, 280(49), pp. 40782–40787. Available at: https://doi.org/10.1074/jbc.M507682200.

Nivarthi, H. *et al.* (2016) 'The ratio of STAT1 to STAT3 expression is a determinant of colorectal cancer growth', *Oncotarget*, 7(32), p. 51096. Available at: https://doi.org/10.18632/ONCOTARGET.9315.

Nkansah, E. *et al.* (2013) 'Observation of unphosphorylated STAT3 core protein binding to target dsDNA by PEMSA and X-ray crystallography', *FEBS Letters*, 587(7), pp. 833–839. Available at: https://doi.org/10.1016/J.FEBSLET.2013.01.065.

Okada, S. *et al.* (2020) 'Human STAT1 Gain-of-Function Heterozygous Mutations: Chronic Mucocutaneous Candidiasis and Type I Interferonopathy', *Journal of Clinical Immunology*, 40(8), pp. 1065–1081. Available at: https://doi.org/10.1007/s10875-020-00847-x.

OShea, J.J., Gadina, M. and Schreiber, R.D. (2002) 'Cytokine signaling in 2002: New surprises in the Jak/Stat pathway', *Cell*. Cell Press, pp. S121–S131. Available at: https://doi.org/10.1016/S0092-8674(02)00701-8.

Ota, N. *et al.* (2004) 'N-domain-dependent nonphosphorylated STAT4 dimers required for cytokine-driven activation', *Nature Immunology*, 5(2), pp. 208–215. Available at: https://doi.org/10.1038/ni1032.

Paine, P.L. and Feldherr, C.M. (1972) 'Nucleocytoplasmic exchange of macromolecules', *Experimental Cell Research*, 74(1), pp. 81–98. Available at: https://doi.org/10.1016/0014-4827(72)90483-1.

Palosaari, H. *et al.* (2003) 'STAT Protein Interference and Suppression of Cytokine Signal Transduction by Measles Virus V Protein', *Journal of Virology*, 77(13), pp. 7635–7644. Available at: https://doi.org/10.1128/jvi.77.13.7635-7644.2003.

Parganas, E. *et al.* (1998) 'Jak2 is essential for signaling through a variety of cytokine receptors', *Cell*, 93(3), pp. 385–395. Available at: https://doi.org/10.1016/S0092-8674(00)81167-8.

Pearse, R.N. *et al.* (1993) 'Interferon γ-induced transcription of the high-affinity Fc receptor for IgG requires assembly of a complex that includes the 91-kDa subunit of transcription factor ISGF3', *Proceedings of the National Academy of Sciences of the United States of America*, 90(9), pp. 4314–4318. Available at: https://doi.org/10.1073/pnas.90.9.4314.

Peritz, L.N. *et al.* (1988) 'The human growth hormone gene contains both positive and negative control elements', *Journal of Biological Chemistry*, 263(11), pp. 5005–5007. Available at: https://doi.org/10.1016/s0021-9258(18)60666-0.

Pettersen, E.F. *et al.* (2004) 'UCSF Chimera—A visualization system for exploratory research and analysis', *Journal of Computational Chemistry*, 25(13), pp. 1605–1612. Available at: https://doi.org/10.1002/JCC.20084.

Pilati, C. *et al.* (2011) 'Somatic mutations activating STAT3 in human inflammatory hepatocellular adenomas', *Journal of Experimental Medicine*, 208(7), pp. 1359–1366. Available at: https://doi.org/10.1084/jem.20110283.

Piston, D.W. and Kremers, G.J. (2007) 'Fluorescent protein FRET: the good, the bad and the ugly', *Trends in Biochemical Sciences*, 32(9), pp. 407–414. Available at: https://doi.org/10.1016/j.tibs.2007.08.003.

Pranada, A.L. *et al.* (2004) 'Real Time Analysis of STAT3 Nucleocytoplasmic Shuttling', *Journal of Biological Chemistry*, 279(15), pp. 15114–15123. Available at: https://doi.org/10.1074/jbc.M312530200.

Qing, Y. and Stark, G.R. (2004) 'Alternative activation of STAT1 and STAT3 in response to interferon-γ', *Journal of Biological Chemistry*, 279(40), pp. 41679–41685. Available at: https://doi.org/10.1074/jbc.M406413200.

Ram, P.A. *et al.* (1996) 'Growth hormone activation of Stat 1, Stat 3, and Stat 5 in rat liver: Differential kinetics of hormone desensitization and growth hormone stimulation of both tyrosine phosphorylation and serine/threonine phosphorylation', *Journal of Biological Chemistry*, 271(10), pp. 5929–5940. Available at: https://doi.org/10.1074/jbc.271.10.5929.

Reich, N.C. (2013) 'STATs get their move on', *Jak-Stat*, 2(4), p. e27080. Available at: https://doi.org/10.4161/jkst.27080.

Ren, Z. *et al.* (2008) 'Crystal structure of unphosphorylated STAT3 core fragment', *Biochemical and Biophysical Research Communications*, 374(1), pp. 1–5. Available at: https://doi.org/10.1016/j.bbrc.2008.04.049.

Renz, M. *et al.* (2012) 'Plasticity of the asialoglycoprotein receptor deciphered by ensemble FRET imaging and single-molecule counting PALM imaging', *Proceedings of*

the National Academy of Sciences of the United States of America, 109(44), pp. E2989–E2997. Available at: https://doi.org/10.1073/pnas.1211753109.

Rodriguez, J.J., Parisien, J.-P. and Horvath, C.M. (2002) 'Nipah Virus V Protein Evades Alpha and Gamma Interferons by Preventing STAT1 and STAT2 Activation and Nuclear Accumulation', *Journal of Virology*, 76(22), pp. 11476–11483. Available at: https://doi.org/10.1128/JVI.76.22.11476-11483.2002.

Rose-John, S. (2020) 'Interleukin-6 signalling in health and disease', *F1000Research*, 9, p. 1013. Available at: https://doi.org/10.12688/f1000research.26058.1.

Sampaio, E.P. *et al.* (2013) 'Signal transducer and activator of transcription 1 (STAT1) gain-of-function mutations and disseminated coccidioidomycosis and histoplasmosis', *Journal of Allergy and Clinical Immunology*, 131(6), pp. 1624-1634.e17. Available at: https://doi.org/10.1016/j.jaci.2013.01.052.

Schindelin, J. *et al.* (2012) 'Fiji: An open-source platform for biological-image analysis', *Nature Methods*, 9(7), pp. 676–682. Available at: https://doi.org/10.1038/nmeth.2019.

Schindler, C. *et al.* (1992) 'Interferon-dependent tyrosine phosphorylation of a latent cytoplasmic transcription factor', *Science*, 257(5071), pp. 809–813. Available at: https://doi.org/10.1126/science.1496401.

Schindler, C. (1995) 'Transcriptional Responses to Polypeptide Ligands: The JAK-STAT Pathway', *Annual Review of Biochemistry*, 64(1), pp. 621–651. Available at: https://doi.org/10.1146/annurev.biochem.64.1.621.

Schindler, C., Levy, D.E. and Decker, T. (2007) 'JAK-STAT signaling: From interferons to cytokines', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, pp. 20059–20063. Available at: https://doi.org/10.1074/jbc.R700016200.

Schroder, K. *et al.* (2004) 'Interferon-γ: an overview of signals, mechanisms and functions', *Journal of Leukocyte Biology*, 75(2), pp. 163–189. Available at: https://doi.org/10.1189/jlb.0603252.

Schulze, H. *et al.* (2000) 'Thrombopoietin induces the generation of distinct Stat1, Stat3, Stat5a and Stat5b homo- and heterodimeric complexes with different kinetics in human platelets', *Experimental Hematology*, 28(3), pp. 294–304. Available at: https://doi.org/10.1016/S0301-472X(99)00154-X.

Sehgal, P.B. (2008) 'Paradigm shifts in the cell biology of STAT signaling', *Seminars in Cell and Developmental Biology*, pp. 329–340. Available at: https://doi.org/10.1016/j.semcdb.2008.07.003.

Seidel, H.M. *et al.* (1995) 'Spacing of palindromic half sites as a determinant of selective STAT (signal transducers and activators of transcription) DNA binding and transcriptional activity', *Proceedings of the National Academy of Sciences of the*

United States of America, 92(7), pp. 3041–3045. Available at: https://doi.org/10.1073/pnas.92.7.3041.

Sievers, F. *et al.* (2011) 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega', *Molecular Systems Biology*, 7(1), p. 539. Available at: https://doi.org/10.1038/MSB.2011.75.

Spangler, J.B. *et al.* (2015) 'Insights into cytokine-receptor interactions from cytokine engineering', *Annual Review of Immunology*, 33, pp. 139–167. Available at: https://doi.org/10.1146/annurev-immunol-032713-120211.

Staab, J., Herrmann-Lingen, C. and Meyer, T. (2013) 'Clinically Relevant Dimer Interface Mutants of STAT1 Transcription Factor Exhibit Differential Gene Expression', *PLoS ONE*, 8(7), p. e69903. Available at: https://doi.org/10.1371/journal.pone.0069903.

Stancato, L.F. *et al.* (1996) 'Preassociation of STAT1 with STAT2 and STAT3 in separate signalling complexes prior to cytokine stimulation', *Journal of Biological Chemistry*, 271(8), pp. 4134–4137. Available at: https://doi.org/10.1074/jbc.271.8.4134.

Stark, G.R. and Darnell, J.E. (2012) 'The JAK-STAT Pathway at Twenty', *Immunity*. NIH Public Access, pp. 503–514. Available at: https://doi.org/10.1016/j.immuni.2012.03.013.

Steen, H.C. and Gamero, A.M. (2013) 'STAT2 phosphorylation and signaling', *Jak-Stat*, 2(4), p. e25790. Available at: https://doi.org/10.4161/jkst.25790.

Tanaka, T., Narazaki, M. and Kishimoto, T. (2014) 'Il-6 in inflammation, Immunity, And disease', *Cold Spring Harbor Perspectives in Biology*, 6(10), pp. 16295–16296. Available at: https://doi.org/10.1101/cshperspect.a016295.

Tisoncik, J.R. *et al.* (2012) 'Into the Eye of the Cytokine Storm', *Microbiology and Molecular Biology Reviews*, 76(1), pp. 16–32. Available at: https://doi.org/10.1128/mmbr.05015-11.

Toubiana, J. *et al.* (2016) 'Heterozygous STAT1 gain-of-function mutations underlie an unexpectedly broad clinical phenotype', *Blood*, 127(25), pp. 3154–3164. Available at: https://doi.org/10.1182/blood-2015-11-679902.

Vidy, A. *et al.* (2007) 'The Nucleocytoplasmic Rabies Virus P Protein Counteracts Interferon Signaling by Inhibiting both Nuclear Accumulation and DNA Binding of STAT1', *Journal of Virology*, 81(8), pp. 4255–4263. Available at: https://doi.org/10.1128/jvi.01930-06.

Vidy, A., Chelbi-Alix, M. and Blondel, D. (2005) 'Rabies Virus P Protein Interacts with STAT1 and Inhibits Interferon Signal Transduction Pathways', *Journal of Virology*, 79(22), pp. 14411–14420. Available at: https://doi.org/10.1128/jvi.79.22.14411-14420.2005.

Vinkemeier, U. *et al.* (1996) 'DNA binding of in vitro activated Stat1α, Stat1β and truncated Stat1: Interaction between NH2-terminal domains stabilizes binding of two dimers to tandem DNA sites', *EMBO Journal*, 15(20), pp. 5616–5626. Available at: https://doi.org/10.1002/j.1460-2075.1996.tb00946.x.

Vinkemeier, U. *et al.* (1998) 'Structure of the amino-terminal protein interaction domain of STAT-4', *Science*, 279(5353), pp. 1048–1052. Available at: https://doi.org/10.1126/science.279.5353.1048.

Vinkemeier, U. (2004) 'Getting the message across, STAT! Design principles of a molecular signaling circuit', *Journal of Cell Biology*, 167(2), pp. 197–201. Available at: https://doi.org/10.1083/jcb.200407163.

Vitor, M.T. *et al.* (2018) 'Tracking the Evolution of Transiently Transfected Individual Cells in a Microfluidic Platform', *Scientific Reports*, 8(1), pp. 1–9. Available at: https://doi.org/10.1038/s41598-018-19483-y.

Vogt, M. *et al.* (2011) 'The role of the N-terminal domain in dimerization and nucleocytoplasmic shuttling of latent STAT3', *Journal of Cell Science*, 124(6), pp. 900–909. Available at: https://doi.org/10.1242/JCS.072520.

Wagner, B.J. *et al.* (1990) 'The SIF binding element confers sis/PDGF inducibility onto the c-fos promoter', *EMBO Journal*, 9(13), pp. 4477–4484. Available at: https://doi.org/10.1002/j.1460-2075.1990.tb07898.x.

van der Wal, J. *et al.* (2001) 'Monitoring Agonist-induced Phospholipase C Activation in Live Cells by Fluorescence Resonance Energy Transfer', *Journal of Biological Chemistry*, 276(18), pp. 15337–15344. Available at: https://doi.org/10.1074/jbc.M007194200.

Wang, Y. *et al.* (2021) 'A virus-induced conformational switch of STAT1-STAT2 dimers boosts antiviral defenses', *Cell Research*, 31(2), pp. 206–218. Available at: https://doi.org/10.1038/s41422-020-0386-6.

Waterhouse, A.M. *et al.* (2009) 'Jalview Version 2—a multiple sequence alignment editor and analysis workbench', *Bioinformatics*, 25(9), pp. 1189–1191. Available at: https://doi.org/10.1093/BIOINFORMATICS/BTP033.

Wedeking, T. *et al.* (2015) 'Single Cell GFP-trap reveals stoichiometry and dynamics of cytosolic protein complexes', *Nano Letters*, 15(5), pp. 3610–3615. Available at: https://doi.org/10.1021/acs.nanolett.5b01153.

Wehinger, J. *et al.* (1996) 'IL-10 induces DNA binding activity of three STAT proteins (Stat1, Stat3, and Stat5) and their distinct combinatorial assembly in the promoters of selected genes', *FEBS Letters*, 394(3), pp. 365–370. Available at: https://doi.org/10.1016/0014-5793(96)00990-8.

Wenta, N. *et al.* (2008) 'Tyrosine phosphorylation regulates the partitioning of STAT1 between different dimer conformations', *Proceedings of the National Academy of*

Sciences of the United States of America, 105(27), pp. 9238–9243. Available at: https://doi.org/10.1073/pnas.0802130105.

Wojciak, J.M. *et al.* (2009) 'Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains', *EMBO Journal*, 28(7), pp. 948–958. Available at: https://doi.org/10.1038/emboj.2009.30.

Wolf, J., Rose-John, S. and Garbers, C. (2014) 'Interleukin-6 and its receptors: a highly regulated and dynamic system', *Cytokine*, 70(1), pp. 11–20. Available at: https://doi.org/10.1016/j.cyto.2014.05.024.

Wu, T.R. *et al.* (2002) 'SHP-2 is a dual-specificity phosphatase involved in Stat1 dephosphorylation at both tyrosine and serine residues in nuclei', *Journal of Biological Chemistry*, 277(49), pp. 47572–47580. Available at: https://doi.org/10.1074/jbc.M207536200.

Xu, X., Sun, Y.L. and Hoey, T. (1996) 'Cooperative DNA binding and sequence-selective recognition conferred by the STAT amino-terminal domain', *Science*, 273(5276), pp. 794–797. Available at: https://doi.org/10.1126/science.273.5276.794.

Yahata, Y. *et al.* (2003) 'Nuclear translocation of phosphorylated STAT3 is essential for vascular endothelial growth factor-induced human dermal microvascular endothelial cell migration and tube formation', *Journal of Biological Chemistry*, 278(41), pp. 40026–40031. Available at: https://doi.org/10.1074/jbc.M301866200.

Yang, E. *et al.* (2002) 'Dissociation time from DNA determines transcriptional function in a STAT1 linker mutant', *Journal of Biological Chemistry*, 277(16), pp. 13455–13462. Available at: https://doi.org/10.1074/jbc.M112038200.

Yang, J. and Stark, G.R. (2008) 'Roles of unphosphorylated STATs in signaling', *Cell Research*, pp. 443–451. Available at: https://doi.org/10.1038/cr.2008.41.

Zhang, L. *et al.* (2006) 'IL-6 signaling via the STAT3/SOCS3 pathway: Functional analysis of the conserved STAT3 N-domain', *Molecular and Cellular Biochemistry*, 288(1–2), pp. 179–189. Available at: https://doi.org/10.1007/s11010-006-9137-3.

Zhang, T. *et al.* (2000) 'The Coiled-Coil Domain of Stat3 Is Essential for Its SH2 Domain-Mediated Receptor Binding and Subsequent Activation Induced by Epidermal Growth Factor and Interleukin-6', *Molecular and Cellular Biology*, 20(19), pp. 7132–7139. Available at: https://doi.org/10.1128/mcb.20.19.7132-7139.2000.

Zhong, M. *et al.* (2005) 'Implications of an antiparallel dimeric structure of nonphosphorylated STAT1 for the activation-inactivation cycle', *Proceedings of the National Academy of Sciences of the United States of America*, 102(11), pp. 3966–3971. Available at: https://doi.org/10.1073/pnas.0501063102.

Zhong, Z., Wen, Z. and Darnell, J.E. (1994a) 'Stat3: A STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6', *Science*, 264(5155), pp. 95–98. Available at: https://doi.org/10.1126/science.8140422.

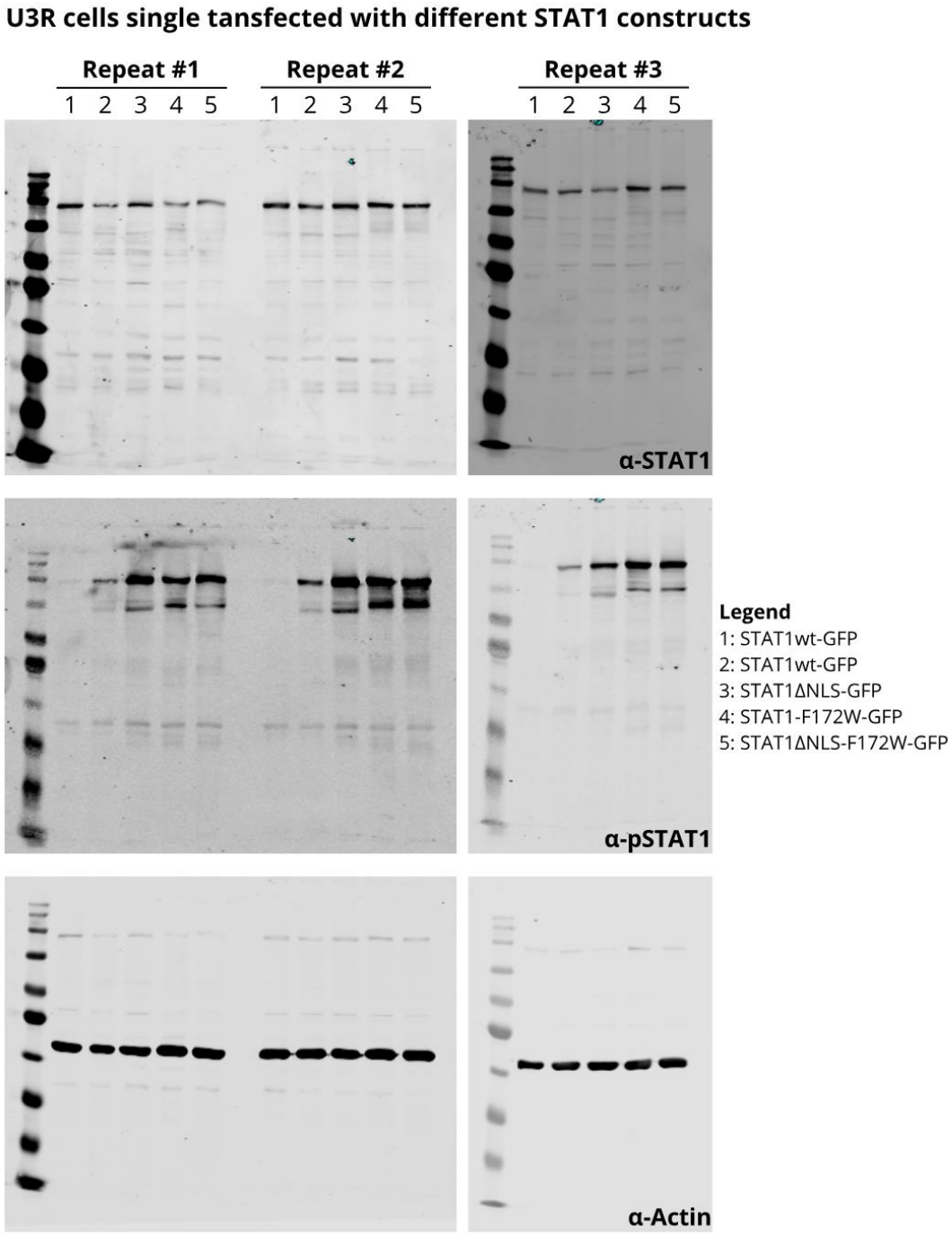
Zhong, Z., Wen, Z. and Darnell, J.E. (1994b) 'Stat3 and Stat4: Members of the family of signal transducers and activators of transcription', *Proceedings of the National Academy of Sciences of the United States of America*, 91(11), pp. 4806–4810. Available at: https://doi.org/10.1073/pnas.91.11.4806.

Zhuo, X. and Knox, B.E. (2022) 'Interaction of human CRX and NRL in live HEK293T cells measured using fluorescence resonance energy transfer (FRET)', *Scientific Reports*, 12(1), pp. 1–13. Available at: https://doi.org/10.1038/s41598-022-10689-9.

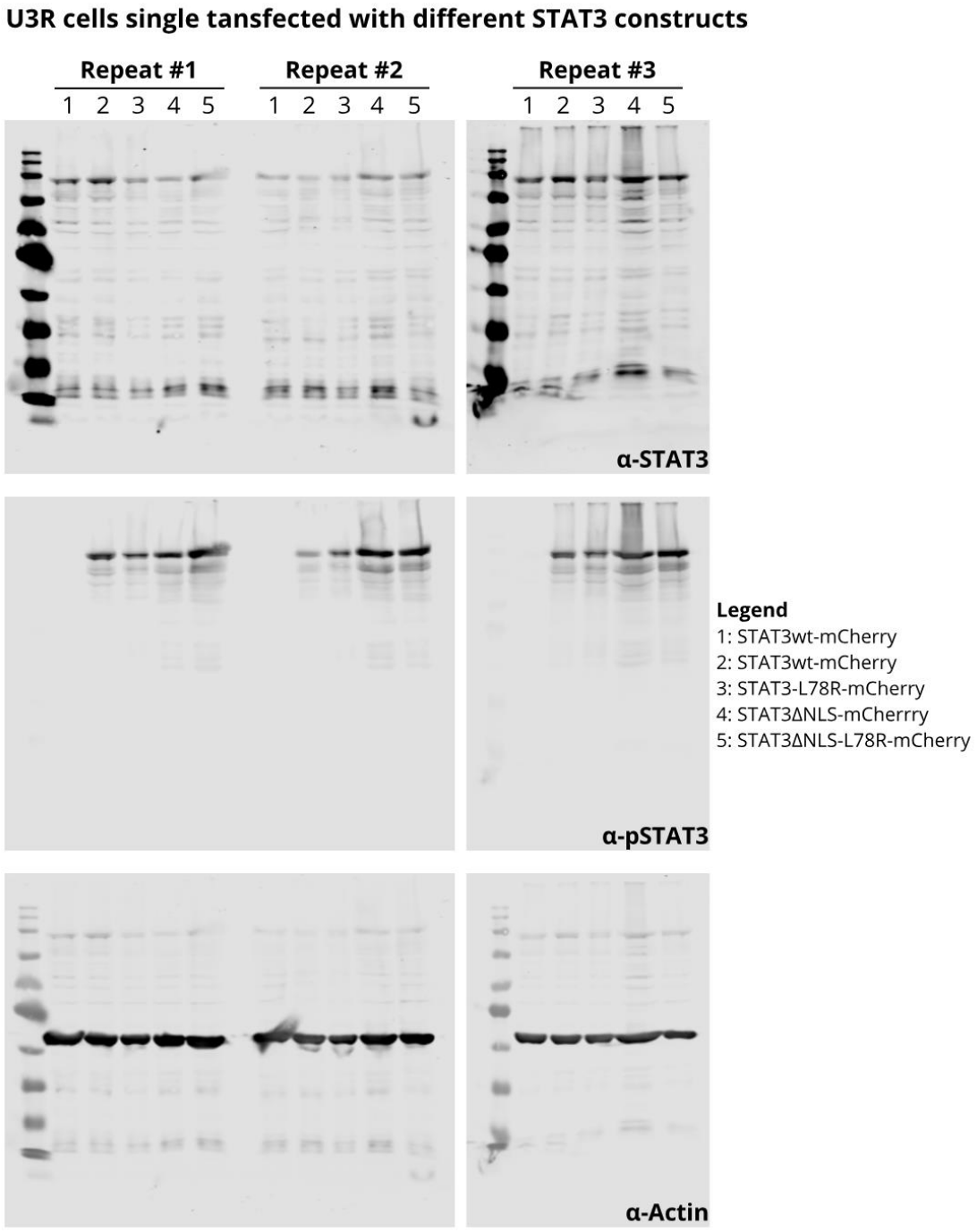
Appendix

A. Raw Western blot and EMSA data

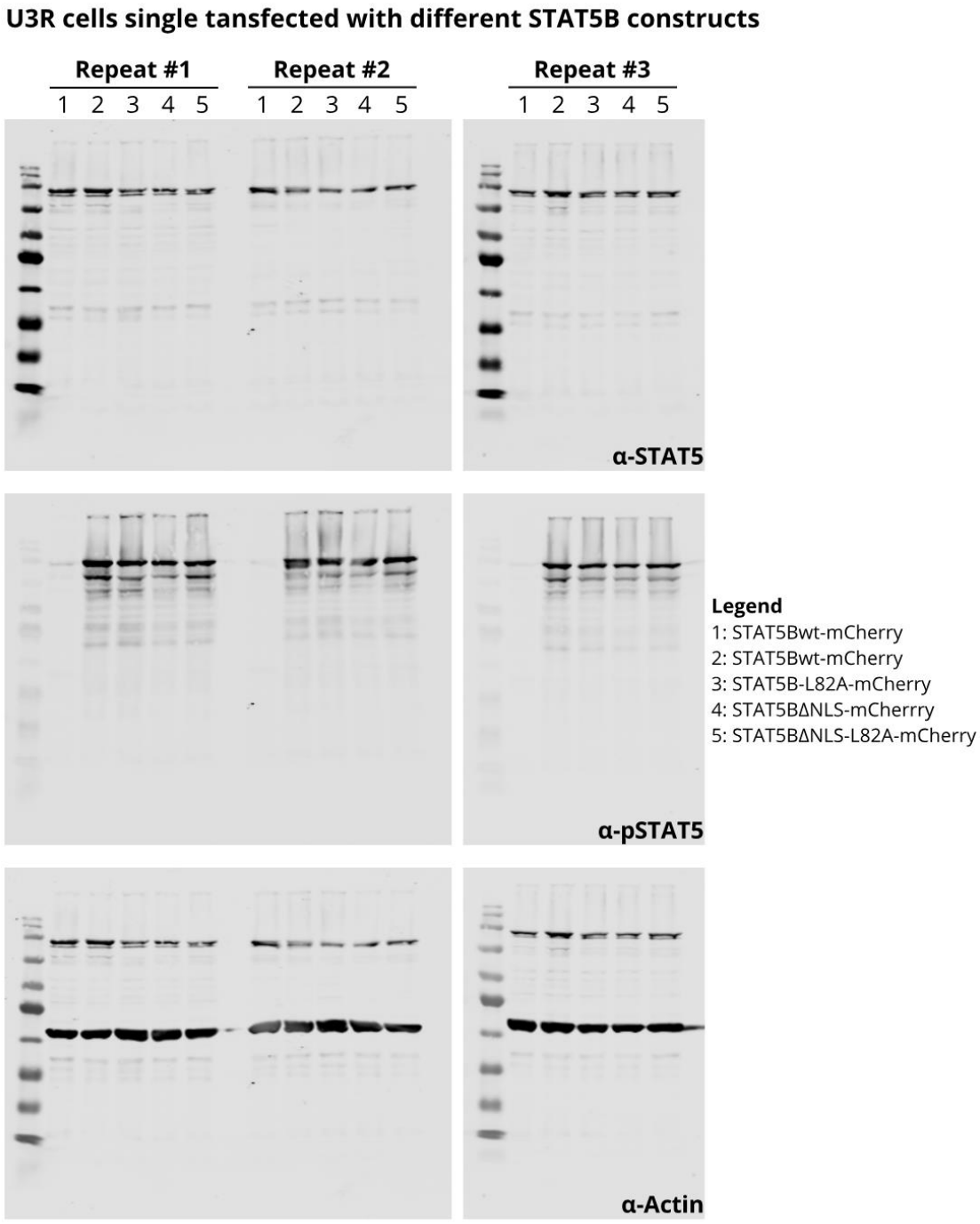
(i) Single transfections



Appendix Figure 1 | Western blots of U3R cells expressing different STAT1 constructs: *Anti-STAT1, anti-pSTAT1 and anti-Actin antibodies were used to detect specific proteins. The first two antibodies were detected simultaneously. Membranes were then incubated with the anti-Actin antibody. All extracts stemmed from IFN-γ treated cells except for number 1 which was left untreated.*

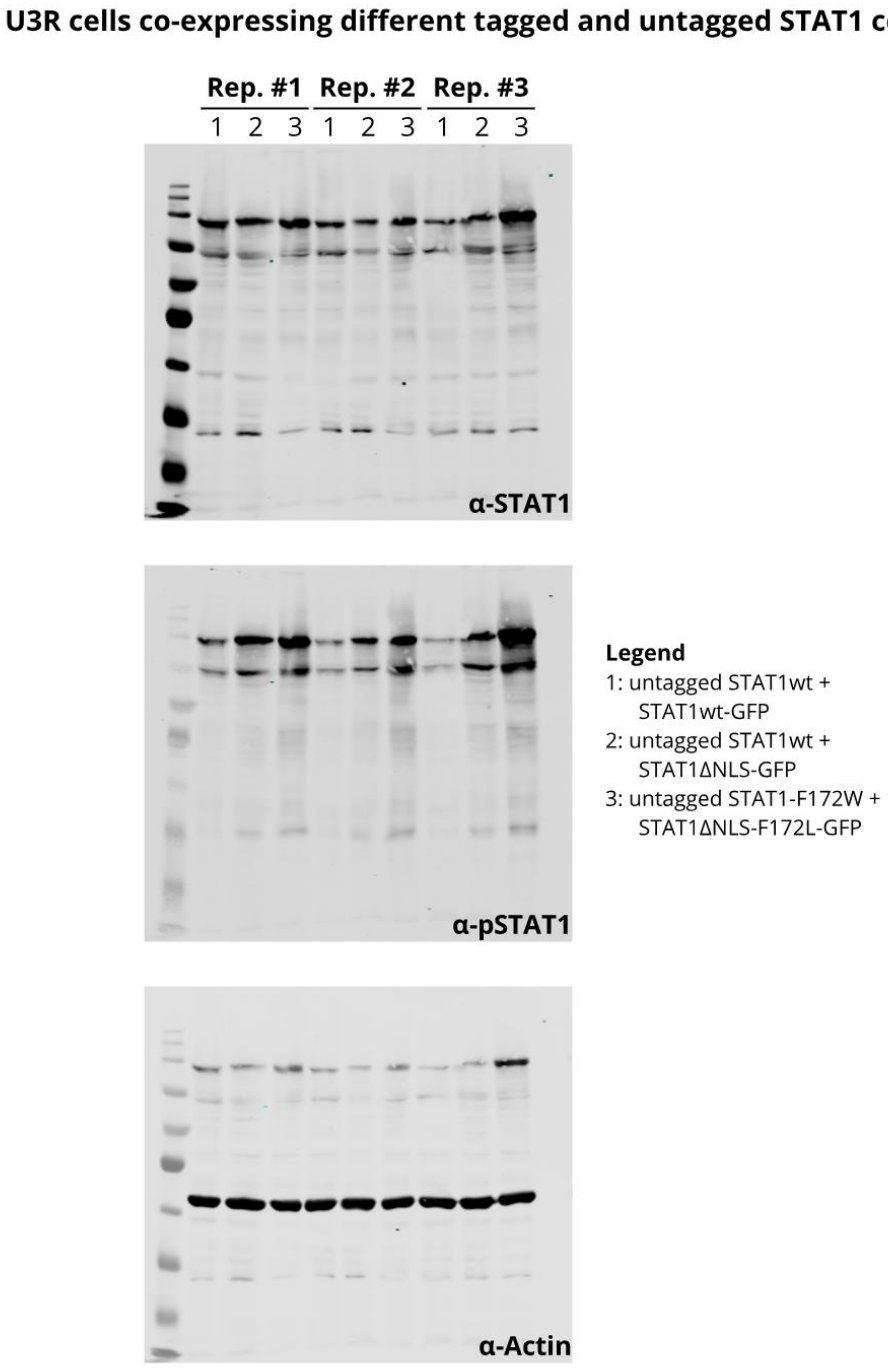


Appendix Figure 2 | Western blots of U3R cells expressing different STAT3 constructs: *Anti-STAT3, anti-pSTAT3 and anti-Actin antibodies were used to detect specific proteins. The first two antibodies were detected simultaneously. Membranes were then incubated with the anti-Actin antibody. All extracts stemmed from IL-6 treated cells except for number 1 which was left untreated.*

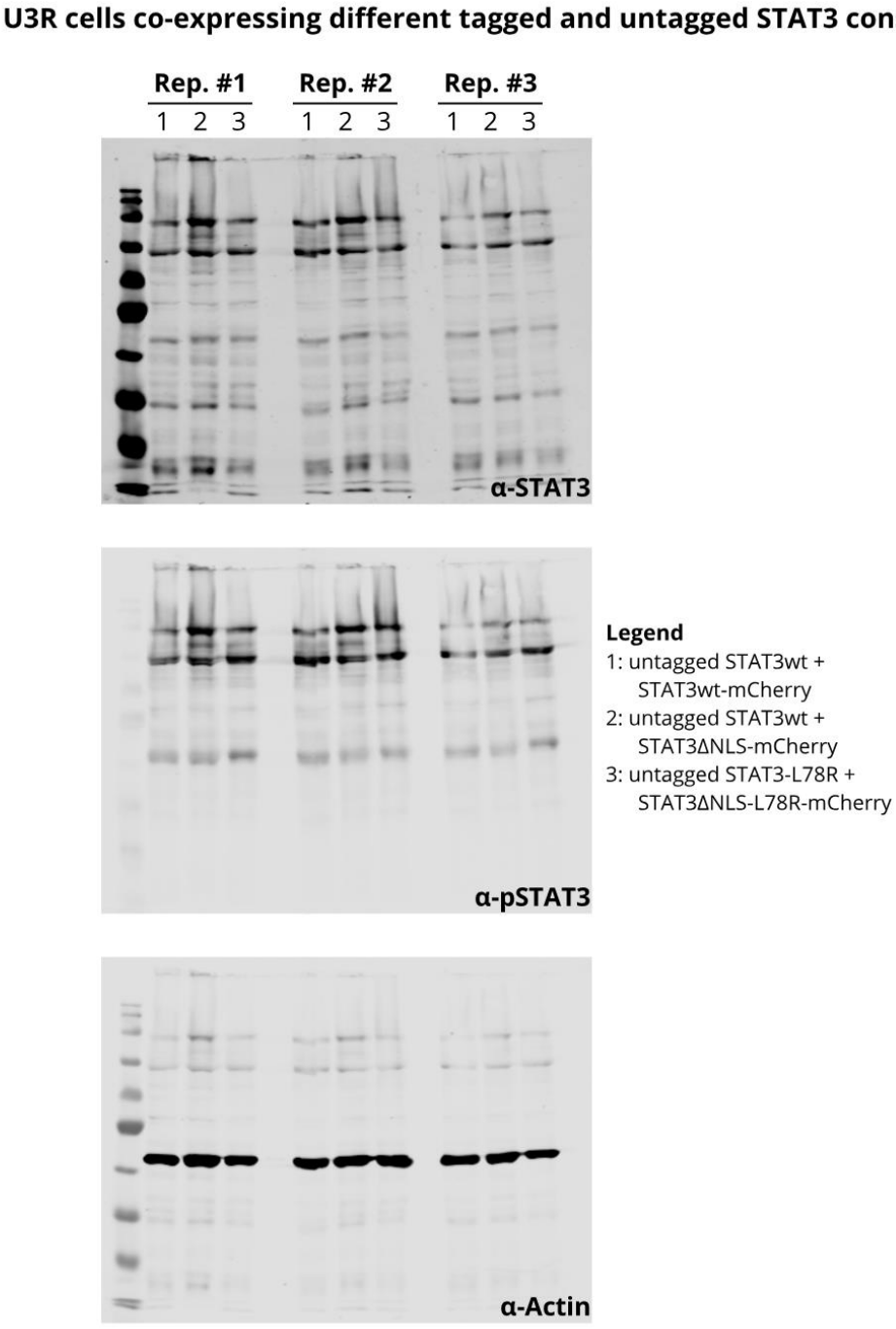


Appendix Figure 3 | Western blots of U3R cells expressing different STAT5B constructs: *Anti-STAT5B, anti-pSTAT5B and anti-Actin antibodies were used to detect specific proteins. The first two antibodies were detected simultaneously. Membranes were then incubated with the anti-Actin antibody. All extracts stemmed from IFN-γ treated cells except for number 1 which was left untreated.*

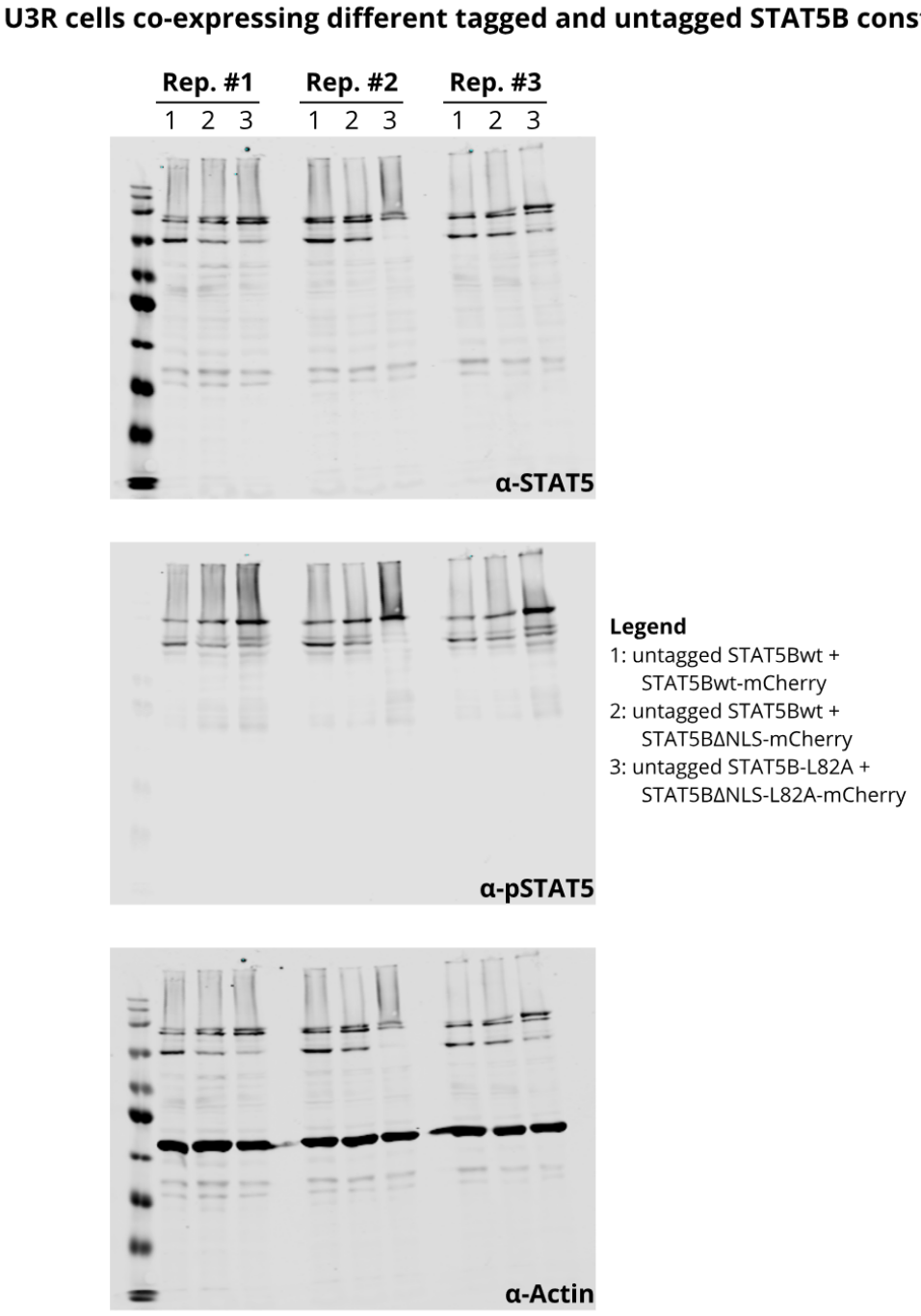
(ii) Western blots of cells co-expression different constructs of the same STAT species



Appendix Figure 4 | Western blots of U3R cells co-expressing different tagged and untagged STAT1 constructs: *Anti-STAT1, anti-pSTAT1 and anti-Actin antibodies were used to detect specific proteins. The first two antibodies were detected simultaneously. Membranes were then incubated with the anti-Actin antibody. All extracts stemmed from IFN-γ treated cells.*

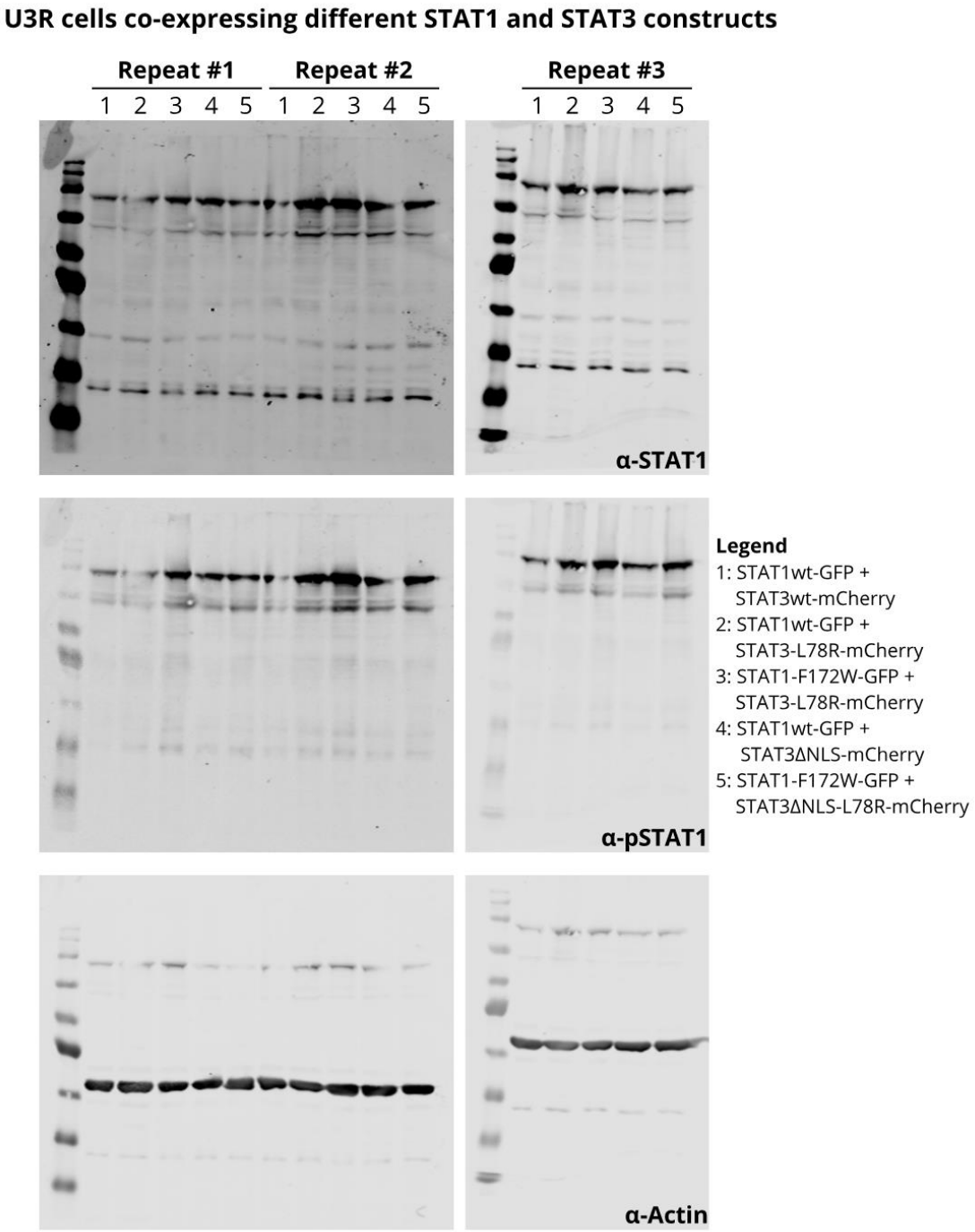


Appendix Figure 5 | Western blots of U3R cells co-expressing different tagged and untagged STAT3 constructs: *Anti-STAT3, anti-pSTAT3 and anti-Actin antibodies were used to detect specific proteins. The first two antibodies were detected simultaneously. Membranes were then incubated with the anti-Actin antibody. All extracts stemmed from IL-6 treated cells.*

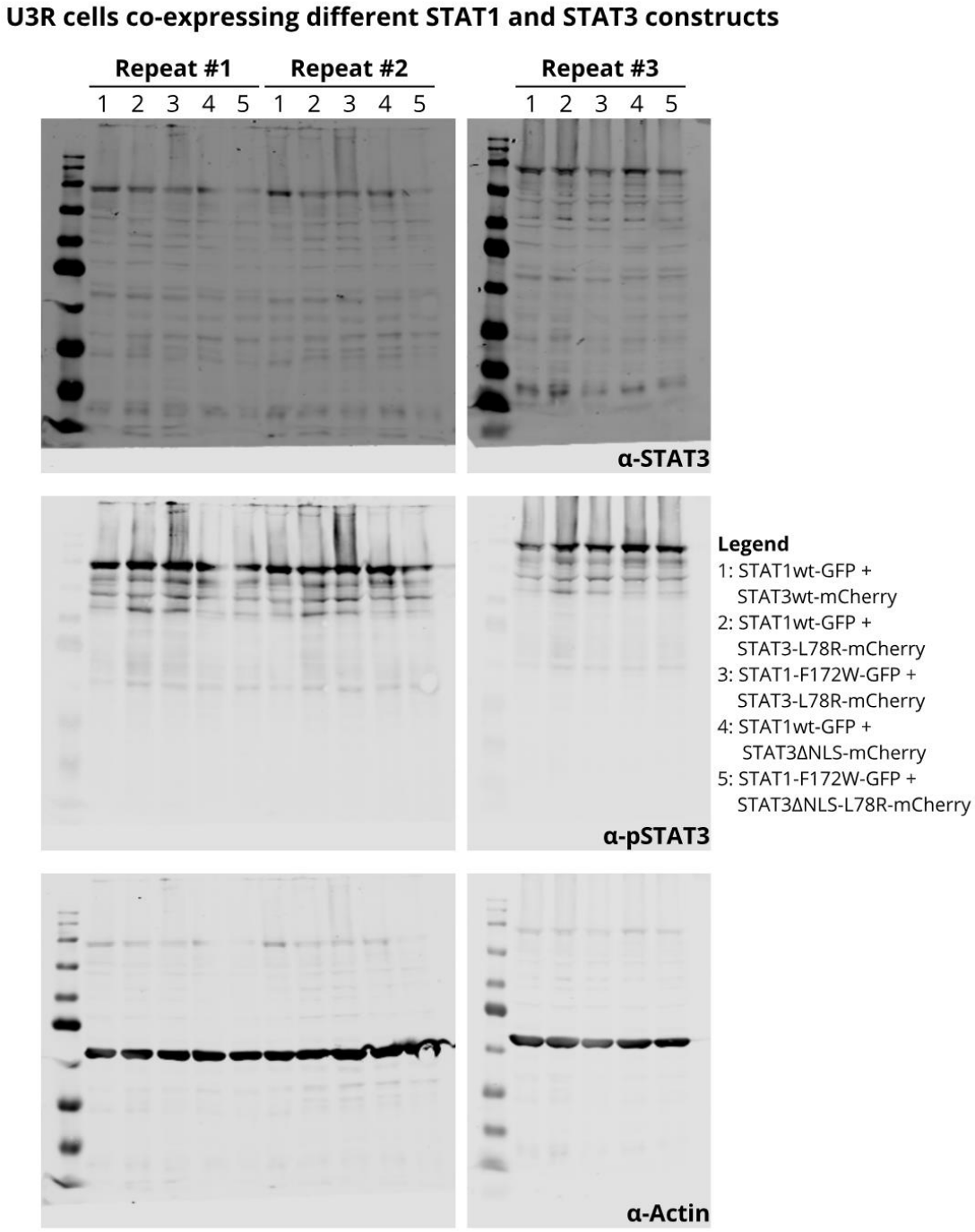


Appendix Figure 6 | Western blots of U3R cells co-expressing different tagged and untagged STAT5B constructs: *Anti-STAT5, anti-pSTAT5 and anti-Actin antibodies were used to detect specific proteins. The first two antibodies were detected simultaneously. Membranes were then incubated with the anti-Actin antibody. All extracts stemmed from IFN-γ treated cells.*

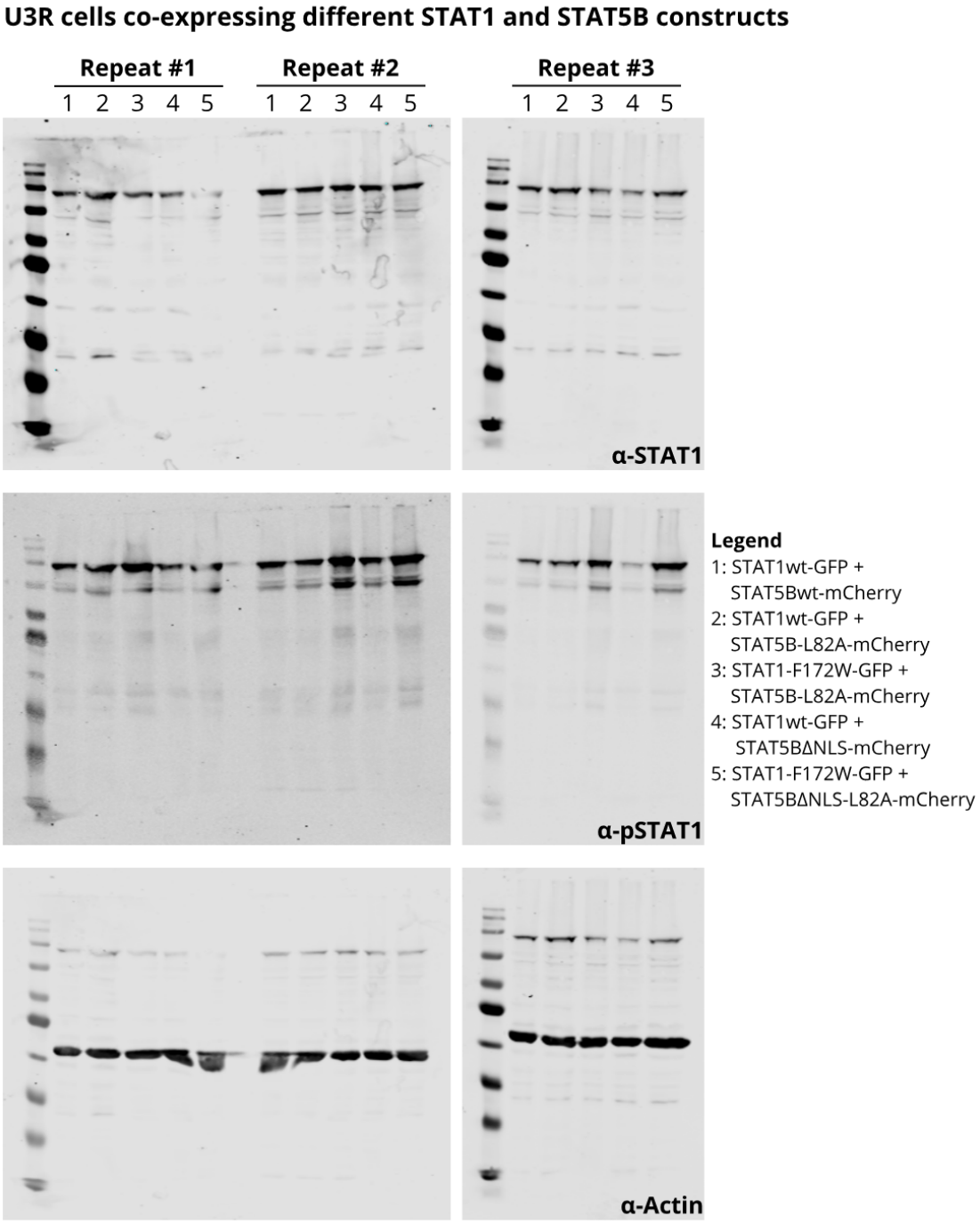
(iii) Western blots of cells co-expressing different STAT1 and STAT3 constructs



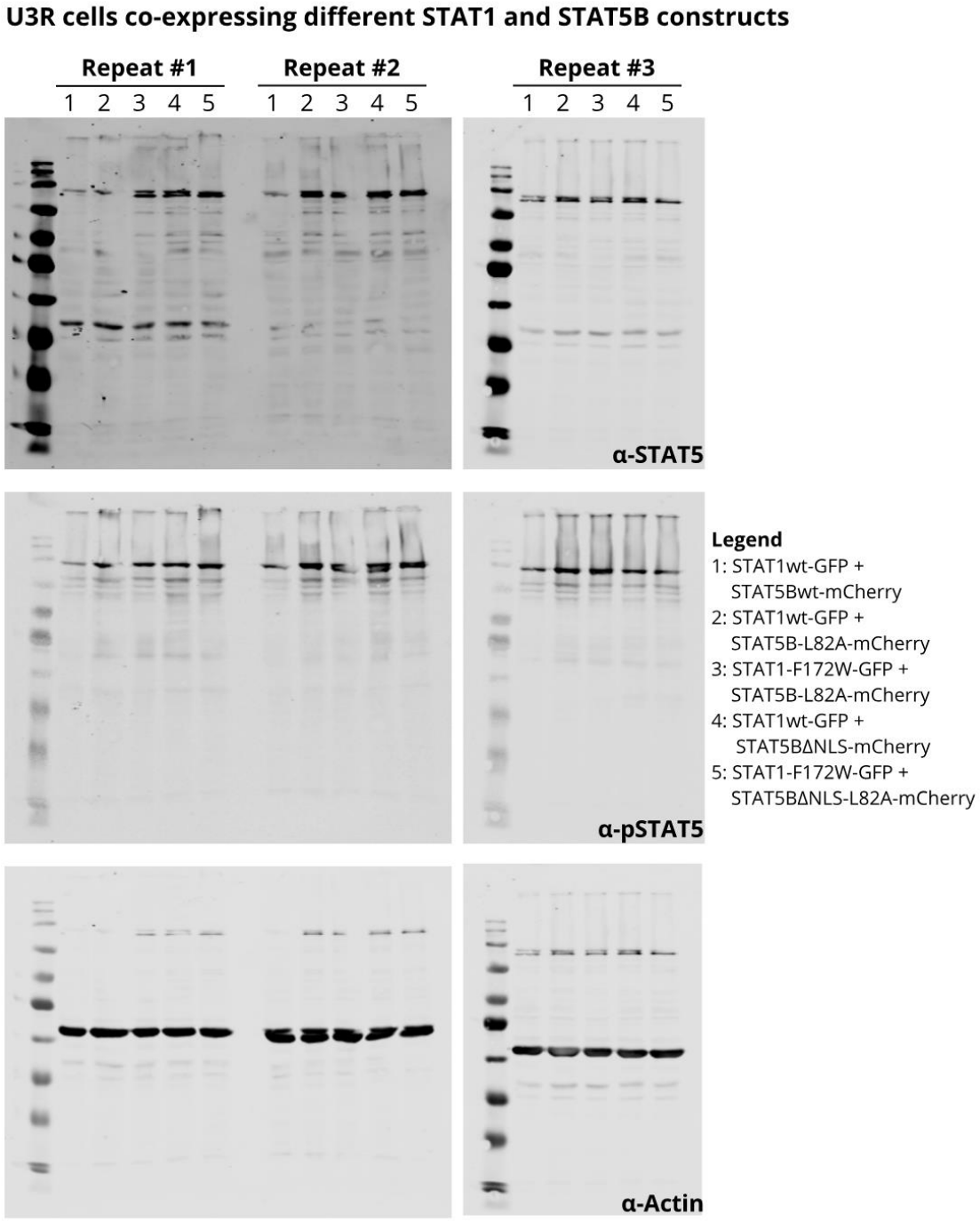
Appendix Figure 7 | Western blots of U3R cells co-expressing different STAT1 and STAT3 constructs: *The same extracts were loaded onto two separate SDS-PAGE gels and blotted onto two separate membranes. This figure depicts the membrane that was incubated with anti-STAT1, anti-pSTAT1 and anti-Actin antibodies to detect specific proteins. The first two antibodies were detected simultaneously. Membranes were then incubated with the anti-Actin antibody. All extracts stemmed from IL-6 & IFN-γ treated cells.*



Appendix Figure 8 | Western blots of U3R cells co-expressing different STAT1 and STAT3 constructs: *The same extracts were loaded onto two separate SDS-PAGE gels and blotted onto two separate membranes. This figure depicts the membrane that was incubated with anti-STAT3, anti-pSTAT3 and anti-Actin antibodies to detect specific proteins. The first two antibodies were detected simultaneously. Membranes were then incubated with the anti-Actin antibody. All extracts stemmed from IL-6 & IFN-γ treated cells.*



Appendix Figure 9 | Western blots of U3R cells co-expressing different STAT1 and STAT5B constructs: *The same extracts were loaded onto two separate SDS-PAGE gels and blotted onto two separate membranes. This figure depicts the membrane that was incubated with anti-STAT1, anti-pSTAT1 and anti-Actin antibodies to detect specific proteins. The first two antibodies were detected simultaneously. Membranes were then incubated with the anti-Actin antibody. All extracts stemmed from GH & IFN-γ treated cells.*



Appendix Figure 10 | Western blots of U3R cells co-expressing different STAT1 and STAT5B constructs: *The same extracts were loaded onto two separate SDS-PAGE gels and blotted onto two separate membranes. This figure depicts the membrane that was incubated with anti-STAT5, anti-pSTAT5 and anti-Actin antibodies to detect specific proteins. The first two antibodies were detected simultaneously. Membranes were then incubated with the anti-Actin antibody. All extracts stemmed from GH & IFN-γ treated cells.*

B. Raw FRET data

Appendix Table 1| FRET measurements for fGFP-mCherry fluorescent standard: *GFP and mCherry intensities were measured for 40 individual regions of interest (ROIs) just before and just after FRET acceptor photobleaching. FRET efficiency E_{FRET} and distance between fluorophores were calculated using the data obtained during measurements.*

ROI No	GFP before acc. Bleach [A.U.]	mCherry before acc. bleach [A.U.]	GFP after acc. bleach [A.U.]	mCherry after acc. bleach [AU]	E_{FRET} [%]	Distance r [nm]
1	1982	14910	2283	1637	13.19	7.23
2	823	1988	807	499	-2.07	NA
3	1009	2717	990	570	-1.99	NA
4	1375	12203	1637	1249	16.04	6.96
5	1323	5794	1294	929	-2.27	NA
6	2271	7825	2051	1202	-10.71	NA
7	1094	3967	1285	671	14.88	7.06
8	1565	5073	1719	814	8.98	7.77
9	940	4279	1078	803	12.83	7.27
10	1970	7218	2003	1353	1.63	10.46
11	1620	7864	2178	1359	25.65	6.3
12	2684	12226	3551	1845	24.42	6.37
13	1304	4813	1653	980	21.13	6.58
14	1715	7043	2205	1159	22.21	6.51
15	1758	6068	2362	959	25.59	6.31
16	1658	5326	2033	1010	18.46	6.76
17	2540	8804	3187	1651	20.28	6.63
18	1376	4552	1680	856	18.11	6.79
19	2225	7449	2641	1475	15.77	6.98
20	1224	3036	1374	647	10.93	7.49
21	6882	45898	10425	7883	33.98	5.9
22	8626	39917	12649	5806	31.8	6
23	2667	6213	3120	1009	14.51	7.1
24	5920	13524	6868	1786	13.81	7.16
25	2723	6095	3108	1005	12.4	7.31
26	1524	2859	1782	613	14.5	7.1
27	4011	10040	4677	1261	14.22	7.12
28	4469	10344	5167	1420	13.52	7.19
29	1508	2798	1631	624	7.54	8.02
30	2453	5154	2791	837	12.1	7.35
31	1768	3685	1960	774	9.8	7.64
32	1493	2820	1676	708	10.91	7.49
33	1569	2985	1830	724	14.28	7.12
34	1767	3491	2183	805	19.05	6.72
35	1101	1853	1237	559	11.01	7.48
36	1610	3264	1906	731	15.56	7

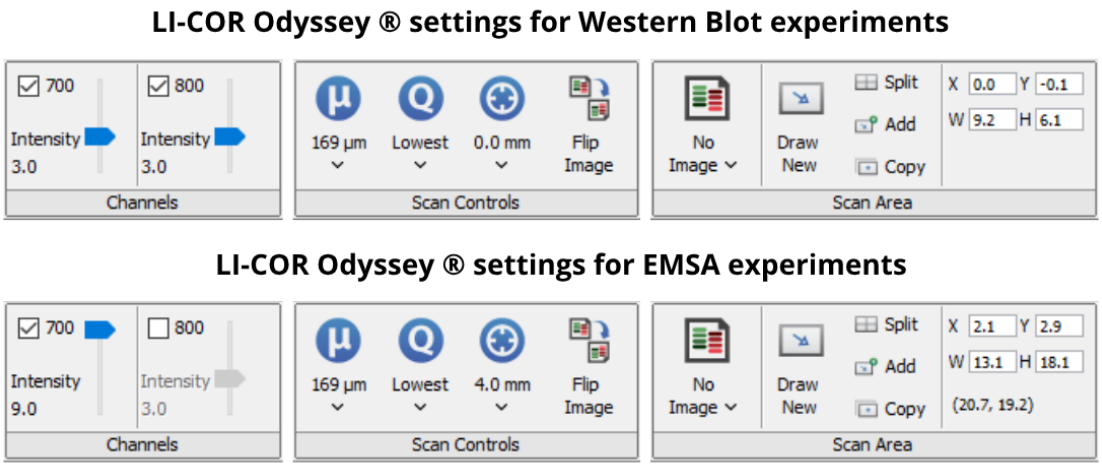
37	1721	3917	1954	839	11.91	7.37
38	2253	5237	2753	884	18.17	6.79
39	1071	1748	1208	558	11.31	7.44
40	3636	10237	4287	1400	15.17	7.03

Appendix Table 2| FRET measurements for rGFP-mCherry fluorescent standard: *GFP and mCherry intensities were measured for 30 individual regions of interest (ROIs) just before and just after FRET acceptor photobleaching. FRET efficiency E_{FRET} and distance between fluorophores were calculated using the data obtained during measurements.*

ROI No	GFP before acc. bleach [A.U.]	mCherry before acc. bleach [A.U.]	GFP after acc. bleach [A.U.]	mCherry after acc. bleach [AU]	E_{FRET} [%]	Distance r [nm]
1	1757	3519	1931	1261	9.02	7.76
2	5042	14353	5662	2849	10.96	7.49
3	4141	11732	4648	3009	10.91	7.49
4	2300	5918	2384	1721	3.53	9.17
5	4652	13173	5107	3659	8.91	7.78
6	6460	17472	7163	4048	9.82	7.64
7	3905	11718	4341	2234	10.03	7.61
8	10189	36875	11958	4992	14.8	7.07
9	3740	11333	4297	2381	12.97	7.25
10	4797	13199	5333	2830	10.06	7.61
11	2818	9312	3462	2492	18.62	6.75
12	1456	3787	1648	1050	11.63	7.4
13	1943	5237	2172	1433	10.58	7.54
14	2480	9364	3042	2213	18.48	6.76
15	3935	15447	4955	2796	20.59	6.61
16	2094	6457	2427	1559	13.74	7.17
17	1847	5494	2126	1346	13.14	7.23
18	1583	4228	1889	1809	16.21	6.94
19	2120	6500	2497	1631	15.08	7.04
20	1759	5277	1944	1167	9.53	7.68
21	2521	7341	2916	2108	13.54	7.19
22	2667	8743	3026	2481	11.85	7.38
23	2566	8566	3090	2070	16.98	6.88
24	5924	23471	6980	5005	15.13	7.04
25	3035	10753	3531	2121	14.05	7.14
26	3117	10521	3517	2083	11.37	7.44
27	2813	9351	3234	1997	13.03	7.24
28	2215	7109	2600	1852	14.82	7.07
29	3020	9936	3520	2065	14.19	7.13
30	1509	4144	1619	1227	6.79	8.17

C. Equipment settings used in this work.

(i) LICOR Odyssey settings for Western blot and EMSA analysis



Appendix Figure 11 | LI-COR Odyssey® Imaging System settings used for Western blot and EMSA experiments in this work.

(ii) Live cell imaging settings for Zeiss LSM710

Zeiss LSM710 settings to image samples containing STAT3 constructs



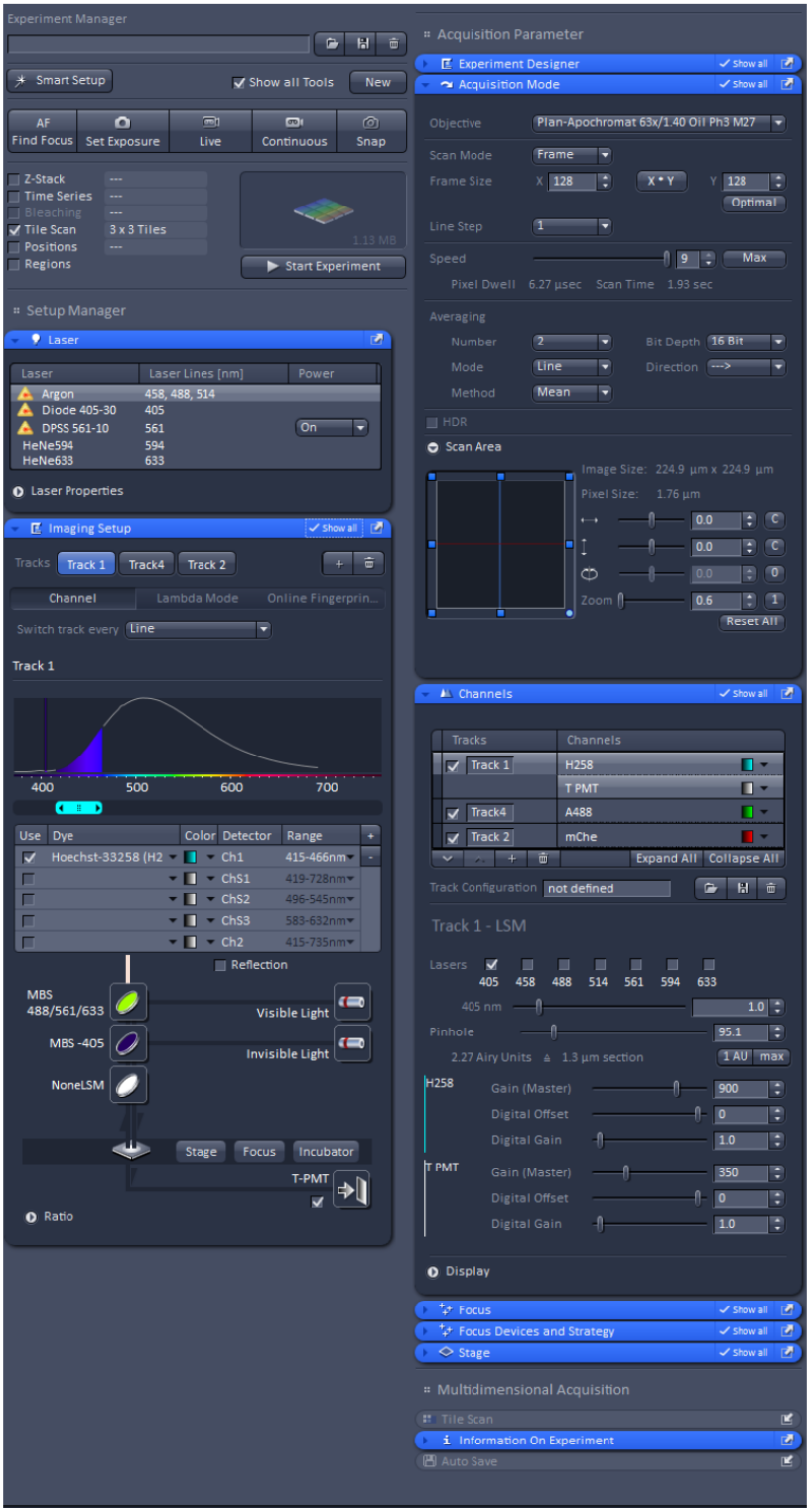
Appendix Figure 13 | Imaging settings used to image cells containing STAT3 constructs using the Zeiss LSM710 microscope in a live-cell imaging setting.

Zeiss LSM710 settings to image samples containing STAT5 constructs



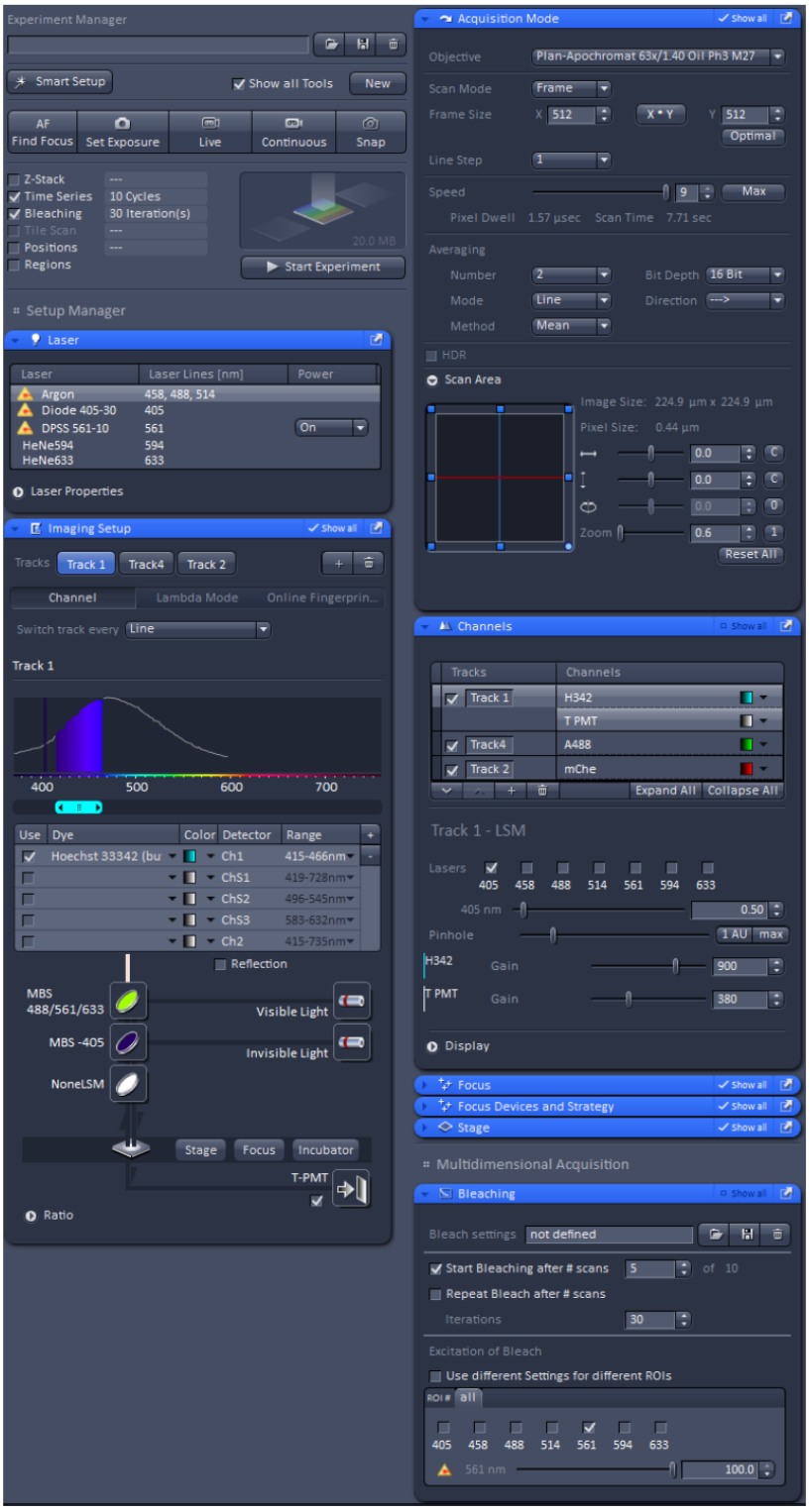
Appendix Figure 14 | Imaging settings used to image cells containing STAT5B constructs using the Zeiss LSM710 microscope in a live-cell imaging setting.

Zeiss LSM710 settings to generate a large low resolution Map image



Appendix Figure 15 | Imaging settings used to generate a large-scale map at low resolution: *The map was used to distinguish cells expressing fluorescent-tagged proteins from non-expressing cells. The position of fluorescent cells was marked for imaging at high resolution based on these maps.*

Zeiss LSM710 settings for FRET experiments



Appendix Figure 16 | Imaging settings used for FRET experiments.

D. Computer programming code used in this work.

(i) ImageJ macros coded in Java.

Whole cell segmentation macro

```
macro "WholeCellSegmentation" {
    run("Z Project...", "projection=[Sum Slices]");
    run("Duplicate...", "duplicate channels=3");
    run("Gaussian Blur...", "sigma=2");
    setAutoThreshold("Default dark");
    //run("Threshold...");
    setThreshold(5600.0000, 1000000000000000000000000000000.0000);//Set for ind. exp.
    setThreshold(5600.0000, 1000000000000000000000000000000.0000);
    setOption("BlackBackground", true);
    run("Convert to Mask");
    run("Analyze Particles...", "size=100-Infinity add");
    roiManager("Select", 0);
}
```

Nuclear segmentation macro

```
macro "NuclearSegnmentation" {
    run("Duplicate...", "duplicate channels=1 slices=5 frames=2");
    run("Gaussian Blur...", "sigma=2");
    setAutoThreshold("Default dark");
    //run("Threshold...");
    setThreshold(6000, 65535); //Set for individual experiments
    setThreshold(6000, 65535);
    setOption("BlackBackground", true);
    run("Convert to Mask");
    run("Analyze Particles...", "size=100.00-Infinity add");
}
```

(ii) Data analysis and graphical plotting in R

Western Blot Data Analysis and Graphing

```
library(tidyverse)

library(ggplot2)

library(ggpubr)

############# Set plot theme

thesis_theme <- theme_light() + theme(

  axis.text = element_text(size = 14),

  axis.title = element_text(size = 16, face = "bold"),

  legend.text = element_text(size = 14),

  legend.title = element_text(size = 16, face = "bold"),

  plot.title = element_text(size = 18, face = "bold"))

# ------- Read data and expand

## Read data

West <- data.frame(read.csv("./Data/QuantU3R.csv",

                            header = TRUE,

                            fileEncoding = 'UTF-8-BOM'))

## Calculate ratios

West$pY_FL_ratio <- West$py / West$FL  # anti-pSTAT/anti-STAT ratio

West$FL_Actin_ratio <- West$FL / West$Actin  # anti-STAT/anti-Actin ratio

# ------- Western blot descriptive statistics
```

```r
## Get descriptive statistics of anti-pSTAT/anti-STAT ratio for each combination
sWest <- West %>%
  group_by(Constructs) %>%
  summarise(
    sd = sd(pY_FL_ratio),  #Get SD
    pY_FL_ratio = mean(pY_FL_ratio),  #Get Mean
    nummer = n_distinct(pY_FL_ratio))  #Get No. of experiments
sWest$se <- sWest$sd/sqrt(sWest$nummer) #Calculate SE
sWest <- sWest %>% mutate_at(vars(roundslen, roundsSE), funs(round(.,2))) # Round
write.csv(sWest,"./Descriptive_statistics/Westerns.csv")
# ------- Single transfections
#### S1wt vs S1dNLS
## Narrow df
WTNLS1 <- West[West$Constructs %in% c("S1wt ","S1dNLS"),]
## Plot
pWTNLS1 <- ggbarplot(WTNLS1, x = "Constructs", y = "pY_FL_ratio",
                     add = c("mean", "jitter"),
                     fill = "#A1C5FF",
                     add.params = list(color = "#5E5EFF", size = 5),
                     size = 1.5,
                     ylab = "anti-pSTAT1/anti-STAT1\n",
                     xlab = "\nConstructs")
pWTNLS1 <- pWTNLS1 + stat_compare_means(method = "t.test",
                                        label = "p.signif",
                                        label.y = 0.9,
                                        label.x = 1.5,
                                        size = 10) +
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15))) +
  scale_x_discrete(labels = c("S1wt", "S1ΔNLS"))
ggsave("./Plots/Singles/pWTNLS1.png", width = 3.5, height = 5.5, dpi = 300)
#### S1-F172W vs S1dNLS-F172L
## Narrow df
S1F172 <- West[West$Constructs %in% c("S1F172W","S1dNLSF172L "),]
## Plot
pS1F172 <- ggbarplot(S1F172, x = "Constructs", y = "pY_FL_ratio",
                     add = c("mean", "jitter"),
                     fill = "#A1C5FF",
                     add.params = list(color = "#5E5EFF", size = 5),
                     size = 1.5,
                     ylab = "anti-pSTAT1/anti-STAT1\n",
                     xlab = "")
```

```
pS1F172 <- pS1F172 + stat_compare_means(method = "t.test",
                                       label = "p.signif",
                                       label.y = 1.2,
                                       label.x = 1.5,
                                       size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15))) +
  scale_x_discrete(labels = c("S1-F172W-\nGFP", "S1ΔNLS-\nF172L-GFP"))
ggsave("./Plots/F172/pS1F172.png", width = 3.5, height = 5.5, dpi = 300)
#### S3wt vs S3dNLS
## Narrow df
WTNLS3 <- West[West$Constructs %in% c("S3wt","S3dNLS"),]
## Plot
pWTNLS3 <- ggbarplot(WTNLS3, x = "Constructs", y = "pY_FL_ratio",
                     add = c("mean", "jitter"),
                     fill = "#80D5CC",
                     add.params = list(color = "#009688", size = 5),
                     size = 1.5,
                     ylab = "anti-pSTAT3/anti-STAT3\n",
                     xlab = "\nConstructs")
pWTNLS3 <- pWTNLS3 + stat_compare_means(method = "t.test",
                                       label = "p.signif",
                                       label.y = 9,
                                       label.x = 1.5,
                                       size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15)))+
  scale_x_discrete(labels = c("S3wt", "S3ΔNLS"))
ggsave("./Plots/Singles/pWTNLS3.png", width = 3.5, height = 5.5, dpi = 300)
#### S3-L78R vs S3dNLS-L78R
## Narrow df
S3L78R <- West[West$Constructs %in% c("S3L78R","S3dNLSL78R"),]
## Plot
pS3L78R <- ggbarplot(S3L78R, x = "Constructs", y = "pY_FL_ratio",
                     add = c("mean", "jitter"),
                     fill = "#80D5CC",
                     add.params = list(color = "#009688", size = 5),
                     size = 1.5,
                     ylab = "anti-pSTAT3/anti-STAT3\n",
                     xlab = "")
pS3L78R <- pS3L78R + stat_compare_means(method = "t.test",
                                       label = "p.signif",
                                       label.y = 10,
```

```
                                        label.x = 1.5,
                                        size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15))) +
  scale_x_discrete(labels = c("S3L78R-\nmCherry", "S3ΔNLS-\nL78R-mCherry"))
ggsave("./Plots/F172/pS3L78R.png", width = 3.5, height = 5.5, dpi = 300)
#### S5Bwt vs S5BdNLS
## Narrow df
WTNLS5 <- West[West$Constructs %in% c("S5Bwt","S5BdNLS"),]
## Plot
pWTNLS5 <- ggbarplot(WTNLS5, x = "Constructs", y = "pY_FL_ratio",
                     add = c("mean", "jitter"),
                     fill = "#E153BB",
                     add.params = list(color = "#96006E", size = 5),
                     size = 1.5,
                     ylab = "anti-pSTAT5/anti-STAT5\n",
                     xlab = "\nConstructs")
pWTNLS5 <- pWTNLS5 + stat_compare_means(method = "t.test",
                                        label = "p.signif",
                                        label.y = 0.9,
                                        label.x = 1.5,
                                        size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15)))+
  scale_x_discrete(labels = c("S5Bwt", "S5BΔNLS"))
ggsave("./Plots/Singles/pWTNLS5.png", width = 3.5, height = 5.5, dpi = 300)
#### S5B-L82A
## Narrow df
S5L82Asolo <- West[West$Constructs %in% c("S5BL82A"),]
## Plot
pS5L82Asolo <- ggbarplot(S5L82Asolo, x = "Constructs", y = "pY_FL_ratio",
                         add = c("mean", "jitter"),
                         fill = "#E153BB",
                         add.params = list(color = "#96006E", size = 5),
                         size = 1.5,
                         ylab = "anti-pSTAT5/anti-STAT5\n",
                         xlab = "")
pS5L82Asolo <- pS5L82Asolo +
  thesis_theme + scale_y_continuous(breaks = seq(0,2,0.25), limits = c(0,1)) +
  theme(axis.title.x = element_blank()) +
  scale_x_discrete(labels = c("S5B-L82A-mCherry"))
ggsave("./Plots/L82A/pS5L82Asolo.png", width = 2.7, height = 4, dpi = 300)
#### S5BdNLS-L82A
```

```
## Narrow df
S5dNLSL82Asolo <- West[West$Constructs %in% c("S5BdNLSL82A"),]
## Plot
pS5dNLSL82Asolo <- ggbarplot(S5dNLSL82Asolo, x = "Constructs", y = "pY_FL_ratio",
                             add = c("mean", "jitter"),
                             fill = "#E153BB",
                             add.params = list(color = "#96006E", size = 5),
                             size = 1.5,
                             ylab = "anti-pSTAT5/anti-STAT5\n",
                             xlab = "")
pS5dNLSL82Asolo <- pS5dNLSL82Asolo +
  thesis_theme + scale_y_continuous(breaks = seq(0,2,0.25), limits = c(0,1)) +
  theme(axis.title.x = element_blank()) +
  scale_x_discrete(labels = c("S5BΔNLS-L82A"))
ggsave("./Plots/L82A/pS5dNLSL82Asolo.png", width = 2.7, height = 4, dpi = 300)
# ------- Homodimer co-transfections
#### S1wt+S1wt vs S1wt+S1dNLS
## Narrow df
WTNLSCo1 <- West[West$Constructs %in% c("S1ut+S1wt",
                                        "S1ut+S1dNLS"),]
## Plot
pWTNLSCo1 <- ggbarplot(WTNLSCo1, x = "Constructs", y = "pY_FL_ratio",
                       add = c("mean", "jitter"),
                       fill = "#A1C5FF",
                       add.params = list(color = "#5E5EFF", size = 5),
                       size = 1.5,
                       ylab = "anti-pSTAT1/anti-STAT1\n",
                       xlab = "\nConstructs")
pWTNLSCo1 <- pWTNLSCo1 + stat_compare_means(method = "t.test",
                                            label = "p.signif",
                                            label.y = 0.9,
                                            label.x = 1.5,
                                            size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15)))+
  scale_x_discrete(labels = c("untag. S1wt +\nS1wt-GFP",
                              "untag. S1wt +\nS1ΔNLS-GFP"))
ggsave("./Plots/HomoCo/pWTNLSCo1.png", width = 4.5, height = 5.5, dpi = 300)
#### S1-F172W+S1dNLS-F172L
## Narrow df
S1F172vsdNLS <- West[West$Constructs %in% c("S1F172Wut+S1dNLSF172W",
                                            "S1dNLSF172W+S1F172Wut"),]
```

```
## Plot
pS1F172vsdNLS <- ggbarplot(S1F172vsdNLS, x = "Constructs", y = "pY_FL_ratio",
                           add = c("mean", "jitter"),
                           fill = "#A1C5FF",
                           add.params = list(color = "#5E5EFF", size = 5),
                           size = 1.5,
                           ylab = "anti-pSTAT1/anti-STAT1\n",
                           xlab = "")
pS1F172vsdNLS <- pS1F172vsdNLS + stat_compare_means(method = "t.test",
                                                    label = "p.signif",
                                                    label.y = 1.2,
                                                    label.x = 1.5,
                                                    size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15))) +
  scale_x_discrete(labels = c("S1-F172W ut +\nS1ΔNLS-\nF172L-GFP",
                              "S1-F172W ut +\nS1ΔNLS-\nF172L-GFP"))
ggsave("./Plots/F172/pS1F172vsdNLS.png", width = 4.5, height = 5.5, dpi = 300)
#### S3wt+S3wt vs S3wt+S3dNLS
## Narrow df
WTNLSCo3 <- West[West$Constructs %in% c("S3ut+S3wt",
                                        "S3ut+S3dNLS"),]
## Plot
pWTNLSCo3 <- ggbarplot(WTNLSCo3, x = "Constructs", y = "pY_FL_ratio",
                       add = c("mean", "jitter"),
                       fill = "#80D5CC",
                       add.params = list(color = "#009688", size = 5),
                       size = 1.5,
                       ylab = "anti-pSTAT3/anti-STAT3\n",
                       xlab = "\nConstructs")
pWTNLSCo3 <- pWTNLSCo3 + stat_compare_means(method = "t.test",
                                            label = "p.signif",
                                            label.y = 2,
                                            label.x = 1.5,
                                            size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15)))+
  scale_x_discrete(labels = c("untag. S3wt +\nS3wt-GFP",
                              "untag. S3wt +\nS3ΔNLS-GFP"))
ggsave("./Plots/HomoCo/pWTNLSCo3.png", width = 4.5, height = 5.5, dpi = 300)
#### S3-L78R+S3dNLS-L78R
## Narrow df
S3L78RvsdNLS <- West[West$Constructs %in% c("S3L78Rut+S3dNLSL78R",
```

```
                                                "S3dNLSF172W+S3L78Rut"),]
## Plot
pS3L78RvsdNLS <- ggbarplot(S3L78RvsdNLS, x = "Constructs", y = "pY_FL_ratio",
                           add = c("mean", "jitter"),
                           fill = "#80D5CC",
                           add.params = list(color = "#009688", size = 5),
                           size = 1.5,
                           ylab = "anti-pSTAT3/anti-STAT3\n",
                           xlab = "")
pS3L78RvsdNLS <- pS3L78RvsdNLS + stat_compare_means(method = "t.test",
                                                    label = "p.signif",
                                                    label.y = 2.5,
                                                    label.x = 1.5,
                                                    size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15))) +
  scale_x_discrete(labels = c("S3-L78R ut +\nS3ΔNLS-\nL78R-mCherry",
                              "S3-L78R ut +\nS3ΔNLS-\nL78R-mCherry"))
ggsave("./Plots/F172/pS3L78RvsdNLS.png", width = 4.5, height = 5.5, dpi = 300)
#### S5Bwt+S5Bwt vs S5Bwt+S5BdNLS
## Narrow df
WTNLSCo5 <- West[West$Constructs %in% c("S5But+S5Bwt",
                                        "S5But+S5BdNLS"),]
## Plot
pWTNLSCo5 <- ggbarplot(WTNLSCo5, x = "Constructs", y = "pY_FL_ratio",
                       add = c("mean", "jitter"),
                       fill = "#E153BB",
                       add.params = list(color = "#96006E", size = 5),
                       size = 1.5,
                       ylab = "anti-pSTAT5/anti-STAT5\n",
                       xlab = "\nConstructs")
pWTNLSCo5 <- pWTNLSCo5 + stat_compare_means(method = "t.test",
                                            label = "p.signif",
                                            label.y = 0.9,
                                            label.x = 1.5,
                                            size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15)))+
  scale_x_discrete(labels = c("untag. S5Bwt +\nS5Bwt-GFP",
                              "untag. S5Bwt +\nS5BΔNLS-GFP"))
ggsave("./Plots/HomoCo/pWTNLSCo5.png", width = 4.5, height = 5.5, dpi = 300)
# ------- Heterodimer co-transfections
##### S1/S3 co-expression
```

```
### Detect STAT1
## Narrow df
dfS1S3 <- West[West$Constructs %in% c("S1wt+S3wt",
                                      "S1wt+S3L78R",
                                      "S1F172W+S3L78R",
                                      "S1wt+S3dNLS",
                                      "S1F172W+S3dNLSL78R"),]
## Plot
pS1S3 <- ggbarplot(dfS1S3, x = "Constructs", y = "pY_FL_ratio",
                   add = c("mean", "jitter"), order = c("S1wt+S3wt",
                                                        "S1wt+S3L78R",
                                                        "S1F172W+S3L78R",
                                                        "S1wt+S3dNLS",
                                                        "S1F172W+S3dNLSL78R"),
                  fill = "#A1C5FF",
                  add.params = list(color = "#5E5EFF", size = 5),
                  size = 1.5,
                  ylab = "anti-pSTAT1/anti-STAT1\n",
                  xlab = "")
pS1S3 <- pS1S3 + stat_compare_means(comparisons = list(c("S1wt+S3wt","S1wt+S3L78R"),
                                                       c("S1wt+S3wt","S1F172W+S3L78R"),
                                                       c("S1wt+S3wt","S1wt+S3dNLS"),
                                                     c("S1wt+S3wt","S1F172W+S3dNLSL78R")),
                                    method = "t.test",
                                    label = "p.signif",
                                    label.y = 0.8,
                                    label.x = 1.5,
                                    size = 10,
                                    step.increase = 0.2)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15))) +
  theme(axis.ticks = element_blank(), axis.text.x = element_blank())
ggsave("./Plots/Hetero/pS1S3.png", width = 11.5, height = 5.5, dpi = 300)
### Detect STAT3
## Narrow df
dfS3S1 <- West[West$Constructs %in% c("S3wt+S1wt",
                                      "S3L78R+S1wt",
                                      "S3L78R+S1F172W",
                                      "S3dNLS+S1wt",
                                      "S3dNLSL78R+S1F172W"),]
## Plot
pS3S1 <- ggbarplot(dfS3S1, x = "Constructs", y = "pY_FL_ratio",
```

```
                    add = c("mean", "jitter"), order = c("S3wt+S1wt",
                                                         "S3L78R+S1wt",
                                                         "S3L78R+S1F172W",
                                                         "S3dNLS+S1wt",
                                                         "S3dNLSL78R+S1F172W"),
                    fill = "#80D5CC",
                    add.params = list(color = "#009688", size = 5),
                    size = 1.5,
                    ylab = "anti-pSTAT3/anti-STAT3\n",
                    xlab = "")
pS3S1 <- pS3S1 + stat_compare_means(comparisons = list(c("S3wt+S1wt", "S3L78R+S1wt"),
                                                       c("S3wt+S1wt", "S3L78R+S1F172W"),
                                                        c("S3wt+S1wt", "S3dNLS+S1wt"),
                                                      c("S3wt+S1wt", "S3dNLSL78R+S1F172W")),
                                    method = "t.test",
                                    label = "p.signif",
                                    label.y = 6,
                                    label.x = 1.5,
                                    size = 10,
                                    step.increase = 0.2)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15)),
                                    labels = function(x) format(x, nsmall = 1)) +
  theme(axis.title.x = element_blank()) +
  scale_x_discrete(labels = c("S1wt-GFP +\nS3wt-mCherry",
                              "S1wt-GFP +\nS3-L78R-mCherry",
                              "S1-F172W-GFP +\nS3-L78R-mCherry",
                              "S1wt-GFP +\nS3ΔNLS-mCherry",
                              "S1-F172W-GFP +\nS3ΔNLS-L78R-mCherry"),
                    guide = guide_axis(n.dodge = 2))
ggsave("./Plots/Hetero/pS3L78R.png", width = 11.5, height = 6.1, dpi = 300)
##### S1/S5B co-expression
#### S1wt+S5Bwt vs S1wt+S5BdNLS
### Detect STAT1
## Narrow df
WTNLSCoS5S1 <- West[West$Constructs %in% c("S1wt+S5Bwt", "S1wt+S5BdNLS"),]
## Plot
pWTNLSCoS5S1 <- ggbarplot(WTNLSCoS5S1, x = "Constructs", y = "pY_FL_ratio",
                          add = c("mean", "jitter"),
                          fill = "#A1C5FF",
                          add.params = list(color = "#5E5EFF", size = 5),
                          size = 1.5,
```

```
                          ylab = "anti-pSTAT1/anti-STAT1\n",
                          xlab = "\nConstructs")
pWTNLSCoS5S1 <- pWTNLSCoS5S1 + stat_compare_means(method = "t.test",
                                                  label = "p.signif",
                                                  label.y = 0.9,
                                                  label.x = 1.5,
                                                  size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15))) +
  xlab("") + scale_x_discrete(labels = c("S1wt-GFP +\nS5Bwt-mCherry",
                                         "S1wt-GFP +\nS5BΔNLS-mCherry"))
ggsave("./Plots/HeteroCoVal/pWTNLSCoS1S5.png", width = 4.8, height = 5.5, dpi = 300)
### Detect STAT5B
## Narrow df
WTNLSCoS1S5 <- West[West$Constructs %in% c("S5Bwt+S1wt", "S5BdNLS+S1wt"),]
## Plot
pWTNLSCoS1S5 <- ggbarplot(WTNLSCoS1S5, x = "Constructs", y = "pY_FL_ratio",
                          add = c("mean", "jitter"),
                          fill = "#E153BB",
                          add.params = list(color = "#96006E", size = 5),
                          size = 1.5,
                          ylab = "anti-pSTAT5/anti-STAT5\n",
                          xlab = "\nConstructs")
pWTNLSCoS1S5 <- pWTNLSCoS1S5 + stat_compare_means(method = "t.test",
                                                  label = "p.signif",
                                                  label.y = 0.9,
                                                  label.x = 1.5,
                                                  size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15))) +
  xlab("") + scale_x_discrete(labels = c("S1wt-GFP +\nS5Bwt-mCherry",
                                         "S1wt-GFP +\nS5BΔNLS-mCherry"))
ggsave("./Plots/HeteroCoVal/pWTNLSCoS5S1.png", width = 4.8, height = 5.5, dpi = 300)
#### S1-F172W+S5B-L82A
### Detect STAT1
## Narrow df
S1F172Wco <- West[West$Constructs %in% c("S1F172W+S5BL82A"),]
## Plot
pS1F172Wco <- ggbarplot(S1F172Wco, x = "Constructs", y = "pY_FL_ratio",
                        add = c("mean", "jitter"),
                        fill = "#A1C5FF",
                        add.params = list(color = "#5E5EFF", size = 5),
                        size = 1.5,
```

```
                          ylab = "anti-pSTAT1/anti-STAT1\n",
                          xlab = "")
pS1F172Wco <- pS1F172Wco + thesis_theme +
  scale_y_continuous(breaks = seq(0,2,0.25), limits = c(0,1)) +
  theme(axis.title.x = element_blank()) +
  scale_x_discrete(labels = c("S5B-L82A +\nS1-F172W"))
ggsave("./Plots/L82A/pS1F172Wco.png", width = 2.0, height = 5.5, dpi = 300)
### Detect STAT5B
## Narrow df
S5dNLSL82Asolo <- West[West$Constructs %in% c("S5BdNLSL82A"),]
## Plot
pS5dNLSL82Asolo <- ggbarplot(S5dNLSL82Asolo, x = "Constructs", y = "pY_FL_ratio",
                             add = c("mean", "jitter"),
                             fill = "#E153BB",
                             add.params = list(color = "#96006E", size = 5),
                             size = 1.5,
                             ylab = "anti-pSTAT5/anti-STAT5\n",
                             xlab = "")
pS5dNLSL82Asolo <- pS5dNLSL82Asolo + thesis_theme +
  scale_y_continuous(breaks = seq(0,2,0.25), limits = c(0,1)) +
  theme(axis.title.x = element_blank()) +
  scale_x_discrete(labels = c("S5BΔNLS-L82A"))
ggsave("./Plots/L82A/pS5dNLSL82Asolo.png", width = 2.7, height = 4, dpi = 300)
#### S1-F172W+S5BdNLS-L82A
### Detect STAT1
## Narrow df
S1F172Wco1 <- West[West$Constructs %in% c("S1F172W+S5BdNLSL82A"),]
## Plot
pS1F172Wco1 <- ggbarplot(S1F172Wco1, x = "Constructs", y = "pY_FL_ratio",
                         add = c("mean", "jitter"),
                         fill = "#A1C5FF",
                         add.params = list(color = "#5E5EFF", size = 5),
                         size = 1.5,
                         ylab = "anti-pSTAT1/anti-STAT1\n",
                         xlab = "")

pS1F172Wco1 <- pS1F172Wco1 + thesis_theme +
  scale_y_continuous(breaks = seq(0,2,0.25), limits = c(0,1)) +
  theme(axis.title.x = element_blank()) +
  scale_x_discrete(labels = c("S5BΔNLS-\nL82A +\nS1-F172W"))
ggsave("./Plots/L82A/pS1F172Wco1.png", width = 2.0, height = 5.5, dpi = 300)
```

```
### Detect STAT5B

## Narrow df

S5dNLSL82Aco <- West[West$Constructs %in% c("S5BdNLSL82A+S1F172W"),]

## Plot

pS5dNLSL82Aco <- ggbarplot(S5dNLSL82Aco, x = "Constructs", y = "pY_FL_ratio",

                           add = c("mean", "jitter"),

                           fill = "#E153BB",

                           add.params = list(color = "#96006E", size = 5),

                           size = 1.5,

                           ylab = "anti-pSTAT5/anti-STAT5\n",

                           xlab = "")


pS5dNLSL82Aco <- pS5dNLSL82Aco + thesis_theme +

  scale_y_continuous(breaks = seq(0,2,0.25), limits = c(0,1)) +

  theme(axis.title.x = element_blank()) +

  scale_x_discrete(labels = c("S5BΔNLS-\nL82A +\nS1-F172W"))

ggsave("./Plots/L82A/pS5dNLSL82Aco.png", width = 2.0, height = 5.5, dpi = 300)
```

Nuclear Translocation Assay data analysis and graphing

```
library(tidyverse)
library(ggplot2)
library(ggpubr)
############## Set plot theme
thesis_theme <- theme_light() + theme(
  axis.text = element_text(size = 14),
  axis.title = element_text(size = 16, face = "bold"),
  legend.text = element_text(size = 14),
  legend.title = element_text(size = 16, face = "bold"),
  plot.title = element_text(size = 18, face = "bold",),
  strip.text = element_text(size = 16, face = "bold"),
  strip.background = element_rect(fill = "black"))
# ------- Read data and expand
df150 <- data.frame(read.csv("./Data/analysisnewimp150.csv",
                             header = TRUE,
                             fileEncoding = 'UTF-8-BOM'))
# ------- Nuclear Translocation Assay descriptive statistics
#### Get descriptive for GFP and mCherry nuclear score
NTAsum <- df150 %>%
  group_by(Constructs) %>%
  summarise(
    GFP_Mean = mean(GFPdNuc0_90), # Get GFP mean
    GFP_SD = sd(GFPdNuc0_90), # Get GFP SD
    GFP_n = n_distinct(GFPdNuc0_90), # Get No
    Cherry_Mean = mean(CherrydNuc0_90), # Get mCherry mean
    Cherry_SD = sd(CherrydNuc0_90), # Get mCherry SD
    Cherry_n = n_distinct(CherrydNuc0_90)) # Get No
NTAsum$GFP_SE <- NTAsum$GFP_SD/sqrt(NTAsum$GFP_n) # Get GFP SE
NTAsum$Cherry_SE <- NTAsum$Cherry_SD/sqrt(NTAsum$Cherry_n) # Get mCherry SE
NTAsum <- NTAsum %>%
  mutate_at(vars(-GFP_n,-Cherry_n,-Constructs), funs(round(.,2))) %>% # Round
  select("Constructs", sort(colnames(.))) # Sort
write.csv(NTAsum,"./Descriptive_statistics/NTA/summaryNTA.csv")
# ------- NTA assay validation
##### S1wt+S1wt vs S1wt+S1dNLS
# Narrow df
df1wtNLS <- df150[df150$Include == "yes",]
df1wtNLS <- df1wtNLS[df1wtNLS$Constructs %in% c("S1wt+S1wt","S1wt+S1dNLS"),]
# Melt
df1wtNLS <- df1wtNLS %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
p1WTNLS <- ggviolin(df1wtNLS, x = "Constructs", y = "value",
```

```
                    facet.by = "variable",
                    panel.labs = list(variable = c("Cherry","GFP")),
                    fill = "variable",
                    width = 1.2,
                    add = "boxplot",
                    add.params = list(width = 0.1, size = 1.05)) +
  stat_compare_means(comparisons = list(c("S1wt+S1wt","S1wt+S1dNLS")),
                     label = "p.signif",
                     label.x = 1.5) + thesis_theme +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  scale_y_continuous(breaks = seq(0,10,1)) +
  theme(legend.position = "none") +
  scale_x_discrete(labels = c("S1wt-GFP +\nS1wt-mCherry",
                              "S1wt-GFP +\nS1ΔNLS-mCherry"))
ggsave("./Plots/NTA/p1WTNLS.png", width = 9, height = 5.5, dpi = 300)
##### S3wt+S3wt vs S3wt+S3dNLS
# Narrow df
df3wtNLS <- df150[df150$Include == "yes",]
df3wtNLS <- df3wtNLS[df3wtNLS$Constructs %in% c("S3wt+S3wt","S3wt+S3dNLS"),]
# Melt
df3wtNLS <- df3wtNLS %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p3WTNLS <- ggviolin(df3wtNLS, x = "Constructs", y = "value",
                    facet.by = "variable",
                    panel.labs = list(variable = c("Cherry","GFP")),
                    fill = "variable",
                    sort.val = "desc",
                    width = 1.2,
                    add = "boxplot",
                    add.params = list(width = 0.1, size = 1.05)) +
  stat_compare_means(comparisons = list(c("S3wt+S3wt","S3wt+S3dNLS")),
                     label = "p.signif",
                     label.x = 1.5) + thesis_theme +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  scale_y_continuous(breaks = seq(0,10,1)) +
  theme(legend.position = "none") +
  scale_x_discrete(limits=rev, labels = c("S3wt-GFP +\nS3wt-mCherry",
                                          "S3wt-GFP +\nS3ΔNLS-mCherry"))
ggsave("./Plots/NTA/p3WTNLS.png", width = 9, height = 5.5, dpi = 300)
##### S5Bwt+S5Bwt vs S5Bwt+S5BdNLS
# Narrow df
df5wtNLS <- df150[df150$Include == "yes",]
df5wtNLS <- df5wtNLS[df5wtNLS$Constructs %in% c("S5Bwt+S5Bwt","S5BdNLS+S5Bwt"),]
# Melt
df5wtNLS <- df5wtNLS %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p5WTNLS <- ggviolin(df5wtNLS, x = "Constructs", y = "value",
                    facet.by = "variable",
                    panel.labs = list(variable = c("Cherry","GFP")),
                    fill = "variable",
                    width = 1.2,
                    add = "boxplot",
                    add.params = list(width = 0.1, size = 1.05)) +
  stat_compare_means(comparisons = list(c("S5Bwt+S5Bwt","S5BdNLS+S5Bwt")),
                     label = "p.signif",
                     label.x = 1.5) + thesis_theme +
  scale fill discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  scale_y_continuous(breaks = seq(0,10,1)) +
  theme(legend.position = "none") +
  scale_x_discrete(labels = c("S5Bwt-GFP +\nS5Bwt-mCherry",
                              "S5Bwt-GFP +\nS5BΔNLS-mCherry"))
ggsave("./Plots/NTA/p5WTNLS.png", width = 9, height = 5.5, dpi = 300)
##### S1wt+S5Bwt vs S1wt+S5BdNLS
# Narrow df
df15wtNLS <- df150[df150$Include == "yes",]
df15wtNLS <- df15wtNLS[df15wtNLS$Constructs %in% c("S1wt+S5wt","S1wt+S5dNLS"),]
# Melt
```

```
df15wtNLS <- df15wtNLS %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p15WTNLS <- ggviolin(df15wtNLS, x = "Constructs", y = "value",
                     facet.by = "variable",
                     panel.labs = list(variable = c("Cherry","GFP")),
                     order = c("S1wt+S5wt","S1wt+S5dNLS"),
                     fill = "variable",
                     width = 1.2,
                     add = "boxplot",
                     add.params = list(width = 0.1, size = 1.05)) +
  stat_compare_means(comparisons = list(c("S1wt+S5wt","S1wt+S5dNLS")),
                     label = "p.signif",
                     label.x = 1.5) + thesis_theme +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  scale_y_continuous(breaks = seq(0,10,1)) +
  theme(legend.position = "none") +
  scale_x_discrete(labels = c("S1wt-GFP +\nS5Bwt-mCherry",
                              "S1wt-GFP +\nS5BΔNLS-mCherry"))
ggsave("./Plots/NTA/p15WTNLS.png", width = 9, height = 5.5, dpi = 300)
# ------- Characterise latent dimer splitting mutations
##### S1-F172W vs S1-F172L
# Narrow df
df1Mono <- df150[df150$Include == "yes",]
df1Mono <- df1Mono[df1Mono$Constructs %in% c("S1F172L+S1F172L",
                                             "S1F172W+S1F172W"),]
# Melt
df1Mono <- df1Mono %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p1Mono <- ggviolin(df1Mono, x = "Constructs", y = "value",
                   facet.by = "variable",
                   panel.labs = list(variable = c("Cherry","GFP")),
                   fill = "variable",
                   width = 1.2,
                   add = "boxplot",
                   add.params = list(width = 0.1, size = 1.05)) +
  stat_compare_means(comparisons = list(c("S1F172L+S1F172L",
                                          "S1F172W+S1F172W")),
                     label = "p.signif",
                     label.x = 1.5) + thesis_theme +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  scale_y_continuous(breaks = seq(0,10,1)) +
  theme(legend.position = "none") +
  scale_x_discrete(labels = c("S1-F172L-GFP +\nS1-F172L-mCherry",
                              "S1-F172W-GFP +\nS1-F172W-mCherry"))
ggsave("./Plots/NTA_F172/p1F172.png", width = 9, height = 5.5, dpi = 300)
##### S3-L78R
# Narrow df
df3L78R <- df150[df150$Include == "yes",]
df3L78R <- df3L78R[df3L78R$Constructs %in% c("S3L78R+S3L78R"),]
# Melt
df3L78R <- df3L78R %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p3L78R <- ggviolin(df3L78R, x = "Constructs", y = "value",
                   facet.by = "variable",
                   panel.labs = list(variable = c("Cherry","GFP")),
                   fill = "variable",
                   width = 1.2,
                   add = "boxplot",
                   add.params = list(width = 0.1, size = 1.05)) +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  thesis_theme +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  scale_y_continuous(breaks = seq(0,10,1)) +
  theme(legend.position = "none") +
  scale_x_discrete(labels = c("S3-L78R-GFP +\nS3-L78R-mCherry"))
ggsave("./Plots/NTA_F172/p3L78R.png", width = 5, height = 5.5, dpi = 300)
```

```
##### S5B-L82A
# Narrow df
df5L82A <- df150[df150$Include == "yes",]
df5L82A <- df5L82A[df5L82A$Constructs %in% c("S1F172W+S5L82A"),]
# Melt
df5L82A <- df5L82A %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p5L82A <- ggviolin(df5L82A, x = "Constructs", y = "value",
                   facet.by = "variable",
                   panel.labs = list(variable = c("Cherry","GFP")),
                   fill = "variable",
                   width = 1.2,
                   add = "boxplot",
                   add.params = list(width = 0.1, size = 1.05)) +
  thesis_theme +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  scale_y_continuous(breaks = seq(0,10,1)) +
  theme(legend.position = "none") +
  scale_x_discrete(labels = c("S1-F172W-GFP +\nS5B-L82A-mCherry"))
ggsave("./Plots/NTA_S1S5/p5L82A.png", width = 5, height = 5.5, dpi = 300)
# ------- Homodimers in the NTA
##### STAT1 homodimers
# Exclude cells
df1Homo <- df150[df150$Include == "yes",]
df1Homo <- df1Homo[df1Homo$Constructs %in% c("S1wt+S1wt", "S1F172W+S1F172W",
"S1wt+S1dNLS","S1F172W+S1dNLS","S1wt+S1dNLSF172L", "S1F172W+S1dNLSF172L"),]
# Melt
df1Homo <- df1Homo %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p1Homo <- ggviolin(df1Homo, x = "Constructs", y = "value",
                   facet.by = "variable", ncol=1,
                   panel.labs = list(variable = c("Cherry","GFP")),
                   order = c("S1wt+S1wt",
                             "S1F172W+S1F172W",
                             "S1wt+S1dNLS",
                             "S1F172W+S1dNLS",
                             "S1wt+S1dNLSF172L",
                             "S1F172W+S1dNLSF172L"),
                   fill = "variable",
                   width = 1.2,
                   add = "boxplot",
                   add.params = list(width = 0.1, size = 1.05)) +
  stat_compare_means(comparisons = list(c("S1F172W+S1dNLS",
                                          "S1wt+S1dNLSF172L"),
                                        c("S1F172W+S1dNLS",
                                          "S1F172W+S1dNLSF172L"),
                                        c("S1wt+S1dNLSF172L",
                                          "S1F172W+S1dNLSF172L")),
                     label = "p.signif",
                     label.x = 1.5,
                     label.y = 5.5,
                     step.increase = 0.3) +
  thesis_theme +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  theme(legend.position = "none") +
  scale_y_continuous(breaks = seq(0,10,1), limits = c(0,10)) +
  scale_x_discrete(labels = c("S1wt-GFP +\nS1wt-mCherry",
                              "S1-F172W-GFP +\nS1-F172W-mCherry",
                              "S1wt-GFP +\nS1ΔNLS-mCherry",
                              "S1-F172W-GFP +\nS1ΔNLS-mCherry",
                              "S1wt-GFP +\nS1ΔNLS-F172L-mCherry",
                              "S1-F172W-GFP +\nS1ΔNLS-F172L-mCherry"),
                   guide = guide_axis(n.dodge = 2))
ggsave("./Plots/NTA_F172/p1Homo.png", width = 11.5, height = 11, dpi = 300)
##### STAT3 homodimers
# Narrow df
df3Homo <- df150[df150$Include == "yes",]
df3Homo <- df3Homo[df3Homo$Constructs %in% c("S3wt+S3wt",
                                             "S3L78R+S3L78R",
```

```
                                          "S3wt+S3dNLS",
                                          "S3L78R+S3dNLS",
                                          "S3wt+S3dNLSL78R",
                                          "S3L78R+S3dNLSL78R"),]
# Melt
df3Homo <- df3Homo %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p3Homo <- ggviolin(df3Homo, x = "Constructs", y = "value",
                   facet.by = "variable", ncol=1,
                   panel.labs = list(variable = c("Cherry","GFP")),
                   order = c("S3wt+S3wt",
                             "S3L78R+S3L78R",
                             "S3wt+S3dNLS",
                             "S3L78R+S3dNLS",
                             "S3wt+S3dNLSL78R",
                             "S3L78R+S3dNLSL78R"),
                   fill = "variable",
                   width = 1.2,
                   add = "boxplot",
                   add.params = list(width = 0.1, size = 1.05)) +
  stat_compare_means(comparisons = list(c("S3L78R+S3dNLS",
                                          "S3wt+S3dNLSL78R"),
                                        c("S3L78R+S3dNLS",
                                          "S3L78R+S3dNLSL78R"),
                                        c("S3wt+S3dNLSL78R",
                                          "S3L78R+S3dNLSL78R"),
                                        c("S3L78R+S3L78R",
                                          "S3wt+S3dNLSL78R"),
                                        c("S3wt+S3wt",
                                          "S3wt+S3dNLSL78R")),
                     label = "p.signif",
                     label.x = 1.5,
                     label.y = 4,
                     step.increase = 0.15) +
  thesis_theme +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  theme(legend.position = "none") +
  scale_y_continuous(breaks = seq(0,10,1), limits = c(0,10)) +
  scale_x_discrete(labels = c("S3wt-GFP +\nS3wt-mCherry",
                              "S3-L78R-GFP +\nS3-L78R-mCherry",
                              "S3wt-GFP +\nS3ΔNLS-mCherry",
                              "S3-L78R-GFP +\nS3ΔNLS-mCherry",
                              "S3wt-GFP +\nS3ΔNLS-L78R-mCherry",
                              "S3-L78R-GFP +\nS3ΔNLS-L78R-mCherry"),
                   guide = guide_axis(n.dodge = 2))
ggsave("./Plots/NTA_F172/p3Homo.png", width = 11.5, height = 11, dpi = 300)
# ------- Heterodimers in the NTA
##### S1wt+S3wt vs S1wt+S3dNLS
# Narrow df
df13wtNLS <- df150[df150$Include == "yes",]
df13wtNLS <- df13wtNLS[df13wtNLS$Constructs %in% c("S1wt+S3wt","S1wt+S3dNLS"),]
# Melt
df13wtNLS <- df13wtNLS %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p13wtNLS <- ggviolin(df13wtNLS, x = "Constructs", y = "value",
                     facet.by = "variable",
                     panel.labs = list(variable = c("Cherry","GFP")),
                     fill = "variable",
                     width = 1.2,
                     add = "boxplot",
                     add.params = list(width = 0.1, size = 1.05)) +
  stat_compare_means(comparisons = list(c("S1wt+S3wt","S1wt+S3dNLS")),
                     label = "p.signif",
                     label.x = 1.5) + thesis_theme +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  scale_y_continuous(breaks = seq(0,10,1)) +
  theme(legend.position = "none") +
  scale_x_discrete(labels = c("S1wt-GFP +\nS3wt-mCherry",
                              "S1wt-GFP +\nS3ΔNLS-mCherry"))
```

```
ggsave("./Plots/NTA_S1S3/p13wtNLS.png", width = 9, height = 5.5, dpi = 300)
##### STAT1/STAT3 heterodimers
# Exclude cells
df13Hetero <- df150[df150$Include == "yes",]
df13Hetero <- df13Hetero[df13Hetero$Constructs %in% c("S1wt+S3wt",
                                                      "S1wt+S3dNLS",
                                                      "S1F172W+S3dNLS",
                                                      "S1wt+S3dNLSL78R",
                                                      "S1F172W+S3dNLSL78R"),]
# Melt
df13Hetero <- df13Hetero %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p13Hetero <- ggviolin(df13Hetero, x = "Constructs", y = "value",
                      facet.by = "variable", ncol=1,
                      panel.labs = list(variable = c("Cherry","GFP")),
                      order = c("S1wt+S3wt",
                                "S1wt+S3dNLS",
                                "S1F172W+S3dNLS",
                                "S1wt+S3dNLSL78R",
                                "S1F172W+S3dNLSL78R"),
                      fill = "variable",
                      width = 1.2,
                      add = "boxplot",
                      add.params = list(width = 0.1, size = 1.05)) +
  stat_compare_means(comparisons = list(c("S1F172W+S3dNLS",
                                          "S1wt+S3dNLSL78R"),
                                        c("S1F172W+S3dNLS",
                                          "S1F172W+S3dNLSL78R"),
                                        c("S1wt+S3dNLSL78R",
                                          "S1F172W+S3dNLSL78R")),
                     label = "p.signif",
                     label.x = 1.5,
                     label.y = 4,
                     step.increase = 0.15) +
  thesis_theme +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  theme(legend.position = "none") +
  scale_y_continuous(breaks = seq(0,10,1), limits = c(0,10)) +
  scale_x_discrete(labels = c("S1wt-GFP +\nS3wt-mCherry",
                              "S1wt-GFP +\nS3ΔNLS-mCherry",
                              "S1-F172W-GFP +\nS3ΔNLS-mCherry",
                              "S1wt-GFP +\nS3ΔNLS-L78R-mCherry",
                              "S1-F172W-GFP +\nS3ΔNLS-L78R-mCherry"),
                   guide = guide_axis(n.dodge = 2))
ggsave("./Plots/NTA_S1S3/p13Hetero.png", width = 11.5, height = 11, dpi = 300)
##### STAT1/STAT5B heterodimers
# Narrow df
df15Hetero <- df150[df150$Include == "yes",]
df15Hetero <- df15Hetero[df15Hetero$Constructs %in% c("S1wt+S5wt",
                                                      "S1F172W+S5L82A",
                                                      "S1wt+S5dNLS",
                                                      "S1F172W+S5dNLSL82A"),]
# Melt
df15Hetero <- df15Hetero %>%
  select("Constructs", "GFPdNuc0_90", "CherrydNuc0_90") %>%
  pivot_longer(-Constructs, names_to = "variable", values_to = "value")
# Plot
p15Hetero <- ggviolin(df15Hetero, x = "Constructs", y = "value",
                      facet.by = "variable", ncol=1,
                      panel.labs = list(variable = c("Cherry","GFP")),
                      order = c("S1wt+S5wt",
                                "S1F172W+S5L82A",
                                "S1wt+S5dNLS",
                                "S1F172W+S5dNLSL82A"),
                      fill = "variable",
                      width = 1.2,
                      add = "boxplot",
                      add.params = list(width = 0.1, size = 1.05)) +
  stat_compare_means(comparisons = list(c("S1wt+S5dNLS",
                                          "S1F172W+S5dNLSL82A"),
                                        c("S1F172W+S5L82A",
                                          "S1F172W+S5dNLSL82A"),
                                        c("S1wt+S5wt",
```

```
                                              "S1F172W+S5dNLSL82A")),
                      label = "p.signif",
                      label.x = 1.5,
                      label.y = 6,
                      step.increase = 0.15) +
  thesis_theme +
  scale_fill_discrete(labels = c("Cherry","GFP"), name = "Fusionprotein") +
  xlab("\nConstructs") +
  ylab("Nuclear Score\n") +
  theme(legend.position = "none") +
  scale_y_continuous(breaks = seq(0,10,1), limits = c(0,10)) +
  scale_x_discrete(labels = c("S1wt-GFP +\nS5Bwt-mCherry",
                              "S1-F172W-GFP +\nS5B-L82A-mCherry",
                              "S1wt-GFP +\nS5BΔNLS-mCherry",
                              "S1-F172W-GFP +\nS5BΔNLS-L82A-mCherry"),
                   guide = guide_axis(n.dodge = 2))
ggsave("./Plots/NTA_S1S5/p15Hetero.png", width = 11.5, height = 11, dpi = 300)
```

Nuclear score time course analysis and graphing

```
library(tidyverse)
library(ggplot2)
############# Set plot theme
thesis_theme <- theme_light() + theme(
  axis.text = element_text(size = 14),
  axis.title = element_text(size = 16, face = "bold"),
  legend.text = element_text(size = 14),
  legend.title = element_text(size = 16, face = "bold"),
  plot.title = element_text(size = 18, face = "bold"))
# ------ STAT1 Timecourse
## Read Data
TC1 <- data.frame(read.csv("./Data/timecourse.csv", header = TRUE, fileEncoding = 'UTF-
8-BOM'))
### Statistics for each timepoint
sTC1 <- TC1 %>%
  group_by(Timepoint) %>%
  summarise(
    sd = sd(S1), # Get SD
    len = mean(S1), # Get mean
    nummer = n_distinct(S1)) # Get No
sTC1$se <- sTC1$sd/sqrt(sTC1$nummer) # Get SE
sTC1$roundslen <- sTC1$len
sTC1$roundsSE <- sTC1$se
sTC1 <- sTC1 %>% mutate_at(vars(roundslen, roundsSE), funs(round(.,2)))
## Plot timecourse
pTC1 <- ggplot(sTC1, aes(x=Timepoint, y=len)) +
  geom_errorbar(aes(ymin=len-se, ymax=len+se),
                width=3, color = "#5E5EFF") +
  geom_line(size=0.8, color = "#A1C5FF") +
  geom_point(size=4, color = "#5E5EFF") +
  geom_text(aes(label = paste(roundslen, "\ub1", roundsSE)),
            size = 4.5,
            vjust = -3,
            fontface = 2) +
  ylab("Nuclear Score\n") + xlab("\nTimepoint [min]") +
  scale_x_continuous(breaks = seq(0,120,15)) +
  ylim(1,4.5) +
  thesis_theme
ggsave("./Plots/Timecourse/pTC1.png", width = 9, height = 5.5, dpi = 300)
# ------ STAT3 Timecourse
## Read Data
TC3 <- data.frame(read.csv("./Data/timecourse3.csv", header = TRUE, fileEncoding =
'UTF-8-BOM'))
### Statistics for each timepoint
sTC3 <- TC3 %>%
  group_by(Timepoint) %>%
  summarise(
    sd = sd(S3), # Get SD
    len = mean(S3), # Get mean
    nummer = n_distinct(S3)) # Get No.
sTC3$se <- sTC3$sd/sqrt(sTC3$nummer) # Get SE
sTC3$roundslen <- sTC3$len
sTC3$roundsSE <- sTC3$se
sTC3 <- sTC3 %>% mutate_at(vars(roundslen, roundsSE), funs(round(.,2)))
### Plot timecourse
pTC3 <- ggplot(sTC3, aes(x=Timepoint, y=len)) +
  geom_errorbar(aes(ymin=len-se, ymax=len+se),
```

```
                width=3, color = "#009688") +
  geom_line(size=0.8, color = "#80D5CC") +
  geom_point(size=4, color = "#009688") +
  geom_text(aes(label = paste(roundslen, "\ub1", roundsSE)),
            size = 4,
            vjust = -3,
            fontface = 2) +
  ylab("Nuclear Score\n") + xlab("\nTimepoint [min]") +
  scale_x_continuous(breaks = seq(0,120,15)) +
  ylim(1,4.5) +
  thesis_theme
ggsave("./Plots/Timecourse/pTC3.png", width = 9, height = 5.5, dpi = 300)
# ------ STAT5 Timecourse
## Read Data
TC5 <- data.frame(read.csv("./Data/timecourse5.csv", header = TRUE, fileEncoding =
'UTF-8-BOM'))
### Statistics for each timepoint
sTC5 <- TC5 %>%
  group_by(Timepoint) %>%
  summarise(
    sd = sd(S5B), # Get SD
    len = mean(S5B), # Get mean
    nummer = n_distinct(S5B)) # Get No.
sTC5$se <- sTC5$sd/sqrt(sTC5$nummer) # Get SE
sTC5$roundslen <- sTC5$len
sTC5$roundsSE <- sTC5$se
sTC5 <- sTC5 %>% mutate_at(vars(roundslen, roundsSE), funs(round(.,2)))
### Plot timecourse
pTC5 <- ggplot(sTC5, aes(x=Timepoint, y=len)) +
  geom_errorbar(aes(ymin=len-se, ymax=len+se),
                width=3, color = "#96006E") +
  geom_line(size=0.8, color = "#E153BB") +
  geom_point(size=4, color = "#96006E") +
  geom_text(aes(label = paste(roundslen, "\ub1", roundsSE)),
            size = 4.5,
            vjust = -3,
            fontface = 2) +
  ylab("Nuclear Score\n") + xlab("\nTimepoint [min]") +
  scale_x_continuous(breaks = seq(0,120,15)) +
  ylim(1,4.5) +
  thesis_theme
ggsave("./Plots/Timecourse/pTC5.png", width = 9, height = 5.5, dpi = 300)
```

Nuclear Translocation assay correlation analysis and graphing.

```
library(tidyverse)
library(ggplot2)
library(ggpubr)
############# Set plot theme
thesis_theme <- theme_light() + theme(
  axis.text = element_text(size = 14),
  axis.title = element_text(size = 16, face = "bold"),
  legend.text = element_text(size = 14),
  legend.title = element_text(size = 16, face = "bold"),
  plot.title = element_text(size = 18, face = "bold"))
# ------- Read data and expand
df150 <- data.frame(read.csv("./Data/analysisnewimp150.csv",header = TRUE, fileEncoding
= 'UTF-8-BOM'))
# ------- STAT1wt vs STAT1dNLS
# Narrow df
df1nls <- df150[df150$Constructs %in% "S1wt+S1dNLS",]
# Exclude cells that were not analysed
df1nls <- df1nls[df1nls$GFPdNuc0_90 > 0,]
# Correlation and test
cor.test(df1nls$PercentExcess, df1nls$GFPdNuc0_90, method = "pearson")
# Linear regression
reg_df1nls <- lm(data=df1nls, GFPdNuc0_90 ~ PercentExcess)
summary(reg_df1nls)
confint(reg_df1nls, "PercentExcess", level = 0.95)
# Plot correlation
p1nls <- ggscatter(df1nls,
                   x = "PercentExcess",
                   y = "GFPdNuc0_90",
                   add = "reg.line",
                   add.params = list(color = "black"),
                   conf.int = TRUE,
                   cor.coef = TRUE,
```

```
                    cor.method = "pearson",
                    cor.coeff.args = list(label.x = 280,
                                          label.sep = "\n",
                                          size = 5),
                    color = "#5E5EFF") +
  thesis_theme +
  xlab("\nSTAT1ΔNLS excess [%]") +
  ylab("STAT1wt nuclear score\n")
ggsave("./Plots/NTA Corr/p1nls.png", width = 9, height = 5.5, dpi = 300)
# S1wt-GFP nuclear score when S1dNLS excess is lower than 1.5x
df1nls2 <- df1nls[df1nls$PercentExcess < 150,]
s1nls2 <- df1nls2 %>%
  group_by(Constructs) %>%
  summarise(
    n_cells = n distinct(GFPdNuc0 90),
    GFP_Mean = mean(GFPdNuc0_90),
    GFP_SD = sd(GFPdNuc0_90),
    Cherry_Mean = mean(CherrydNuc0_90),
    Cherry_SD = sd(CherrydNuc0_90))
s1nls2$GFP_SE <- s1nls2$GFP_SD/sqrt(s1nls2$n_cells)
s1nls2$Cherry_SE <- s1nls2$Cherry_SD/sqrt(s1nls2$n_cells)
s1nls2 <- s1nls2 %>%
  mutate_at(vars(-n_cells,-Constructs), funs(round(.,2))) %>%
  select("Constructs", sort(colnames(.)))
# ------- STAT3wt vs STAT3dNLS
# Narrow df
df3nls <- df150[df150$Constructs %in% "S3wt+S3dNLS",]
df3nls <- df3nls[df3nls$GFPdNuc0_90 <= 5,]
df3nls <- df3nls[df3nls$CherrydNuc0_90 <= 3,]
df3nls <- df3nls[df3nls$GFPdNuc0_90 > 0,]
# Correlation S3dNLS excess and S3wt nuclear score
cor.test(df3nls$PercentExcess, df3nls$GFPdNuc0_90, method = "pearson")
# Plot correlation
p3nls <- ggscatter(df3nls,
                   x = "PercentExcess",
                   y = "GFPdNuc0_90",
                   add = "reg.line",
                   add.params = list(color = "black"),
                   conf.int = TRUE,
                   cor.coef = TRUE,
                   cor.method = "pearson",
                   cor.coeff.args = list(label.x = 850, label.sep = "\n", size = 5),
                   color = "#009688") +
  thesis_theme + xlab("\nSTAT3ΔNLS excess [%]") + ylab("STAT3wt nuclear score\n")
ggsave("./Plots/NTA_Corr/p3nls.png", width = 9, height = 5.5, dpi = 300)
# Regression descriptive statistics
reg_df3nls <- lm(data=df3nls, GFPdNuc0_90 ~ PercentExcess)
summary(reg_df3nls)
confint(reg_df3nls, "PercentExcess", level = 0.95)
# S3wt-GFP nuclear score when S3dNLS excess is lower than 1.5x
df3nls2 <- df3nls[df3nls$PercentExcess < 150,]
s3nls2 <- df3nls2 %>%
  group_by(Constructs) %>%
  summarise(
    n_cells = n_distinct(GFPdNuc0_90),
    GFP_Mean = mean(GFPdNuc0_90),
    GFP_SD = sd(GFPdNuc0_90),
    Cherry_Mean = mean(CherrydNuc0_90),
    Cherry_SD = sd(CherrydNuc0_90))
s3nls2$GFP_SE <- s3nls2$GFP_SD/sqrt(s3nls2$n_cells)
s3nls2$Cherry_SE <- s3nls2$Cherry_SD/sqrt(s3nls2$n_cells)
s3nls2 <- s3nls2 %>%
  mutate_at(vars(-n_cells,-Constructs), funs(round(.,2))) %>%
  select("Constructs", sort(colnames(.)))
# ------- STAT5wt vs STAT5dNLS
# Narrow df
df5nls <- df150[df150$Constructs %in% "S5BdNLS+S5Bwt",]
# Correlation S5BdNLS excess and S5Bwt nuclear score
cor.test(df5nls$PercentExcess, df5nls$GFPdNuc0_90, method = "pearson")
# Plot correlation
p5nls <- ggscatter(df5nls,
                   x = "PercentExcess",
                   y = "GFPdNuc0_90",
                   add = "reg.line",
                   add.params = list(color = "black"),
                   conf.int = TRUE,
                   cor.coef = TRUE,
```

```
                  cor.method = "pearson",
                  cor.coeff.args = list(label.x = 450, label.sep = "\n", size = 5),
                  color = "#96006E") +
  thesis_theme + xlab("\nSTAT5BΔNLS excess [%]") + ylab("STAT5Bwt nuclear score\n")
ggsave("./Plots/NTA_Corr/p5nls.png", width = 9, height = 5.5, dpi = 300)
```

GFP-Cherry FRET graphing and analysis.

```
library(tidyverse)
library(ggplot2)
############# Set plot theme
thesis_theme <- theme_light() + theme(
  axis.text = element_text(size = 14),
  axis.title = element_text(size = 16, face = "bold"),
  legend.text = element_text(size = 14),
  legend.title = element_text(size = 16, face = "bold"),
  plot.title = element_text(size = 18, face = "bold"))
# ------- New rGFP-Cherry
# Read data
GCNewFRET <- data.frame(read.csv("./Data/GFPCherryNew.csv",
                                 header = TRUE,
                                 fileEncoding = 'UTF-8-BOM',
                                 sep = "\t"))
# Plot GFP fluorescence over time
GCplotGFP <- ggplot(GCNewFRET) +
  geom_line(aes(x = Time, y = GFP, group = No, ), color="#7FC87F") +
  geom_point(aes(x = Time, y = GFP, group = No), color="#095E09") +
  xlab("\nTime [s]") + ylab("Intensity GFP [A.U.]\n") + thesis_theme
ggsave("./Plots/FRET/GCplotGFP.png", width =5.5, height = 5.5, dpi = 300)
# Plot mCherry fluorescence over time
GCplotCherry <- ggplot(GCNewFRET)+
  geom_line(aes(x = Time, y = Cherry, group = No, ), color="#EE6666") +
  geom_point(aes(x = Time, y = Cherry, group = No), color="#8C2626") +
  xlab("\nTime [s]") + ylab("Intensity Cherry [A.U.]\n") + thesis_theme
ggsave("./Plots/FRET/GCplotCherry.png", width =5.5, height = 5.5, dpi = 300)
### Calculate FRET using the images just before and after bleaching
# Select last image before bleaching
sBGC1 <- GCNewFRET[c(seq(5,300,10)),]
# Select first image after bleaching
sAGC1 <- GCNewFRET[c(seq(6,300,10)),]
# Merge data frames
sGC1 <- merge(sBGC1, sAGC1, by = "No")
# Delete unnecessary columns
sGC1 <- subset(sGC1, select = -c(Time.x, Time.y, Bleach.x, Bleach.y))
# Calculate FRET efficiency
sGC1$EFRET <- ((1-(sGC1$GFP.x/sGC1$GFP.y))*100)
# Calculate distance between fluorophores
sGC1$Radius <- (5.28*(((1/(sGC1$EFRET/100))-1)**(1/6)))
# Round
sGC1 <- sGC1 %>% mutate_at(vars(-EFRET, -Radius), funs(round(.,0))) %>%
  mutate_at(vars(EFRET, Radius), funs(round(.,2)))
# Rename columns
sGC1 <- sGC1 %>% rename(Cell_No = No,
                        GFP_before_bleach = GFP.x,
                        Cherry_before_bleach = Cherry.x,
                        GFP_after_bleach = GFP.y,
                        Cherry_after_bleach = Cherry.y)
write.csv(sGC1, "./Descriptive_statistics/sGC1.csv")
# ------- Old fGFP-Cherry
# Read data
GCOldFRET <- data.frame(read.csv("./Data/GFPCherryOld.csv",
                                 header = TRUE,
                                 fileEncoding = 'UTF-8-BOM'))
# Plot GFP fluorescence over time
fGCplotGFP <- ggplot(GCOldFRET) +
  geom_line(aes(x = Time, y = GFP, group = No, ), color="#7FC87F") +
  geom_point(aes(x = Time, y = GFP, group = No), color="#095E09") +
  xlab("\nTime [s]") + ylab("Intensity GFP [A.U.]\n") + thesis_theme
ggsave("./Plots/FRET/fGCplotGFP.png", width =5.5, height = 5.5, dpi = 300)
# Plot mCherry fluorescence over time
fGCplotCherry <- ggplot(GCOldFRET)+
  geom_line(aes(x = Time, y = Cherry, group = No, ), color="#EE6666") +
  geom_point(aes(x = Time, y = Cherry, group = No), color="#8C2626") +
  xlab("\nTime [s]") + ylab("Intensity Cherry [A.U.]\n") + thesis_theme
ggsave("./Plots/FRET/fGCplotCherry.png", width =5.5, height = 5.5, dpi = 300)
### Calculate FRET using the images just before and after bleaching
# Select last image before bleaching
```

```
sBfGC1 <- GCOldFRET[c(seq(5,400,10)),]
# Select first image after bleaching
sAfGC1 <- GCOldFRET[c(seq(6,400,10)),]
# Merge data frames
sfGC1 <- merge(sBfGC1, sAfGC1, by = "No")
# Delete columns
sfGC1 <- subset(sfGC1, select = -c(Time.x, Time.y, Bleach.x, Bleach.y))
# Calculate FRET efficiency
sfGC1$EFRET <- ((1-(sfGC1$GFP.x/sfGC1$GFP.y))*100)
# Calculate distance between fluorophores
sfGC1$Radius <- (5.28*(((1/(sfGC1$EFRET/100))-1)**(1/6)))
# Round
sfGC1 <- sfGC1 %>% mutate_at(vars(-EFRET, -Radius), funs(round(.,0))) %>%
  mutate_at(vars(EFRET, Radius), funs(round(.,2)))
# Rename columns
sfGC1 <- sfGC1 %>% rename(Cell_No = No,
                          GFP_before_bleach = GFP.x,
                          Cherry_before_bleach = Cherry.x,
                          GFP_after_bleach = GFP.y,
                          Cherry_after_bleach = Cherry.y)
write.csv(sfGC1, "./Descriptive_statistics/sfGC1.csv")
```

GFP-Cherry linear regression analysis.

```
library(tidyverse)
library(ggplot2)
library(janitor)
library(broom)
library(gsubfn)
############# Set plot theme
thesis_theme <- theme_light() + theme(
  axis.text = element_text(size = 14),
  axis.title = element_text(size = 16, face = "bold"),
  legend.text = element_text(size = 14),
  legend.title = element_text(size = 16, face = "bold"),
  plot.title = element_text(size = 18, face = "bold"))
# ------- Read dataa
df150 <- data.frame(read.csv("./Data/analysisnewimp150.csv",
                             header = TRUE,
                             fileEncoding = 'UTF-8-BOM'))
# Filter for GFP-Cherry only
dG <- df150[df150$GFPCherry == "G",]
# ------- Rigid linker GFP-Cherry
##### Plot all rigid linker GFP-Cherry experiments
# Comparison vector
valids <- c("G1r","G3r","G5r")
# Filter for G1r and G3r
dGr <- dG[dG$Constructs %in% valids,]
# Plot
plotdGr <- ggplot(dGr) +
  geom_point(aes(x=GFPCorrIntDenW, y=CherryCorrIntDenW, color=Constructs)) +
  xlab("\nTotal intensity GFP [A.U.]") +
  ylab("Total intensity Cherry [A.U.]\n") +
  xlim(0,5e+08) +
  ylim(0,5e+08) +
  thesis_theme +
  scale_color_discrete(labels = c("S1 settings",
                                  "S3 settings",
                                  "S5 settings"))
ggsave("./Plots/GFPCherry/plotGr.png", width = 9, height = 5.5, dpi = 300)
##### Plot rigid GFP-Cherry imaged with STAT1 settings
# Filter for STAT1 settings
dG1r <- dG[dG$Constructs == "G1r",]
# Extract imaging dates
dG1r$Dates <-strapplyc(dG1r$Transfection, "[0-9/]{6,}", simplify = TRUE)
# Do linear regression and tidy
tG1r<- dG1r %>%
  group_by(Dates) %>%
  do(tidy(lm(CherryCorrIntDenW ~ GFPCorrIntDenW, data = .)))
# Remove statistics for intersect
tG1r <- tG1r %>%
  filter(term == "GFPCorrIntDenW")
# Do linear regressiona and glance
gG1r <- dG1r %>%
  group_by(Dates) %>%
  do(glance(lm(CherryCorrIntDenW ~ GFPCorrIntDenW, data = .)))
# Merge tidy and glance
```

```
mG1r <- merge(tG1r,gG1r, by = "Dates")
# Remove unnecessary columns
mG1r <- subset(mG1r, select = c(Dates,
                                estimate,
                                std.error,
                                p.value.x,
                                r.squared,
                                adj.r.squared))
# Rename columns
mG1r <- mG1r %>% rename(Dates = Dates,
                        Slope = estimate,
                        Slope_SEM = std.error,
                        p.value = p.value.x)
write.csv(mG1r,"./Descriptive_statistics/G1rRegression.csv")
# Plot STAT1 settings
plotdG1r <- ggplot(dG1r, aes(x=GFPCorrIntDenW,
                             y=CherryCorrIntDenW,
                             color=Dates)) +
  geom_point() +
  xlab("\nTotal intensity GFP [A.U.]") +
  ylab("Total intensity Cherry [A.U.]\n") +
  xlim(0,5e+08) +
  ylim(0,5e+08) +
  stat_smooth(method = "lm", level = FALSE) + thesis_theme
ggsave("./Plots/GFPCherry/plotG1r.png", width = 9, height = 5.5, dpi = 300)
##### Plot rigid GFP-Cherry imaged with STAT3 settings
# Filter for STAT3 settings
dG3r <- dG[dG$Constructs == "G3r",]
# Extract imaging dates
dG3r$Dates <-strapplyc(dG3r$Transfection, "[0-9/]{6,}", simplify = TRUE)
# Do linear regression and tidy
tG3r<- dG3r %>%
  group_by(Dates) %>%
  do(tidy(lm(CherryCorrIntDenW ~ GFPCorrIntDenW, data = .)))
# Remove statistics for intersect
tG3r <- tG3r %>%
  filter(term == "GFPCorrIntDenW")
# Do linear regressiona and glance
gG3r <- dG3r %>%
  group_by(Dates) %>%
  do(glance(lm(CherryCorrIntDenW ~ GFPCorrIntDenW, data = .)))
# Merge tidy and glance
mG3r <- merge(tG3r,gG3r, by = "Dates")
# Remove unnecessary columns
mG3r <- subset(mG3r, select = c(Dates,
                                estimate,
                                std.error,
                                p.value.x,
                                r.squared,
                                adj.r.squared))
# Rename columns
mG3r <- mG3r %>% rename(Dates = Dates,
                        Slope = estimate,
                        Slope_SEM = std.error,
                        p.value = p.value.x)
write.csv(mG3r,"./Descriptive_statistics/G3rRegression.csv")
# Plot STAT3 settings
plotdG3r <- ggplot(dG3r, aes(x=GFPCorrIntDenW,
                             y=CherryCorrIntDenW,
                             color=Dates)) +
  geom_point() +
  xlab("\nTotal intensity GFP [A.U.]") +
  ylab("Total intensity Cherry [A.U.]\n") +
  xlim(0,5e+08) +
  ylim(0,5e+08) +
  stat_smooth(method = "lm", level = FALSE) + thesis_theme
ggsave("./Plots/GFPCherry/plotG3r.png", width = 9, height = 5.5, dpi = 300)
##### Plot rigid GFP-Cherry imaged with STAT5 settings
# Filter for STAT5 settings
dG5r <- dG[dG$Constructs == "G5r",]
# Extract imaging dates
dG5r$Dates <-strapplyc(dG5r$Transfection, "[0-9/]{6,}", simplify = TRUE)
# Do linear regression and tidy
tG5r<- dG5r %>%
  group_by(Dates) %>%
  do(tidy(lm(CherryCorrIntDenW ~ GFPCorrIntDenW, data = .)))
# Remove statistics for intersect
```

```
tG5r <- tG5r %>%
  filter(term == "GFPCorrIntDenW")
# Do linear regressiona and glance
gG5r <- dG5r %>%
  group_by(Dates) %>%
  do(glance(lm(CherryCorrIntDenW ~ GFPCorrIntDenW, data = .)))
# Merge tidy and glance
mG5r <- merge(tG5r,gG5r, by = "Dates")
# Remove unnecessary columns
mG5r <- subset(mG5r, select = c(Dates,
                                estimate,
                                std.error,
                                p.value.x,
                                r.squared,
                                adj.r.squared))
# Rename columns
mG5r <- mG5r %>% rename(Dates = Dates,
                        Slope = estimate,
                        Slope_SEM = std.error,
                        p.value = p.value.x)
write.csv(mG5r,"./Descriptive_statistics/G5rRegression.csv")
# Plot STAT5 settings
plotdG5r <- ggplot(dG5r, aes(x=GFPCorrIntDenW,
                             y=CherryCorrIntDenW,
                             color=Dates)) +
  geom_point() +
  xlab("\nTotal intensity GFP [A.U.]") +
  ylab("Total intensity Cherry [A.U.]\n") +
  xlim(0,5e+08) +
  ylim(0,5e+08) +
  stat_smooth(method = "lm", level = FALSE) + thesis_theme
ggsave("./Plots/GFPCherry/plotG5r.png", width = 9, height = 5.5, dpi = 300)
# ------- Flexible linker GFP-Cherry
##### Plot all flexible linker GFP-Cherry experiments
# Comparison vector
valids <- c("G1f","G3f")
# Filter for G1f and G3f
dGf <- dG[dG$Constructs %in% valids,]
# Plot
plotdGf <- ggplot(dGf) +
  geom_point(aes(x=GFPCorrIntDenW, y=CherryCorrIntDenW, color=Constructs)) +
  xlab("\nTotal intensity GFP [A.U.]") +
  ylab("Total intensity Cherry [A.U.]\n") +
  xlim(0,5e+08) +
  ylim(0,5e+08) +
  thesis_theme +
  scale_color_discrete(labels = c("S1 settings", "S3 settings", "S5 settings"))
ggsave("./Plots/GFPCherry/plotGf.png", width = 9, height = 5.5, dpi = 300)
##### Plot flexible GFP-Cherry imaged with STAT1 settings
# Filter for STAT1 settings
dG1f <- dG[dG$Constructs == "G1f",]
# Extract imaging dates
dG1f$Dates <-strapplyc(dG1f$Transfection, "[0-9/]{6,}", simplify = TRUE)
# Do linear regression and tidy
tG1f<- dG1f %>%
  group_by(Dates) %>%
  do(tidy(lm(CherryCorrIntDenW ~ GFPCorrIntDenW, data = .)))
# Remove statistics for intersect
tG1f <- tG1f %>%
  filter(term == "GFPCorrIntDenW")
# Do linear regressiona and glance
gG1f <- dG1f %>%
  group_by(Dates) %>%
  do(glance(lm(CherryCorrIntDenW ~ GFPCorrIntDenW, data = .)))
# Merge tidy and glance
mG1f <- merge(tG1f,gG1f, by = "Dates")
# Remove columns
mG1f <- subset(mG1f, select = c(Dates,
                                estimate,
                                std.error,
                                p.value.x,
                                r.squared,
                                adj.r.squared))
# Rename columns
mG1f <- mG1f %>% rename(Dates = Dates,
                        Slope = estimate,
                        Slope_Std.Error = std.error,
```

```
                          p.value = p.value.x)
write.csv(mG1f,"./Descriptive_statistics/G1fRegression.csv")
# Plot STAT1 settings
plotdG1f <- ggplot(dG1f, aes(x=GFPCorrIntDenW,
                             y=CherryCorrIntDenW,
                             color=Dates)) +
  geom_point() +
  xlab("\nTotal intensity GFP [A.U.]") +
  ylab("Total intensity Cherry [A.U.]\n") +
  xlim(0,5e+08) +
  ylim(0,5e+08) +
  stat_smooth(method = "lm", level = FALSE) + thesis_theme
ggsave("./Plots/GFPCherry/plotG1f.png", width = 9, height = 5.5, dpi = 300)
##### Plot flexible GFP-Cherry imaged with STAT3 settings
# Filter for STAT3 settings
dG3f <- dG[dG$Constructs == "G3f",]
# Extract imaging dates
dG3f$Dates <-strapplyc(dG3f$Transfection, "[0-9/]{6,}", simplify = TRUE)
# Do linear regression and tidy
tG3f<- dG3f %>%
  group_by(Dates) %>%
  do(tidy(lm(CherryCorrIntDenW ~ GFPCorrIntDenW, data = .)))
# Remove statistics for intersect
tG3f <- tG3f %>%
  filter(term == "GFPCorrIntDenW")
# Do linear regression and glance
gG3f <- dG3f %>%
  group_by(Dates) %>%
  do(glance(lm(CherryCorrIntDenW ~ GFPCorrIntDenW, data = .)))
# Merge tidy and glance
mG3f <- merge(tG3f,gG3f, by = "Dates")
# Remove columns
mG3f <- subset(mG3f, select = c(Dates,
                                estimate,
                                std.error,
                                p.value.x,
                                r.squared,
                                adj.r.squared))
# Rename columns
mG3f <- mG3f %>% rename(Dates = Dates,
                        Slope = estimate,
                        Slope_Std.Error = std.error,
                        p.value = p.value.x)
write.csv(mG3f,"./Descriptive_statistics/G3fRegression.csv")
# Plot STAT3 settings
plotdG3f <- ggplot(dG3f, aes(x=GFPCorrIntDenW,
                             y=CherryCorrIntDenW,
                             color=Dates)) +
  geom_point() +
  xlab("\nTotal intensity GFP [A.U.]") +
  ylab("Total intensity Cherry [A.U.]\n") +
  xlim(0,5e+08) +
  ylim(0,5e+08) +
  stat_smooth(method = "lm", level = FALSE) + thesis_theme
ggsave("./Plots/GFPCherry/plotG3f.png", width = 9, height = 5.5, dpi = 300)
```

Modified Nuclear Translocation Assay investigations into latent STAT dimers.

```
library(tidyverse)
library(ggplot2)
library(ggpubr)
############## Set plot theme
thesis_theme <- theme_light() + theme(
  axis.text = element_text(size = 14),
  axis.title = element_text(size = 16, face = "bold"),
  legend.text = element_text(size = 14),
  legend.title = element_text(size = 16, face = "bold"),
  plot.title = element_text(size = 18, face = "bold"))
# ------- Read data
dfLatent <- data.frame(read.csv("./Data/LatentDimers.csv",
                                header = TRUE,
                                fileEncoding = 'UTF-8-BOM'))
# ------- S3wt+S3-NES vs S3-L78R+S3-NES
# Narrow df
dfS3wt_L78R <- dfLatent[dfLatent$Combo %in% c("S3_NES_wt","S3_NES_L78R"),]
# Plot
pS3wt_L78R <- ggbarplot(dfS3wt_L78R, x = "Combo", y = "Corr",
```

```
                      add = c("mean", "jitter"),
                      fill = "#80D5CC",
                      add.params = list(color = "#009688", size = 2),
                      size = 1.5,
                      ylab = "Pearson r",
                      xlab = "",
                      label = TRUE,
                      lab.vjust = 2,
                      lab.nb.digits = 2,
                      lab.size = 6,
                      lab.col = "white")

pS3wt_L78R <- pS3wt_L78R + stat_compare_means(method = "kruskal",
                                              label = "p.signif",
                                              label.y = 1,
                                              label.x = 1.5,
                                              size = 10) +
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15)),
                                    limits = c(-1,1)) +
  scale_x_discrete(labels = c("S3-NES-GFP +\nS3wt-mCherry",
                              "S3-NES-GFP +\nS3-L78R-mCherry"))
ggsave("./Plots/Val/pS3wt_L78R.png", width = 5.5, height = 5.5, dpi = 300)
# Get statistics
sS3wt_L78R <- dfS3wt_L78R %>%
  group_by(Combo) %>%
  summarise(
    sd = sd(Corr), # Get SD
    av = mean(Corr), # Get mean
    nummer = n()) # Get No.
sS3wt_L78R$se <- sS3wt_L78R$sd/sqrt(sS3wt_L78R$nummer) # Get SE
write.csv(sS3wt_L78R,"./Descriptive_statistics/Val/Validation.csv")
# ------- S1wt+S1-NLS vs S1-F77A+S1-NLS
# Narrow df
dfS1wt_F77A <- dfLatent[dfLatent$Combo %in% c("S1_NLS_wt","S1_NLS_F77A"),]
# Plot
pS1wt_F77A <- ggbarplot(dfS1wt_F77A, x = "Combo", y = "Corr",
                        order = c("S1_NLS_wt","S1_NLS_F77A"),
                        add = c("mean", "jitter"),
                        fill = "#A1C5FF",
                        add.params = list(color = "#5E5EFF", size = 2),
                        size = 1.5,
                        ylab = "Pearson r",
                        xlab = "",
                        label = TRUE,
                        lab.vjust = 2,
                        lab.nb.digits = 2,
                        lab.size = 6,
                        lab.col = "white")
pS1wt_F77A <- pS1wt_F77A + stat_compare_means(method = "kruskal",
                                              label = "p.signif",
                                              label.y = 1,
                                              label.x = 1.5,
                                              size = 10) +
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15)),
                                    limits = c(-1,1)) +
  scale_x_discrete(labels = c("S1-NLS-GFP +\nS1wt-mCherry",
                              "S1-NLS-GFP +\nS1-F77A-mCherry"))
ggsave("./Plots/F77A/pS1wt_F77A.png", width = 5.5, height = 5.5, dpi = 300)
# Get statistics
sS1wt_F77A <- dfS1wt_F77A %>%
  group_by(Combo) %>%
  summarise(
    sd = sd(Corr), # Get SD
    av = mean(Corr), # Get mean
    nummer = n()) # Get No.
sS1wt_F77A$se <- sS1wt_F77A$sd/sqrt(sS1wt_F77A$nummer) # Get SE
write.csv(sS1wt_F77A,"./Descriptive_statistics/Val/S1wt_F77A.csv")
# ------- S5Bwt+S5B-NES vs S5B-L82A+S5B-NES
## Narrow df
dfS5wt_L82A <- dfLatent[dfLatent$Combo %in% c("S5B_NES_wt","S5B_NES_L82A"),]
## Plot
pS5wt_L82A <- ggbarplot(dfS5wt_L82A, x = "Combo", y = "Corr",
                        order = c("S5B_NES_wt","S5B_NES_L82A"),
                        add = c("mean", "jitter"),
                        fill = "#E153BB",
                        add.params = list(color = "#96006E", size = 2),
                        size = 1.5,
```

```
                        ylab = "Pearson r",
                        xlab = "",
                        label = TRUE,
                        lab.vjust = 2,
                        lab.nb.digits = 2,
                        lab.size = 6,
                        lab.col = "white")
pS5wt_L82A <- pS5wt_L82A + stat_compare_means(method = "kruskal",
                                              label = "p.signif",
                                              label.y = 1,
                                              label.x = 1.5,
                                              size = 10)+
  thesis_theme + scale_y_continuous(expand = expansion(mult = c(0.05, 0.15)),
                                    limits = c(-1,1)) +
  scale x discrete(labels = c("S5B-NES-GFP +\nS5Bwt-mCherry",
                              "S5B-NES-GFP +\nS5B-L82A-mCherry"))
ggsave("./Plots/F77A/pS5wt_L82A.png", width = 5.5, height = 5.5, dpi = 300)
# Get statistics
sS5wt_L82A <- dfS5wt_L82A %>%
  group_by(Combo) %>%
  summarise(
    sd = sd(Corr), # Get SD
    av = mean(Corr), # Get mean
    nummer = n()) # Get No.
sS5wt_L82A$se <- sS5wt_L82A$sd/sqrt(sS5wt_L82A$nummer) # Get SE
write.csv(sS5wt_L82A,"./Descriptive_statistics/Val/S5wt_L82A.csv")
```